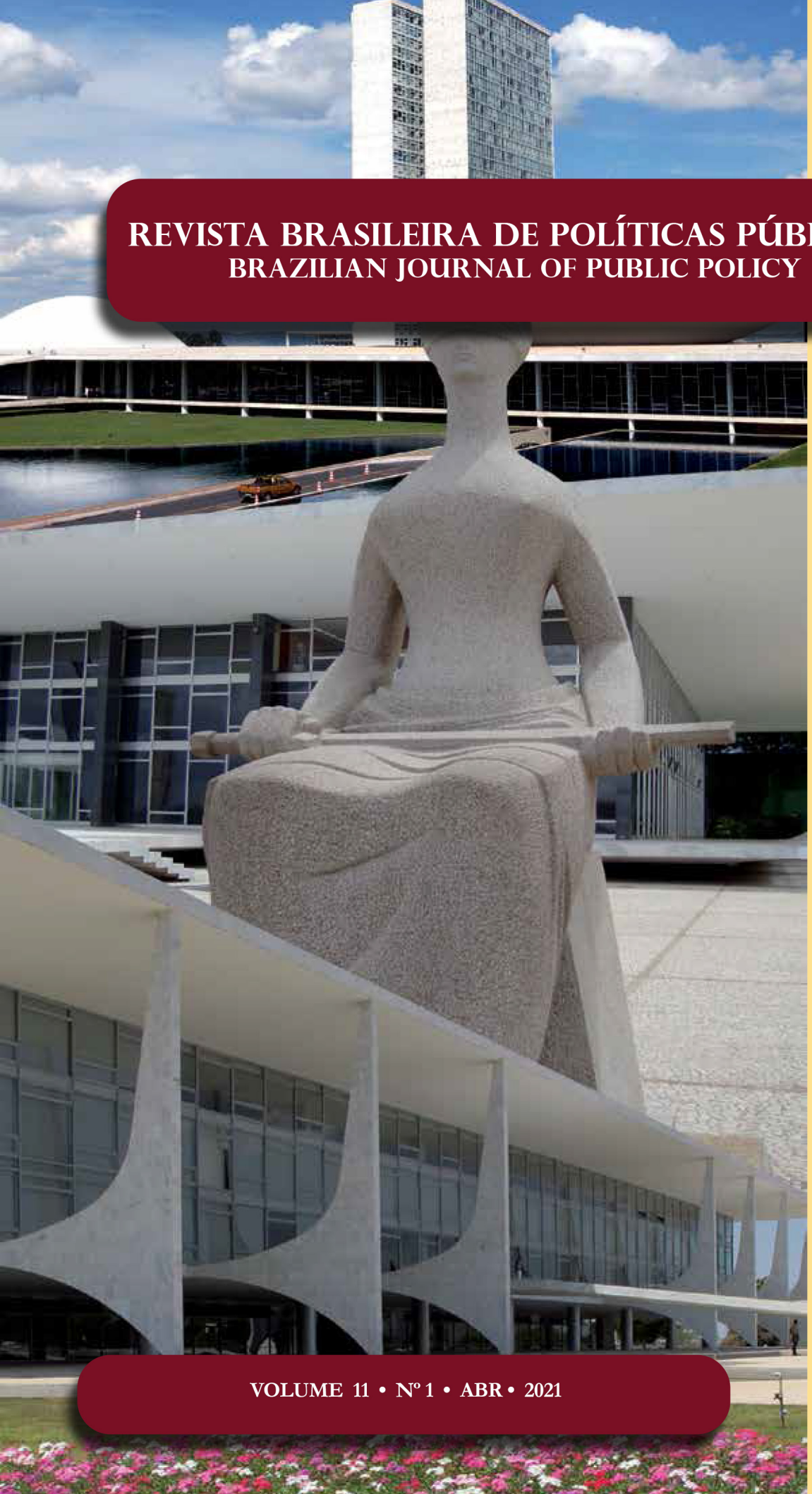

UniCEUB
ISSN 2236-1677

REVISTA BRASILEIRA DE POLÍTICAS PÚBLICAS
BRAZILIAN JOURNAL OF PUBLIC POLICY

"Let the algorithm decide": is human dignity at stake?

"Deixe o algoritmo decidir": dignidade humana em risco?

Marcela Mattiuzzo

VOLUME 11 • Nº 1 • ABR • 2021

Sumário

Parte 1: Políticas Públicas 17

1. Políticas públicas: aspectos gerais 18

Um modelo político de implementação para as políticas públicas: os papéis do direito e dos juristas 20
William H. Clune III

Evaluación de las obras públicas en gobiernos locales en México: desafíos de las políticas públicas de participación ciudadana 83
Louis Valentin Mballa e Arturo Bermúdez Lara

Paternalismo libertário e políticas públicas: intervenção e transparência 105
Marcia Carla Pereira Ribeiro e Victor Hugo Domingues

2. Políticas públicas e COVID-19 121

Limites e possibilidades para o uso secundário de dados pessoais no poder público: lições da pandemia 123
Miriam Wimmer

Eficiência das políticas de inovação nos setores industriais brasileiros: sugestões para a crise da COVID-19 144
Michelle Márcia Viana Martins e Chrystian Soares Mendes

Compliance em tempos de calamidade pública: análise sobre a flexibilização da transparência de dados e informações durante o enfrentamento da COVID-19 no Brasil 169
Luciana Cristina da Conceição Lima, Alcindo Fernandes Gonçalves, Fernando Cardoso Fernandes Rei e Cláudio Benevenuto de Campos Lima

3. Políticas públicas e accountability 188

Accountability e desenho institucional: um "ponto cego" no direito público brasileiro 190
Danielle Hanna Rached

Estratégias regulatórias de combate à corrupção 211
Eduardo Jordão e Luiz Carlos Penner Rodrigues da Costa

O controle e a avaliação pelo Tribunal de Contas da União das políticas públicas implementadas por desonerações tributárias no Brasil 243
Vinicius Garcia e Carlos Araújo Leonetti

4. Políticas públicas em matéria de saúde 266

A livre opção pela cesariana: um "nudge às avessas" 268
Bruna Menezes Gomes da Silva e Júlio Cesar de Aguiar

Autismo: aspectos jurídicos da acessibilidade e respeito 283
Fabiana Barrocas Alves Farah e Danilo Fontenele Sampaio Cunha

Saúde e doenças raras: análise da judicialização do acesso ao tratamento e suas limitações 301
Danilo Henrique Nunes e Lucas de Souza Lehfeld

5. Outras políticas públicas em espécie 318

Regulação das águas: uma análise empírica da produção normativa dos órgãos reguladores federais 320
Bianca Borges Medeiros Pavão, Natasha Schmitt Caccia Salinas e Thauany do Nascimento Vigar

"Let the algorithm decide": is human dignity at stake? 343
Marcela Mattiuzzo

Das acepções dos direitos dos refugiados às vozes silenciadas nas políticas públicas 371
Thaís Araújo Dias e Monica Mota Tassigny

Planejamento familiar: "inimigo" a ser combatido, "aliado" libertador ou falso "amigo"? 395
Vinicius Ferreira Baptista

A ausência de políticas públicas para a juventude como ofensa aos direitos humanos 419
William Timóteo e Ilzver de Matos Oliveira

Análise cênica dos feminicídios em Curitiba: propostas preventivas e repressivas 433
Ticiane Louise Santana Pereira, Octahydes Ballan Junior e Antonio Henrique Graciano Suxberger

Origin and consequences of the war on drugs. from the united states to Andean countries 451
Silvio Cuneo e Nicolás Oxman

Trabalho Decente: comportamento ético, política pública ou bem juridicamente tutelado? 471
Silvio Beltramelli Neto e Mônica Nogueira Rodrigues

El final de una política pública: análisis del ciclo político del proyecto destinos inductores para el desarrollo turistico regional (DIDTR) – Brasil 496
María Belén Zambrano Pontón, Magnus Luiz Emmendoerfer e Suely de Fátima Ramos Silveira

Alternativa tecnológica para compensação de créditos de ICMS: estudo de caso da viabilidade do uso de DLT em nota fiscal eletrônica 520
Danielle Mendes Thame Denny, Roberto Ferreira Paulo e Fernando Crespo Queiroz Neves

Parte 2: Temas Gerais 549

A construção do Direito Humano ao Alimento no plano internacional 551
Tatiana de A. F. R. Cardoso Squeff

Grupos vulnerables de especial protección por parte del Instituto Nacional de Derechos Humanos (INDH) ¿en quién podría y debería enfocarse en base a la doctrina y a la experiencia comparada iberoamericana? 571
Juan Pablo Díaz Fuenzalida

El sufragio electrónico como alternativa al sufragio tradicional: luces y sombras de un debate recurrente 595
David Almagro Castro, Felipe Ignacio Paredes Paredes e Edgardo Lito Andres Cancino

Cognoscibilidade e controle social da transparência pública sob a égide da demodiversidade: estudo empírico de portais eletrônicos ministeriais latino-americanos 621
Ana Carolina Campara Verdum, Leonardo Fontana Trevisan e Rosane Leal da Silva

Desafios e benefícios da inteligência artificial para o direito do consumidor 655
Sthéfano Bruno Santos Divino

Quem tem medo da responsabilização subjetiva? As teorias da conduta e da imputação, para um direito administrativo sancionador constitucionalizado 690
Sandro Lúcio Dezan e Paulo Afonso Cavichioli Carmona

A insuficiência de tributação como fundamento para o afastamento da reserva do possível na garantia do mínimo existencial e da dignidade humana 711
Dione J. Wasilewski e Emerson Gabardo

doi: 10.5102/rbpp.v11i1.6784

"Let the algorithm decide": is human dignity at stake?*

"Deixe o algoritmo decidir": dignidade humana em risco?

Marcela Mattiuzzo**

Abstract

The goal of this article is to argue that the debate regarding algorithmic decision-making and its impact on fundamental rights can be better addressed in order to allow for adequate regulatory policies regarding recent technological developments in automation. Through a review of the literature on algorithms and an analysis of Articles 6, IX and 20 of the Brazilian Federal Law n° 13.709/2018 (LGPD) this article concludes that claims that algorithmic decisions are unlawful because of profiling or because they replace human analysis are imprecise and could be better framed. Profiles are nothing more than generalizations, largely accepted in legal systems, and there are many kinds of decisions based on generalizations which algorithms can adequately make with no human intervention. In this context, this article aims at restating the debate about automated decisions and fundamental rights focusing on two main obstacles: (i) the potential for discrimination by algorithmic systems and (ii) accountability of their decision-making processes. Lastly, the arguments put forward are applied to the current case of the covid-19 pandemic to illustrate the challenges ahead.

Keywords: Algorithms. Automated decisions. Decision-making. Human rights. Fundamental rights. Human dignity.

Resumo

O objetivo deste artigo é demonstrar que o debate sobre decisões algorítmicas e seus impactos em direitos fundamentais pode ser melhor colocado a fim de permitir respostas regulatórias adequadas aos desenvolvimentos tecnológicos recentes em automação. Por meio de uma revisão de literatura sobre algoritmos e da análise dos artigos 6°, IX e 20 da Lei Federal n° 13.709/2018 (Lei Geral de Proteção de Dados, ou LGPD), o artigo conclui que alegações de que decisões algorítmicas são ilícitas porque se baseiam na criação de perfis (*profiling*) ou porque substituem a análise humana são imprecisas e poderiam ser melhor colocadas. Perfis são apenas generalizações, largamente aceitas em ordenamentos jurídicos em múltiplas situações, e há vários tipos de decisões baseadas em generalizações que podem ser tomadas por algoritmos sem necessidade de análise humana. Nesse contexto, este artigo recoloca o debate sobre decisões automatizadas e direitos fundamentais com foco em dois principais obstáculos: (i) o potencial discriminatório de

* Recebido em 27/05/2020
Aprovado em 06/11/2020

** Doutoranda em direito comercial, mestre em direito do estado, com dissertação focada em debates sobre algoritmos, discriminação e políticas públicas, e bacharel em direito pela Universidade de São Paulo (USP). Realizou intercâmbio acadêmico na University of Zurich (Suíça), com estudos focados em direito econômico europeu, e foi visiting researcher na Yale University. Participou do curso da "Escola de Formação" da Sociedade Brasileira de Direito Público (SBDP), e também dos cursos de "Constituição e Política", "Metodologia Científica" e "Direito Constitucional" da instituição. Completou o curso jurídico da Escola de Governança da Internet do CGI.br e o curso CopyrightX, oferecido em parceria pela Universidade de Harvard e pelo ITS Rio. É membro do grupo Constituição, Política e Instituições da USP e coordenadora do Núcleo de Direito Concorrencial e Economia Digital da mesma instituição. Foi professora e coordenadora da primeira edição do curso "Proteção de Dados", organizado pelo Centro de Estudos de Direito Econômico e Social (Cedes). Foi vencedora do Prêmio Ibrac-Tim de monografias (2014) para artigos de estudantes de graduação, com o texto "Propaganda online e privacidade – o varejo de dados pessoais na perspectiva antitruste", posteriormente publicado também pela Competition Policy International (CPI).
E-mail marcela.mattiuzzo@gmail.com

sistemas algorítmicos e (ii) *accountability* em seus processos decisórios. Por fim, os argumentos levantados neste artigo são aplicados ao contexto da atual pandemia de covid-19 para ilustrar os desafios à frente.

Palavras-chave: Algoritmos. Decisões automatizadas. Processo decisório. Direitos humanos. Direitos fundamentais. Dignidade da pessoa humana.

1 Introduction

Decision-making carried out by algorithms is no longer science fiction. Beyond the automation scenario where computers simply execute tasks following detailed instructions given by human programmers, recent developments in artificial intelligence, specifically machine learning, and the emergence of Big Data[1] have made it possible for computers to learn from large databases and "program themselves". Through advanced software and processors, machines are able to handle extensive data and draw conclusions from it without explicit instructions on what to look for and how, making inferences, spotting correlations and identifying patterns.[2]

This kind of algorithmic systems is now largely applied to problem-solving in the most relevant areas of our individual and social lives: for example, an algorithm may decide whether we get a job or whether we have access to a line of credit. In these situations, algorithms are used to analyze a variety of data from hundreds of thousands – or even millions – of people: all present and past employees of large companies, or all borrowers from a commercial bank, for instance. Based on this data, the algorithm learns to differentiate between top-performing and uncommitted employees, or reliable payers and defaulters, aggregating the identified traits to create profiles of what good employees or good payers look or behave like, and which characteristics they usually have. Then, in reference to these profiles, algorithms analyze new job or credit applicants, thus being a key element in the decision-making process of many companies and financial institutions.

