

INTELIGÊNCIA ARTIFICIAL E ALGORITMOS ENVIESADOS

Fabiana Aparecida dos Reis Silva (IT) e Marcelo Chiavassa de Mello Paula Lima (Orientador)

Apoio: Santander

RESUMO

Com as transformações digitais e o aumento crescente do uso de tecnologias de Inteligência Artificial (IA) para auxiliar, otimizar e melhorar os diversos aspectos que moldam o cotidiano da sociedade, algumas questões relacionadas a confiabilidade nas decisões tomadas por algoritmos autônomos começam a surgir. Essas tecnologias são alimentadas por uma série de informações – bases de dados – que determinam o seu funcionamento direcionado a um resultado específico. Uma vez que esses dados introduzidos na IA for repleto de vieses, existe uma grande probabilidade de os resultados apresentados pelos *softwares* serem enviesados, violando direitos humanos como a liberdade, a não discriminação e a privacidade. Com o aumento dessas tecnologias, torna-se necessária a adoção de medidas que garantam o uso de algoritmos neutros com decisões objetivas, ou que ao menos limitem os vieses em suas decisões. É preciso que se realizem mais testes com as tecnologias de IA que apresentam dificuldade de distinção entre determinados grupos, como é o exemplo dos carros autônomos, que são imprecisos na detecção de pessoas negras em comparação com pessoas brancas. É necessária a criação de bases éticas nas quais os programadores e desenvolvedores de tecnologias de IA devam se fundar para criação desses sistemas, de modo a garantir a neutralidade da máquina. Ainda, que existam formas constantes de auditoria nas empresas públicas e privadas que se valham desses instrumentos, para constante comprovação de que seus algoritmos estão livres de vieses.

Palavras-Chave: Algoritmos enviesados, Discriminação Algorítmica, Inteligência Artificial.

ABSTRACT

With digital transformations and the increasing of the use of Artificial Intelligence (AI) technologies to assist, optimize and improve the various aspects that shape society's daily life, some issues related to the reliability of decisions made by autonomous algorithms are beginning to arise. These technologies are fed by a series of information - databases - that determine their operation directed to a specific result. Once this data introduced in AI is full of biases, there is a high probability that the results presented by the software will be biased, violating human rights such as freedom, non-discrimination and privacy. With the increase in these technologies, it is necessary to adopt measures that guarantee the use of neutral algorithms with objective decisions, or that at least limit the biases in their decisions. It is necessary to carry out more tests on AI courses that show difficulty in distinguishing between certain groups, such as the example of autonomous cars, which are inaccurate in detecting black people compared to white people. It is necessary to create ethical bases on which AI technology programmers and developers must base themselves to create these systems, in order to guarantee the neutrality of the machine. Also, the implementation of constant audits in public and private companies that use these instruments, to constantly prove that their algorithms are free from bias.

Keywords: Bias Algorithms, Algorithmic discrimination, Artificial Intelligence.

1. Introdução

Por diversas razões históricas e sociais, condutas discriminatórias fazem parte da sociedade. E com a era digital e da automação, em que tecnologias de Inteligência Artificial se entrelaçam cada vez mais com o cotidiano, com sistemas algorítmicos de decisões automatizadas, passa-se a ter um risco de essas tecnologias reproduzirem certos comportamentos preconceituosos que são características do ser humano.

A tomada de decisão dos algoritmos é definida por sua programação, e se o conteúdo que alimenta a IA for discriminatório, o resultado será discriminatório. Portanto, quando a inteligência artificial é programada com um algoritmo enviesado, para selecionar um candidato a uma vaga de emprego, identificar suspeitos de um crime ou realizar tarefas simples nas redes sociais, quais são os impactos disso? Certamente este é um problema que as novas tecnologias trouxeram e que devem ser analisadas, a fim de entender como isso ocorre e como lidar com essas situações.

Hoje temos *softwares* capazes de identificar se uma pessoa é negra ou branca pela análise de seu nome, carros autônomos que utilizam câmeras para detectar pedestres, sistemas de reconhecimento facial para identificar suspeitos etc. O uso dessas tecnologias, quando programadas com algoritmos enviesados, oferecem um altíssimo risco de discriminação.

Assim, o objetivo dessa pesquisa é analisar alguns casos em que restou demonstrado o perigo de dados enviesados alimentando sistemas algorítmicos de decisões automatizadas. E muitas vezes essa discriminação acontece pela falta de treinamento adequado à IA, ou pela ausência de preceitos éticos no momento da programação.

2. Problema

Com a era digital, tornou-se cada vez mais comum em todas as áreas e setores o uso de tecnologias para o auxílio na execução das tarefas, melhoria no desempenho e agilidade. E as ferramentas de inteligência artificial passaram a ocupar um importante espaço no cotidiano, principalmente quando o assunto

é a tomada de decisão algorítmica, que agilizou a execução das atividades em muitos aspectos, sobretudo quanto a trabalhos repetitivos. O uso dessa tecnologia tem sido cada vez mais utilizada para aprimorar a segurança pública, as empresas no momento da seleção de candidatos, em monitoramento privado e melhoria da experiência dos usuários em redes sociais. E diante desse cenário, em que praticamente tudo é organizado e funciona com base em programas de inteligência artificial com decisões automatizadas, levanta-se a seguinte pergunta:

Existe riscos de as decisões tomadas por sistemas automatizados levarem à discriminação? Existe uma forma de garantir que os vieses sejam eliminados ou limitados?

3. Inteligência Artificial e Algoritmos discriminatórios

De um modo geral, a sociedade está em constante processo de transformação, e nas últimas décadas esse processo ocorre de forma cada vez mais acelerado, com uma rapidez nunca vista. E a era digital é concebida hoje por uma vasta quantidade de tecnologias que dão vida a um conjunto de instrumentos que automatizam os mais diversos setores existentes. Toda essa mudança é decorrente da 4ª Revolução Industrial¹.

Tem-se por revolução industrial todo processo de mudanças abruptas e radicais, que são associados ao surgimento de novas tecnologias que modificam o modo de viver de uma sociedade, principalmente no setor político, econômico e social². E a 4ª Revolução Industrial surge exatamente nesse sentido, marcando a nova era a partir de um conjunto de tecnologias disruptivas, como por exemplo “a robótica, realidade aumentada, o *big-data*, nanotecnologia, internet das coisas, inteligência artificial entre outras muitas”³. O que marca

¹ CHIAVASSA, Marcelo de Mello Paula Lima; ANDRADE, Victor Moraes. **Manual de direito digital**. Tirant: São Paulo, 2020. p. 17.

