

A utilização da inteligência artificial no processo de tomada de decisões

Por uma necessária *accountability*

HENRIQUE ALVES PINTO

Resumo: Com o desenvolvimento dos sistemas de inteligência artificial, novas possibilidades surgiram para os operadores do Direito na busca por mais eficiência no exercício de suas atividades. Apesar de a economia de tempo ser fator relevante na promoção da justiça, esta não pode abdicar de critérios éticos no processo de sua produção. Nesse contexto, este estudo propõe algumas discussões que devem ser feitas dada a crescente introdução dos sistemas de inteligência artificial no universo jurídico, de modo a submetê-lo a um indispensável filtro ético para garantir a sua legitimidade.

Palavras-chave: Decisão. Inteligência artificial. *Accountability*. *Machine learning*.

The use of artificial intelligence in the decision making process: for a necessary accountability

Abstract: With the development of Artificial Intelligence systems, new possibilities have emerged for Law Operators in the search for more efficiency in the exercise of their activities. Although time saving is a relevant factor in promoting justice, it cannot abdicate ethical criteria during the process of its production. This study proposes some discussions that should be made to the increasing introduction of Artificial Intelligence systems in the legal universe, in order to subject it to an indispensable ethical filter in order to guarantee its legitimacy.

Keywords: Decision. Artificial intelligence. *Accountability*. *Machine learning*.

Recebido em 24/10/19

Aprovado em 22/11/19

1 Introdução

Tem aumentado a percepção de que as transformações promovidas pelos avanços da inteligência artificial (IA) alterarão decisivamente os rumos tomados pela sociedade moderna. Apesar disso, mais que nunca, é necessário que a sociedade que se beneficia do avanço tecnológico saiba também lidar com as novas questões éticas que tal avanço tem implicado, com o intuito de amenizar os efeitos colaterais provocados pela era digital.

Esse novo avanço tecnológico – que pode ser comparado, sem exagero algum, com o início da Revolução Industrial, quando se trocou a força animal pelas máquinas a vapor – é, de fato, um caminho sem volta na história contemporânea da humanidade. Se na era da Revolução Industrial se delineou o que se conhece na atualidade como o homem moderno e urbano, hoje se vive na era de idealização do que poderá ser chamado, num futuro não muito distante, de cidadão digital.

Sob essa perspectiva, um dos impactos mais profundos do novo contexto social é a formação de um novo tipo de ser humano dotado de novas sensibilidades cognitivas que, ao trafegar por uma incrementada via tecnológica, produz grandes quantidades de informações numa escala de crescimento contínua. Com efeito, permitem-se às pessoas várias formas de conexões entre elas e de acesso a outros países, culturas e novas experiências; em contrapartida, existem também vários sinais de uma tendência à homogeneização nas formas de pensar, organizar a vida e planejar o futuro. De certo modo – e até contraditoriamente –, quanto mais variadas são as informações adquiridas nesse processo de alargamento dos horizontes, mais homogêneos se tornam os comportamentos experimentados pela sociedade.

Apesar de essas discussões provocarem um aprofundado debate em todas as áreas do co-

nhecimento científico, pretende-se discutir neste breve estudo algumas formas que os operadores do Direito podem empregar para lidar melhor com as questões advindas das transformações proporcionadas pela IA no campo jurídico.

Diante desse contexto e sem pretensão de exaurir o tema, discute-se aqui que as naturais resistências encontradas na comunidade jurídica a respeito do uso da via tecnológica no processo de tomada de decisões pelo Poder Judiciário – justificadas pelo fato, sobretudo, de o Direito ser uma ciência social aplicada – não necessariamente se configuram como um dos pontos mais cruciais nesse processo. Em vez disso, discute-se sob que bases éticas tais decisões são tomadas e de que forma elas podem passar por um rigoroso processo de *accountability*.

2 Inteligência artificial: alguns esclarecimentos necessários

Como todo conceito científico, várias são as definições encontradas na tentativa de determinar o que pode ser compreendido como IA. Algumas dessas definições salientam semelhanças comportamentais (*behavior*), ao passo que outras concentram seus estudos na forma humana de pensar (*reasoning*).

Para Russell e Norvig (1995, p. 5, tradução nossa), a IA está na precisão das decisões tomadas pelo sistema (*rationality*), estando elas sistematizadas em quatro categorias básicas: “I – um sistema que pensa como os seres humanos; II – um sistema que age como seres humanos; III – um sistema que pensa racionalmente; e IV – um sistema que age racionalmente”. De acordo com Bellman (c1978, p. 2, tradução nossa), que tem uma concepção matemática a respeito do assunto, a IA é “a automatização de atividades que associamos à cognição humana, tais como a tomada de decisões, resolução de

problemas e o aprendizado”. Genérica e sinteticamente, a IA é a tentativa de transpor a capacidade humana de cognição para sistemas artificiais¹.