In decision-making, algorithms are but one of the tools available for making predictions, i.e. using information we have to generate information we do not have. They help to infer the probability of given subjects behaving in a specific way, e.g. their chances of paying their bills on time or being a responsible employee, based on the information we feed them. The main advantages they bear is better and faster prediction in many more scenarios at cheaper prices.[3] In making predictions, algorithms commonly use proxies. A proxy refers to a piece of information which is considered to be representative of another piece of information, because it has been observed that those two pieces vary in similar patterns. It becomes useful in scenarios where the main information we are after is hard or even impossible to obtain. For example, one may be able to anticipate the rates of flu infestation based on search queries about flu symptoms, which means that search queries may be a good proxy for the flu. Since observing the real spread of an infectious disease in

[1] As the FTC notes, the essential aspects that define Big Data are volume, velocity, and variety. FEDERAL TRADE COMMISSION. *Big Data, A Tool for Inclusion or Exclusion? – Understanding the Issues*. p. 1-2, Jan. 2016. Available at: https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf. Access: 1 Mayo 2020. Another fundamental aspect of Big Data worth highlighting is that it renders into data many aspects of the world never quantified before. MAYER-SCHÖNBERGER, V.; CUKIER, K. *The Rise of Big Data: How It's Changing the Way We Think*. Foreign Affairs, v. 92, n. 3, p. 29, Mayo/June, 2013. Available at: https://www.jstor.org/stable/23526834?seq=1#page_scan_tab_contents. Access: 1 Mayo 2020.

[2] DOMINGOS, P. *Master Algorithm*. Basic Books Inc. New York, 2018. p. xi. Algorithms can be seen as a gender with two main species: systems that are entirely programmed by humans to execute specific and detailed tasks, and artificial intelligence systems (the most striking example being machine learning algorithms), which work in close resemblance with biological neural networks insofar as they learn by observing and experiencing. KNIGHT, W. The Dark Secret at the Heart of AI. *MIT Technology Review*. April 11, 2017. Available at: https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/. Access: 3 Mayo 2020.

[3] AGRAWAL, A.; GANS, J. and GOLDFARB, A. Prediction Machines: The Simple Economics of Artificial Intelligence. *Harvard Business Press*, p. 24, April 17, 2018.

its early stages is hard for doctors and public health agents, this proxy may be of help for decision-makers. Naturally, this result is essentially probabilistic, since there is no causal relationship between the variables.[4]

There are several examples of this mechanism in place. The recruitment system designed by HireVue, for instance, ranks job candidates based on an analysis of their facial movements, word choice and speaking voice during an interview in front of a camera, and then compares the score with the ones obtained by current well-performing employees. Even though it remains possible for employers to pursue applicants poorly ranked, they mostly rely on the analysis made by the algorithm, since companies with automated hiring processes are ultimately seeking to reduce the costly human scrutiny of the numerous applications they regularly receive.[5]

When it comes to credit, algorithms like the ones used in the newly launched Apple Card are responsible for assessing the level of risk of different applicants and then deciding the credit limit that will be offered in each case. As with recruitment, there is usually no human participation in individual decisions: in search for reduced costs and efficiency, lenders are delegating credit assessment to machine learning systems.[6]

Healthcare is another field where artificial intelligence is thriving. Hundreds of thousands of patient records are fed into machine learning algorithms which then become able to determine, with extreme accuracy, patients' risk of developing certain diseases like diabetes, schizophrenia and cancer[7], or their chances of dying 24 hours after being hospitalized.[8] These algorithms, also called learners, are used to help doctors in hospitals decide which patients have the most intensive healthcare needs and should get more doctor visits, for instance.[9] Ultimately, this kind of automated system may even replace human judgment altogether: recent research shows that algorithms are as accurate as human doctors in interpreting medical images, which has led policy experts in public health to suggest the possibility of letting algorithms make diagnoses in places that lack specialists.[10]

---

[4] Mayer-Schönberer & Cukier emphasize that "[c]orrelations let us analyze a phenomenon not by shedding lights on its inner workings but by identifying a useful proxy for it." MAYER-SCHÖNBERGER, V.; CUKIER, K. *Big Data:* A Revolution That Will Transform How We Live, Work, Think. Houghton Mifflin Harcourt, 2013. p. 53.

[5] As explained by Stephen Buranyi, "the idea is that a good prospective employee looks a lot like a good current employee, just not in any way a human interviewer would notice." BURANYI, S. How to persuade a robot that you should get the job. *The Guardian*. March 4, 2018. Available at: https://www.theguardian.com/technology/2018/mar/04/robots-screen-candidates-for-jobs-artificial-intelligence. Access: 1 Mayo 2020. For additional information, see: HARWELL, D. A face-scanning algorithm increasingly decides whether you deserve the job. *The Washington Post*. November 6, 2019. Available at: https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/. Access: 1 Mayo 2020.

[6] NATARAJAN, S.; NASIRIPOUR, S. Viral Tweet About Apple Card Leads to Goldman Sachs Probe. *Bloomberg*. November 9, 2019. Available at: https://www.bloomberg.com/news/articles/2019-11-09/viral-tweet-about-apple-card-leads-to-probe-into-goldman-sachs. Access: 1 Mayo 2020.

[7] That is the case of Deep Patient, a system created at the Mount Sinai Hospital in New York, trained with a database of over 700,000 patient records. For more information, see: MIOTTO, R.; LI, L.; KIDD, B.; DUDLEY, J. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Scientific Reports* 6, Article n° 26094, 2016. Available at: https://www.nature.com/articles/srep26094. Access: 1 Mayo 2020.

[8] Algorithmic systems used in academic medical centers in the United States of America were able to predict, within 24 hours of a patient's hospitalization, their odds of dying, with over 90% accuracy. For more information, see: RAJKOMAR, A.; OREN, E.; CHEN, K. et al. Scalable and accurate deep learning with electronic health records. *npj Digital Medicine* 1, Article n° 18, 2018. Available at: https://www.nature.com/articles/s41746-018-0029-1. Access: 4 Mayo 2020. Other examples of similar algorithms can be found in: BURT, A. and VOLCHENBOUM, S. How Health Care Changes When Algorithms Start Making Diagnoses. *Harvard Business Review*, May 08, 2018. Available at: https://hbr.org/2018/05/how-health-care-changes-when-algorithms-start-making-diagnoses. Access: 1 Mayo 2020.

[9] That is the case of a system sold by Optum, a UnitedHealth Group-owned service, largely used in hospitals around the United States of America. For more information, see: PAUL, K. Healthcare algorithm used across America has dramatic racial biases. *The Guardian*. October 25, 2019. Available at: https://www.theguardian.com/society/2019/oct/25/healthcare-algorithm-racial-biases-optum. Access: 1 Mayo 2020.

[10] Xiaoxuan Liu, from the University Hospitals Birmingham NHS Foundation Trust and the lead author of the study comparing deep learning performance against healthcare professionals, said artificial intelligence systems could be useful in places with no experts to interpret images. DAVIS, N. AI equal with human experts in medical diagnosis, study finds. *The Guardian*. September 24, 2019. Available at: https://www.theguardian.com/technology/2019/sep/24/ai-equal-with-human-experts-in-medical-diagnosis-study-finds. Access: 1 Mayo 2020.

The radical changes artificial intelligence entails in healthcare may be a closer reality in Brazil today than ever before. The current coronavirus pandemic has pressured public health systems around the world and resulted in a scenario where there are not enough beds in hospitals for everyone who needs it.[11] Anticipating the possible shortage of Intensive Care Units (ICUs) in Brazil as the virus spreads, healthcare specialists started debating the possibility of using algorithms to decide which patients should occupy the available beds.[12] Among the suggested criteria for establishing priority are the chances of survival or how fast the recovery process will be – which, as seen in the examples above, machine learning algorithms can be extremely precise in predicting.

These developments have naturally raised several concerns regarding their impact on human dignity and human rights. Since algorithms mainly work by making predictions based on proxies and statistics about the behavior of groups of people, some argue that algorithmic decision-making hinders the dignity of individuals, since it fails to consider each person as an individual, rather making inferences and taking them to be part of groups.

This kind of concern has worried specialists and regulators for decades. Lee Bygrave, for instance, analyzed the 1995 European Commission Directive on data protection and noted it was already worried that the registered data-images of persons (their "data-shadow") are often misleading, and if taken as the sole basis for decisions may "usurp the constitutive authority of the physical self" and deprive individuals of the capacity to influence decision-making processes within public and private institutions.[13]-[14]

Similarly, a 2017 resolution by the United Nations Human Rights Council noted that, in a context where Big Data and artificial intelligence enable States and business enterprises to make inferences about peoples' physical and mental characteristics and create detailed personality profiles, "profiling may lead to discrimination or decisions that have the potential to affect the enjoyment of human rights, including economic, social and cultural rights."[15-16]

Others have concerns beyond the use of statistics and the creation of profiles: they see a violation to human dignity insofar as decisions made by algorithms forego human judgment and delegate important aspects of our lives to machines. Peter Asaro, affiliate scholar at Stanford Law School's Center for Internet and Society and co-founder of the International Committee for Robot Arms Control, argues that individuals and states have the moral and legal duty not to delegate to any automated process the authority or capability

[11] In March 11th, 2020, the World Health Organization declared the situation regarding the spread of the COVID-19 virus a pandemic. The Director-General's opening remarks at the media briefing that made this assessment public can be found at: WORLD HEALTH ORGANIZATION. *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. March 11, 2020. Available at: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020. Access: 1 Mayo 2020. In Italy, the spread has led to a collapse of hospitals in the most affected regions. For more information, see: HOROWITZ, J. and KIRKPATRICK, D. Dip in Italy's Cases Does Not Come Fast Enough for Swamped Hospitals. *The New York Times*. March 23, 2020. Available at: https://www.nytimes.com/2020/03/23/world/europe/italy-coronavirus-hospitals.html. Access: 1 Mayo 2020.

[12] COLLUCCI, C. Algoritmos e inteligência artificial podem ajudar Brasil a decidir sobre leitos de UTI. *Folha de S. Paulo*. March 24, 2020. Available at: https://www1.folha.uol.com.br/equilibrioesaude/2020/03/algoritmos-e-inteligencia-artificial-podem-ajudar-brasil-a-decidir-sobre-leitos-de-uti.shtml. Access: 1 Mayo 2020.

[13] BYGRAVE, L. Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling. *Privacy Law & Policy Reporter*, v. 7, p. 67–76, 2000.

[14] Other work that presents similar concerns include: BROWNSWORD, R. From Erewhon to AlphaGo: for the sake of human dignity, should we destroy the machines? *Law, Innovation and Technology*. V. 9, Issue 1, p. 117-153, 2017. Available at: https://www.tandfonline.com/doi/abs/10.1080/17579961.2017.1303927. Access: July 11, 2020.

[15] UNITED NATIONS. Resolution adopted by the Human Rights Council on 23 March 2017. April 7, 2017. p. 3. Available at: https://documents-dds-ny.un.org/doc/UNDOC/GEN/G17/086/31/PDF/G1708631.pdf?OpenElement. Access: 1 Mayo 2020.

[16] Likewise, other authors from different fields have pointed towards the risks that algorithms and automation can present to human rights. See: SEAVER, N. Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*. V. 24, Issue 4, p. 421-436, Dec. 2018. Available at: https://doi.org/10.1177%2F1359183518820366. Access: July 9, 2020; ROYAKKERS et al. Societal and ethical issues in digitization. Ethics and Information Technology. V. 20, March 2018, p. 127-142. Available at: https://link.springer.com/article/10.1007/s10676-018-9452-x. Access: July 9, 2020.

to initiate the use of lethal force independently of human analysis in each and every case. Expanding from the military use of automated systems and the fundamental right to life, he further suggests that a similar duty could exist regarding any automated decisions that override human rights, in cases such as arrests and restrictions of movement, surveillance and tracking, deportation, eviction, denial of healthcare and other civil, political, economic, social, and cultural rights.[17-18]

Following a similar line of reasoning as the one presented above, the European General Data Protection Regulation (GDPR), which came into force in May 2018, states a general prohibition of any decision-making based solely on automated processing which significantly affects individuals.[19]

In Brazil, the same concern motivated an intense debate involving the data protection federal law.[20] The initial proposal, approved by Congress in August 2018, stated a right to request human revision of automated decisions, which became the object of disagreement between the Legislative and the Executive. After two attempts by the Executive to override Congress' decision and remove the mandatory participation of humans in the revision process[21], the Presidential veto finally prevailed, by a margin of one Senator's vote.[22]

The tight score which settled the controversy demonstrates there is still little clarity over how algorithmic systems and automated decision-making should be regulated. To contribute to better policy designs, this article focuses on improving the diagnosis of the issue at hand. More specifically, it aims to address two different but related inquiries: (i) if the use of algorithmic systems in decision-making is indeed harming human dignity, and if so, (ii) whether that is a result of profiling, or rather a problem related to the removal of human judgment from the equation.

In order to achieve such goal, I will aim at repositioning the debate regarding algorithmic decision-making, in order to hopefully help identify the nuances in complex scenarios, and thus improve policy choices on this matter. I intend to contribute to the discussion by demonstrating that some concerns with algorithms are misplaced and should be reframed. In doing so, I intend to answer the question of how exactly human rights, and specifically the fundamental rights established in the Brazilian Constitution, are affected by algorithms in decision-making.

Part II of this article demonstrates algorithms are new decision-making tools that still resemble traditional decision methods in relevant aspects: they operate through inferences, profiles and generalizations, which are long-established and legally valid in most jurisdictions, including the Brazilian legal system, for a myriad of situations. Part II also highlights that, although there are some kinds of decisions that algorithms are unsuited to make – ones that require case-by-case assessment and the exercise of prudence, for instance –, there are several others they could adequately make with no necessary harm to human dignity. Part III goes on to sistematize the debate on algorithms and fundamental rights: it identifies discriminatory misfunction and accountability as two important obstacles in the search for lawful algorithmic decisions, and analyzes the associated regulatory challenges. Part IV recaps the main assertions of this article by means of

17 ASARO, P. On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross*. V. 94, Issue 886, p. 687709, Jun. 2012. Available at: https://international-review.icrc.org/sites/default/files/irrc-886-asaro.pdf. Access: 1 Mayo 2020.

