REDE
DE PESQUISA EM GOVERNANÇA DA INTERNET

Anais do Encontro Virtual da Rede de Pesquisa em Governança da Internet

Outubro de 2021

IV ENCONTRO DA REDE DE PESQUISA EM GOVERNANÇA DA INTERNET - REDE 2021

FICHA TÉCNICA

ANAIS DA REDE DE PESQUISA EM GOVERNANÇA DA INTERNET-VOL. 4, ISSN 2675-1690

Editores
Gustavo Ramos Rodrigues
Hemanuel Jhosé Veras Alves Veras
Kimberly de Aguiar Anastácio
Nathan Paschoalini Ribeiro Batista

Autores
Alexandre Arns Gonzales
André Ramiro
Carolina Fiorini Ramos Giovanini
Hemanuel Jhosé Alves Veras
João Paulo Aguiar
Luis Gustavo de Souza Azevedo
Marcos César M. Pereira
Maria Vitoria Pereira de Jesus
Mark William Datysgeld
Nathalia Sautchuk Patrício
Nathan Paschoalini Ribeiro Batista
Pedro Amaral
Sergio Marcos Carvalho de Ávila Negri

Revisores
Alexandre Arns Gonzales
Carolina Israel
Hemanuel Jhosé Alves Veras
Luis Gustavo de Souza Azevedo
Maria Vitoria Pereira de Jesus
Nathan Paschoalini Ribeiro Batista

COMITÊ ORGANIZADOR

Núcleo de Coordenação da Rede de Pesquisa em Governança da Internet
Carolina Batista Israel
Diego Jair Vicentin
Fernanda R. Rosa
Gustavo Ramos Rodrigues
Hemanuel Jhosé Alves Veras
Kimberly de Aguiar Anastácio
Maria Vitoria Pereira de Jesus
Nathan Paschoalini Ribeiro Batista

Comitê Científico
Ana Paula Camelo
Daniela Araújo
Diego Jair Vicentin
Edison Spina
Fernanda Rosa
Flávio Rech Wagner
Gills Vilar-Lopes
Jean Santos
Leonardo Ribeiro da Cruz
Lisandro Granville
Maria Mónica Arroyo
Marta Mourão Kanashiro
Olga Cavalli
Raquel Gatto
Rodolfo Avelino
Rosemary Segurado

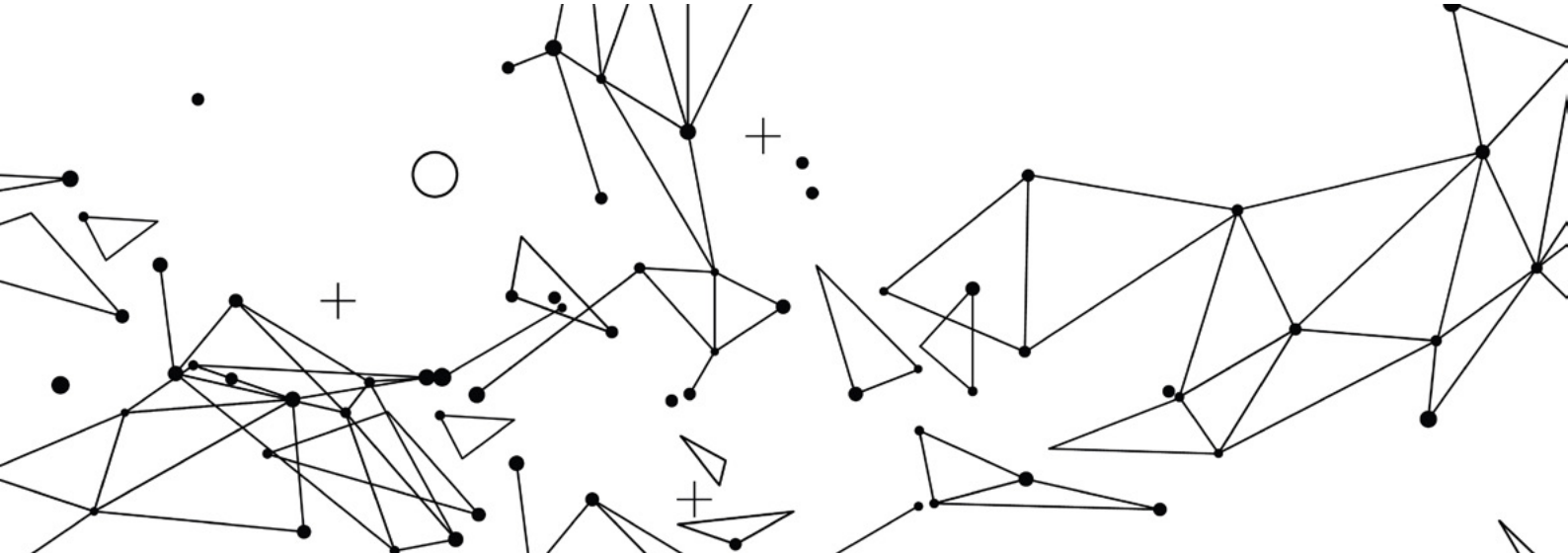
Moderadores
Ana Paula Camelo
Ana Claudia Farranha
Flávio Rech Wagner
Leonardo Ribeiro da Cruz
Marisa von Bulow
Maria Alexandra Viegas Cortez da Cunha
Rodolfo Avelino
Rodolfo Avelino
Rosemary Segurado

Debatedores
Hemanuel Jhosé Alves Veras
Luis Gustavo de Souza Azevedo
André Ramiro
Nathalia Sautchuk Patrício
Alexandre Arns Gonzales
Carolina Giovanini
Nathan Paschoalini Ribeiro Batista
Mark William Datysgeld
João Paulo Aguiar
Maria Vitoria Pereira de Jesus

NEUTROS, OBJETIVOS E EFICIENTES? O USO DE ALGORITMOS EM PROCESSOS INSTITUCIONAIS DE TOMADAS DE DECISÃO

Maria Vitoria Pereira de Jesus

SUMÁRIO



NEUTROS, OBJETIVOS E EFICIENTES? O USO DE ALGORITMOS EM PROCESSOS INSTITUCIONAIS DE TOMADAS DE DECISÃO

Maria Vitoria Pereira de Jesus[1]

RESUMO

Neste trabalho propomo-nos a analisar as controvérsias em torno da neutralidade e da eficiência dos algoritmos em tecnologias de informação presentes em instituições públicas e, de modo geral, em processos de tomadas de decisão. Em países como Brasil, Estados Unidos e Nova Zelândia, o uso de algoritmos já é realidade em procedimentos que visam à neutralidade, a objetividade e a modernização através do uso de tecnologias e sistemas como a Inteligência Artificial. Na Nova Zelândia, por exemplo, um sistema foi utilizado para prever a probabilidade de que crianças recém-nascidas sofram ou não maus-tratos com base em dados como a idade dos pais, saúde mental e situação financeira. Já no Brasil, constata-se o uso de algoritmos em setores do judiciário, nos quais sistemas de Inteligência Artificial realizam a triagem de processos e sugerem a tomada de decisões, e da administração pública, que se vale de tecnologias de coleta e cruzamento de dados para a melhoria dos serviços e concessão de benefícios sociais. Nesse sentido, poderiam os algoritmos nos fornecer análises e sugestões neutras e objetivas? Para tanto, apoiamo-nos nos Estudos Sociais da Ciência e Tecnologia e, para fins de fundamentação empírica, utilizamos de estudos e matérias jornalísticas sobre os algoritmos e as polêmicas em torno da sua atuação. Nossas análises levam-nos a questionar sobre suas qualidades políticas e como, em alguns casos, as análises e sugestões dadas por eles podem emitir vieses políticos, econômicos e socioculturais que corroboram opiniões e atitudes baseadas em discriminação e preconceito.