² CHIAVASSA, Marcelo de Mello Paula Lima; ANDRADE, Victor Moraes. **Manual de direito digital**. Tirant: São Paulo, 2020. p. 17.

³ MIASATO, Alessandra; SILVA, Fabiana Reis. Artificial Intelligence as an Instrument of Discrimination in Workforce Recruitment. **Acta Univ. Sapientiae**, Cluj-Napoca, v. 8, n. 2, p. 191-212, 2019.

então essa sociedade digital, é a convergência entre o mundo real, com as tecnologias digitais e os fatores biológicos.

3.1. Inteligência artificial e os algorítmicos

A Inteligência Artificial (IA) é a ciência que torna as máquinas inteligentes; é a ciência que transporta para a máquina algumas das principais capacidades humanas; uma ferramenta de *software* que pode ser integrada à um robô e que é, até certo ponto, inteligente⁴. Não há uma definição exata, mas pode-se dizer que a inteligência artificial:

[...] é um termo abrangente que inclui uma variedade de técnicas e processos associados dedicados a melhorar a capacidade das máquinas em fazer coisas que requerem inteligência, como reconhecimento de padrões, visão computacional e processamento de linguagem (tradução nossa)⁵.

As mudanças decorrentes da 4ª revolução industrial tornam difícil imaginar algo que não envolva o uso dessa tecnologia, ou que não seja resultado dela. Cabe lembrar que a inteligência artificial existe desde o século passado, sendo o seu desenvolvimento definido em 4 ondas⁶: IA da internet, IA dos negócios, IA de percepção e IA autônoma.

As ondas da IA da internet e dos negócios já são perceptíveis, fazem parte da nossa realidade. Inúmeros serviços utilizam algoritmos a fim de aprimorar suas ferramentas, como os provedores de busca – ex. Google – que personalizam o sistema de busca conforme os dados coletados pela própria plataforma somados aos dados fornecidos pelo usuário; ou empresas que oferecem consultoria de negócios de big data para empresas e governos⁷.

A terceira onda, a IA de percepção, está atualmente analisando o mundo ao seu redor, por exemplo, por meio do reconhecimento facial, ferramenta esta

⁴ ANDRÉS, Moisés Barrio. **Manual de derecho digital**. Valencia: Tirant lo blanch, 2020. p. 56.

⁵ RASO, Filippo et al. **Artificial Intelligence & Human Rights: Opportunities & Risks**. Cambridge: Berkman Klein Center, 2018. 63 p. Disponível em: <<https://cyber.harvard.edu/publication/2018/artificial-intelligence-human-rights>>. Acesso em: 20 abr. 2020.

⁶ LEE, Kai-Fu. **Inteligência Artificial: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos**. Rio de Janeiro: Globo Livros, 2019. p. 131.

⁷ LEE, Kai-Fu. **Inteligência Artificial: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos**. Rio de Janeiro: Globo Livros, 2019. p. 137.

que tem gerado inúmeras discussões quanto ao seu uso extensivo, como o caso do americano Robert Williams preso erroneamente, que será melhor abordado nessa pesquisa. E a quarta e última onda, a IA autônoma, será a que causará maior impacto, como aponta Kai-Fu Lee:

À medida que carros autônomos tomem as ruas, drones autônomos tomes os céus e robôs inteligentes tomem as fábricas, eles vão transformar tudo, da agricultura orgânica a viagens por autoestradas e o fast-food⁸.

Os algoritmos são um conjunto de instruções ou comandos para a realização de uma determinada tarefa⁹, ou seja, um conjunto de informações que ditam o comportamento da máquina. A IA recebe um comando inicial chamado de *input*, que deverá gerar um resultado, chamado de *output*. Para atingir esse resultado, a Inteligência artificial é “alimentada” por dados em sua programação, que auxiliarão no alcance daquilo que é pretendido.

No caso desta pesquisa, a análise será feita com base nos algoritmos que podem ser lidos por dispositivos computacionais, que sejam computáveis – *softwares* de seleção de currículos, câmeras de carros autônomos, sistemas de reconhecimento facial e plataformas de redes sociais.

3.2. Discriminação algorítmica

Conforme expresso no artigo I, item 1 da Convenção Internacional Sobre a Eliminação de Todas as Formas de Discriminação Racial,

Discriminação consiste em “qualquer distinção, exclusão, restrição ou preferência fundadas na raça, cor, descendência ou origem nacional ou étnica que tenha por fim ou efeito anular ou comprometer o reconhecimento, o gozo ou o exercício, em igualdade de condições, dos direitos humanos e das liberdades fundamentais nos domínios político, econômico, social, cultural ou em qualquer outro domínio da vida pública¹⁰.

Existem diversos instrumentos normativos que conceituam o ato de discriminar, e em essência todos identificam a discriminação como sendo uma

⁸ LEE, Kai-Fu. **Inteligência Artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Rio de Janeiro: Globo Livros, 2019. p. 131.

⁹ MIASATO, Alessandra; SILVA, Fabiana Reis. Artificial Intelligence as an Instrument of Discrimination in Workforce Recruitment. **Acta Univ. Sapientiae**, Cluj-Napoca, v. 8, n. 2, p. 191-212, 2019.

¹⁰ ORGANIZAÇÃO DAS NAÇÕES UNIDAS - ONU. **Declaração universal dos direitos humanos**. Paris, 1948.

forma de tratar as pessoas de modo diferente, levando em consideração suas características físicas ou pessoais. De uma forma geral, as diferenças são utilizadas para justificar que pessoas diferentes devem ser tratadas de forma desigual, e que os mesmos direitos não devem ser garantidos a elas em igual proporção.

Juridicamente é legítima a discriminação positiva, a qual consiste em um tratamento diferenciado entre as pessoas quando o fundamento for o de melhoria das condições de um determinado grupo que, histórica, social ou economicamente estão em desvantagem quando comparados a outros grupos¹¹. Em contrapartida, é vedado qualquer prática discriminatória que não tenha como fundamento a ação positiva, ou seja, sempre que uma conduta visar o tratamento desigual arbitrário, por crenças ou estigmas, existirá uma discriminação ilegítima – negativa.