18 For a broader analysis of Asaro's work in light of his concern for human dignity in automated-weapons systems, see: SHARKEY, A. Autonomous weapons systems, killer robots and human dignity. *Ethics and Information Technology*. V. 21, p. 75-87, 2019. Available at: https://link.springer.com/article/10.1007/s10676-018-9494-0. Access: 10 Jul. 2020.

19 EUROPEAN UNION. General Data Protection Regulation (Regulation EU 2016/679, May 25, 2018). Article 22.

20 BRAZIL. Lei Geral de Proteção de Dados (Federal Law n° 13.709, August 14, 2018).

21 The first attempt was the issue of the Executive Order 869, in December 27, 2018 (known as Medida Provisória in Portuguese, a type of act issued by the Presidency that comes into force immediately but must be confirmed by Congress before officially becoming law, otherwise its effects are reversed), and the second was the Presidential veto to some items of Federal Law n° 13.853/2019, which modified the data protection federal law.

22 The record of the voting can be found at: CONGRESSO NACIONAL. Votação do dispositivo 24.19.001 - § 3º do art. 20 da Lei nº 13.709, de 14 de agosto de 2018, com a redação dada pelo art. 2º do projeto. October 2, 2019. Available at: https://www.congressonacional.leg.br/materias/vetos/-/veto/detalhe/12445/1. Access: 1 Mayo 2020.

a concrete case analysis: it applies the findings to decision-making scenarios related to the ongoing covid-19 pandemic.

2 Striving for individualization and human scrutiny

Algorithmic decision-making is usually questioned by critics for its use of statistics about the behavior of groups of people to create profiles and make decisions, and/or because it often replaces human judgement. This part of the article argues this is not an entirely accurate framing for the problems raised by algorithms. First, profiles are nothing more than generalizations, and decisions based on generalizations are largely accepted in the legal system. Therefore, algorithms cannot be considered unlawful solely for employing the same methodology already used in traditional decision-making mechanisms. Second, when it comes to human analysis, general claims against automated decisions are inaccurate because they disregard the varied nature of situations that need addressing: not every decision related to human rights requires human intervention to be lawfully made. In fact, many of them are already based on generalizations and objective criteria which would be even more efficiently applied if algorithms were in place.

Regarding the first proposition, profiles are no different than generalizations. In a very broad sense, they both can be understood as any practice that associates things, people, or situations on the basis of certain characteristics that we are interested in observing or in taking into consideration. It is equally relevant to observe that generalizations and profiles often make use of proxies. As explained, a proxy is a characteristic inferred from another, given observation of both traits through time and the conclusion that they tend to vary similarly; in other words, the traits are correlated. Because many traits are unobservable, proxies allow us to infer the behavior of an invisible trait by analyzing the behavior of another.

Decisions based on generalizations are incredibly frequent in the legal system. They are, for example, the foundation of any legal statute that establishes minimum age for the performance of certain activities, such as voting, drinking alcoholic beverages or driving a car. When Brazilian law determines that only people over eighteen years old will be allowed to drive, it is assuming that, at this age – and only then –, people become physically and mentally able to safely conduct a motor vehicle – which is surely an assumption disproved in many cases. Even with safeguards in place, like the mandatory education courses and tests that precede the issuance of a driver license, we know that our method is imperfect and will prevent skilled drivers from obtaining a license while allowing less skilled ones to do so. Even if in the aggregate the 18 years-old rule holds true, it certainly does not in all scenarios. For instance, professional kart drivers who are 15 years old are probably more apt to drive a car than their non-kart driver friends who just turned 18.

Generalizations are also the basis for several public policies. When the government organizes flu vaccination campaigns, for example, it establishes risk groups that will be the main – if not the only – target of the immunization effort[23], even if the highlighted characteristics (such as age, underlying health conditions, pregnancy) do not cover all the possible cases of serious complications and lethality.[24] In profiles and gene-

[23] According to the Brazilian government, the definition of priority groups for immunization campaigns is based on scientific research of mortality rates and other exposure risks regarding the illness in different populations. For more information, see: MINISTÉRIO DA SAÚDE. *Vacinação contra a gripe*: entenda se você faz parte do grupo prioritário. 3 Mayo 2016. Available at: http://www.blog.saude.gov.br/alh5s1. Access: 1 Mayo 2020.

[24] For examples of influenza deaths outside the risk groups, see: MORTES por gripe H1N1 na região têm 33% de vítimas fora do grupo de risco. *G1 Campinas e Região*. May 13, 2016. Available at: http://g1.globo.com/sp/campinas-regiao/noticia/2016/05/mortes-por-gripe-h1n1-na-regiao-tem-33-de-vitimas-fora-do-grupo-de-risco.html. Access: 1 Mayo 2020. As pointed by José Geraldo Leite Ribeiro, an epidemiology professor at the Minas Gerais School of Medical Sciences, the risk groups are only patterns of people with greater chances of dying after contracting the H1N1 virus, established in Brazil and around the world by health organizations as a form of optimizing vaccination campaigns, since there are not enough vaccines available for the entire population. It does not mean that people outside the established groups are completely immune to the virus. KIEFER, S. Morte de pessoas fora da faixa de risco acende alerta para o H1N1. *Estado de Minas*. May 27, 2016. Available at: https://www.em.com.br/app/noticia/

ralizations currently in place, therefore, some individuals are included and some are excluded, and society accepts that this method for allocating resources is flawed.

But why, then, do we generalize, when we know the result of such generalization is imperfect? Simply put, we face a trade-off between efficiency and accuracy: we live in a world of imperfect information and limited resources where the cost of knowing for sure if a certain citizen is in fact a responsible driver is inestimable, since it would involve extensive assessment of several individual characteristics and behavior for a long period of time. Even then, we would still generalize, for the observed behavior of past conduct is not an infallible indicator of future conduct. In the case of immunization campaigns, the cost of seasonally vaccinating every human being for different types of the flu is considered too high, and even establishing new priority groups would possibly require further information about other risk factors that current research may not have yet discovered. Generalizing is, therefore, one way of turning decision-making possible: it allows us to work with the characteristics we can observe (age, approval in a test, chronic diseases etc.) and infer the information we do not access (responsibility in driving a car, risk of flu complications etc.).

On the background of concerns about profiling, there seems to be a belief that individualization is more accurate and thus always a superior form of assessment. This, however, is an assumption that should be revisited. Frederick Schauer points that this understanding is often misguided because the process by which one reaches an individualized decision is usually flawed. We often lose sight of the fact that much of what we consider to be individualized decision-making is in fact based on generalizations from past experiences. He concludes that our preference for this kind of decision-making is based on two mistakes:

> an overconfidence in the empirical reliability and even the very directness of direct evidence, and an underappreciation of the essential continuity between so-called indirect or statistical evidence and evidence that on its face appears to be more individualized and thus less statistical.[25]

This means that, if we were to designate a public servant to individually analyze each and every citizen interested in obtaining a driver's license, with no reference to minimum age, she would still base the decision in proxies and generalizations, referring to her own understanding of what being a good driver is, based on characteristics she perceive in drivers she has known during her life. Healthcare is another example, since most individual diagnoses are based on the association of the observed symptoms with the most common illnesses for that kind of patient (age, gender etc.) in that time and location (city, climate etc.). Even when doctors resort to specific and thorough exams to fully understand the situation of that single person, they still request an exam based on the probability of that individual having a certain type of disease or complication.

If individualization was indeed always possible and if it was not riddled with the defects pointed out by Schauer, it could be advantageous to rethink and maybe expand its use in decision-making. However, because it is not (and cannot be at present circumstances), generalizing is a valid form of decision-making in many contexts. Society understands and accepts the necessity to use proxies and make inferences and generalizations, enduring a lesser degree of accuracy in a great many situations in order to turn decision--making into a feasible effort.

Bringing the debate to algorithmic decision-making, if we understand generalizations are common and tolerated by the legal system, we must admit the problem with algorithms cannot be the simple act of generalizing. Claiming algorithms fail to account for the dignity of human beings because they use proxies and make inferences based on groups of individuals is an argument that requires further enrichment: it is necessary to identify in which cases generalizations through algorithms pose a problem to the legal system, and why.

The first kind of situation in which the use of algorithms and profiling is unlawful is actually a scenario

gerais/2016/05/27/interna_gerais,766671/alerta-fora-do-grupo-de-risco.shtml. Access: 1 Mayo 2020.

[25] SCHAUER, F. *Profiles, Probabilities, and Stereotypes*. Belknapp Press. April, 2016. p. 106.

where any kind of generalization is unlawful – be it human-made or algorithmically generated. Whenever the Constitution bans inferences and demands decision based on direct evidence of specific individual behavior, algorithms are clearly unsuited, since they work with proxies and information of groups of people.

The most evident case is criminal law, where individual assessment of illegal conduct is mandatory. Many jurisdictions, including Brazil, forbid a person from being judged by someone else's crimes[26], demanding a ruling according to the person's own actions or omissions in concrete cases, and not according to profiles or proxies. In these legal systems, generalization and the efficiency that may derive from it in public security policies face the limitations imposed by fundamental rights: due to the seriousness of sanctions (great restrictions to the freedom of movement), the degree of accuracy required is extremely high. Therefore, it is deemed illegal to send someone to prison based on the notion that the family to which that specific person pertains has committed crimes in the past, even if that person has never been involved in any act of the sort. It is also illegal to abstain from imprisoning someone who has committed a crime because they belong to a group deemed less prone to such actions.

That is one of the problems with the use of profiling algorithms by the police, for example. Law enforcement authorities in many cities in the United States have used "predictive policing" algorithms to create a map of the most likely places to have criminal occurrences[27] or a list of the people most likely to commit crimes. Therefore, a person may be a more frequent target of police frisk because she lives in a "dangerous" neighborhood or because she is friends with people with prior criminal history[28], which poses a challenge to fundamental rights. It is crucial to note that, in these situations, human generalizations were a problem long before algorithmic systems were in place: empirical research shows that police officers have a longstanding practice of selecting targets for intensive action based on common traits they perceive among criminals during their time in the police force, thus using profiles to focus scrutiny.[29] As pointed by Andrew Ferguson, "while these predictive technologies are excitingly new, the concerns underlying them remain frustratingly old-fashioned".[30] Recognizing that, in some fields, generalization is not suitable as a tool for decision-making is also admitting that algorithms should not be employed in those same fields.

The second kind of situation in which the use of algorithms is unlawful is every scenario that demands the exercise of judgment or prudence, i.e. perceiving the nuances of specific situations and being able to evaluate pros and cons, interpret and balance values or provide adequate responses to situations based on ethical or moral assessments. One of the most striking examples is precisely the one highlighted by Peter Asaro: warfare. Determinations of whether one should enter diplomatic negotiations with an opponent or initiate a war, or whether a few thousand soldiers should be sent to their deaths in order to save hundreds of thousands of others, are some scenarios where political, ethical and moral dilemmas impose themselves on decision-making.

However, the exercise of judgment is not restricted to those extreme situations. In routine judicial activity, judges are required to interpret open-ended concepts and legal provisions which often involve values. For instance, the Brazilian Criminal Code establishes that, if a crime is committed by any heinous or cruel

[26] BRAZIL. Constituição da República Federativa do Brasil de 1988. Article 5º, XLV.

[27] PredPol, for example, is a startup in California specialized in making these kinds of maps. For more information, see: HUET, E. Server And Protect: Predictive Policing Firm PredPol Promises To Map Crime Before It Happens. *Forbes*. February 11, 2015. Available at: https://www.forbes.com/sites/ellenhuet/2015/02/11/predpol-predictive-policing/#229530d34f9b. Access: 1 Mayo 2020.

[28] That is the case of an algorithm used by the Chicago Police Department. For details on how it works, see: FERGUSON, A. *The rise of big data policing:* surveillance, race, and the future of law enforcement. New York: New York University Press, 2017. p. 47.

[29] SINHORETTO, J. et al. A filtragem racial na seleção policial de suspeitos: segurança pública e relações raciais. In: LIMA, C.; BAPTISTA, G.; FIGUEIREDO, I. (org.). *Segurança Pública e Direitos Humanos*: Temas Transversais. Brasília: Ministério da Justiça, 2014. p. 133. Coleção Pensando a Segurança Pública. v. 5. A study of this practice in the United States of America can be found in: HARRIS, D. U.S. experiences with racial and ethnic profiling: history, current issues, and the future. *Critical Criminology*. V. 14, Issue 3, p. 213–239, Sep. 2006.

[30] FERGUSON, A. *The rise of big data policing:* surveillance, race, and the future of law enforcement. New York: New York University Press, 2017. p. 44.

means, the sentence must be aggravated.[31] Evidently, the application of that rule requires interpretation of values related to what is considered cruel. While law can take the form of predictable ex ante instructions for behavior, it can also take the form of standards, which demand ex post judgments that should tailor an outcome to facts, and demand decision-makers to articulate their choices.[32]

In healthcare, similarly, the need for analysis of pros and cons and interpretation of values and ethical principles is evident. Whenever a doctor has to advise a patient about the possibility of an experimental and risky treatment for an illness, or whenever clinical trials, which must observe bioethical guidelines[33], are submitted for approval, judgment and prudence are exercised. In these cases, the decision should not be solely left to algorithms of any kind, since they are not able to make the required evaluations. Machine learning has enabled computers to learn patterns, but not to criticize patterns according to an order of values, much less to make decisions based on an appraisal of what is right or wrong.

Currently, programming algorithms to exercise judgment is far beyond the reach of technology: we are unable to teach machines how to reach a fair decision.[34] In a recent study, Sandra Wachter, Brent Mittelstadt and Chris Russell analyzed jurisprudence from the European Court of Justice and national courts in the European Union and demonstrated that the assessment of discrimination cases is far too contextual to be standardized for automated application by algorithms. European courts lack well-defined patterns for illegal disparity which would hold across different cases. According to the authors, "fairness is defined by judicial intuition, not statistics", which in turn makes it nearly impossible to automate fairness and non-discrimination.[35]

Not surprisingly, the discussions of ethics and algorithms often focus not on "teaching" algorithms values, but on erecting barriers to prevent them from taking certain actions: we either program the fair decision into their systems or we exclude that decision from their purview.[36] Once we understand algorithms are not actually "intelligent", nor do they provide answers for each and every question that has long haunted humanity, we are able to identify situations where the use of algorithms in decision-making pose a problem to the legal system: whenever judgement or prudence are considered necessary.