PALAVRAS-CHAVE

Algoritmos, Inteligência artificial, Tomadas de decisão.

ABSTRACT

In this work, we propose to analyze the controversies surrounding the neutrality and efficiency of algorithms in information technologies present in public institutions and, in

[1] Graduanda em Ciências Sociais pela Universidade Estadual de Montes Claros (UNIMONTES). E-mail: maria.vi.toriap959@gmail.com

general, in decision-making processes. In countries like Brazil, the United States and New Zealand, the use of algorithms is already a reality in procedures that aim at neutrality, objectivity and modernization through the use of technologies and systemssuch as Artificial Intelligence. In New Zealand, for example, a system was used to predict the likelihood that newborn children will suffer or will not mistreatment based on data such as parental age, mental health and financial status. Already in Brazil, found the use of algorithms in sectors of the judiciary, in which Artificial Intelligence systems screen processes and suggest decision-making, and in the public administration, which uses technologies for collecting and crossing data to improve services and social benefits. In this sense, could algorithms provide us with neutral and objective analysis and suggestions? Therefore, we rely on Social Studies of Science and Technology and, for the purpose of empirical foundation, we use studiesand journalistic material about algorithms and controversies surrounding their performance. Our analyses lead us to question about their political qualities and how, in some cases, the suggestions given by them can issue political, economic and sociocultural biases that corroborate opinions and attitudes based on discrimination and prejudice.

KEY WORDS
Algorithms, Artificial inteligence, Decisions making.

INTRODUÇÃO

Em países como Brasil, Estados Unidos e Nova Zelândia, o uso de algoritmos já é realidade em procedimentos que buscam a modernização e a neutralidade através de tecnologias e sistemas como a Inteligência Artificial (IA). Na Nova Zelândia, por exemplo, este fato se fez evidente na criação de um sistema utilizado para prever a possibilidade de que os recém-nascidos sofram ou não maus-tratos com base em variáveis como a idade dos pais, saúde mental e situação financeira (PASCUAL, 2019). Já no Brasil, o uso de algoritmos foi evidenciado em setores do judiciário, nos quais sistemas de Inteligência Artificial realizam a triagem de processos e sugerem a tomada de decisões com base na digitalização de processos e sentenças dadas anteriormente. Deste modo, embora sejam mais conhecidos em mecanismos de busca e em redes sociais online, torna-se evidente a presença de algoritmos em instituições públicas e em tomadas de decisões importantes, sendo adotados em processos e por instituições que visam o alcance de soluções “neutras” e objetivas.

No entanto, recentemente, são vários os estudos, notícias e matérias jornalísticas em que se evidenciam as controvérsias em torno da neutralidade e da eficiência dos algoritmos presentes em tecnologias e sistemas automatizados, que, ao realizar análises e sugestões com base em perfis e predições algorítmicas, promovem a discriminação de indivíduos e corroboram os vieses e atitudes baseadas em preconceitos.

Dessa forma, tal fato leva-nos a refletir: “tecnologias têm qualidades políticas?” (WINNER, 1986, p. 1). Ou ainda, é possível alcançar resultados neutros e eficientes quando empregamos os algoritmos em procedimentos socioinstitucionais? Diante disso, as análises realizadas por Langdon Winner (1986) e Andrew Feenberg (2010) nos são de suma importância, pois revelam a política e as contradições existentes nas aplicações práticas das tecnologias e do conhecimento científico.

Propomo-nos a analisar as controvérsias em torno da neutralidade e da eficiência dos algoritmos presentes em tecnologias e sistemas de Inteligência Artificial, que cada vez mais tem ganhado espaço em instituições públicas e processos de tomadas de decisões relevantes em diferentes setores da sociedade, que buscam prover-se de soluções e resultados neutros e objetivos. Poderiam os algoritmos nos fornecer soluções neutras, sem emitir qualquer discriminação e preconceito? Poderiam eles ser objetivos e eficientes ao indicar intervenções com base em perfis e predições provenientes da mineração de dados?

Para tanto, realizamos leituras no âmbito dos Estudos Sociais da Ciência e Tecnologia, considerando a política dos artefatos tecnológicos, o contexto socioeconômico de sua construção e a possibilidade de que eles incorporem “diferentes graus de poder assim como diferentes níveis de consciência” (WINNER, 1986, p. 9). Também consultamos estudos como os realizados por Danilo Doneda e Virgílio Almeida, Fernanda Bruno, e Shoshana Zuboff, que, em certa medida, tratam direta ou indiretamente da operação dos algoritmos, possibilitando-nos uma melhor compreensão das controvérsias em torno do seu funcionamento. Além disso, visando fundamentar empiricamente a nossa discussão, foram acessadas notícias e reportagens jornalísticas, que compreendemos terem sido de suma importância, poisnos forneceram dados relevantes sobre o tema abordado.

A PROMESSA DE UMA INTELIGÊNCIA ARTIFICIAL

Com o surgimento da cibernética e o desenvolvimento do behaviorismo radical, cujos experimentos se dedicam ao controle social do comportamento humano, assim como a natureza, o homem e o comportamento humano tornaram-se alvo da racionalidade do conhecimento científico. Sendo assim, depois de nos tornarmos “os mestres e senhores da natureza” (FEENBERG, 2010, p. 53), restar- nos-ia desvendar ou resolver as “equações” da comunicação e do comportamento humano, combinando-os ao pro-

gresso e as exigências do desenvolvimento tecnológico. Na década de 1940, ao propor o campo de estudos denominado cibernética, Norbert Wiener e seus colegas apresentam-nos a descrição de um sistema eletromecânico que, segundo ele, seria capaz de desempenhar funções exclusivamente humanas (KIM, 2004) e, para isso, Wiener parte da ideia de que “certas funções de controle e de processamento de informações semelhantes em máquinas e seres vivos [...] são, de fato, equivalentes e redutíveis aos mesmos modelos e [...] leis matemáticas” (KIM, 2004, p. 200). Tal proposição estabelece um fundamento importante para os experimentos em favor da inteligibilidade das máquinas, que passariam a se comunicar e a realizar funções e tarefas, antes executadas por humanos, por meio de um complexo cálculo de operações matemáticas. Dessa forma, poderíamos então, ter, de fato, funções e tarefas realizadas de modo “neutro” e “objetivo” que, em consonância com a ciência e a tecnologia, confirmariam as promessas e os valores da sociedade moderna.