A discriminação pode se manifestar de diferentes formas, podendo ser com base na raça, cor, sexo, idade, religião, gênero, antecedentes criminais etc. E a utilização dessas características para diferenciar as pessoas pode acontecer nos mais diversos aspectos do cotidiano, como por exemplo nas relações empregatícias, em questões envolvendo segurança pública, sistemas de saúde, sistemas bancários e inúmeros outros.

As tecnologias que funcionam a partir de algoritmos, e são responsáveis por tomadas de decisões, dependem de um banco de dados que fornecerá as informações necessárias para a ferramenta atingir o resultado desejado. E a discriminação algorítmica pode acontecer quando os dados que alimentam a IA estiverem enviesados. Ou seja, quando o *input* estiver repleto de informações preconceituosas, o *output* será preconceituoso.

Quando o programador determina quais dados serão inseridos na tecnologia, por mais neutra que seja a sua análise quanto o que é “certo” e o que é “errado”, há que se levar em consideração que trata-se de uma pessoa humana que por si só é repleta de vieses. E em sendo o programador enviesado,

¹¹ MOREIRA, Adilson José. **O que é discriminação?** Belo Horizonte: Letramento: Casa do Direito: Justificando, 2017. p. 35.

certamente transportará para o algoritmo suas crenças e preconceitos, fazendo com que o risco de a IA tomar decisões discriminatórias aumente¹².

A frase *garbage in, garbage out*¹³ é um bom tradutor para esse problema. Na tradução significa “entra lixo, sai lixo”, o que é o mesmo que dizer: se uma sistema algorítmico é desenhado a partir de dados sujos, ruins, discriminatórios (entrou lixo), o único resultado possível dessa equação será uma série de decisões preconceituosas (saiu lixo).

Marcelo Chiavassa explica isso associando a Inteligência Artificial a uma criança de 5 anos, dizendo que se bem observado, não existe muita diferença entre as duas: uma criança não nasce racista, machista, homofóbica, por exemplo, mas se ela conviver com pessoas que proferem comentários com esse teor, existe uma grande probabilidade de ela refletir esses preconceitos com os quais cresceu¹⁴. De mesmo modo acontece com a IA – ela não tem preconceitos, mas quando algoritmos enviesados são introduzidos nela, muito provavelmente ter-se-á uma máquina reproduzindo comportamentos racistas, machistas, homofóbicos etc.

Com as mudanças tecnológicas que transformam a dinâmica cotidiana, é possível perceber um padrão discriminatório que se manifesta pelo uso de ferramentas automatizadas.

3.3. Casuística

O crescente uso de tecnologias de inteligência artificial e a adoção de sistema algorítmicos, viabilizou o surgimento de alguns problemas no tocante às decisões decorrentes desses recursos, sobretudo quando essas decisões são discriminatórias.

¹² MIASATO, Alessandra; SILVA, Fabiana Reis. Artificial Intelligence as an Instrument of Discrimination in Workforce Recruitment. **Acta Univ. Sapientiae**, Cluj-Napoca, v. 8, n. 2, p. 191-212, 2019.

¹³ NEFF-MALLON, Nathan. **Garbage in, garbage out**: machines learn to discriminate from discriminatory data. 2019. Disponível em: <<https://www.extractsystems.com/govnews-blog/2019/3/5/garbage-in-garbage-out-machines-learn-to-discriminate-from-discriminatory-data>>. Acesso em: 04 abr. 2020.

¹⁴ CHIAVASSA, Marcelo. Podcast: Algoritmos enviesados no processo de recrutamento. Spotify. Episódio 2.

Passa-se então a analisar alguns casos que evidenciam como o uso de tecnologias com tomada de decisão algorítmica podem ser discriminatórias.

3.3.1. Discriminação algorítmica na seleção de currículos

Com o tempo tornou-se muito comum que empresas se utilizassem de *softwares* que auxiliassem no processo de contratação de seus candidatos. Além de diminuir o tempo despendido nessa procura, as empresas podem se valer da rapidez que a Inteligência artificial terá para analisar um banco de dados com inúmeros currículos. Basta colocar no sistema as informações e características que o funcionário deve ter, e em minutos a IA fornecerá os currículos que estejam de acordo com os interesses da empresa.

A gigante Amazon é um bom exemplo do uso dessa tecnologia, que em 2014 utilizou um programa de IA para contratar engenheiros. Em 2015 a empresa percebeu que o seu sistema algorítmico ignorava os currículos de engenheiras¹⁵. O programa tinha como objetivo classificar os melhores candidatos para aquela vaga, e após algumas contratações, a Amazon percebeu que seu novo sistema não classificava candidatos para o emprego de maneira neutra em termos de gênero. A discriminação acontecia porque:

Os modelos de computador da Amazon foram treinados para examinar os candidatos observando padrões em currículos enviados à empresa durante um período de 10 anos. A maioria veio de homens, um reflexo do domínio masculino em toda a indústria de tecnologia¹⁶.

A partir desse comando, o algoritmo treinado para repetir o padrão dos últimos 10 anos de contratação entendeu que os candidatos homens deveriam ter preferência sobre as mulheres. A discriminação aconteceu com base no gênero. Após o problema ter sido evidenciado, a Amazon alterou a programação

¹⁵ THE GUARDIAN (Reino Unido). **Amazon ditched AI recruiting tool that favored men for technical jobs.** 2018. Disponível em: <<https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine>>. Acesso em: 10 fev. 2020.

¹⁶ THE GUARDIAN (Reino Unido). **Amazon ditched AI recruiting tool that favored men for technical jobs.** 2018. Disponível em: <<https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine>>. Acesso em: 10 fev. 2020.

do *software* para que ele tomasse decisões mais neutras e que não houvesse discriminação por gênero¹⁷.

Mas o gênero não é a única característica que pode ser utilizada para escolha de um currículo em detrimento de outro.