That is one of the problems with some uses of algorithms in judicial matters. Around the world and also in Brazil, there have been many applications of artificial intelligence in the judicial process aiming at reducing costs and making court proceedings faster. For instance, in Pernambuco, a system called Elis lear-

[31] BRAZIL. Código Penal (Decreto-Lei n° 2.848, December 7, 1940). Article 61, II, d).

[32] CITRON, D. Technological Due Process. *Washington University Law Review*. V. 85, Issue 6, p. 1303, 2008. Available at: http://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2. Access: 1 Mayo 2020.

[33] The Resolution n° 466/2012 from the National Board of Health (in Portuguese, Conselho Nacional de Saúde) regulates research on humans and establishes in section III.1, for instance, the need for "social relevance" and "balance between risks and benefits". Institutions which conduct research involving humans must also have an ethics committee to approve and supervise the studies. For more information, see: CONSELHO NACIONAL DE SAÚDE. *Manual operacional para comitês de ética em pesquisa*. 4. ed. Brasília: Ed. Ministério da Saúde, 2008. Available at: http://bvsms.saude.gov.br/bvs/publicacoes/manual_operacional_comites_pesquisa_4ed.pdf. Access: 1 Mayo 2020.

[34] In the context of military use of automated systems, there are researchers who investigate the possibility of embedding ethical analysis in algorithms from the onset (and not as a post hoc activity) to make war machines respect the principles established in international treaties (e.g. the Geneva Conventions). For an example, see: ARKIN, R.; ULAM, P.; WAGNER, A. Moral Decision-making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust and Deception. *Proceedings of the IEEE*. V. 100, Issue 3, p. 571-589, Mar. 2012. Suffice to say there is still no evidence of that goal being achieved, and the majority of researchers argue it most likely will not be accomplished through current technology.

[35] According to the authors, definitions regarding who can be considered part of a disadvantaged group and the severity and type of harm suffered, for example, require judges to make normative and political choices on a case-by-case basis. WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*. March 3, 2020. Available at https://ssrn.com/abstract=3547922. Access: 1 Mayo 2020. p. 44, 64-65.

[36] INTERNATIONAL ASSOCIATION OF PRIVACY PROFESSIONALS; THE UNITED NATIONS GLOBAL PULSE. *Building Ethics into Privacy Frameworks for Big Data and AI*. October 2018. Available at: https://iapp.org/media/pdf/resource_center/BUILDING-ETHICS-INTO-PRIVACY-FRAMEWORKS-FOR-BIG-DATA-AND-AI-UN-Global-Pulse-IAPP.pdf. Access: 1 Mayo 2020.

ned from a database of more than four hundred thousand previous disputes and now decides which tax execution cases were initiated within the limitation period and correctly follow procedural rules, being able to move forward for trial.[37] This use of Elis seems appropriate not only because what the program analyzes is a clearly and lawfully defined concept – the rules of proceedings and the limitation period – but also because it is not responsible for trial itself, it simply helps judges more speedily analyze one aspect of decision-making. But in the United States of America, algorithmic systems have also been used for sentencing in criminal cases: the Correctional Offender Management Profiling for Alternative Sanctions, or COMPAS, a tool developed by the company Northpointe (later renamed Equivant) and adopted by the justice system in states such as Florida and Wisconsin, was used to aggravate the sentence of defendants who were considered by the algorithm more prone to committing future violent crimes.[38]

Giving algorithms the power to determine legal decisions means we are removing judgement and prudence from the decision process, which can pose a problem related to the application of legal standards and values – which are especially crucial in criminal sentencing. As Danielle Citron puts it, "The emergence of automation threatens to overwhelm this debate by giving rules a huge, and often decisive, advantage on the basis of cost and convenience rather than the desirability of the substantive results they produce."[39]

The necessary conclusion is that algorithms are extremely useful tools that must be used without losing sight of their limitations. The same way using a hammer to cut paper will likely be ineffective, using algorithms to determine what is "fair" is doomed to failure, or at the very least will yield suboptimal results.

Differentiating between problems is crucial, since it allows us to more accurately address the concerns raised by authors like Peter Asaro. Given the intrinsic limitations in current technology relating to judgement and prudence, decisions regarding the use of lethal force cannot be delegated to algorithms without impairing the enforcement of principles and values established in domestic legislation and international treaties, since automated systems are not able to judge situations according to those standards. However, that does not mean algorithmic systems are unable to decide in every situation that involves fundamental rights, like suggested by Asaro.

With that in mind, we have so far established two scenarios where algorithms are inadequate: whenever generalizations are forbidden or whenever prudence is required. To highlight that difference, it is worth noting that there are several situations where fundamental rights are lawfully restricted through decisions based on generalizations with no examination of values or principles. For instance, in Brazil, a person who suffers from a chronic disease for which the treatment is an expensive medication to be used on a regular basis can resort to the public health system (in Portuguese, Sistema Único de Saúde, or SUS). The system imposes several criteria for the supply of this kind of medication: a standardized medical report needs to be filled and clinical protocols and therapeutic directives previously established by the Ministry of Health for the disease need to be observed. The required medication also needs to be listed in the national catalog of essential pharmaceuticals.[40]

Now imagine that the government decides to automate this process. Automation involves nothing more

[37] It remains possible for judges to review the decision made by the algorithm. FERREIRA, F. Inteligência artificial atua como juiz, muda estratégia de advogado e 'promove' estagiário. *Folha de S. Paulo*. March 10, 2020. Available at: https://www1.folha.uol.com.br/poder/2020/03/inteligencia-artificial-atua-como-juiz-muda-estrategia-de-advogado-e-promove-estagiario.shtml. Access: 1 Mayo 2020.

[38] ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine Bias. *ProPublica*, May 23, 2016. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Access: 1 Mayo 2020.

[39] CITRON, D. Technological Due Process. *Washington University Law Review*. V. 85, Issue 6, p. 1303, 2008. Available at: http://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2. Access: 1 Mayo 2020.

[40] The requirements that need to be observed by applicants are discriminated under the Specialized Component of Pharmaceutical Assistance (Componente Especializado da Assistência Farmacêutica, or CEAF, in Portuguese) established in regulations from the Ministry of Health (Portaria de Consolidação nº 2/GM/MS, enacted in September 28, 2017). The clinical protocols and therapeutic directives have standards for diagnosis, eligibility of patients and specific procedures to be observed in the use of pharmaceuticals for treatment.

than turning the already existing criteria applicable by an algorithm. If generalizations existed before, they will remain there and not be novel. The main concern in this case, apparently, is not with the criteria used by the government, because they were already well-established in legislation. The question is making sure the algorithm correctly applies the criteria to specific real-case requests.

It must be said that the establishment of generalizations in the case of access to pharmaceuticals is subject to a longstanding debate. Case law throughout the country understands that everyone should have their medical needs fully contemplated by public policy, given that the right to health is set forth in the Brazilian Constitution, and therefore the establishment of generalizations through which some may be excluded from specific kinds of treatment would be unlawful. In May 2019, the Brazilian Supreme Court (Supremo Tribunal Federal, or STF, in Portuguese) ruled that medicines not registered before the National Agency for Sanitary Vigilance (in Portuguese, Agência Nacional de Vigilância Sanitária, or ANVISA) for rare and ultrarare diseases must still be supplied by the public health system for those who need it. The decision undermines the criteria established in ANVISA's current regulation, which determines special proceedings for registration of such medications, including the establishment of a price limit. The debate is still open, as other cases on similar matters await ruling by the Supreme Court.[41]

The STF's reasoning is useful as an indicator of a third kind of situation where algorithms could pose a problem to the legal system: again, like the first scenario, the problem is not with algorithms themselves, but with generalizations and the exclusion that results from it. In this case, the issue is not that individual or ethical assessment is required, but rather that the exclusion of anyone from the enjoyment of a given right can in itself be a legal violation. As ruled by the STF, people who suffer from rare and ultrarare diseases have the right to access medication despite the general requirements established by ANVISA. Regardless of whether we agree with this line of thought, the relevant conclusion to this article is that the debate is prior to the emergence of automation, and is therefore a problem regarding generalizations and not algorithms.

A tentative answer to the question set forth in this section – whether algorithms present a problem because they decide based on generalizations and forgo human judgement – would therefore go in the lines of arguing algorithmic decision-making will certainly be a problem if it delegates value-based decisions to machines, given that the substance of such decisions is not something these systems are well-suited to handle. It will not be a problem in itself, however, if it automatizes generalizations already accepted and carried out by society before. As stated, there are several decisions regarding public health currently made by the use of proxies and generalizations. Therefore, the employment of algorithms in these fields is not always a problem and can actually make our decision-making process better.

Simply stating that algorithms violate human dignity because they lack human participation is imprecise. The problem is not with algorithms themselves, but with the question of how we want to deal with certain situations: do we allow generalizations in order to obtain efficiency, or do we think the situation demands an individualized assessment of concrete cases? Do we want the decision to be based on objective and detailed criteria, or do we want there to be room for judgment and principle-based decisions? If we opt for objective criteria and proxies, leaving human judgment aside can actually be better for the outcome.

Establishing that the problem with algorithmic decision-making is not simply that it is based on profiling or that it excludes human judgment does not mean that algorithms pose no challenges. Saying that profiling is allowed in several cases is not saying that every profile established by every automated system is admissible. Profiling through algorithms can entail different problems than other forms of profiling. These problems will be the subject of the next section of this article.

[41] The decision was made in the Extraordinary Appeal (in Portuguese, Recurso Extraordinário) n° 657.718, in May 22, 2019 (reporting Justice: Marco Aurélio). It undermines the current proceedings established by Anvisa in Resolution RDC n° 205, issued in December 28, 2017. The other cases on similar matters pending ruling by the Supreme Court are the Extraordinary Appeal n° 566.471 (reporting Justice: Marco Aurélio), where the discussion is about the obligation of the state to provide high-cost medication, and the Extraordinary Appeal n° 1.165.959 (reporting Justice: Marco Aurélio), which discusses the obligation to provide unregistered medication.

3 Discrimination and accountability

A large group of decisions fall outside the scope of value-based decision-making and have been subjected to generalizations long before algorithms were applied to them. In this context, if the use of algorithms is not in itself a problem, it is also not entirely devoid of concerns. This section highlights two important issues related to automated decisions: (i) algorithmic systems' misfunctions that result in discrimination and (ii) accountability of the decision-making process. It also analyzes the impact of the identified concerns on regulatory efforts.

The legal issues that may emerge out of algorithmic decision-making and that have received the most attention from academics so far are connected to the mechanics and functioning of the algorithmic system.[42] In these scenarios what is questionable is the model, the dataset, or the methods by which inferences are drawn from the combination of algorithm and data, which in turn may lead to biases.

There can be mistakes in the data that is captured or used by the algorithmic system – incorrect or outdated data are the most common sources of problems. For instance, if the patient records fed into an algorithm used for diagnostic purposes were incorrectly filed by doctors, the analyses produced by the algorithm would likely be wrong. There can also be errors in the algorithm itself, which include faulty coding, unintentional failures to account for part of the database, and so on. If the algorithm used for diagnosis was wrongly programmed to dismiss instead of flag signs of tumors, for example, the results would naturally be unreliable. When the algorithm somehow ignores data due to an engineering mistake, such as failure to include relevant information in its analysis, unsoundness arises.

The algorithm may also rely on biased samples: the data is not outdated or incorrect, but rather the dataset as a whole is somehow compromised, be it because it misrepresents the intended population, or because the algorithm is programmed to select only part of the dataset, thus generating questionable results.

Algorithms designed for recruitment are clearly subject to this risk. If a tech firm creates an algorithm aimed at choosing suitable candidates for a data scientist position, and that algorithm is programmed to look for people whose profiles are similar to those already hired by the firm, it will likely give enormous preference to men instead of women. The algorithm itself is not mistakenly programmed, nor is the data available "wrong", for it accurately depicts the staff of that company. But the sample is biased, for there are more men than women in data science jobs. Even though women are no less suited for these positions, the fact that current employees are mostly male taints the sample and leads to an undesired result.

This problem is particularly cumbersome when dealing with machine learning algorithms, for in this case the dataset is not merely an input, it is also what trains the program and allows for the creation or modification of the algorithm, which means distortions may be multiplied at each iteration. This type of failure can be hard to identify, for the issue lays in the target variable selected by the algorithmic system, which for an outside observer can be extremely laborious to recognize.

An illustration of this issue can be found in the controversy surrounding the already mentioned Apple Card in the United States of America – the credit card offered by a joint venture between Apple and Goldman Sachs –, which became target of an investigation by the New York's Department of Financial Services regarding gender bias.[43] Women who apply for the card are supposedly being offered less credit than men in the same economic and financial situation, and although Goldman Sachs has stated that its algorithm

[42] See, for example, BAROCAS, S.; SELBST, A. Big Data's Disparate Impact. *California Law Review*. V. 104, n. 3, p. 671-732, Jun. 2016. Available at: http://www.californialawreview.org/wp-content/uploads/2016/06/2Barocas-Selbst.pdf. Access: 1 Mayo 2020.

[43] NATARAJAN, S.; NASIRIPOUR, S. Viral Tweet About Apple Card Leads to Goldman Sachs Probe. *Bloomberg*. November 9, 2019. Available at: https://www.bloomberg.com/news/articles/2019-11-09/viral-tweet-about-apple-card-leads-to-probe-into-goldman-sachs. Access: 1 Mayo 2020.

does not factor gender into its decision[44], it remains possible that, given the historically unequal access to the banking system, the learner algorithm concluded by itself from the data it was trained with that women should be given less credit than men and used proxies to discover the gender of each applicant.[45]

Issues may equally arise when the algorithm stumbles upon a misleading correlation. In this case, the model works well, both in terms of the dataset and the algorithm, and there are no inherent biases in the data, but the careless use of correlations leads to misclassification. It is worth noting that this use of an algorithmic system does not reproduce a bias, it simply ignores some of the characteristics of a given person, usually because those characteristics are not available in the database, and takes the characteristics that are available as the complete picture. More often than not, this results in people being included in categories they do not belong to or that do not accurately describe them, which may have significant impact in the exercise of their fundamental rights.