A Inteligência Artificial, além da ideia de modernização, surge com a promessa de contornar a subjetividade e os valores socioculturais fortemente imbricados nas ações e no comportamento humano. Nas instituições, empresas e em processos de decisão, os sistemas de Inteligência Artificial são solicitados por sua “neutralidade”, “objetividade” e agilidade na realização de trabalhos e procedimentos até então executados por seres humanos. Em diversos setores de nossas sociedades, o uso de Inteligência Artificial já é realidade em instituições públicas e privadas, assim como também em processos de decisão. No Brasil, por exemplo, elas estão embutidas em propostas de “modernização” e de um governo digital, que deseja “ampliar o acesso e a qualidade dos serviços públicos e promover a transformação digital da gestão e dos serviços” (BRASIL, 2021, p. 2).

No decreto 10.332 publicado no dia 28 de abril de 2020, o governo federal institui a Estratégia de um Governo Digital para os órgãos e demais setores da administração pública, no qual propõe a oferta de “serviços públicos digitais simples e intuitivos”; além disso, pretende “disponibilizar a identificação digital ao cidadão”; realizar a promoção de “políticas públicas baseadas em dados e evidências; e oferecer serviços preditivos e personalizados, com a utilização de tecnologias emergentes” (BRASIL, 2020, p. 3) como o Big Data e a Inteligência Artificial que vem sendo amplamente utilizada nos serviços de atendimento ao público. Já no que tange à Política Nacional de Modernização do Estado, o “Moderniza Brasil”, publicado no dia 21 de janeiro de

2021, os objetivos apontam para a “articulação, o monitoramento e a avaliação de políticas, programas e iniciativas de modernização do Poder Executivo” (BRASIL, 2020, p. 1), que pretende promover a simplificação das relações entre Cidadão e Estado, a agilidade dos serviços e a eficiência da gestão pública por meio da implementação de um “governo e uma sociedade digital” (BRASIL, 2021, p. 2). Além disso, através de tecnologias digitais, ambas as propostas buscam promover a “desburocratização” do Estado, tornando o atendimento, os serviços e as políticas públicas mais ágeis, eficientes, econômicas (ARAGÃO; BENEVIDES, 2019) e menos burocráticas.

Ainda no caso brasileiro, nota-se o esforço no desenvolvimento de programas “inteligentes” em setores e instituições responsáveis pela jurisdição do Estado. Segundo Bruno Bioni, Mariana Rielli e Marina Kitayama (2021, p. 62), “a Defensoria Pública [...] está empenhada em um projeto de sistematização e geração de inteligência sobre os processos de todos os seus órgãos distribuídos pelo estado”. Tal projeto pretende “gerar inteligência para a propositura de ações civis públicas” (BIONI; RIELLI; KITAYAMA, 2021, p. 62), com base no conjunto de informações sobre os litígios individuais. Além disso, o programa teria o objetivo de servir de argumento convincente, para a firmação de acordos, e aumento das chances de sucesso das ações da instituição no judiciário (BIONI; RIELLI; KITAYAMA, 2021). Sendo assim, o uso de “Inteligência” auxiliaria na triagem e na sugestão de andamento com os processos, visando à maximização de ganhos das causas e dos processos protocolados pela Defensoria Pública.

Em ambos os casos, percebe-se a tentativa de tornar eficientes e econômicos os serviços e procedimentos realizados no âmbito do Estado. Com a conexão digital “inteligente” estabelecida entre governo e sociedade, o governo poderia assim, “enxugar” os gastos com o serviço público e os cidadãos seriam beneficiados com a desburocratização do Estado (ARAGÃO; BENEVIDES, 2019). O uso da Inteligência Artificial no serviço de atendimento ao público, por exemplo, possibilitaria a obtenção de respostas rápidas, quase automáticas, com base em algoritmos que realizam a predição e a personalização dos serviços. Já na Defensoria Pública, os algoritmos do programa „inteligente“ deveriam identificar os padrões e regularidades dos processos contidos nas bases de dados da instituição, orientando aos advogados sobre as chances de sucesso e a melhor maneira de dar prosseguimento às causas.

OS ALGORITMOS EM PROCESSOS DE ANÁLISE E TOMADAS DE DECISÃO

Os algoritmos são componentes essenciais das tecnologias de informação e sistemas de inteligência artificial que paulatinamente vêm sendo adotadas por instituições públicas e privadas que buscam, através dessas tecnologias, a neutralidade e a eficiência na realização de suas análises e procedimentos. Eles são conhecidos por sua agilidade em extrair padrões e regularidades em meio a um conjunto de informações que são obtidas através dos rastros digitais (BRUNO, 2013) deixados, por exemplo, quando acessamos notícias ou clicamos em anúncios e sites. Desta forma, após serem colhidas as nossas informações digitais, tudo é traduzido em dados que, depois de serem codificados, são colocados à venda, ficando à disposição de empresas de publicidade, órgãos e entidades que desejam vender e aprimorar seus serviços (ZUBOFF, 2019), com base em perfis e predições algorítmicas.

Na internet, com o crescimento do comércio e da publicidade online, os algoritmos têm sido cada vez mais solicitados por empresas e grupos publicitários a fim de estabelecer produtos e práticas e direcioná-las aos internautas, potenciais compradores e usuários das redes sociais online. Neste contexto, “os algoritmos de recomendação mapeiam nossas preferências em relação a outros usuários, trazendo ao nosso encontro sugestões” (GILLESPIE, 2018, p. 97) de produtos, hábitos e informações consideradas relevantes para nós com base no conhecimento já acumulado sobre a nossa atividade e a de outros usuários considerados “parecidos” conosco em termos preferenciais, probabilísticos e demográficos (BEER apud GILLESPIE, 2018).

Nas instituições públicas, a implementação de tecnologias de informação para a realização de diferentes procedimentos fez com que os algoritmos se tornassem elementos constitutivos importantes, por exemplo, dos sistemas de policiamento preditivo, câmeras de reconhecimento facial e programas utilizados na previsão de maus-tratos contra crianças e bebês recém-nascidos. Predizer, reconhecer e prever, nesses casos, só é possível depois de submetermos a máquina a um processo de aprendizagem. O processo de aprendizagem da máquina, também conhecido como Machine Learning, ocorre quando a “alimentamos” com um volume significativo de dados, tornando assim possível, que ela forneça padrões e perfis sobre os indivíduos e as situações da realidade. Portanto, no uso de sistemas de policiamento preditivo e de previsão de comportamento, são os padrões e perfis algorítmicos a darem as sugestões de intervenção e de como proceder em processos institucionais de tomadas de decisão.

Aparentemente, a tendência à racionalização crescente também parece perpassar pelos procedimentos de análise de dados que existem em diferentessetores e várias instituições da sociedade. O processo de análise de dados, sobretudo para e em tomadas decisão, que até então ficava a cargo da atividade e interpretação humana, agora, é feito por sistemas de Inteligência Artificial, deixando dessa forma, as decisões por conta da classificação e da seleção algorítmica. Nas instituições, os resultados fornecidos pelos algoritmos oferecem bases sólidas para atomada de decisões de forma eficiente, neutra e objetiva, na medida em que “as informações geradas são analisadas de modo matemático” (ARAGÃO; BENEVIDES,2019, p. 7) e inviabiliza a ocorrência de vieses e interpretações subjetivas, marcandoassim, o sucesso do pensamento racional sobre a interpretação humana, permeada de preconceitos e vieses econômicos, sociais e políticos (ROUVROY, 2012).