Anexar foto no currículo deixou de ser uma prática social, justamente como uma forma de se evitar a seleção de um candidato com base em suas características físicas. Mas um estudo realizado em 2004 nos Estados Unidos mostrou que muitas empresas utilizam programas de inteligência artificial para conseguir descobrir determinadas informações sobre os seus candidatos a partir de seus nomes. De acordo com a pesquisa, nomes que “soam branco” têm 50% mais chances de serem selecionados em comparação com os candidatos que têm nomes que “soam afro americanos”¹⁸. Na pesquisa, nomes como Brian e Emily são nomes que “soam brancos”, enquanto nomes como Jamal e Lakisha remetem a pessoas de origem afro descendentes¹⁹.

Desse modo, para realizar essa seleção, à inteligência artificial é dito quais os nomes são “bons” e quais os nomes são “ruins”, fazendo com que ela exclua parcela dos currículos de acordo com a base de dados que alimentou seu sistema.

Esse é um exemplo em que a discriminação se deu de forma proposital, mas esse nem sempre é o caso. A depender dos dados que são inseridos na IA, a discriminação pode acontecer que forma não intencional.

Quando da definição do candidato que a IA precisará selecionar, será preciso informar ao programa quais as características esse candidato deverá ter. Em classificando um bom candidato para a empresa como sendo “aquele que

¹⁷ THE GUARDIAN (Reino Unido). **Amazon ditched AI recruiting tool that favored men for technical jobs.** 2018. Disponível em: <<https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine>>. Acesso em: 10 fev. 2020.

¹⁸ BERTRAND, Marianne; MULLAINATHAN, Sendhil. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. **The American Economic Review**, Pittsburgh, v. 94, n. 4, p. 991-1013, set. 2004.

¹⁹ BERTRAND, Marianne; MULLAINATHAN, Sendhil. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. **The American Economic Review**, Pittsburgh, v. 94, n. 4, p. 991-1013, set. 2004.

nunca se atrasa”, pode acontecer de o algoritmo da IA entender que todos aqueles que moram afastados do local de trabalho sejam candidatos ruins, pois teriam maior propensão a se atrasarem; o mesmo pode acontecer com aqueles que dependem de transporte público para se locomover²⁰.

Fatores como depender de transporte público ou residir em determinada localidade não são suficientes para determinar a pontualidade de uma pessoa. Mas essas características seriam suficientes para ensejar em uma decisão enviesada do algoritmo, ainda que não intencional.

3.3.2. Provedores de busca, aplicativos e Redes Sociais

Um empresa que já teve seu algoritmo acusado de racismo algumas vezes foi a Google. Um dos casos aconteceu em 2015 envolvendo o aplicativo de fotos da empresa – Google Photos. Uma das funcionalidades do algoritmo é agrupar as fotos com conteúdo semelhantes. Um usuário percebeu que o aplicativo classificava todos os seus amigos negros como gorilas em decorrência da cor da pele, sendo o algoritmo incapaz de distinguir a pele de um gorila com a pele negra²¹.

O algoritmo não tinha informações, dados suficientes, que o ajudassem a fazer essa distinção, levando-o a fazer essa classificação enviesada. Diante das acusações a empresa se desculpou com o usuário e prometeu que resolveria o problema.

Outro problema no mesmo sentido que a Google teve envolvia seu provedor de buscas que classificou a foto de uma mulher negra com “cabelo feio”. A jovem, após dar uma entrevista em que afirmava que cabelo crespo não é feio, o algoritmo de reconhecimento de imagens do Google passou a

²⁰ COUNCIL OF EUROPE. **Discrimination, artificial intelligence, and algorithmic decision-making**. Disponível em: <<https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>>. Acesso em: 18 mar. 2020.

²¹ CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. Twitter, youtube: seriam os algoritmos discriminatórios? **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/twitter-youtube-seriam-os-algoritmos-discriminatorios>>. Acesso em: 17 de out. 2020.

apresentar sua foto como resultado de busca para “cabelo feio”²². Após tentar contato com a empresa para a desassociação da sua imagem à essa correspondência de palavras, não obteve resposta.

No dia 16 de outubro de 2020 um usuário da plataforma do Instagram resolveu fazer um teste do algoritmo de recomendações da plataforma. O algoritmo sugere conteúdo aos usuários a partir daquilo que está em alta e dos interesses de cada um²³. O jovem negro, que sempre posta fotos de seu rosto, recebe em média 18 mil alcances por dia em suas fotos, ou seja, o algoritmo recomenda o conteúdo para aproximadamente 18 mil pessoas em um dia. O teste que fizera envolveu a postagem de uma imagem de seu rosto, como de costume, mas dessa vez mudou o tom de pele para que se assemelhasse a uma pessoa branca. Em apenas 3 horas sua postagem já havia obtido mais de 34 mil interações – o algoritmo de recomendações do Instagram entendeu que a imagem era relevante o suficiente para ter mais impulsionamento²⁴. A única diferença entre as imagens era o tom da pele, o que foi suficiente para que o algoritmo entendesse por dar maior visibilidade à foto.

Com relação a rede social Twitter, uma série de acusações sobre o algoritmo ser racista começaram a acontecer no dia 19 de setembro de 2020, guardando relação com o sistema de corte de fotos da plataforma. Os usuários da rede identificaram que o algoritmo sempre destacava pessoas brancas em detrimento das negras²⁵.

O algoritmo recebe o comando para identificar em uma imagem pontos na imagem com maior probabilidade de serem olhados por uma pessoa, e para

²² GZH – Ciência e Tecnologia. **O que leva o Google a mostrar fotos de mulheres negras para a busca de “cabelo feio”**. Brasil, 2020. Disponível em: <<https://gauchazh.clicrbs.com.br/tecnologia/noticia/2020/07/o-que-leva-o-google-a-mostrar-fotos-de-mulheres-negras-para-a-busca-de-cabelo-feio-ckcq9jbmx002k013g2f9hw2xa.html>>. Acesso em: 10 out. 2020.

²³ TECHTUDO. **Instagram testa recomendação de posts no feed**. Brasil, 2018. Disponível em: <<https://www.techtudo.com.br/noticias/2018/08/instagram-testa-recomendacao-de-posts-no-feed.ghml>>. Acesso em: 20 abr. 2020.

²⁴ INSTAGRAM. **Perfil do usuário “hawk.jpg” na rede social Instagram**. Disponível em: <https://www.instagram.com/p/CGa1ButHepT/?utm_source=ig_web_copy_link>. Acesso em: 17 out. 2020.