That seems to be the problem with an algorithm largely used in hospitals within the United States of America to help doctors decide which patients have the most need for intensive healthcare. Recent studies have suggested that the system sold by Optum, a UnitedHealth Group-owned service, uses – as a proxy for a patient's medical needs – the patient's past medical expenses. The assumption is that these two variables walk hand-in-hand, and the more the patient has spent on medical care in the past, the higher the likelihood of her needing further medical attention in the future. But because of social inequalities, black people in the US have overall lower health spending than white people with similar health conditions, meaning the proxy is not well-suited for the intended purposes in case of this population and leads to a misleading correlation, which in turn leads the algorithm to flag healthier white patients for extra assistance over sicker black patients.[46] This is a typical case of an algorithm that was not at all intended to reproduce or create biases, but that, because of the potential limitations of the proxy chosen to generate results, could lead to unwanted discrimination.

The goal of this article is not to delve deeper into the specifics of how algorithms may lead to biases by way of faulty data collection, mistakes in training data and so forth,[47] but rather to claim that, although there are situations where profiling or algorithms replacing humans in decision-making constitute a problem per se, there are several other scenarios where the issue is not precisely defined by the general statement put forth by some critics of algorithmic systems. They are better stated as a problem of discrimination, i.e. a violation of the right to equality, resulting from a failure in the profiles created, which once addressed would allow the algorithm to soundly operate. As a result of the errors in the use of the algorithm, black people may get less access to healthcare, women may be offered less high-qualified jobs, and so on.

One aspect that must be highlighted is that, in Brazil, this violation to the fundamental right to equality is present both in the case of an algorithm used by the public health system for medical diagnosis – i.e.

[44] The bank's statement regarding the controversy can be found in the company's Twitter account. GOLDMAN SACHS BANK SUPPORT. We hear you #AppleCard. Twitter: @gsbanksupport. Available at: https://twitter.com/gsbanksupport/status/1194022629419704320. Access: 1 Mayo 2020.

[45] As noted by Alisha Gupta in The New York Times, until 1975, single, divorced or widowed women in the United States of America needed a man to co-sign their credit card applications, and it was not until 2013 that the Consumer Financial Protection Bureau urged banks and credit card issuers to consider shared incomes in the applications of stay-at-home spouses. GUPTA, A. Are Algorithms Sexist? *The New York Times*. November 15, 2019. Available at: https://www.nytimes.com/2019/11/15/us/apple-card-goldman-sachs.html. Access: 1 Mayo 2020. Will Knight wrote on Wired that a "gender-blind algorithm could end up biased against women as long as it's drawing on any input or inputs that happen to correlate with gender". For instance, where a person shops might be a proxy of their gender. KNIGHT, W. The Apple Card Didn't 'See' Gender — and That's the Problem. *Wired*. November 19, 2019. Available at: https://www.wired.com/story/the-apple-card-didnt-see-genderand-thats-the-problem/. Access: 1 Mayo 2020.

[46] OBERMEYER, Z.; POWERS, B.; VOGELI, C.; MULLAINATHAN, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. v. 366, Issue 6464, p. 447-453, Oct. 25, 2019. Available at: https://science.sciencemag.org/content/366/6464/447. Access: 1 Mayo 2020.

[47] Other authors are working on uncovering these new problems. For an example, see: KROLL, J. et al. Accountable Algorithms. *University of Pennsylvania Law Review*. V. 165, Issue 3, 2017. Available at: https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3/. Access: 3 Mayo 2020.

public-private relations – and in the case of a recruitment algorithm used by a private company – i.e. private relations. Unlike other jurisdictions[48], the Brazilian legal system recognizes that the fundamental rights set forth in the Constitution directly apply to horizontal relations: the STF has repeatedly decided in favor of the applicability of such rights to private relations.[49] For instance, in 1996, a ruling by the STF obligated a French multinational company in Brazil to equally apply its internal regulations to all employees (including benefits), regardless of nationality, in light of the prohibition to discrimination which runs from the fundamental right to equality.[50] Fundamental rights immediately confer individuals rights against other individuals – which is an important statement given that, as illustrated, threats to fundamental rights in algorithmic decision-making often emerge from interaction among private parties.

While human decision-making has long been the subject of extensive research regarding biases[51], we are still unveiling the discriminatory outcomes and other problems related to algorithmic decisions. Many of the aforementioned discriminatory results in the context of automation have already been recognized in human decision-making long before algorithms were in operation: gender bias in hiring policies is not novel, nor is the fact that the criminal justice system is biased against black people in many jurisdictions. Therefore, the mechanisms to deal with human biases are more developed when compared to solutions in automated systems. In the legal justice system, for example, one of the methods designed to contain the harmful effects of distortions in human judgement are courts of appeal, which give the parties a chance to review and correct any mistakes that the judge might have made. Similar safety mechanisms for algorithmic decision-making have yet to be devised.

As already pointed, the way we state our concerns has a direct impact on policy design. If we believe the main problem with algorithms relates to the absence of human scrutiny, we will consequently set up regulations banning most automated decisions or requiring each and every automated result to be reviewed by human beings. That seems to be what the European Union's Regulation n. 2016/679, commonly referred to as the General Data Protection Regulation (GDPR)[52], chose to do: in Article 22, it states that individuals have a right not to be subject to a decision based solely on automated processing whenever such decision has legal effects regarding those individuals or similarly significantly affects him or her. It is not difficult to conclude that most decisions indeed have legal effects upon someone or affect that person in some significant way, which means that the European regulation in practice establishes an ex ante ban on many instances of algorithmic use. There are exceptions: (a) if the automated decision is necessary for a contract between

[48] The prevailing understanding of comparative constitutional law is that, in the United States of America, fundamental rights protections are not applicable to private relations. In the words of Mark Tushnet, "standard U.S. constitutional doctrine is that constitutional provisions do not have horizontal effect". TUSHNET, M. The issue of state action/horizontal effect in comparative constitutional law. *International Journal of Constitutional Law*. V. 1, Issue 1, p. 81, Jan. 2003. Available at: https://academic.oup.com/icon/article/1/1/79/671955. Access: 1 Mayo 2020. However, there are several exceptions under the "state action doctrine" that allow the application of fundamentally established values to some private scenarios. For a study on those exceptions, see: GARDBAUM, S. The 'Horizontal Effect' of Constitutional Rights. *Michigan Law Review*, v. 102, 2003. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=437440. Access: 1 Mayo 2020.

[49] Paula Gorzoni investigated the STF cases where horizontal applicability of fundamental rights was expressly debated and concluded that the Court often recognizes fundamental rights in private relations. GORZONI, P. *Supremo Tribunal Federal e a Vinculação dos Direitos Fundamentais nas Relações entre Particulares*. 2007. p. 7-11. Available at: http://www.sbdp.org.br/publication/supremo-tribunal-federal-e-a-vinculacao-dos-direitos-fundamentais-nas-relacoes-entre-particulares/. Access: 1 Mayo 2020.

[50] Recurso Extraordinário n° 161.243-6/DF. Relator: Min. Carlos Velloso. Data do acórdão: 29 de outubro de 1996.

[51] As noted by Ricardo Horta, the question of how to ensure democratic control of judicial decision-making is relatively old in the legal debate, and many empirical studies have already showed how humans are influenced by gender, race, politics and many other factors that go way beyond the established rational parameters of legal decision-making. HORTA, R. Por que existem vieses cognitivos na Tomada de Decisão Judicial? A contribuição da Psicologia e das Neurociências para o debate jurídico. *Revista Brasileira de Políticas Públicas*. V. 9, n° 3, Dec. 2019. Available at: https://www.publicacoes.uniceub.br/RBPP/article/view/6089/pdf. Access: 3 Mayo 2020.

[52] It must be highlighted that the European GDPR and the Brazilian LGPD are legislations that discipline the use of personal data: they do not regulate the use of algorithms in general. Regardless, they are relevant as legal landmarks for the debate carried out in this article, since the most significant uses of algorithms involve personal data. Therefore, the rules brought about by those statutes are largely applicable in algorithmic decision-making.

the individual and the decision-maker; (b) if it is authorized by law, as long as there are safeguards protecting the individual's rights and legitimate interests; or (c) if the individual explicitly consented to the decision process. Even in cases (a) and (c), the GDPR ensures a right to obtain human revision of the outcome.

Since there are multiple situations where algorithmic decision-making is not a problem in itself, as exemplified in Part II of this article, the general prohibition of automated decisions adopted by the GDPR will likely create an unnecessary obstacle to efficiency, without always guaranteeing greater protection for individuals. For instance, if a company wishes to use an algorithm to speed up its recruitment process and analyze thousands of résumés received to select the ones that fulfill minimum requirements (such as level of education and years of experience)[53], it will need a specific legal authorization to do so. In other words, the premise is that automation is not allowed, and it can only be carried out exceptionally. This solution undermines the efficiency gains, which were the goal of the automation in the first place, and does not address the real problem, that is, the lack of clarity and transparency in hiring policies.

Likewise, if the European rationale was applied to the example of public funding of medication in Brazil, and a decision was made for an algorithm to replace humans in the operation of the established legal criteria, a law would have to specifically authorize such automation, even though the algorithm in question would only apply the already sanctioned rules.

The striking misconception here is that efficiency is sacrificed with no guaranteed win to equality or any other fundamental right. The main issue with algorithms identified by the GDPR is the risk of discrimination, which is certainly a noble and justifiable goal. But such risk does not cease with the simple recourse to human scrutiny of each individual decision, and could be better addressed through different measures without impairment of the efficiency gained with automation.

A better policy choice seems to have been made in Brazil, with the enactment of the Federal Law n. 13,709/2018 (in Portuguese, Lei Geral de Proteção de Dados, or LGPD). In Article 6, IX, it states a general prohibition of illicit or abusive discrimination – regardless of whether it is a result of automation or any other kind of decision-making. Because the law only prohibits discriminatory behavior that results in illicit or abusive conduct, it also tacitly allows discriminatory conduct that does not lead to such results. Furthermore, Article 20, which was inspired by the GDPR, sets forth a right to revision of fully automated decisions that affect one's interests, but establishes no ex ante ban, or a requirement that such revision be made by a human. Thus, it appears that a scrutiny of the programming itself would suffice for compliance purposes.[54] Clear information about the decision process and the criteria applied by the algorithm should be provided whenever required, and if that information cannot be shared due to commercial or industrial secrecy reasons, a specialized government agency has the power to audit the algorithm in order to verify any discriminatory outcomes.

Insofar as the LGPD does not create a general ban on automation, it recognizes the possibility that discrimination may take place lawfully and plans to address it through a specialized audit of the algorithm[55], which presents an opportunity to target the problem more accurately than the European regulation. Better than generally forbidding automated decisions is guaranteeing the decision process is understandable and explainable and can be audited by authorities in order to eliminate biases and reduce failures.

This regulatory solution connects with another set of problems in algorithmic decision-making: in face of the possibly faulty nature of algorithms, identifying and correcting the flaw in concrete cases can be hard. In order for someone to verify if the algorithm is working properly, what automation must ensure is some

[53] Many companies use applicant-tracking systems for recruitment and hiring that tend to reject up to 75% of CVs before any human sees them. HOW an algorithm may decide your career. *The Economist*. June 21, 2018. Available at: https://www.economist.com/business/2018/06/21/how-an-algorithm-may-decide-your-career. Access: 3 Mayo 2020.

[54] Since the legislation is very recent, the way courts will interpret and enforce the mentioned legal provisions is still an open question.

[55] The state authority in charge of implementing the federal legislation and conducting the referred audit is yet to be created.

form of transparency or accountability – which is specially challenging when machine learning is involved, since learner algorithms identify correlations beyond human understanding and ultimately program themselves. The case of Deep Patient, the algorithm used at Mount Sinai Hospital in New York to determine patients' risk of developing certain diseases, illustrates this scenario. It anticipated the onset of psychiatric disorders notoriously difficult for doctors to predict, but the designers of the algorithm had no clue how it was able to do that, which severely impairs treatment efforts. As Will Knight wrote, "as the technology advances, we might soon cross some threshold beyond which using AI requires a leap of faith".[56]

Most of today's machine learning uses neural networks to execute decision-making, which can be extremely difficult for non-trained humans to understand. In these cases, transparency is not enough, and thus the literature has been focusing on the concept of explainability as a potential substitute. As Doshi-Velez et al. put it, the idea of explanation or explainability in decision-making refers to the reasons or justifications behind a particular outcome, rather than a general description of the decision-making process. Therefore, what they consider to be an explanation is a "human-interpretable description of the process by which a decision-maker took a particular set of inputs and reached a particular conclusion", which is different than transparency, for being able to understand the process by which a decision was made is not the same as knowing every step taken to reach it.[57]

Many jurisdictions already impose an obligation of transparency and explainability upon government decisions. The Brazilian Constitution states in Article 37 that the public administration must follow the principle of publicity, and the Federal Administrative Procedure Act (Federal Law n. 9,784/1999) states that all decisions by public agents must clarify its motivations and ensure that the individuals affected are able to contest the rulings. As to the Judiciary, Article 93, IX of the Brazilian Constitution requires all rulings and their reasonings to be public. Given these requirements, the use of algorithms by the State in its decision--making processes is from the outset required to follow strict transparency and accountability criteria.

In this context, Danielle Citron presents suggestions to ensure what she calls "technological due process" in the case of governmental use of algorithms. First, she argues government agencies should refrain from automating policies which have not undergone formal or informal rulemaking. Should there be automation, the public ought to participate as much as possible in the building of the system, and government agencies should maintain audit trails to provide individuals with the reasons supporting the system's decision on a case-by-case basis. She also suggests public hearings to clarify automated systems' fallibility, and that systems' source code should be public and tested by independent agents.[58]

The case of the aforementioned tool called COMPAS provides a clear illustration of the problems raised by algorithmic transparency and accountability. The tool was used in 2013 by the State of Wisconsin in the United States of America to sentence a man to six years in prison, based on the algorithm's assessment that he posed a high risk of re-offending. However, the inner workings of the system were not disclosed to the defendant, and were not even accessible to the judge, it being a proprietary algorithm of a private company. Such use of algorithms clearly poses a problem for judicial due process, since it undermines the ability of defendants to fully counter the sentence and question the reasoning behind it.[59]

---

[56] KNIGHT, W. The Dark Secret at the Heart of AI. *MIT Technology Review*. April 11, 2017. Available at: https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/. Access: 3 Mayo 2020.

[57] They go on to say that an explanation should be able to answer at least one of the three following questions: (i) what were the main factors in a decision? (ii) Would changing a certain factor have changed the decision? (iii) Why did two similar-looking cases yield different decision, or vice-versa? DOSHI-VELEZ, F.; KORTZ, M. *Accountability of AI Under the Law: The Role of Explanation*. Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society working paper, 2017, p. 2-3. Available at: https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf?sequence=3. Access: 3 Mayo 2020.