Apesar de as técnicas de IA serem discutidas desde 1950, depois de um longo hiato em suas inovações, na última década houve considerável avanço nesse campo. Hoje, vários ingredientes têm contribuído de forma decisiva para o aperfeiçoamento dos sistemas de IA, como a grande massa de dados que passou a ser produzida e disponibilizada em ambientes de rede, a habilidade de aprendizado pelas máquinas diante das informações a elas transferidas, além do avanço dos computadores, com a redução de seus custos e a criação de novas modalidades de algoritmos voltados à compreensão e à simulação da capacidade humana de cognição.

À medida que se aprofunda o processo de *digitização* da sociedade, um imenso número de dados é produzido. Fotos, vídeos, textos, além de quaisquer outros tipos de informação, são transmitidos através do espaço virtual por meio do uso de computadores, *smartphones* ou outros meios tecnológicos. Os dados transmitidos – que impressionam não só pela grande quantidade, mas sobretudo pela velocidade com que são transmitidos – não podem ser analisados e compreendidos somente com base numa análise de correlações concebidas pela mente humana, sem a ajuda de programas tecnológicos mais inteligentes. E é exatamente nesse contexto que se inserem os sistemas de IA.

Sob tal aspecto, dois campos operacionais da IA merecem ser destacados: um deles é o *analytics*, palavra inglesa que se refere aos algoritmos que fazem a análise de dados e seus cruzamentos; o outro é o sistema do *machine learning*, cujos algoritmos “são capazes de prever ou generalizar padrões apreendidos a partir de um conjunto de dados utilizados para treinar o sistema” (WOLKART, 2019, p. 706).

No sistema de algoritmo de análise, os dados já estão estruturados e ajudam o seu usuário a fazer correlações na busca de padrões comportamentais diante da amostra por ele pesquisada: “Tanto os dados quanto os parâmetros possíveis de tratamento de dados são dados *a priori*, ficando ao operador do algoritmo a possibilidade de manipulá-lo dentro de um contexto específico e com algumas limitações” (GUTIERREZ, 2019, p. 85). Uma planilha simplificada do *Excel* é um bom e simples exemplo de sua aplicação.

¹ Sem desconsiderar a ideia geral em torno do conceito de IA, mencione-se também outra importante vertente do conceito, proposta por Kaplan e Haenlein (2019), que a definem como a capacidade de um sistema de interpretar e de aprender os dados obtidos fora do seu próprio sistema operacional de maneira correta, com a finalidade de utilizá-los para atingir objetivos e tarefas específicas por meio de uma adaptação flexível. No original: “we define AI as a system’s ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation” (KAPLAN; HAENLEIN, 2019, p. 17).

Por sua vez, os sistemas baseados em *machine learning* têm maior grau de complexidade, se comparados aos sistemas de algoritmo de análise, pois são capazes de prever ou generalizar padrões apreendidos com base num conjunto de dados utilizados para treinar o sistema. A construção algorítmica nesse sistema não depende de dados previamente escolhidos por seus operadores, isto é, o sistema aprende baseado na sua interação com um ambiente externo e dinâmico, por meio do qual realiza correlações com o objetivo de reconhecer padrões. Uma das diferenças entre o sistema de *machine learning* e o de *analytics* é que o primeiro é dotado de capacidade para analisar, fazer correlações e buscar padrões com base em dados não organizados, tais como fotos, vídeos ou textos coletados de *smartphones*.

De forma geral, é possível classificar o sistema de *machine learning* em dois subgrupos: os supervisionados e os não supervisionados. No grupo supervisionado, o aprendizado da máquina ocorre por meio de correlações iniciais definidas pelos seres humanos; ou seja, por esse grupo, mapeia-se um conjunto de informações para um dado conjunto de resultados, incluindo-se aí métodos como a árvore de classificação, as redes neurais ou a regressão linear. Em ambientes dinâmicos, são necessárias várias interações iniciais, com o intuito de ajustar o sistema de IA por alguém que, de fato, tenha um domínio específico da área de aplicação do sistema até que ele consiga produzir resultados mais precisos e minimamente satisfatórios.

No grupo não supervisionado, as informações que alimentam o sistema são rotuladas, mas os resultados não; isso significa que o algoritmo precisa inferir a estrutura subjacente dos próprios dados, com o intuito de agrupar seus elementos em categorias similares sem que se conheçam previamente a quantidade e a estrutura dos dados. Nesse sistema, é dispensado o ajuste inicial elaborado pelo sujeito que detém o domínio específico da área em que se deseja desenvolver a tecnologia de IA.