²⁵ CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. Twitter, youtube: seriam os algoritmos discriminatórios? **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/twitter-youtube-seriam-os-algoritmos-discriminatorios>>. Acesso em: 17 de out. 2020.

isso, deve identificar os pontos de maior saliência/contraste. Mas o problema desse comando é que o algoritmo tem uma limitação para identificar contrastes nos rostos negros em comparado ao rosto branco, uma vez que os rostos negros apresentam bem menos pontos contrastantes. O resultado dessa limitante é na IA da plataforma acabar priorizando rostos brancos ao invés de rostos negros, por ser capaz de identificar pontos de saliência nesses e não nestes.

Neste caso, o algoritmo do Twitter fora programado para identificar os pontos de maior saliência da imagem, mas o problema guarda relação com o fato do algoritmo ter dificuldade em identificar contrastes em peles negras, resultando em decisões algorítmicas enviesadas, ainda que a intenção não seja a de discriminar.

3.3.4. Riscos dos Carros Autônomos

Não muito diferente das questões envolvidas no algoritmo enviesado da rede social Twitter, os carros autônomos possuem problemas para identificação de pessoas negras. Os carros autônomos geralmente utilizam câmeras e sensores que são responsáveis por detectar o pedestre, como é o exemplo dos carros autônomos desenvolvidos pela Tesla²⁶.

A programação para detecção de pedestres e objetos é fundamental para evitar acidentes, devendo o desenho do algoritmo ser feito com muito cuidado.

Em 2018, durante testes da Uber, o carro autônomo matou uma pedestre por não identificá-la como pessoa²⁷. Em decorrência da programação que não previa a possibilidade de a pessoa estar fora da faixa de pedestre, o carro autônomo o identificou como “outros” e acabou por atropelá-la.

Outro problema envolve a dificuldade que os carros autônomos têm de identificar pessoas negras. Segundo pesquisa realizada no início de 2019 pelo Instituto de Tecnologia da Universidade de Georgia, alguns dos sistemas

²⁶ FORBES. **Whats happens when self-driving cars kill people?** Disponível em: <<https://www.forbes.com/sites/cognitiveworld/2019/09/26/what-happens-with-self-driving-cars-kill-people/#12d09849405c>>. Acesso em 20 ago. 2020.

²⁷ TILT. **Carro autônomo da Uber mata pedestre nos EUA; empresa interrompe testes.** Disponível em: <<https://www.uol.com.br/tilt/noticias/redacao/2018/03/19/carro-autonomo-da-uber-mata-pedestre-nos-eua-empresa-interrompe-testes.htm>>. Acesso em: 14 fev. 2020.

utilizados por carros autônomos para detecção de pedestres tem dificuldades de identificar pessoas de pele mais escura²⁸. Isso as tornam mais suscetíveis de serem atropeladas.

São diversas as pesquisas que demonstram esse problema com a programação que é falha ao distinguir os tons de pele. De forma generalizada, o problema dessa limitação algorítmica decorre da falta de treino da IA para detecção de pessoas com tons de pele mais escuro.

3.3.5. Tecnologias de Reconhecimento Facial

Nos últimos 10 anos houve um grande aumento no uso de reconhecimento facial em todo o mundo, sendo um dos principais motivos para a adesão da tecnologia de vigilância, a segurança pública. Uma pesquisa realizada pelo grupo Surfshark concluiu que dos 196 países analisados, 109 já fazem uso da tecnologia, sendo que 3 países baniram seu uso (Luxemburgo, Marrocos e Bélgica)²⁹. Os principais motivos que levaram esses países a banirem o uso de reconhecimento facial são a falta de regulamentação adequada para garantia de privacidade e a liberdade do cidadão.

Ao redor do mundo o uso de reconhecimento facial é crescente, sobretudo como ferramenta para auxiliar investigações policiais, monitoramento público e privado e controle de acesso em aeroportos, por exemplo. Mas diversos são as notícias de que essas tecnologias podem cometer erros, como por exemplo, erros de identificação de pessoas negras³⁰.

Um caso que expressa bem esse problema aconteceu nos Estados Unidos, e é o que parece ser a primeira prisão errada no país com fundamento em reconhecimento facial, envolvendo um homem negro. Em janeiro de 2020 o americano Robert Williams foi preso pela polícia de Michigan depois que o

²⁸ WILSON, Benjamin; HOFFMAN, Judy; MORGENSTERN, Jamie. Predictive Inequity in Object Detection. **Arxiv**, Cornell University, 2019. Disponível em: <<https://arxiv.org/pdf/1902.11097.pdf>>. Acesso em: 27 jul. 2020.

²⁹ SURFSHARK. **The facial recognition world map**. Disponível em: <<https://surfshark.com/facial-recognition-map>>. Acessado em: 21.07.2020.

³⁰ CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. Reconhecimento facial, racismo e privacidade. **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/reconhecimento-facial-racismo-e-privacidade>>. Acesso em: 17 de out. 2020.

sistema de reconhecimento facial o identificou como o responsável por um crime de furto³¹. Após horas preso, e tendo sido realizado interrogatório em que o homem alegou não ser o autor do crime, e após a coleta de DNA, ele fora liberado com a comprovação de que a IA havia errado – o algoritmo não conseguiu fazer a distinção de seu rosto com o rosto do real autor do crime.

Na mesma linha dos problemas apresentados quando do uso de câmeras de detecção de pedestres em carros autônomos, as tecnologias de reconhecimento facial tendem a apresentar problemas na identificação e diferenciação entre pessoas negras. Segundo pesquisa realizada no Brasil, 90% das pessoas presas com base em sistemas de reconhecimento facial são negras³². Levando-se em consideração a propensão que a tecnologia tem em cometer erros com pessoas negras, esse alto número de detenções demonstra o porquê que a atenção para o uso de reconhecimento facial deve ser muito maior; ainda mais levando-se em consideração o histórico carcerário brasileiro que é majoritariamente negro.

Não à toa que *softwares* de reconhecimento facial para fins policiais estão começando a serem banidas, como é o exemplo de São Francisco – cidade berço do Vale do Silício, que proibiu o uso de reconhecimento facial para fins policiais³³.