[58] CITRON, D. Technological Due Process. *Washington University Law Review*. V. 85, Issue 6, p. 1303, 2008. Available at: http://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2. Access: 1 Mayo 2020.

[59] UNITED STATES OF AMERICA. Loomis v. Wisconsin: 881 N.W.2d 749 (Wis. 2016), cert. denied, 137 S.Ct. 2290 (2017). Despite the difficulties related to due process, the Wisconsin Supreme Court ruled against Loomis, arguing the sentence would

When it comes to the use of algorithms in private relations, transparency will be a relevant aspect of decision-making in order to ensure individuals can verify whether or not their rights are being observed. The question subsequently becomes when a right effectively exists in a concrete case, because it is then, and only then, that transparency will come into play. In the case of Optum's healthcare algorithm, for instance, New York's Department of Financial Services is currently investigating UnitedHealth Group for discriminatory practices and demanded the company demonstrates the algorithm is not racially biased.[60]

However, there are several scenarios where automation is present, but the existence of a right is questionable at best. For instance, in the case of advertisement, we do not need to go very far to understand that much of the decision-making regarding ads was always a process to which individuals outside of specific Madison Avenue offices were not privy to. Advertisement agencies in fact specialized in finding the best way to convey the messages their clients wished, without ever asking for anyone's consent. They already based their decision-making on clients' data - though naturally the amount of information available about each person was severely limited. Ads for cars, for example, varied a lot depending on what the target audience was perceived to be, both in terms of gender and social status. The process was the opposite of transparent, precisely because the method by which decisions were reached usually represented the real asset of the firm. But the goal was always the same: find your target audience and use the best language to reach it.

Coming back to 2020, when advertisement is largely delivered online by automated algorithms, we see a similar pattern. Ads are still targeted to specific audiences, and they are still based on what these people are perceived to enjoy. The main difference is that the amount of data available is vastly larger, which allows for different targeting strategies, more precise and more tailored. Also, rather than letting Mad Men decide who the best audience for a given product is, companies delegate that decision to algorithms.

The question that remains is if this change in process requires any adaptations in terms of procedure. The answer will be positive when we think about, for example, data collection. The way by which companies collect individuals' data to feed their algorithms is relevant, and needs to be scrutinized, which is why data protection legislation is largely focused on mechanisms for consent. But assuming such process was carried out properly, according to novel data protection legislation, is there still a right for the individual to require transparency from algorithmic decision-making?

The answer to that question can be negative for one simple reason: the lack of transparency in many advertising contexts will not correspond to a right's violation. If a search engine creates a profile to track my behavior and direct advertisements to me, and for whatever reason its algorithms come to the conclusion I enjoy spending my holidays at the beach, as long as the only outcome is an increased number of advertisements for trips to Rio de Janeiro showing up on my browser, rather than ads for trips to Campos do Jordão, it is very hard to sustain a right was violated. Despite that, as mentioned earlier, according to Article 20 of the LGPD, individuals may require revision of automated decisions which affect their interests, including the ones that establish consumer profiles.

The scenario could be different, however, if the advertisement in question related to job opportunities. Due to the relevance of the matter advertised – employment –, mistakes in the algorithm are more likely to entail rights violations, in which case transparency and explainability become crucial to assess if the decision-making process is lawful. In conclusion, there is a close relationship between how understandable an

have been the same had COMPAS never been consulted. However, it established that "risk scores may not be used to determine whether an offender is incarcerated [or] to determine the severity of the sentence" and that it could never be the determining factor in sentencing. For an analysis of the decision, see: STATE V. LOOMIS, 881 N.W.2d 749 (Wis. 2016). *Harvard Law Review*. V. 130, March 10, p. 1530-1537, 2017. Available at: https://harvardlawreview.org/wp-content/uploads/2017/03/1530-1537_online.pdf. Access: 3 Mayo 2020.

[60] EVANS, M and MATHEWS, A. New York Regulator Probes UnitedHealth Algorithm for Racial Bias. *The Wall Street Journal*. October 26, 2019. Available at: https://www.wsj.com/articles/new-york-regulator-probes-unitedhealth-algorithm-for-racial-bias-11572087601. Access: 3 Mayo 2020.

algorithm must be and how relevant the automated decision is to the individuals affected. Whenever there is a right at stake, explanations will be needed regarding the inner workings of the system in order to ensure it does not result in biases or inequalities.

In this context, Wachter, Mittelstadt and Russell have pointed out that algorithmic discrimination is significantly different than human discrimination: it can be more widespread and at the same time more abstract, subtle and indirect, often based on complex patterns and correlations drawn by the computer. This reality requires new methods of detection and remedy. To bridge the "accountability gap", the authors suggest a specific method of producing statistical evidence to be used by developers in pre-emptive audits and also before European regulators and courts in investigations and rulings regarding discrimination claims.[61]

Coming to the end of this section, it is now hopefully clearer what the problems with algorithmic decisions usually are and what they are not. The next section will summarize the main arguments put forward in this article through an illustrative analysis of a concrete situation: the covid-19 pandemic.

4 **Conclusion:** algorithms and the case of ICU beds during the covid-19 pandemic

To cement the assertions made in this article, this section will recap the main arguments in light of the present challenges to decision-making that are emerging from the current covid-19 pandemic.

In late 2019/early 2020, a new virus emerged in the planet. It spread quickly, and though it did not cause equally severe symptoms on all contaminated individuals, many started to go into acute respiratory failure that led to the need of hospitalization and, in many cases, resulted in death. The disease caused by the new virus, later named Sars-CoV-2, is now called covid-19, and gave rise to a pandemic that, as of July 11, 2020, had claimed the lives of over 560 thousand people around the globe, and contaminated over 12.5 million.[62]

The group most affected by covid-19 is the elderly. Given that their immunity is already more compromised, they are leading the numbers of deaths in all countries: over 95% of the deceased are older than 60 years; more than half of all deaths occur among people aged 80 years or older. But as it has become abundantly clear, covid-19 is not exclusively dangerous for the elderly: 8 out of 10 deaths are occurring in individuals with chronic underlying conditions, particularly those with cardiovascular diseases/hypertension and diabetes. It is estimated that 10% to 15% of people under 50 years old who become infected have moderate to severe infection.[63] The data in the United States of America – which in early July 2020 was the country

[61] The proposed method is called "Conditional Demographic Disparity" (CDD) and it works as a baseline statistical measurement designed to align with European Union standards – though the authors make it clear that it does not replace judicial interpretation of discrimination claims. For details about the method, see: WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*. March 3, 2020. Available at https://ssrn.com/abstract=3547922. Access: 1 Mayo 2020. p. 54-64.

[62] The numbers are tracked on a global scale by the World Health Organization (WHO) and can be found at: WORLD HEALTH ORGANIZATION. *WHO Health Emergency Dashboard*. Available at: https://covid19.who.int/. Access: July 11, 2020. The WHO's statement regarding the compared virus lethality can be found at: WORLD HEALTH ORGANIZATION. *WHO Director-General's opening remarks at the media briefing on COVID-19 - 13 April 2020*. April 13, 2020. Available at: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19--13-april-2020. Access: 3 Mayo 2020. The H1N1 pandemic, which first started as an outbreak in April 2009, resulted in over 18,000 deaths worldwide as of August 2010, and moved to the post-pandemic period in September 2010. WHO's reports on the H1N1 pandemic can be found at: WORLD HEALTH ORGANIZATION. *Situation updates - Pandemic (H1N1) 2009*. Available at: https://www.who.int/csr/disease/swineflu/updates/en/. Access: 3 Mayo 2020.

[63] WORLD HEALTH ORGANIZATION'S REGIONAL OFFICE FOR EUROPE. *Statement – Older people are at highest risk from COVID-19, but all must act to prevent community spread*. Copenhagen, April 2, 2020. Available at: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/statements/statement-older-people-are-at-highest-risk-from-covid-19,-but-all-must-act-to-prevent-community-spread. Access: 3 Mayo 2020.

with the greatest numbers of cases and deaths – indicates that hospitalization of younger people is relevant: between March 1st, 2020, and April 11th, 2020, the cumulative hospitalization rate among individuals aged 18 to 49 years old was 10 per 100,000, compared to an overall rate of 20 per 100,000.[64]

Because covid-19 impacts the lungs, it is common for patients to need help breathing and oxygenating. Therefore, hospitalization usually means admission in intensive care units (ICUs), where ventilators are available. Globally, the need for intensive care unit admission is present in approximately 10% of all infection cases.

The reality of this scenario first became clear in Italy, the European epicenter of the pandemic. The country failed to implement sufficient measures to slow down the spread of the virus in its early stages, which caused severe damage. The region of Lombardy, where the city of Milan is located, was the first to be gravely impacted. In March 3rd, 2020, there were already reports of the collapse of the healthcare system.[65] Twenty days later, healthcare professionals were faced with a tragic challenge, having to choose which patients would receive care and which would not, given that there were not enough intensive care beds available for all.[66]

The problem is not exclusive of Europe. Given that the virus is new and populations have yet to develop immunity to it, and considering its high transmission rate, a study by the Imperial College London warned that, without any control measures to slow down the spread, the demand for ICU beds would be, at its peak, thirty times greater than the number of beds available both in the United States of America and in Great Britain.[67] In other words, there would be a collapse, and the countries' health systems would simply be unable to handle the patient intake. The problem is even more serious if we consider that the existing beds are already used to provide assistance to patients who need it for reasons other than covid-19.

Since some countries have failed to take sufficient measures to minimize contagion and thus hospitalization, their healthcare systems have collapsed, and became unable to admit all patients in need of care. Such was the case in Italy, but it is also the case of some regions in Brazil[68], and possibly of the entire country, depending on how the scenario evolves in the coming months.

Amidst this scenario, health care professionals are faced with the need to decide to whom they will provide assistance. In Brazil, the situation is even worse, because there was already a shortage of ICU beds prior to the spread of the coronavirus.[69] As Daniel Neves Forte, president of the Bioethics Committee at Sírio-

[64] According to the numbers available on April 21, 2020. The data are constantly updated as more information from hospitalizations become available. For the most recent number in the United States of America, see: CENTERS FOR DISEASE CONTROL AND PREVENTION. U.S. Department of Health & Human Services. *Laboratory-Confirmed COVID-19-Associated Hospitalizations*. Available at: https://gis.cdc.gov/grasp/covidnet/COVID19_3.html. Access: April 21, 2020.

[65] AL-ARSHANI, S. The healthcare system in Italy's Lombardy region is so strained from the new coronavirus that officials are asking doctors to come out of retirement and nursing students are being fast-tracked to graduation. *Business Insider*. March 3, 2020. Available at: https://www.businessinsider.com/italys-lombardy-regions-healthcare-system-is-crumbling-to-covid-19-2020-3. Access: 3 Mayo 2020.

[66] According to reports of doctors in the city of Bergamo, the health system went "far beyond the tipping point": most intensive care unit beds were being reserved for coronavirus patients with "a reasonable chance to survive", while older patients stopped being resuscitated. HOROWITZ, J.; KIRKPATRICK, D. Dip in Italy's Cases Does Not Come Fast Enough for Swamped Hospitals. *The New York Times*. March 23, 2020. Available at: https://www.nytimes.com/2020/03/23/world/europe/italy-coronavirus-hospitals.html. Access: 3 Mayo 2020.

[67] FERGUSON, N.; LAYDON, D. et al. *Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. Imperial College London. March 16, 2020. p. 7. Available at: https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf. Access: 3 Mayo 2020.

[68] Even though efforts are being made to expand the number of beds available, the states of Amazonas, Pará, Rio de Janeiro and Pernambuco already have occupancy rates of over 90%. PITOMBO, J.; BARBON, J. et al. Mesmo com novos leitos para Covid-19, UTIs brasileiras têm alta taxa de ocupação. *Folha de S. Paulo*. April 28, 2020. Available at: https://www1.folha.uol.com.br/equilibrioesaude/2020/04/mesmo-com-novos-leitos-para-covid-19-utis-brasileiras-tem-alta-taxa-de-ocupacao.shtml. Access: 4 Mayo 2020.

[69] RACHE, B.; ROCHA, R. et al. *Necessidades de Infraestrutura do SUS em Preparo à COVID-19:* Leitos de UTI, Respiradores e Ocupação Hospitalar. Instituto de Estudos para Políticas de Saúde. Nota Técnica n° 3. March 2020. Available at: https://ieps.org.br/wp-content/uploads/2020/04/IEPS-NT3.pdf. Access: 5 Mayo 2020.

-Libanês, one of the leading healthcare complexes in Brazil, has stated, every doctor who works in Brazilian ICUs makes "Sophie's choices" everyday; sometimes, there are ten or even twenty patients pledging for a single hospital bed.[70]

In trying to address the problem, one of the suggested alternatives has been the development of algorithms that help triage and rank patients, determining priorities in ICU admissions. According to Forte, he was part of a research conducted by the Medical School at the University of São Paulo that aimed at developing one such tool, with the goal of bringing transparency to the decision-making process. As reported by the newspaper Folha de São Paulo, the proposed tool uses guidelines and criteria approved by the Federal Board of Medicine (in Portuguese, Conselho Federal de Medicina) and the Brazilian Association for Intensive Medicine (in Portuguese, Associação Brasileira de Medicina Intensiva), and works from a standard questionnaire filled by the doctor requesting the ICU bed, which includes, for example, information on whether the patient has any underlying conditions.

In light of the questions raised in the previous sections, the first problem that appears is not truly a problem with algorithms, but will likely be indistinctively merged with the discussion of algorithmic use, and therefore should be urgently identified and separated, if it is to be correctly addressed. Reports from the covid-19 pandemic in Europe revealed scenarios where a patient with heart disease and on the wait for a kidney transplant contracted the coronavirus and had his death precipitated by a doctor's decision to give the available ventilator to a younger coronavirus patient who had better chances of surviving. The family members of people who faced similar situations – and ultimately died from not receiving medical attention – raised the question of whether citizens were not wrongfully abandoned by the state, which should have been more prepared.[71]

On the one hand, doctors highlight the scarcity of beds and medical supplies and the need to established guidelines to employ the existing resources as well as possible. Dr. Marco Vergano, for example, an anesthesiologist based in Turin and chairman of an ethics committee for his medical specialty, was responsible for drafting emergency recommendations for Italian doctors allocating scarce beds among the increasing number of patients, and stressed that the focus on the needs of the community, rather than on one single patient, is a recommendation many doctors are following. On the other hand, people with disabilities have contested existing regulations in the United States that allows for medical care to be rationed in detriment of their survival.[72]

This debate brought about by the covid-19 pandemic is, first and foremost, a dispute on whether there is an individual right to treatment, and what this right would entail in concrete scenarios. Like the controversy related to the public funding of medication in Brazil, the problem is not with the possible use of algorithms in decision-making, but with the generalizing decision itself, which excludes individuals from the needed

[70] COLLUCCI, C. Algoritmos e inteligência artificial podem ajudar Brasil a decidir sobre leitos de UTI. *Folha de S. Paulo.* March 24, 2020. Available at: https://www1.folha.uol.com.br/equilibrioesaude/2020/03/algoritmos-e-inteligencia-artificial-podem-ajudar-brasil-a-decidir-sobre-leitos-de-uti.shtml. Access: 1 Mayo 2020.