No Brasil, os resultados de correlações algorítmicas já são bases relevantes para a concessão e cortes de benefícios sociais do governo federal, tal como oBolsa Família. Segundo Grossmann (2018), desde o lançamento da plataforma Govdata, do governo federal, em abril de 2018, já haviam sido cancelados 5,2milhões de benefícios do bolsa família por meio do cruzamento de dados. O GovData abriga informações de diferentes setores e órgãos públicos do Estado, e, por meio de ferramentas de análise de dados, promete auxiliar no fornecimento de resultados objetivos e estratégicos para a formulação de políticas públicas e a concessão de benefícios sociais. No Govdata, são os algoritmos os responsáveis por realizar o cruzamento de dados e, por conseguinte, identificar a correlação e a veracidade das informações contidas nas bases do governo federal. Portanto, neste caso, nas situações em que há cortes de benefícios, as análises sobre a veracidade das informações declaradas pelos beneficiários seriam feitas por algoritmos que, com base no Big Data do governo federal, daria permissão ou indicaria as possíveis fraudes na concessão dos benefícios.

Com a crescente digitalização da vida social e o armazenamento digital de dados e informações pessoais, os algoritmos são instrumentos importantes na análise e no processamento de dados contidos em Big Data. O Big Data, por sua vez, caracterizado pela variedade, volume e velocidade com que os dados são armazenados, tem se tornado fonte de alimentação importante para os algoritmos (ALMEIDA; DONEDA, 2018) de Inteligência Artificial. As tecnologias “Inteligentes”, tal como as conhecemos nas plataformas digitais e nos sistemas preditivos, só as são devido aos dados que as

“alimentam” e garantem a sua aprendizagem no reconhecimento e na identificação dos padrões e regularidades.

Os algoritmos alimentados por Big Data buscam a similaridade das informações armazenadas e com isso, visam extrair perfis, realizar predições e dar sugestões que possibilitem antecipar e intervir em ações e comportamentos que, de algum modo, se assemelham com os de outros sujeitos. Seria através de induções algorítmicas, por exemplo, que o Govdata indicaria os possíveis indivíduos que recebiam o benefício do programa Bolsa Família de forma indevida. Também vale ressaltar que são através das predições e induções algorítmicas que os sistemas de previsão de crimes e reincidência criminal indicam os prováveis criminosos e reincidentes.

“TREINANDO” OS ALGORITMOS: A PREOCUPAÇÃO EM TORNO DOSRISCOS DE PRECONCEITO E DISCRIMINAÇÃO

Embora se compreenda a importância dos dados utilizados para a “alimentação” dos algoritmos e a sua essencialidade na aprendizagem da Inteligência Artificial (IA), tanto a coleta, quanto o seu uso resultam em implicações relevantes, passíveis de discussões, sobretudo em relação à vigilância e ao direito à privacidade e à proteção de dados pessoais. No entanto, aqui destacaremos somente o seu uso na aprendizagem da máquina, pois ao minerarem um grande conjunto de dados, com as características do Big Data, os algoritmos podem produzir predições que resultam em discriminação com base em induções estatísticas (ALMEIDA; DONEDA, 2018; PASQUINELLI, 2017).

Os algoritmos são treinados para reconhecer os padrões implícitos no conjunto de dados utilizados para a aprendizagem da IA. No entanto, se os dados utilizados para o seu treinamento contarem com vieses socioeconômicos, „raciais“ e de gênero, é provável que a máquina os reproduza (PASQUINELLI, 2017) e até discrimine com base em inferências estatísticas. Nos Estados Unidos, por exemplo, foi identificado que um sistema utilizado na prevenção da reincidência de presos no país emitia vieses racistas ao induzir que “os acusados negros eram duas vezes mais propensos a serem mal rotulados como prováveis reincidentes” (SALAS, 2017,

p. 3) do que os acusados brancos. Matteo Pasquinelli (2017) também salienta sobre as falhas ocorridas nos sistemas de reconhecimento facial que, ao serem treinados com dados enviesados – por exemplo, dados que contemplem apenas os rostos de pessoas brancas – fracassam consideravelmente em reconhecer pessoas negras.

O contrário também acontece nos sistemas de reconhecimento facial utilizados pela polícia que insistem em „reconhecer‟ os negros como prováveis criminosos, aumentando com isso, os riscos de prisões injustas e perseguições racistas.

Os erros e vieses emitidos pela IA ocorrem devido ao tratamento dado à coleta de dados utilizados em seu treinamento ou, em alguns casos, é possível que eles estejam implícitos na maneira com que os algoritmos são programados. Quando programamos um algoritmo para que ele desempenhe uma função que geralmente é feita por um ser humano, definimos um problema e o inscrevemos em uma sequência de passos para que ele o resolva (LUCENA, 2019). Resta-nos saber se uma série de passos inscritos por um sujeito localizado em um lugar social específico poderia sugerir soluções que contemple os interesses e a heterogeneidade de um grupo ou de uma população.

Na maioria dos casos, a produção dos artefatos é feita sem muita preocupação com os interesses e riscos de discriminação e violação de direitos de mulheres, negros e demais sujeitos tido como “minorias”. Homens e mulheres, embora sejam de uma mesma cultura, interagem de maneira diferente com os ambientes sociais e “à medida que se ocupam de diferentes tipos de atividade, desenvolverão e manterão padrões distintos de conhecimento” (HARDING, 2007, p. 167) sobre os problemas e a realidade. Tal proposição nos faz pensar em como seria o desenvolvimento de uma IA e, por conseguinte, a programação de algoritmos feita por mulheres, sobretudo nos casos em que o uso dessas tecnologias é feito por públicos diversos em termos de “raça” e gênero. Teríamos menos casos de discriminação por contarmos com percepções de pessoas com conhecimentos distintos e situados, em lugares sociais diferentes? Como seria o funcionamento e a eficácia de um sistema cujos algoritmos foram programados e treinados por mulheres para prever probabilisticamente, por exemplo, a chance de pais e mães cuidarem bem ou não de seus filhos?

O uso de sistemas como a Inteligência Artificial para análise e tomadas de decisões tem se apoiado no cruzamento de informações pessoais contidas nas bases de dados de diferentes setores e órgãos públicos. Nessas situações, os algoritmos realizam o cruzamento das informações pessoais e, com isso, emitem sugestões de procedimentos e tomadas de decisão. Atualmente, com o aumento da capacidade de armazenamento digital dos dados em nuvens, os algoritmos mineram informações an-

tigas que, embora não condigam com a condição presente do indivíduo, pode prejudicá-lo em determinada situação.