Tecnologias de reconhecimento facial têm maior dificuldade em reconhecer e identificar pessoas negras. E esse problema é decorrente da falta de treinos. No momento do treinamento da IA, as amostras são majoritariamente feitas com pessoas brancas, tornando, naturalmente, a tecnologia mais precisa quanto à identificação de pessoas brancas. De igual modo, para melhoria desse

³¹ NY TIMES. **Wrongfully accused by an algorithm**. 2020. Disponível em: <<https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>>. Acesso em: 13 ago. 2020.

³² SINDICATO DOS EMPREGADOS DE AGENTES AUTÔNOMOS DO COMÉRCIO – SEAAC. **90% das pessoas presas por reconhecimento facial no país são negras**. Brasil, 2019. Disponível em: <<https://www.seaaccampinas.org.br/90-das-pessoas-presas-por-reconhecimento-facial-no-pais-sao-negras/>>. Acesso em 16 jun. 2020.

³³ INSTITUTO HUMANISTAS UNISINOS. **A pandemia acelerou em uma década e o “novo” preocupa, pois não é necessariamente algo bom. Entrevista especial com Marcelo Chiavassa**. Brasil, 2020. Disponível em: <<http://www.ihu.unisinos.br/159-noticias/entrevistas/599424-a-pandemia-acelerou-o-mundo-em-uma-decada-e-o-novo-nao-necessariamente-sera-um-ganho-social-entrevista-especial-com-marcelo-chiavassa>>. Acesso em 03 jun. 2020.

problema, é necessário que existam maiores testes e treinos para que o algoritmo adquira a capacidade de diferenciação.

Nesse sentido, buscando aprimorar a precisão de *softwares* de reconhecimento facial, o relatório da *New America Foundation* apresentou que algumas empresas de tecnologia chinesas estão realizando testes em massa em Zimbábue para aperfeiçoar suas tecnologias de reconhecimento facial. Um dos objetivos, segundo relatório, é o de melhorar a precisão dos algoritmos de reconhecimento, principalmente no que diz respeito às pessoas negras³⁴.

4. O exemplo da Nova Zelândia

Nos últimos anos restou-se evidenciado diversas manifestações de algoritmos de decisões automatizadas repletos de preconceitos, e com isso cresce as tentativas de regulamentação para o uso dessas tecnologias. E a ausência de modelos regulatórios para o uso da IA tornam o controle dessas ferramentas mais ardiloso.

Existem diversas tentativas para controle do uso de IA no sentido de garantir a não violação de direitos, tais como questões relacionadas aos direitos fundamentais, à liberdade individual, garantia de regras e respeito aos princípios éticos, responsabilização pelos danos causados etc. Mas até o momento não existe nenhum Estado que conseguiu regular de fato o uso da IA³⁵. E com o intuito de mitigar os riscos trazidos pela tomada de decisão algorítmica, a Nova Zelândia tornou-se um exemplo.

Em 2017, o país iniciou um projeto de 3 anos que tinha como objetivo a análise do impacto que o uso de inteligência artificial teria no país, sobretudo

³⁴ THE NATIONAL INTEREST. **Chinese companies use Zimbabweans as guinea pigs to identify black faces.** 2019. Disponível em: <<https://nationalinterest.org/blog/buzz/chinese-companies-use-zimbabweans-guinea-pigs-identify-black-faces-report-101447>>. Acesso em: 18 ago. 2020.

³⁵ CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. A tutela jurídica da IA: o caso da Nova Zelândia. **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/a-tutela-juridica-da-ia-o-caso-da-nova-zelandia>>. Acesso em: 17 de out. 2020.

quanto à ética da tecnologia quando utilizada em sistemas de segurança (público e privado), transparência e danos³⁶.

Após os três anos de análises, fora lançado um documento que expressa um rol de princípios que devem ser utilizados pelos programadores e utilizadores de IA, de modo a garantir que a tecnologia não oferecerá riscos aos cidadãos do país. O intuito é garantir uma IA atuando de forma ética, sendo assim uma ferramenta confiável, que garanta a segurança, privacidade, transparência, o bem-estar da população e que seja responsável³⁷.

Na mesma linha, o país lançou em meados de 2020 o *Algorithm Charter for Aotearoa New Zealand*³⁸ - Carta de Algoritmo para Aotearoa Nova Zelândia para o setor público. O objetivo do documento é garantir que o uso da IA com tomadas de decisões algorítmicas sejam livres de vieses, respeitando direitos como a privacidade e transparência. Para tanto, a criação, programação e uso das tecnologias de IA deverão ser baseadas nas regras expressas na Carta, sendo elas: Transparência, Parceria, Pessoas, Dados, Privacidade, Ética, Direitos Humanos, bem como ter a Supervisão Humana.

E como uma forma da organização que faz uso de IA garantir que sua tecnologia não oferece riscos de danos aos cidadãos, a Carta apresenta uma matriz de risco que deve ser utilizada. A Matriz classifica em nível alto, médio e baixo a probabilidade de a decisão tomada pelo algoritmo causar danos a alguém, de modo que:

A classificação “Baixo” indica que o impacto da decisão baseada em IA não gera consequências graves ou trata-se de um impacto isolado, podendo a Carta ser ou não aplicada; quando “Moderado”, significa que a decisão da IA pode atingir uma certa quantidade de pessoas gerando danos médios em suas vidas, e nesse caso a Carta deveria ser aplicada (embora não seja obrigatória); quando “Alto”, significa que a decisão da IA atingirá muitas pessoas, lhes causando graves consequências, e assim

³⁶ UNIVERSITY OF OTAGO. **Major new Otago study to tackle artificial intelligence law and policy**. 2017. Disponível em: <<https://www.otago.ac.nz/news/news/otago633498.html>>. Acesso em: 8 fev. 2020.

³⁷ AI FORUM NEW ZEALAND. **Trustworthy AI in Aotearoa: AI principles**. 2020. Disponível em: <<https://aiforum.org.nz/wp-content/uploads/2020/03/Trustworthy-AI-in-Aotearoa-March-2020.pdf>>. Acesso em: 05 abr. 2020.

³⁸ NEW ZEALAND GOVERNMENT. **Algorithm charter for Aotearoa New Zealand**. 2020. Disponível em: <https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-Charter-2020_Final-English-1.pdf>. Acesso em: 08 ago. 2020.

sendo, será obrigatória a aplicação das regras da Carta de forma integral³⁹.