[71] As reported by the New York Times, the Spanish sports director of a Honda motorcycling racing team, Oscar Haro, released a video in which he said that his father had died after not being given a respirator: "I cannot understand how a person like my father who had been working since he was 15, paying into the system, died because there are no respirators, because they could not treat him anymore," Mr. Haro said. He blamed the Spanish authorities for failing to prepare adequately, even after witnessing the spread of the coronavirus in China and Italy. HOROWITZ, J.; KIRKPATRICK, D. Dip in Italy's Cases Does Not Come Fast Enough for Swamped Hospitals. *The New York Times*. March 23, 2020. Available at: https://www.nytimes.com/2020/03/23/world/europe/italy-coronavirus-hospitals.html. Access: 1 Mayo 2020.

[72] Ariella Barker, a fellow at the Harvard Kennedy School of Government, states that "the existence of draconian laws that call for medical rationing during a public health emergency makes it clear that I will certainly be left to die if our hospitals reach capacity". She argues against a Tennessee regulation that allows for people with Ehlers-Danlos syndrome to be legally denied critical care if they require assistance with activities of daily living. "It must be clear to us and the nation that disabled lives matter. Anything less is an undeniable statement that our lives are expendable and our mental health is irrelevant". BARKER, A. Those with disabilities have a right to survive the coronavirus pandemic. *The Boston Globe*. March 30, 2020. Available at: https://www.bostonglobe.com/2020/03/30/opinion/those-with-disabilities-have-right-survive-coronavirus-pandemic/. Access: 4 Mayo 2020.

medical attention. The public health systems around the globe and in Brazil are already taking emergency measures to multiply the number of available beds[73], but ultimately, as with the STF's ruling for patients with rare and ultrarare medical conditions, the Judiciary could decide that the government is obliged to make room for all patients who need care, which could mean the construction of more field hospitals, for example.

Aside from this first problem, other issues arise. Dr. Vergano claims, for example, that admitting a person with fewer chances of surviving in the ICU in detriment of healthier patients would be immoral. On the other hand, Roger Severino, the director of the Office for Civil Rights at the United States Department of Health and Human Services, claims medical providers cannot discriminate on the basis of disabilities, race or age when deciding who would receive lifesaving medical care during the coronavirus pandemic, and that civil rights laws "protect the equal dignity of every human life from ruthless utilitarianism".[74]

Here, the main issue is what ethical standard is embraced by the legal system and should be applied to solve the current crisis. Because this debate is absolutely central, this question ought to be answered before any algorithm is put in place. If we come to the conclusion that the utilitarian approach is to be adopted, then algorithms like the one suggested by Daniel Forte could be successfully employed, since they would simply use the information provided by the questionnaire regarding the patient's health condition and allow for a faster and more standardized calculation of the patient's odds of surviving. In terms of precision, machine learning would probably be even more efficient than Forte's algorithm: as already mentioned, learner algorithms have proved successful in predicting deaths in hospitalized patients with extreme accuracy, beyond human doctors' understanding. If, however, we reject the utilitarian notion that the available resources should be used to maximize the number of survivors, regardless of the costs to individual groups, and opt for a case-by-case judgement call by the doctor in the field, for instance, algorithms would be unsuited, since prudence and fairness cannot be taught to computers.

There is yet another observation to be made regarding this issue. Whether or not the ethical debate results in the adoption of algorithms, it is important to highlight that the decision may not be even for all covid-19 patients. In Brazil, the healthcare system is both public and private – public care is handled by the SUS, whereas private care is handled by hospital and facilities not managed by the state. In a scenario where poorer people are subject to strictly utilitarian calculations and richer patients are able to obtain the needed medical treatment in a case-by-case analysis of their individual needs and desires, there would be a severe violation to the fundamental right to equality, likely resulting in discrimination against the poor.[75]

A third issue arises. As the virus spreads and death rates rises, it becomes clearer that an individual's chances of surviving the infection are closely related to race. In the state of Illinois, where African-Americans make only 15% of the total population and 28% of the people infected by the coronavirus, they correspond to 43% of covid-19 deaths so far. In Michigan, African-Americans are 14% of the State's population, but correspond to nearly 33% of the infected and 40% of the deceased. Another example is Louisiana: a third

[73] Many State governments in Brazil have constructed field hospitals, hired more healthcare professionals and purchased ventilators and other necessary medical equipment to temporarily expand the capacity of the public health system to deal with the virus spread. For an example of these measures, see: RUPRECHT, T. Hospitais de campanha: como vão funcionar e por que são tão importantes. *Saúde*. April 9, 2020. Available at: https://saude.abril.com.br/medicina/hospitais-de-campanha-como-vao-funcionar/. Access: May 5, 2020.

[74] FINK, S. U.S. Civil Rights Office Rejects Rationing Medical Care Based on Disability, Age. *The New York Times*. March 28, 2020. Available at: https://www.nytimes.com/2020/03/28/us/coronavirus-disabilities-rationing-ventilators-triage.html. Access: 4 Mayo 2020.

[75] Virgina Eubanks highlights that "we all inhabit this new regime of digital data, but we don't all experience it in the same way". She argues marginalized groups are much more vulnerable to automation that reinforces exclusion, in a "feedback loop of injustice". While she calls for a deep change in the cultural understandings and political responses to poverty, she also proposes some palliative principles to orient technological designs for now. Engineers and data scientists involved in the creation of automated systems should answer two questions to assess social and economic implications of their designs: "Does the tool increase the self-determination and agency of the poor? Would the tool be tolerated if it was targeted at non-poor people?". EUBANKS, V. *Automating Inequality: how high-tech tools profile, police and punish the poor.* New York: St. Martin's Press, 2018. p. 10-11, 168.

of its population is African-American, but up to 70% of the people how died from covid-19 are African--Americans.[76] In Brazil, the statistics are also disturbing: in a two-week period in April, there was a fivefold rise in the number of deaths by covid-19 among black people, compared to a threefold increase in deaths among white people in the same interval. A similar discrepancy was found in the number of hospitalizations caused by the virus in that timespan.[77]

Although the numbers are preliminary, many specialists have argued that they reveal the inequalities that have historically structured the American and Brazilian societies. Jason Silverstein, lecturer and writer-in--residence in the department of global health and social medicine at Harvard Medical School, highlights that "black people in the United States have, since slavery, been systematically overexposed to health risks while also being deprived of healthcare".[78] According to him, many studies point to racial health disparities in chronic illnesses such as asthma, arthritis, aggressive breast cancer, kidney disease, heart failure, maternal mortality, lung disease, chemical exposures, and overall life expectancy – and many of those conditions severely reduce the patient's ability to fight the covid-19.

Given this scenario, a supposedly "neutral" algorithm designed to save as many lives as possible could ultimately result in racial bias in many regions, since black people have overall less access to healthcare than white people, and thus are more likely to have underlying health conditions that would motivate the refusal of an ICU bed, when compared to white patients. Similarly to the case of the algorithm provided by Optum, even when the algorithm does not explicitly considers race, the correlations it is based on might result in racial disparity, which in turn possibly results in a violation to equality.

The final problem relates to transparency. When it comes to the algorithm proposed by Forte, this issue seems to be correctly addressed, since the criteria set forth are the same already put forward by the CFM and the AMIB, which are currently used by attending physicians, and the objective is precisely bringing transparency to the decision-making process. However, machine learning algorithms – which as mentioned could be even better at determining the most efficient way of employing existing resources to save the higher possible number of lives –, would definitely pose a challenge to transparency and explainability. When this technology is applied, as already pointed, computers program themselves, and humans, even specialists, can have trouble explaining the reasoning adopted by the machine. This obstacle could severely impair scrutiny efforts to verify whether or not it reproduces any biases or violates equality, unless explainability is designed into the system. The effort to establish "explainability by design" has been discussed elsewhere,[79] but faces a curious challenge. Some authors argue that there is a trade-off between complexity and accuracy, meaning that the more complex a system is, the more accurate its decision-making process. In other words, making a system understandable and preserving accuracy requires a method other than simplification.[80]

The main goal of this article was to identify and detail the concerns around algorithmic decision-making, highlighting their specificities so that regulation and the public debate regarding the employment of algori-

[76] NEGROS enfrentam índices alarmantes de contaminação pelo coronavírus nos EUA. *Folha de S. Paulo*. April 8, 2020. Available at: https://www1.folha.uol.com.br/mundo/2020/04/negros-enfrentam-indices-alarmantes-de-contaminacao-pelo-coronavirus-nos-eua.shtml. Access: 4 Mayo 2020.

[77] MUNIZ, B.; FONSECA, B.; PINA, R. Em duas semanas, número de negros mortos por coronavírus é cinco vezes maior no Brasil. *Pública*. May 6, 2020. Available at: https://apublica.org/2020/05/em-duas-semanas-numero-de-negros-mortos-por-coronavirus-e-cinco-vezes-maior-no-brasil/. Access: 26 Mayo 2020.

[78] SILVERSTEIN, J. Being Black in America Is a Health Risk. It's Time for Reparations. *Vice*. June 19, 2019. Available at: https://www.vice.com/en_us/article/5973yq/the-healthcare-case-for-reparations-hr40?utm_content=1586356292&utm_medium=social&utm_source=VICE_twitter. Access: 4 Mayo 2020.

[79] For example, LEE et al. An explainable deep-learning algorithm for the detection of acute intracranial hemorrhage from small datasets. *Nature Biomedical Engineering*. V. 3, Issue 3, p. 173-182, 2019. Available at: https://www.nature.com/articles/s41551-018-0324-9.pdf?origin=ppub. Access: July 11, 2020.

[80] KLEINBERG, J.; MULLAINATHAN, S. Simplicity Creates Inequity: Implications for Fairness, Stereotypes, and Interpretability. *arXiv Cornell University*. Article n° 1809.04578v1 [cs.LG]. August 2018. Available at: https://arxiv.org/abs/1809.04578v1. Access: 4 Mayo 2020.

thms can adequately address them. In healthcare and in policy design more broadly, successful treatments are largely dependent on solid diagnoses. Hopefully, this article will contribute to better policies in the present covid-19 crisis but also in upcoming challenges related to the evermore pervasive use of algorithmic systems for decision-making.

References

AGRAWAL, A.; GANS, J. and GOLDFARB, A. Prediction Machines: The Simple Economics of Artificial Intelligence. *Harvard Business Press*, Apr. 17, 2018.

AL-ARSHANI, S. The healthcare system in Italy's Lombardy region is so strained from the new coronavirus that officials are asking doctors to come out of retirement and nursing students are being fast-tracked to graduation. *Business Insider*. Mar. 3, 2020. Available at: https://www.businessinsider.com/italys-lombardy-regions-healthcare-system-is-crumbling-to-covid-19-2020-3. Access: 3 Mayo 2020.

ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. Machine Bias. *ProPublica*, Mayo 23, 2016. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Access: 1 Mayo 2020.

ARKIN, R.; ULAM, P. and WAGNER, A. Moral Decision-making in Autonomous Systems: Enforcement, Moral Emotions, Dignity, Trust and Deception. *Proceedings of the IEEE*. v. 100, Issue 3, p. 571-589, Mar. 2012.

ASARO, P. On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross*. v. 94, Issue 886, p. 687709, June 2012. Available at: https://international-review.icrc.org/sites/default/files/irrc-886-asaro.pdf. Access: 1 Mayo 2020.

BARKER, A. Those with disabilities have a right to survive the coronavirus pandemic. *The Boston Globe*. March 30, 2020. Available at: https://www.bostonglobe.com/2020/03/30/opinion/those-with-disabilities-have-right-survive-coronavirus-pandemic/. Access: Mayo 4, 2020.

BAROCAS, S. and SELBST, A. Big Data's Disparate Impact. *California Law Review*. V. 104, n. 3, p. 671-732, Jun. 2016. Available at: http://www.californialawreview.org/wp-content/uploads/2016/06/2Barocas-Selbst.pdf. Access: 1 Mayo 2020.

BURANYI, S. How to persuade a robot that you should get the job. *The Guardian*. March 4, 2018. Available at: https://www.theguardian.com/technology/2018/mar/04/robots-screen-candidates-for-jobs-artificial-intelligence. Access: 1 Mayo 2020.

BURT, A. and VOLCHENBOUM, S. How Health Care Changes When Algorithms Start Making Diagnoses. *Harvard Business Review*, Mayo 08, 2018. Available at: https://hbr.org/2018/05/how-health-care-changes-when-algorithms-start-making-diagnoses. Access: 1 Mayo 2020.

BYGRAVE, L. Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling. *Privacy Law & Policy Reporter*, v. 7, p. 67–76, 2000.

CENTERS FOR DISEASE CONTROL AND PREVENTION. U.S. Department of Health & Human Services. *Laboratory-Confirmed COVID-19-Associated Hospitalizations*. Available at: https://gis.cdc.gov/grasp/covidnet/COVID19_3.html. Access: April 21, 2020.

CITRON, D. Technological Due Process. *Washington University Law Review*. Volume 85, Issue 6, 2008. Available at: http://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2. Access: 1 Mayo 2020.

COLLUCCI, C. Algoritmos e inteligência artificial podem ajudar Brasil a decidir sobre leitos de UTI. *Fo-*

lha de S. Paulo. March 24, 2020. Available at: https://www1.folha.uol.com.br/equilibrioesaude/2020/03/algoritmos-e-inteligencia-artificial-podem-ajudar-brasil-a-decidir-sobre-leitos-de-uti.shtml. Access: 1 Mayo 2020.