Em uma entrevista à Rede Lavits (Rede Latino-americana de estudos sobre vigilância, tecnologia e sociedade), Cristina Plamadeala (2021) salienta a situação de mulheres grávidas e em pós-parto que preferem não revelar à equipe médica que estão passando por conflitos que interferem em sua saúde mental com medo de serem consideradas como "mães ruins ou incapazes de cuidar dos filhos" (p. 4) e com isso, terem eles levados pelos serviços sociais. Em um contexto de adoção de tecnologias de cruzamento e coleta de dados, é possível que dados sensíveis, como os relacionados à saúde mental, sejam minerados pelos algoritmos, que fornecem sugestões e resultados para a tomada de decisões com base na definição do problema a ser solucionado. Numa situação em que programas de Inteligência Artificial, como o da Nova Zelândia, são utilizados para analisar e prever a capacidade dos pais cuidar bem ou não de seus filhos, os algoritmos mineram dados de diferentes bases do governo, e, se definimos a saúde mental e a condição financeira, por exemplo, como fatores determinantes que ditam a capacidade de mães e pais cuidarem bem ou não de suas crianças, é provável que o sistema sugira o acompanhamento e a tutela do exercício da maternidade nos casos de mulheres em que se verifica o histórico de problemas relacionados à saúde mental.

Alega-se que as classificações feitas pelos algoritmos são definidas de forma objetiva e levam em consideração os padrões e regularidades observadas em cada caso. No entanto, é preciso que nos mantenhamos atentos às posições e intenções dos sujeitos responsáveis pela programação, que, embora não sejam (ou nem sempre são) conscientes, podem fixar modelos discriminatórios que não compreendem os indivíduos e as personalidades em questão. Para o algoritmo dizer que uma mulher com o histórico de problemas de saúde mental será uma má mãeou incapaz de cuidar dos seus filhos, é necessário que os forneçamos um modelo "ideal" de uma mulher-mãe e das condições necessárias ao exercício da maternidade.

Em uma sociedade cujos padrões de atitudes específicos são atribuídos às mulheres, opiniões sexistas podem ser "amplamente sustentadas por instituições e pela sociedade como um todo" (HARDING, 2007, p. 165) e, com isso, serem "ensinados" aos algoritmos como sendo formas de conhecimento e classificações "objetivas", baseadas em observações, cálculos matemáticos, testes e revisões por pares. No entan-

to, precisamos nos questionar de que lugar se observa, "quem revisa" e "para quem é feito", pois num contexto em que as mulheres ainda são minorias nos laboratórios destinados a produção ciência e tecnologia, sobretudo nas áreas de computação e engenharia, é possível que tenhamos modelos e classificações algorítmicas definidas por homens com base em visões masculinas sobre as condições "ideais" ou qualidades essenciais necessárias ao bom desempenho e exercício da maternidade.

Além da programação, a coleta e o tratamento dado aos dados utilizados parao treinamento dos sistemas são fatores consideráveis quando se trata dos vieses implícitos nos resultados e sugestões algorítmicas. Os algoritmos treinados a partirde um conjunto de dados totalmente "homogêneos" ou "enviesados", tendem a fixar padrões extremamente específicos quanto ao contexto e à população sobre a qualirá atuar. Por outro lado, os algoritmos treinados a partir de um conjunto de dados "heterogêneos", cujo volume e variedade compreendam as características do Big Data, tendem a operar em uma lógica infra-individual e supra-individual (ROUVROY, 2012; BRUNO, 2008), que, embora baseiem em fragmentos de informações pessoais de vários sujeitos, distribuídas e fragmentadas em redes e relações interpessoais (BRUNO, 2008), continuam a replicar discriminações e preconceitos.

Valendo-se da primeira opção, isso se faz evidente, por exemplo, nas situações em que se constata que "os programas usados nos departamentos de contratação de algumas empresas mostram uma inclinação por nomes usados por brancos e rejeitam os dos negros" (SALAS, 2017, p. 4). O mesmo ocorre nos casos em que se identifica uma disparidade significativa na seleção de homens ao invés demulheres em vagas de emprego. Nesses casos, se o sistema foi treinado com dados "homogêneos" ou que não representam proporcionalmente a população, é provável que ele vá discriminar indivíduos que destoem do padrão ou perfil já fixado de homem branco à procura de um emprego. Já se valermos da segunda opção, a discriminação do sistema ocorre devido à "heterogeneidade" dos dados que fazem com que os algoritmos operem em uma lógica infra ou supra-individual, selecionando padrões e produzindo predições que dizem pouco a respeito de um indivíduo ou de sua pessoa identificável (BRUNO, 2013). Nesses casos, os algoritmos podem produzir resultados completamente aleatórios e prejudicar indivíduos que se submetem a processos de seleção que consideram importantes para a sua trajetória pessoal e profissional. Recentemente, no Reino Unido, por exemplo, um algoritmo utilizado para julgar a nota dos estudantes foi acusado de

fornecer resultados injustos e inesperados que fez com inúmeros jovens perdessema chance de entrar na Universidade (BBC, 2020).

Portanto, em tais situações, não seria de se estranhar a existência de pontos de vista opostos em relação à instalação de sistemas algorítmicos em análises e procedimentos de tomadas de decisão. No caso do Brasil, os sistemas de Inteligência Artificial (IA) já atuam tribunais, como o Superior Tribunal de Justiça (STJ), sugerindo decisões e fornecendo súmulas das ações judiciais (SAKAI; GALDINO; BURG, 2021). No entanto, para este fim, o ex-desembargador do Tribunal de Justiça [de São Paulo], José Roberto Neves Amorim (apud FERREIRA, 2020) salienta os possíveis riscos de processos criminais e de guardas serem analisados pela IA, pois, para ele, causas como essas "jamais poderão passar por máquinas" (p. 3), uma vez que envolvem uma série de circunstâncias que podem ser negligenciadas na mineração dos dados ou não corresponderem aos profilings identificados pelos algoritmos.

TORNANDO OS ALGORITMOS AUDITÁVEIS PARA DECISÕES MAIS JUSTAS, TRANSPARENTES E RESPONSÁVEIS

É certo que as opiniões em torno da implementação de sistemas de Inteligência Artificial em processos de tomada de decisão tem sido diferentes e suscitado debates fortemente profícuos sobre a sua eficiência, neutralidade e objetividade. Se por um lado, temos inúmeras propostas como a instituição de uma "Estratégia de um Governo Digital" e iniciativas como o uso de Inteligência Artificial em diferentes setores do judiciário, com o objetivo de "contribuir com a agilidade e coerência do processo de tomada de decisão" (CNJ, 2020, p. 1), por outro, temos a preocupação com a regulação dos algoritmos que, além de produzirem predições que resultam em discriminação e preconceito, emitem resultados e sugestões para a tomada de decisões pouco explicáveis.

Diante das últimas polêmicas que apontam para a existência de vieses e possibilidade de erros em decisões automatizadas, somos cada vez mais convidados a pensar sobre os constrangimentos sofridos pelos indivíduos que se submetem ou são submetidos às análises e predições feitas pelos algoritmos. Sabemos que os erros e vieses emitidos pela IA podem não só ocasionar emsentenças e demissões injustas e colaborar para que estudantes percam a oportunidade de entrar em uma Universidade, como também fazer com que pais se sintam amedrontados pela possibilidade de

perderem a tutela dos seus filhos ao serem acusados como prováveis cometedores de maus-tratos. Nesses casos, geralmente, quando são fornecidos os resultados ou tomadas às decisões, não se deixa claro quais foram os critérios e os dados utilizados para se chegar àquele resultado ou tomar àquela decisão.