A classificação define se a organização deverá ou não rever seus algoritmos, visando a eliminação do risco. A matriz deve ser constantemente utilizada, mantendo assim uma fiscalização quanto ao comportamento do algoritmo, assegurando que exista confiabilidade nas decisões proferidas e, consequentemente, que nenhum direito será desrespeitado e nenhum dano será causado.

Ainda que não se aplique a todos os usuários de IA, a Matriz de Risco é uma importante estratégia adotada pela Nova Zelândia como forma de evitar ao máximo decisões que causem, de algum modo, dano aos cidadãos neozelandeses.

5. Diretrizes Éticas da União Europeia

A preocupação com o teor dos dados que são inseridos nas tecnologias é grande, e cresce também a preocupação com a forma que o programador desenha o algoritmo e quais os dados ele entende serem os mais adequados para a programação daquela tecnologia – correndo-se o risco do programador transmitir suas ideologias e crenças à sua criação. Nesse sentido, a União Europeia lançou em 2018 um guia chamado *Ethics guidelines for trustworthy AI*. O documento sugere ao programador que, ao criar um algoritmo, sejam respondidas algumas perguntas que objetivam garantir que a IA funcionará de modo a evitar vieses, mitigando os riscos de decisões discriminatórias⁴⁰.

O guia é composto por 4 grupos de perguntas principais que os programadores devem fazer para garantirem a objetividade de seus algoritmos e a obtenção de um *output* livre de vieses. São exemplos de perguntas que

³⁹ CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. A tutela jurídica da IA: o caso da Nova Zelândia. **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/a-tutela-juridica-da-ia-o-caso-da-nova-zelandia>>. Acesso em: 17 de out. 2020.

⁴⁰ EUROPEAN COMMISSION. **Ethics guidelines for trustworthy AI**. 2019. Disponível em: <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477>. Acesso em: 25 abr. 2020.

devem ser respondidas: “Você estabeleceu uma estratégia ou um conjunto de procedimentos para evitar a criação ou o reforço de preconceitos injustos no sistema de IA, tanto no uso de dados de entrada quanto no design do algoritmo?”; também: “Você considerou a diversidade e a representatividade dos usuários nos dados? Você testou populações específicas ou casos de uso problemáticos?”⁴¹.

O objetivo da União Europeia ao desenvolver esse *check list* para o programador é estabelecer um conjunto de diretrizes éticas que sirva como base para a programação das IA. Em sendo a base ética, existe maior probabilidade de a programação ser mais objetiva e neutra, de modo a não transmitir preconceitos pessoais aos algoritmos, tornando-os menos tendenciosos e menos propensos à discriminação.

6. Conclusão

Na grande maioria das vezes, as tecnologias de IA que funcionam a partir de sistemas algorítmicos de decisões automatizadas tendem a oferecer resultados preconceituosos pela ausência de treinamento adequado da Inteligência Artificial. Problemas relacionados à carros autônomos que são menos precisos na identificação de pedestres negros, poderiam ser resolvidos com o aumento de testes com pessoas negras; esse comportamento racista do “automóvel” poderia ser mitigado com um estudo melhorado, uma vez que os *softwares* têm maior dificuldade em identificar a pele negra em detrimento da branca.

A mesma solução é válida para as tecnologias de reconhecimento facial, que têm maior dificuldade em diferenciar pessoas negras quando comparadas a pessoas brancas. Dessa forma, existiria uma necessidade da realização de mais testes com peles negras buscando o aprimoramento do *software*.

Outra questão diz respeito às bases de dados utilizadas para alimentar os sistemas de IA. Existe um problema ético por trás da programação dos

⁴¹ EUROPEAN COMMISSION. **Ethics guidelines for trustworthy AI**. 2019. Disponível em: <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477>. Acesso em: 25 abr. 2020.

softwares, tendo em vista que estes são programados por pessoas, que naturalmente são repletas de vieses, que podem ser transportados para a sua criação. Dessa forma, é necessário que medidas de prevenção sejam implementadas para garantir que as bases de dados sejam o mais objetivas e neutras possíveis, buscando que a IA tenha uma tomada de decisões livre de discriminação – eliminando o viés algorítmico.

Duas medidas seriam válidas para garantia da criação e funcionamento de tecnologias de IA com comportamentos mais éticos: i) o uso de uma base ética padronizada na qual o programador deve se basear no momento da criação do algoritmo; ii) auditorias constantes realizadas pelas empresas privadas e públicas que fazem uso de IA, com análises de riscos de dano pelas decisões algorítmicas de suas tecnologias, de modo a garantir que seus *softwares* estão tomando decisões objetivas e neutras.

7. Referências Bibliográficas

AI FORUM NEW ZEALAND. **Trustworthy AI in Aotearoa: AI principles**. 2020. Disponível em: <<https://aiforum.org.nz/wp-content/uploads/2020/03/Trustworthy-AI-in-Aotearoa-March-2020.pdf>>. Acesso em: 05 abr. 2020.

ANDRÉS, Moisés Barrio. **Manual de derecho digital**. Valencia: Tirant lo blanch, 2020.

BERTRAND, Marianne; MULLAINATHAN, Sendhil. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. **The American Economic Review**, Pittsburgh, v. 94, n. 4, p. 991-1013, set. 2004.

CHIAVASSA, Marcelo de Mello Paula Lima; ANDRADE, Victor Moraes. **Manual de direito digital**. Tirant: São Paulo, 2020. p. 17.

CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. Twitter, youtube: seriam os algoritmos discriminatórios? **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/twitter-youtube-seriam-os-algoritmos-discriminatorios>>. Acesso em: 17 de out. 2020.

CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. Twitter, youtube: seriam os algoritmos discriminatórios? **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/twitter-youtube-seriam-os-algoritmos-discriminatorios>>. Acesso em: 17 de out. 2020.

CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. Reconhecimento facial, racismo e privacidade. **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/reconhecimento-facial-racismo-e-privacidade>>. Acesso em: 17 de out. 2020.

CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. A tutela jurídica da IA: o caso da Nova Zelândia. **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/a-tutela-juridica-da-ia-o-caso-da-nova-zelandia>>. Acesso em: 17 de out. 2020.

CHIAVASSA, Marcelo de Mello Paula Lima; SILVA, Fabiana Aparecida dos Reis. A tutela jurídica da IA: o caso da Nova Zelândia. **Empório do Direito**, São Paulo, 2020. Disponível em: <<https://emporiiododireito.com.br/leitura/a-tutela-juridica-da-ia-o-caso-da-nova-zelandia>>. Acesso em: 17 de out. 2020.