CONGRESSO NACIONAL. *Votação do dispositivo 24.19.001 - § 3º do art. 20 da Lei nº 13.709, de 14 de agosto de 2018, com a redação dada pelo art. 2º do projeto*. October 2, 2019. Available at: https://www.congressonacional.leg.br/materias/vetos/-/veto/detalhe/12445/1. Access: 1 Mayo 2020.

CONSELHO NACIONAL DE SAÚDE. *Manual operacional para comitês de ética em pesquisa*. 4. ed. Brasília: Editora do Ministério da Saúde, 2008. Available at: http://bvsms.saude.gov.br/bvs/publicacoes/manual_operacional_comites_pesquisa_4ed.pdf. Access: 1 Mayo 2020.

CORMEN, T. H., Algorithms Unlocked. MIT Press, 2013.

DAVIS, N. AI equal with human experts in medical diagnosis, study finds. *The Guardian*. September 24, 2019. Available at: https://www.theguardian.com/technology/2019/sep/24/ai-equal-with-human-experts-in-medical-diagnosis-study-finds. Access: 1 Mayo 2020.

DOMINGOS, P. *Master Algorithm*. Basic Books Inc. New York, 2018.

DOSHI-VELEZ, F. and KORTZ, M. *Accountability of AI Under the Law: The Role of Explanation*. Berkman Klein Center Working Group on Explanation and the Law, Berkman Klein Center for Internet & Society working paper, 2017. Available at: https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf?sequence=3. Access: Mayo 3, 2020.

EUBANKS, V. *Automating Inequality: how high-tech tools profile, police and punish the poor*. New York: St. Martin's Press, 2018.

EVANS, M and MATHEWS, A. New York Regulator Probes UnitedHealth Algorithm for Racial Bias. *The Wall Street Journal*. October 26, 2019. Available at: https://www.wsj.com/articles/new-york-regulator-probes-unitedhealth-algorithm-for-racial-bias-11572087601. Access: Mayo 3, 2020.

FEDERAL TRADE COMMISSION. *Big Data, A Tool for Inclusion or Exclusion? – Understanding the Issues*. January, 2016. Available at: https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf. Access: 1 Mayo 2020.

FERGUSON, A. *The rise of big data policing: surveillance, race, and the future of law enforcement*. New York: New York University Press, 2017.

FERGUSON, N.; LAYDON, D. et al. *Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. Imperial College London. March 16, 2020. Available at: https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf. Access: Mayo 3, 2020.

FERREIRA, F. Inteligência artificial atua como juiz, muda estratégia de advogado e 'promove' estagiário. *Folha de S. Paulo*. March 10, 2020. Available at: https://www1.folha.uol.com.br/poder/2020/03/inteligencia-artificial-atua-como-juiz-muda-estrategia-de-advogado-e-promove-estagiario.shtml. Access: 1 Mayo 2020.

FINK, S. U.S. Civil Rights Office Rejects Rationing Medical Care Based on Disability, Age. *The New York Times*. March 28, 2020. Available at: https://www.nytimes.com/2020/03/28/us/coronavirus-disabilities-rationing-ventilators-triage.html. Access: Mayo 4, 2020.

GARDBAUM, S. The 'Horizontal Effect' of Constitutional Rights. *Michigan Law Review*, v. 102, 2003. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=437440. Access: 1 Mayo 2020.

GORZONI, P. *Supremo Tribunal Federal e a Vinculação dos Direitos Fundamentais nas Relações entre Particulares*. 2007. Available at: http://www.sbdp.org.br/publication/supremo-tribunal-federal-e-a-vinculacao-dos-direitos-fundamentais-nas-relacoes-entre-particulares/. Access: 1 Mayo 2020.

GUPTA, A. Are Algorithms Sexist? *The New York Times*. November 15, 2019. Available at: https://www.nytimes.com/2019/11/15/us/apple-card-goldman-sachs.html. Access: 1 Mayo 2020.

HARRIS, D. U.S. experiences with racial and ethnic profiling: history, current issues, and the future. *Critical Criminology*. v. 14, Issue 3, p. 213–239, Sep. 2006.

HARWELL, D. A face-scanning algorithm increasingly decides whether you deserve the job. *The Washington Post*. November 6, 2019. Available at: https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/. Access: 1 Mayo 2020.

HOROWITZ, J.; KIRKPATRICK, D. Dip in Italy's Cases Does Not Come Fast Enough for Swamped Hospitals. *The New York Times*. March 23, 2020. Available at: https://www.nytimes.com/2020/03/23/world/europe/italy-coronavirus-hospitals.html. Access: 1 Mayo 2020.

HORTA, R. Por que existem vieses cognitivos na Tomada de Decisão Judicial? A contribuição da Psicologia e das Neurociências para o debate jurídico. *Revista Brasileira de Políticas Públicas*. v. 9, n. 3, Dec. 2019. Available at: https://www.publicacoes.uniceub.br/RBPP/article/view/6089/pdf. Access: Mayo 3, 2020.

HOW an algorithm may decide your career. *The Economist*. June 21, 2018. Available at: https://www.economist.com/business/2018/06/21/how-an-algorithm-may-decide-your-career. Access: Mayo 3, 2020.

HUET, E. Server And Protect: Predictive Policing Firm PredPol Promises To Map Crime Before It Happens. *Forbes*. February 11, 2015. Available at: https://www.forbes.com/sites/ellenhuet/2015/02/11/predpol-predictive-policing/#229530d34f9b. Access: 1 Mayo 2020.

INTERNATIONAL ASSOCIATION OF PRIVACY PROFESSIONALS; THE UNITED NATIONS GLOBAL PULSE. *Building Ethics into Privacy Frameworks for Big Data and AI*. October 2018. Available at: https://iapp.org/media/pdf/resource_center/BUILDING-ETHICS-INTO-PRIVACY-FRAMEWORKS-FOR-BIG-DATA-AND-AI-UN-Global-Pulse-IAPP.pdf. Access: 1 Mayo 2020.

KIEFER, S. Morte de pessoas fora da faixa de risco acende alerta para o H1N1. *Estado de Minas*. Mayo 27, 2016. Available at: https://www.em.com.br/app/noticia/gerais/2016/05/27/interna_gerais,766671/alerta-fora-do-grupo-de-risco.shtml. Access: 1 Mayo 2020.

KLEINBERG, J.; MULLAINATHAN, S. Simplicity Creates Inequity: Implications for Fairness, Stereotypes, and Interpretability. *arXiv Cornell University*. Article n° 1809.04578v1 [cs.LG]. August 2018. Available at: https://arxiv.org/abs/1809.04578v1. Access: 4 Mayo 2020.

KNIGHT, W. The Apple Card Didn't 'See' Gender — and That's the Problem. *Wired*. November 19, 2019. Available at: https://www.wired.com/story/the-apple-card-didnt-see-genderand-thats-the-problem/. Access: 1 Mayo 2020.

KNIGHT, W. The Dark Secret at the Heart of AI. *MIT Technology Review*. April 11, 2017. Available at: https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/. Access: 3 Mayo 2020.

KROLL, J. et al. Accountable Algorithms. *University of Pennsylvania Law Review*. v. 165, Issue 3, 2017. Available at: https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3/. Access: Mayo 3, 2020.

MAYER-SCHÖNBERGER, V. and CUKIER, K. *Big Data: A Revolution That Will Transform How We Live, Work, Think*. Houghton Mifflin Harcourt, 2013.

MAYER-SCHÖNBERGER, V. and CUKIER, K. *The Rise of Big Data: How It's Changing the Way We Think*. Foreign Affairs, v. 92, n. 3, Mayo/Jun. 2013. Available at: https://www.jstor.org/stable/23526834?seq=1#page_scan_tab_contents. Access: 1 Mayo 2020.

MINISTÉRIO DA SAÚDE. *Vacinação contra a gripe: entenda se você faz parte do grupo prioritário*. Mayo 3, 2016. Available at: http://www.blog.saude.gov.br/alh5s1. Access: 1 Mayo 2020.

MIOTTO, R.; LI, L.; KIDD, B.; DUDLEY, J. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Scientific Reports* 6, Article n° 26094, 2016. Available at: https://www.nature.com/articles/srep26094. Access: 1 Mayo 2020.

MORTES por gripe H1N1 na região têm 33% de vítimas fora do grupo de risco. *G1 Campinas e Região*. Mayo 13, 2016. Available at: http://g1.globo.com/sp/campinas-regiao/noticia/2016/05/mortes-por-gripe-h1n1-na-regiao-tem-33-de-vitimas-fora-do-grupo-de-risco.html. Access: 1 Mayo 2020.

MUNIZ, B.; FONSECA, B.; PINA, R. Em duas semanas, número de negros mortos por coronavírus é cinco vezes maior no Brasil. *Pública*. Mayo 6, 2020. Available at: https://apublica.org/2020/05/em-duas-semanas-numero-de-negros-mortos-por-coronavirus-e-cinco-vezes-maior-no-brasil/. Access: 26 Mayo 2020.

NATARAJAN, S.; NASIRIPOUR, S. Viral Tweet About Apple Card Leads to Goldman Sachs Probe. *Bloomberg*. November 9, 2019. Available at: https://www.bloomberg.com/news/articles/2019-11-09/viral-tweet-about-apple-card-leads-to-probe-into-goldman-sachs. Access: 1 Mayo 2020.

NEGROS enfrentam índices alarmantes de contaminação pelo coronavírus nos EUA. *Folha de S. Paulo*. April 8, 2020. Available at: https://www1.folha.uol.com.br/mundo/2020/04/negros-enfrentam-indices-alarmantes-de-contaminacao-pelo-coronavirus-nos-eua.shtml. Access: Mayo 4, 2020.

OBERMEYER, Z.; POWERS, B.; VOGELI, C.; MULLAINATHAN, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. v. 366, Issue 6464, p. 447-453, Oct. 2019. Available at: https://science.sciencemag.org/content/366/6464/447. Access: 1 Mayo 2020.

PAUL, K. Healthcare algorithm used across America has dramatic racial biases. *The Guardian*. October 25, 2019. Available at: https://www.theguardian.com/society/2019/oct/25/healthcare-algorithm-racial-biases-optum. Access: 1 Mayo 2020.

PITOMBO, J.; BARBON, J. et al. Mesmo com novos leitos para Covid-19, UTIs brasileiras têm alta taxa de ocupação. *Folha de S. Paulo*. April 28, 2020. Available at: https://www1.folha.uol.com.br/equilibrioesaude/2020/04/mesmo-com-novos-leitos-para-covid-19-utis-brasileiras-tem-alta-taxa-de-ocupacao.shtml. Access: 4 Mayo 2020.

RACHE, B.; ROCHA, R. et al. *Necessidades de Infraestrutura do SUS em Preparo à COVID-19:* Leitos de UTI, Respiradores e Ocupação Hospitalar. Instituto de Estudos para Políticas de Saúde. Nota Técnica n° 3. March 2020. Available at: https://ieps.org.br/wp-content/uploads/2020/04/IEPS-NT3.pdf. Access: Mayo 5, 2020.

RAJKOMAR, A.; OREN, E.; CHEN, K. et al. Scalable and accurate deep learning with electronic health records. *npj Digital Medicine* 1, Article n° 18, 2018. Available at: https://www.nature.com/articles/s41746-018-0029-1. Access: Mayo 4, 2020.

RUPRECHT, T. Hospitais de campanha: como vão funcionar e por que são tão importantes. *Saúde*. April 9, 2020. Available at: https://saude.abril.com.br/medicina/hospitais-de-campanha-como-vao-funcionar/. Access: Mayo 5, 2020.

SCHAUER, F. *Profiles, Probabilities, and Stereotypes*. Belknapp Press. April, 2016.

SILVERSTEIN, J. Being Black in America Is a Health Risk. It's Time for Reparations. *Vice*. June 19, 2019. Available at: https://www.vice.com/en_us/article/5973yq/the-healthcare-case-for-reparations-hr40?utm_content=1586356292&utm_medium=social&utm_source=VICE_twitter. Access: Mayo 4, 2020.

SINHORETTO, J. et al. A filtragem racial na seleção policial de suspeitos: segurança pública e relações raciais. In: LIMA, C.; BAPTISTA, G. and FIGUEIREDO, I. (org.). *Segurança Pública e Direitos Humanos*: Temas Transversais. Brasília: Ministério da Justiça, 2014. *Coleção Pensando a Segurança Pública*. v. 5.

STATE V. LOOMIS, 881 N.W.2d 749 (Wis. 2016). *Harvard Law Review*. v. 130, March 10, p. 1530-1537, 2017.

Available at: https://harvardlawreview.org/wp-content/uploads/2017/03/1530-1537_online.pdf. Access: 3 Mayo 2020.

TUSHNET, M. The issue of state action/horizontal effect in comparative constitutional law. *International Journal of Constitutional Law*. v. 1, Issue 1, Jan. 2003. Available at: https://academic.oup.com/icon/article/1/1/79/671955. Access: 1 Mayo 2020.

UNITED NATIONS. Resolution adopted by the Human Rights Council on 23 March 2017. April 7, 2017. Available at: https://documents-dds-ny.un.org/doc/UNDOC/GEN/G17/086/31/PDF/G1708631.pdf?OpenElement. Access: 1 Mayo 2020.

WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*. March 3, 2020. Available at https://ssrn.com/abstract=3547922. Access: 1 Mayo 2020.

WORLD HEALTH ORGANIZATION. *Situation updates - Pandemic (H1N1) 2009*. Available at: https://www.who.int/csr/disease/swineflu/updates/en/. Access: 3 Mayo 2020.

WORLD HEALTH ORGANIZATION. *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. March 11, 2020. Available at: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020. Access: 1 Mayo 2020.

WORLD HEALTH ORGANIZATION. *WHO Director-General's opening remarks at the media briefing on COVID-19 - 13 April 2020*. April 13, 2020. Available at: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19--13-april-2020. Access: Mayo 3, 2020.

WORLD HEALTH ORGANIZATION. *WHO Health Emergency Dashboard*. Available at: https://covid19.who.int/. Access: Mayo 3, 2020.

WORLD HEALTH ORGANIZATION'S REGIONAL OFFICE FOR EUROPE. *Statement – Older people are at highest risk from COVID-19, but all must act to prevent community spread*. Copenhagen, April 2, 2020. Available at: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/statements/statement-older-people-are-at-highest-risk-from-covid-19,-but-all-must-act-to-prevent-community-spread. Access: Mayo 3, 2020.

Para publicar na revista Brasileira de Políticas Públicas, acesse o endereço eletrônico www.rbpp.uniceub.br
Observe as normas de publicação, para facilitar e agilizar o trabalho de edição.