Em uma reportagem recente, publicada no jornal *El País*, Miquel Echarri (2021) relata, por exemplo, a situação de trabalhadores demitidos por programas de Inteligência Artificial que cada vez mais são adotados por empresas em processos de tomadas de decisões, seja para realizar possíveis contratações, ou demitir funcionários. Esse foi o caso de um dos funcionários da Amazon, que, em 2019, foi demitido por um algoritmo que o considerou com um desempenho insatisfatório para o trabalho. No entanto, ele considerou essa decisão injusta e pouco clara, pois ninguém o esclareceu sobre os critérios utilizados no resultado emitido pela máquina(ECHARRI, 2021). Nos Estados Unidos, um sistema semelhante foi utilizado para realizar a avaliação de professores e, a partir disso, sugerir demissões com base no perfil esperado de um “bom” professor. Nesse caso, ao fixar uma pontuação específica para o que seria um “bom” professor, o sistema demitiu centenas de profissionais cujas pontuações ficaram abaixo do padrão estabelecido (O‟NEIL, 2016). Entretanto, os indivíduos que receberam os comunicados de demissão com base nos resultados fornecidos pelos algoritmos, os consideraram inexplicáveis diante das opiniões positivas da comunidade escolar em relação ao trabalho desempenhado por eles.

A opacidade e a crença na eficiência e na objetividade dos algoritmos baseados em cálculos e operações matemáticas fazem com que os seus resultados sejam tidos como “incontestáveis” e funcionem como bases relevantes para a tomada de decisões. Segundo Canclíni (2020), vivemos em um contexto de informatização, em que cada vez mais confiamos nas decisões e na autoridade dos algoritmos de macrodados. Confiamos tanto nos resultados algorítmicos que não pensamos duas vezes em aceitar as suas sugestões de como prosseguir em determinados procedimentos ou realizar a tomada de decisões de forma neutra e objetiva. No entanto, quando os vemos fornecer sugestões e resultados completamente aleatórios ou que destoam da realidade, provocando constrangimentos, causando danos e promovendo a perda de oportunidades, é necessário que os seus resultados sejam revistos e o passo a passo dado nos processos de decisão se tornem auditáveis.

No Brasil, o artigo 20 da Lei Geral de Proteção de Dados (Lei nº 13.709, publicada no dia 14 de agosto de 2018), define que, nas situações em que os indivíduos são alvos análises discriminatórias ou recebem resultados que comprometem os seus interesses e o alcance de oportunidades, "o titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais" (BRASIL, 2018, p. 10). Além disso, o artigo também inclui os casos em que as decisões são tomadas com base em predições, que insistem em prever perfis de crédito, pessoais e profissionais com base nos padrões e tendências retiradas de um grande conjunto de dados. O dispositivo reconhece as possibilidades de falhas e vieses virem implícitos nas predições e sugestões de tomadas de decisão fornecidas pelos algoritmos, por isso, garante aos sujeitos, alvos de discriminações e injustiças, que tenham não só osseus resultados revistos por humanos, como também determina que sejam dadas "informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada" (BRASIL, 2018, p. 10).

Cathy O"Neil (2016), em seu livro *"Weapons of math destruction: how big data increases inequality and threatens democracy"*, defende que precisamos exigir a transparência dos resultados e predições feitas pelos algoritmos. Temos o direito de saber quais dados e informações pessoais estão sendo utilizadas para que o algoritmo realize a análise ou sugira determinada tomada de decisão. Nesse sentido, a criação de legislações e órgãos de regulação é de suma importância para que se determine a auditoria e a revisão dos processos de tomadas de decisão. Segundo a autora, nos casos em que sistemas de Inteligência Artificial são utilizados pararealizar a concessão de crédito aos consumidores, os EUA dispõe de legislações específicas para que se alcance resultados justos e igualitários. Dessa forma, enquanto o FCRA (Fair Credit Reporting Act) garante que o consumidor tenha acesso e verifique a veracidade dos dados analisados pelos algoritmos, o ECOA (Equal Credit Opportunity Act) proíbe que eles façam a correlação de dados como raça, gênero (O"NEIL, 2016) e estado civil para indicar os indivíduos aptos a receberem os empréstimos.

No caso do Brasil, é a LGPD (Lei Geral de Proteção de Dados) que garante a transparência das informações processadas pelos algoritmos e impedem o uso de dados que, por algum motivo, possam prejudicar os indivíduos em determinadas análises. A ANPD (Autoridade Nacional de Proteção de Dados) estabelece que, nos casos em que se verificam vieses e incoerências nos resultados e predições feitas

pelos algoritmos, é necessário que sejam realizadas auditorias para a “verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais” (BRASIL, 2018, p. 10). Portanto, nessas situações, a existência de uma legislação e a criação de órgãos de supervisão, como a ANPD, é fundamental para que possamos pensar na possibilidade de governança dos algoritmos e em como promover decisões justas, transparentes e explicáveis.

A reivindicação de posse de um segredo econômico e industrial não pode continuar a servir de argumento para que empresas e instituições impeçam o prosseguimento de ações que solicitam que os algoritmos sejam auditáveis. Da mesma forma, entendemos que o argumento de que os sistemas de Inteligência Artificial são “autônomos” e “incontestáveis”, não pode continuar funcionando como pretexto para que os seus algoritmos permaneçam opacos, inexplicáveis e incontroláveis (PASQUINELLI; JOLER, 2020). A ideia de que os algoritmos são “neutros” e “objetivos” não deve permitir com que fechemos os olhos ou desconsideremos a possibilidade de falhas e constrangimentos serem gerados pelo uso da IA. Portanto, cientes disso, precisamos estar atentos ao processo de sua construção, a base de dados com a qual foi treinada e a forma com que os algoritmos foram programados. Quem inscreve ao algoritmo o “passo a passo” para definir se um sujeito é um “bom” professor ou um funcionário inapto para o trabalho? Em que conceito, visão de mundo ou experiências se baseia a definição de “bom” ou “mal” professor ou mesmo a aptidão necessária para um bom desempenho no emprego? Além disso, quem define se a IA é ou não eficiente para a realização dessas análises? A expectativa é a de que as respostas a essas perguntas nos possibilitem esclarecer e tornar mais justas as decisões baseadas nas sugestões emitidas pela máquina.