CHIAVASSA, Marcelo. Podcast: Algoritmos enviesados no processo de recrutamento. Spotify. Episódio 2.

COUNCIL OF EUROPE. **Discrimination, artificial intelligence, and algorithmic decision-making**. Disponível em: <<https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>>. Acesso em: 18 mar. 2020.

EUROPEAN COMMISSION. **Ethics guidelines for trustworthy AI**. 2019. Disponível em: <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477>. Acesso em: 25 abr. 2020.

EUROPEAN COMMISSION. **Ethics guidelines for trustworthy AI**. 2019. Disponível em: <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477>. Acesso em: 25 abr. 2020.

FORBES. **Whats happens when self-driving cars kill people?** Disponível em: <<https://www.forbes.com/sites/cognitiveworld/2019/09/26/what-happens-with-self-driving-cars-kill-people/#12d09849405c>>. Acesso em 20 ago. 2020.

GZH – Ciência e Tecnologia. **O que leva o Google a mostrar fotos de mulheres negras para a busca de “cabelo feio”**. Brasil, 2020. Disponível em: <<https://gauchazh.clicrbs.com.br/tecnologia/noticia/2020/07/o-que-leva-o-google-a-mostrar-fotos-de-mulheres-negras-para-a-busca-de-cabelo-feio-ckcq9jbm002k013g2f9hw2xa.html>>. Acesso em: 10 out. 2020.

INSTAGRAM. **Perfil do usuário “hawk.jpg” na rede social Instagram**. Disponível em: <https://www.instagram.com/p/CGa1ButHepT/?utm_source=ig_web_copy_link>. Acesso em: 17 out. 2020.

INSTITUTO HUMANISTAS UNISINOS. **A pandemia acelerou em uma década e o “novo” preocupa, pois não é necessariamente algo bom. Entrevista especial com Marcelo Chiavassa**. Brasil, 2020. Disponível em: <<http://www.ihu.unisinos.br/159-noticias/entrevistas/599424-a-pandemia-acelerou-o-mundo-em-uma-decada-e-o-novo-nao-necessariamente-sera-um-ganho-social-entrevista-especial-com-marcelo-chiavassa>>. Acesso em 03 jun. 2020.

LEE, Kai-Fu. **Inteligência Artificial: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos**. Rio de Janeiro: Globo Livros, 2019.

MIASATO, Alessandra; SILVA, Fabiana Reis. Artificial Intelligence as an Instrument of Discrimination in Workforce Recruitment. **Acta Univ. Sapientiae**, Cluj-Napoca, v. 8, n. 2, p. 191-212, 2019.

MOREIRA, Adilson José. **O que é discriminação?** Belo Horizonte: Letramento: Casa do Direito: Justificando, 2017.

NEFF-MALLON, Nathan. **Garbage in, garbage out: machines learn to discriminate from discriminatory data**. 2019. Disponível em: <<https://www.extractsystems.com/govnews-blog/2019/3/5/garbage-in-garbage-out-machines-learn-to-discriminate-from-discriminatory-data>>. Acesso em: 04 abr. 2020.

NEW ZEALAND GOVERNMENT. Algorithm charter for Aotearoa New Zealand. 2020. Disponível em: <<https://data.govt.nz/assets/data->

ethics/algorithm/Algorithm-Charter-2020_Final-English-1.pdf>. Acesso em: 08 ago. 2020.

NY TIMES. **Wrongfully accused by an algorithm.** 2020. Disponível em: <<https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>>. Acesso em: 13 ago. 2020.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS - ONU. **Declaração universal dos direitos humanos.** Paris, 1948.

RASO, Filippo et al. **Artificial Intelligence & Human Rights: Opportunities & Risks.** Cambridge: Berkman Klein Center, 2018. 63 p. Disponível em: <<https://cyber.harvard.edu/publication/2018/artificial-intelligence-human-rights>>. Acesso em: 20 abr. 2020.

SINDICATO DOS EMPREGADOS DE AGENTES AUTÔNOMOS DO COMÉRCIO – SEAAC. **90% das pessoas presas por reconhecimento facial no país são negras.** Brasil, 2019. Disponível em: <<https://www.seaaccampinas.org.br/90-das-pessoas-presas-por-reconhecimento-facial-no-pais-sao-negras/>>. Acesso em 16 jun. 2020.

SURFSHARK. **The facial recognition world map.** Disponível em: <<https://surfshark.com/facial-recognition-map>>. Acessado em: 21.07.2020.

TECHTUDO. **Instagram testa recomendação de posts no feed.** Brasil, 2018. Disponível em: <<https://www.techtudo.com.br/noticias/2018/08/instagram-testa-recomendacao-de-posts-no-feed.ghtml>>. Acesso em: 20 abr. 2020.

THE GUARDIAN (Reino Unido). **Amazon ditched AI recruiting tool that favored men for technical jobs.** 2018. Disponível em: <<https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine>>. Acesso em: 10 fev. 2020.

THE NATIONAL INTEREST. **Chinese companies use Zimbabweans as guinea pigs to identify black faces.** 2019. Disponível em: <<https://nationalinterest.org/blog/buzz/chinese-companies-use-zimbabweans-guinea-pigs-identify-black-faces-report-101447>>. Acesso em: 18 ago. 2020.

TILT. **Carro autônomo da Uber mata pedestre nos EUA; empresa interrompe testes.** Disponível em: <<https://www.uol.com.br/tilt/noticias/redacao/2018/03/19/carro-autonomo-da-uber-mata-pedestre-nos-eua-empresa-interrompe-testes.htm>>. Acesso em: 14 fev. 2020.

UNIVERSITY OF OTAGO. **Major new Otago study to tackle artificial intelligence law and policy.** 2017. Disponível em: <<https://www.otago.ac.nz/news/news/otago633498.html>>. Acesso em: 8 fev. 2020.

WILSON, Benjamin; HOFFMAN, Judy; MORGENSTERN, Jamie. Predictive Inequity in Object Detection. **Arxiv**, Cornell University, 2019. Disponível em: <<https://arxiv.org/pdf/1902.11097.pdf>>. Acesso em: 27 jul. 2020.