Os vieses e a controvérsia em torno da imparcialidade dos algoritmos podem estar contidos já no momento de seleção dos dados que servirão como base de treinamento. “O ato de selecionar uma fonte de dados em vez de outra é a marca profunda da intervenção humana” (PASQUINELLI; JOLER, 2020, p. 8) no processo de construção da IA e uma das principais causas de discriminação e preconceito produzidas pelos algoritmos. Por exemplo, se ao treinarmos um sistema de Inteligência Artificial, criado para a seleção de candidatos a vagas de emprego, utilizamos uma base de dados em que os últimos contratados do ano foram homens brancos, com idade entre 25 a 40 anos e estado civil solteiro, não poderíamos ficar surpresos se, ao ser posto em

funcionamento, o sistema discriminar homens negros, mulheres e demais indivíduos cuja idade seja inferior a 25 ou superior a 40 anos ou cujo estado civil seja o de casado, separado ou divorciado. Portanto, nessa etapa, prestar atenção na base de dados utilizada no treinamento da IA é fundamental para que os algoritmos para que os algoritmos forneçam análises e sugestões justas e menos enviesadas. Uma base de dados pode conter vieses históricos (PASQUINELLI; JOLER, 2020) e estar permeada de preconceitos e discriminações de cunho racista, sexista e misógino, e, nesses casos, os algoritmos de *Machine Learning* podem não só aumentar, como também contribuir para que tais estruturas continuem de forma permanente em nossas sociedades.

A etapa de rotulagem dos dados e programação dos algoritmos também pode interferir nos resultados e predições que obtemos de suas análises. Quando treinamos uma máquina e programamos os seus algoritmos para que sugiram decisões de forma automática, o processo de classificação dos dados e de atribuição do que seria uma decisão ideal ou um comportamento adequado para determinada situação não é de forma alguma um processo imparcial (PASQUINELLI; JOLER, 2020). Classificamos os dados e fornecemos aos algoritmos modelos ideais de comportamento e tomadas de decisão com base em conceitos e pontos de vista que, além de reduzirem aspectos da realidade, comprometem a neutralidade e objetividade de suas análises. Os conceitos com os quais os algoritmos são programados podem não só esconder escolhas morais (O"NEIL, 2016), como também corresponder a perfis muito específicos, por exemplo, de indivíduos esperados para que ocupem determinados postos de trabalho. Porisso, a tentativa de regulamentar o uso de sistemas de Inteligência Artificial em instituições públicas buscam incorporar certa de diversidade nas diferentes etapasde seu desenvolvimento até o momento de sua implementação.

No judiciário brasileiro, a Resolução n° 332, publicada no dia 21 de agosto de 2020, já defende que a diversidade e "a participação representativa deverá existir em todas as etapas do processo, tais como planejamento, coleta e processamento de dados, construção, verificação, validação e implementação" (CNJ, 2020, p. 4) da IA. A participação de um grupo diversificado em termos raça, gênero, profissão e etnia pretende possibilitar um treinamento mais cuidadoso da IA, que considere a pluralidade de indivíduos e as inúmeras realidades e pontos de vista, para que, assim, os algoritmos forneçam resultados mais justos e comprometidos com o princípio de equidade e justiça. O conceito ou ponto de vista econômico de eficiência em termos

de custo e agilidade não pode continuar a servir como base para a implementação e uso da IA. Precisamos, ao invés disso, considerar que existem conceitos múltiplos e que o conceito de eficiência deve observar a necessidade de equidade e justiça e de compreensão das particularidades presentes em cada caso em que se solicita o uso de sistemas de Inteligência Artificial em processos deanálise e tomadas de decisão.

Se quisermos que os algoritmos nos forneçam decisões justas e responsáveis, precisamos exigir a prestação de contas e a transparência em todas as etapas do processo de desenvolvimento da IA. Temos o direito de saber o porquê ou com base em quê estamos sendo desconsiderados (O"NEIL, 2016) quando solicitamos o recebimento de um benefício ou somos demitidos por sermos considerados "maus" funcionários. Portanto, nesses casos, trata-se de tornar os algoritmos auditáveis para que tenhamos decisões justas e compreensíveis.

CONSIDERAÇÕES FINAIS

Vimos que a promessa de uma inteligência artificial surge sob a proposta de uma modernização e na tentativa de inviabilizar a interferência de valores e vieses subjetivos, sobretudo em análises e procedimentos de tomada de decisão. Nasinstituições públicas e privadas, os resultados e sugestões fornecidas pelos algoritmos, ao passo em que se baseiam em cálculos matemáticos, aparecem como bases sólidas para proposições e tomadas de decisões de modo "neutro" e objetivo. Porém, consideramos que em seu treinamento, uma simples falha na seleção de dados ou o uso de uma base de dados "heterogênea" ou "enviesada" pode fazercom que eles reproduzam discriminações como a racial e de gênero.

Nesse sentido, inserir a Inteligência Artificial em procedimentos de análises e tomadas de decisão acaba se tornando um desafio, pois os algoritmos, concebidos como isentos de valores e a serviço da eficiência e do progresso tecnológico (FEENBERG, 2010), revelam-se cada vez menos "neutros" e objetivos, ao discriminar indivíduos e fixar padrões específicos. Portanto, se por um ladocompreender o resultado ou a omissão dos algoritmos em determinadas análises requer o conhecimento dos dados utilizados na aprendizagem da máquina (*machine learning*) e a observância dos conceitos que os fornecemos para que realizem a análise, por outro, compreendemos a relevância da criação legislações e órgãos governamentais de regulação e supervisão (ALMEIDA; DONEDA, 2018), que estejam comprometidos com a responsabilidade

social e tecnológica frente à sua atuação e com a transparência dos resultados e sugestões que fornecem para a tomada de decisão.

"Tecnologias têm qualidades políticas?" (WINNER, 1986, p. 1). Essa foi uma das perguntas que nos fizemos quando pesquisamos os casos de discriminação e controvérsias decorrentes das técnicas e análises algorítmicas. Deste modo, salientamos serem de suma importância os questionamentos realizados sobre a possível política implícita nas tecnologias e nos objetos técnicos que se apresentam controversos em relação aos objetivos a que são propostos e diante do contexto e da situação em que são aplicados. Compreendemos que tais reflexões se tornam fundamentais se o objetivo é fazer com que tenhamos decisões justas e imparciais frente à digitalização crescente dos processos sociais realizados em setores e por instituições que cada vez mais buscam por melhorias através das inovações sociotécnicas.

REFERÊNCIAS

ARAGÃO, Francisca. Alana. Araújo; BENEVIDES, Pablo Severiano. Governamentalidade algorítmica e Big data: o uso da correlação de dados como critério de tomada de decisão. In: SIMPÓSIO INTERNACIONAL LAVITS: "ASSIMETRIAS E (IN)VISIBILIDADES: VIGILÂNCIA, GÊNERO E RAÇA", 6, 2019, Salvador. Disponível em: <https://lavits.org/wp-content/uploads/2019/12/Araujo_Benevides-2019-LAVITS.pdf>. Acesso em: 20 de março de 2021.

BBC. 2020. "Algoritmo roubou meu futuro": solução para "Enem britânico" na pandemia provoca escândalo. Disponível em: <https://www.bbc.com/portuguese/internacional-53853627>. Acesso em 16 de maio de 2021.

BIONI, Bruno Ricardo; RIELLI, Mariana; KITAYAMA, Marina. O legítimo interesse na LGPD: quadro geral e exemplos de aplicação. São Paulo: Associação Data Privacy Brasil de Pesquisa, 2021.

BRASIL. Decreto nº 10.332, de 28 de abril de 2020, Brasília – DF, 2020. Disponível em: <https://www.in.gov.br/web/dou/-/decreto-n-10.332-de-28-de-abril-de-2020-254430358>. Acesso em: 14 de maio de 2021.

BRASIL. Decreto nº 10.609, de 26 de janeiro de 2021, Brasília — DF, 2021. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2021/Decreto/D10609.htm#art18>. Acesso em: 21 de março de 2021.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018, Brasília — DF, 2021. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709compilado.htm>. Acesso em 31 de outubro de 2021.

BRUNO, Fernanda. **Máquinas de ver, modos de ser:** vigilância, tecnologia esubjetividade. Porto Alegre: Sulina, 2013.

BRUNO, Fernanda. Monitoramento, classificação e controle nos dispositivos de vigilância digital. **Revista FAMECOS**. Porto Alegre, n. 36, p. 10-16, agosto, 2008.

CANCLINI, Nestor Garcia. Ciudadanos reemplezados por algoritmos. **CALAS**. México, 2020.

CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em 02 de novembro de 2021

DONEDA, DANILO; ALMEIDA, VÍRGILIO AUGUSTO FERNANDES. O que é a governança dos algoritmos? In: BRUNO, Fernanda; CARDOSO, B; KANASHIRO, M; GUILHON, L; MELGAÇO, L. (orgs). **Tecnopolíticas da vigilância.** São Paulo: Boitempo, 2018, p. 141-148.

ECHARRI, Miquel. 150 demissões em um segundo: os algoritmos que decidem quem deve ser mandado embora. **El País**. Barcelona, 10 de outubro de 2021.Disponível em: <https://brasil.elpais.com/tecnologia/2021-10-10/150-demissoes-em-um-segundo-assim-funcionam-os-algoritmos-que-decidem-quem-deve-ser-mandado-embora.html>. Acesso em: 29 de outubro de 2021.

FEENBERG, Andrew. Racionalização democrática, poder e tecnologia. In: NEEDER, Ricardo T. (org.). **Ciclo de conferências Andrew Feenberg**, 2010.

FERREIRA, Flávio. Inteligência Artificial atua como juiz, muda estratégia de advogado e „promove" estagiário. **Folha de S. Paulo**. São Paulo, 10 de março de2020. Disponível em: <https://www1.folha.uol.com.br/poder/2020/03/inteligencia- artificial-atua-como-juiz-muda-estrategia-de-advogado-e-promove-estagiario.shtml>. Acesso em: 19 de setembro de 2020.
GILLESPIE, Talerton. A relevância dos algoritmos. **Parágrafo**. São Paulo, v. 6, n. 1, p. 95-121, janeiro/abril, 2018.

GROSSMAN, Luís Osvaldo. Big Data do Governo Federal levou ao corte de 5milhões do Bolsa Família. **Convergência Digital**. São Paulo, 16 de abril de 2018. Disponível em: <https://ciab.convergenciadigital.com.br/cgi/cgilua.exe/sys/start.htm?infoid=47765&sid=11>. Acesso em: 15 de maio de 2021.

HARDING, Sandra. Gênero, democracia e filosofia da ciência. **RECIIS – Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**. Rio de Janeiro, v. 1, n. 1, p. 163-168, janeiro-junho, 2007.

KIM, Joon Ho. Cibernética, ciborgues e ciberespaço: notas sobre as origens da cibernética e sua reinvenção cultural. **Horiz. antropol.**, Porto Alegre , v. 10, n. 21, p. 199-219, Junho, 2004. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104- 71832004000100009&lng=en&nrm=iso>. Acesso em 25 de março de 2021.

LUCENA, Pedro, Arthur Capelari de. Policiamento preditivo, discriminação algorítmica e racismo: potencialidade e reflexos no Brasil. In: SIMPÓSIO INTERNACIONAL LAVITS: "ASSIMETRIAS E (IN)VISIBILIDADES: VIGILÂNCIA, GÊNERO E RAÇA", 6,

2019, Salvador. Disponível em: <https://lavits.org/wp-content/uploads/2019/12/Lucena--2019-LAVITSS.pdf>. Acesso em: 15 de maio de2021.

O'NEIL, Cathy. **Weapons of math destruction: how big data increases inequality and threatens democracy**. New York: Crown Publishers, 2016.

PASCUAL, Manuel González. Quem vigia os algoritmos para que não sejam racistas ou sexistas? **El País**. Madrid, 17 de março de 2019. Disponível em: <https://brasil.elpais.com/brasil/2019/03/18/tecnologia/1552863873_720561.html>. Acesso em: 22 de junho de 2020.

PASQUINELLI, Matteo. Machines that Morph Logic: Neural Networks and the Distorted Automation of Intelligence as Statistical Inference. **Logic Gate: the Politicsof the Artifactual Mind (Online)**. 2017. Disponível em: <https://www.glass- bead.org/article/machines-that-morph-logic/?lang=enview>. Acesso em: 20 de marçode 2021.

PASQUINELLI, Matteo; JOLER, Vladan. O manifesto Nooscópio: InteligênciaArtificial como Instrumento de Extrativismo do Conhecimento. **Lavits**, Rio de Janeiro, 30 de julho de 2020. Disponível em: <https://lavits.org/o-manifesto-nooscopio-inteligencia--artificial-como-instrumento-de-extrativismo-do-conhecimento/?lang=pt>. Acesso em 02 de novembro de 2021.

PLAMADEALA, Cristina. Vigilância de dossiê e a produção de arquivos como ferramentas de controle. [Entrevista concedida a] Paulo Faltay. **Lavits**, Rio de Janeiro, 23 de abril 2021. Disponível em: <https://lavits.org/entrevista-cristina- plamadeala/?lang=pt>. Acesso em: 06 de maio de 2021.

ROUVROY, Antoinette. The end(s) of critique: data-behaviourism vs. due-process.In: ROUVROY, Antoinette. **Privacy Due Process and the Computational Turn. Philosophers of Law Meet Philosophers of Technology**. Routledge, 2012.

SAKAI, Juliana; GALDINO, Manoel; BURG, Tamara. Governance recommendations: Use of Artificial Intelligence by public authorities. **Transparência Brasil**. São Paulo, 2021. Disponível em: <https://www.transparencia.org.br/downloads/publicacoes/Governance_Recommendations.pdf>. Acesso em: 26 de maio de 2021.

SALAS, Javier. Se está na cozinha é uma mulher: como os algoritmos reforçam preconceitos. **El País**. Madrid, 23 de setembro de 2017. Disponível em: <https://brasil.elpais.com/brasil/2017/09/19/ciencia/1505818015_847097.html>. Acesso em: 16 de setembro de 2020.

WINNER, Langdon. **Artefatos têm política?** Traduzido por Fernando Manso.Chicago: The University of Chicago Press, 1986, p. 19-39.

ZUBOFF, SHOSHANA. **The age of surveillance capitalismo: The fight for a human future at the new frontier of power**. New York: Public Affairs, 2018. Resenha de: EVANGELISTA, Rafael. Surveilance & Society, New York, v. 17, n. 1/2, p. 246-251, march, 2019.