Os algoritmos do sistema não supervisionado aprendem com uma vasta quantidade de dados que estão disponibilizados imediatamente na internet ou em qualquer outra fonte (*big data*). Isso só é possível em virtude do desenvolvimento de novas tecnologias, como as redes neurais advindas de um desdobramento do *machine learning*: o *deep learning*. Pelo *deep learning*, o sistema passa a ser capaz não só de criar, mas também de estabelecer padrões de correlações próprias, desligados do raciocínio intelectual humano. E isso só é alcançado pelo sistema por meio de uma forma não linear de aprendizado por ele mesmo desenvolvida em várias camadas – algo similar ao que supostamente ocorre no cérebro humano por sua rede neurológica, na qual uma rede múltipla de unidades condutoras de dados se retroalimenta. A maioria dos *softwares* de reconhecimento de voz, de identificação de faces, de tradução, de reconhecimento de

objetos (COPELAND, 2016), entre outros, são bons exemplos de sistemas tecnológicos que já operam com o *deep learning* e que dependem de uma grande quantidade de dados disponíveis na rede virtual para que possam funcionar. Se funcionam bem ou mal, é o que precisa ser investigado.

Como este estudo não tem a pretensão de analisar exaustivamente o assunto, por ora, registra-se apenas que o estudo da IA é domínio muito amplo e que apresenta diferentes vertentes e aspectos a serem analisados, assim como a capacidade de conhecimento e raciocínio humano.

3 A utilização da inteligência artificial pelo Poder Judiciário

Ainda que cause certo desconforto, já não é mais novidade a utilização de tecnologia na tomada de decisões pelo Poder Judiciário. De acordo com Sartor e Branting (1998)², existem várias possibilidades de emprego da IA que podem ser ou já são utilizadas por esse Poder no exercício de suas atividades típicas. Tais possibilidades são voltadas ao aperfeiçoamento das atividades e podem ser aplicadas para auxiliar o raciocínio casuístico relacionado ao aprimoramento da performance argumentativa, associativa e discricionária dos magistrados. Para esses autores, tais sistemas trazem maior flexibilidade ao processo decisório, além de oferecer vantagens a toda a equipe de apoio ligada ao Poder Judiciário, de modo a amenizar as consequências dos excessos de litígios diante das limitações de recursos. Nesse sentido, a IA tornaria mais rápido, barato e previsível o acesso à Justiça, sem comprometer a sua fundamentação intelectual.

Ao mapear a aplicabilidade de sistemas inteligentes no exercício da atividade judicial, Chittenden (2017) apontou os seguintes sistemas: (i) *Public legal education*: desenvolvido pela Universidade de Cambridge,

²“As a reply to those challenges, AI & law has developed increasingly sophisticated models and techniques that address many of the concerns of the critics of earlier AI models, including case-based reasoning, formal dialectics, theory construction, neural networks, formal argumentation and negotiation, intelligent document assembly, and tools for supporting discretionary decision-making. These achievements are useful and important for the judiciary, because they provide a deeper and clearer understanding of some aspects of judicial problem-solving, and an effective support to the judges and their collaborators. Judges, squeezed between tightened budgets and increasing demands for justice, are desperately trying to maintain the quality of their decision-making process while coping with time and resource limitations. Flexible AI tools for decision support may promote a sufficient degree of uniformity and efficiency in judicial practice, while supporting a rational exercise of judicial discretion (and so possibly help to prevent, for example, the draconian rigidity of compulsory sentencing guidelines). In the same way, AI may help to reconcile flexibility, efficiency and accuracy in complementary tasks, such as the drafting of various judicial documents. In conclusion, we believe that the judiciary is in the early stages of a transformation in which AI technology will make the judicial process faster, cheaper, and more predictable without compromising the integrity of judges’ discretionary reasoning” (SARTOR; BRANTING, 1998, p. 110).

ajuda pessoas comuns a compreenderem melhor complexos problemas judiciais associados ao direito criminal inglês para que possam encontrar a solução mais adequada para os seus casos. Esse programa, mais recentemente, passou a ser usado também para os casos de divórcio (*DivorceBot*); (ii) *Case outcome prediction*: desenvolvido com base em pesquisas das Universidades de Londres e da Pensilvânia, tal sistema aplicou um algoritmo baseado em IA a 584 casos julgados pela Corte Europeia de Direitos Humanos com o intuito de pesquisar termos padronizados nas argumentações envolvidas em tais decisões. Esse sistema atingiu um índice de acertos de 79% por conta de sua capacidade de leitura de padrões não estritamente jurídicos, como delineamentos circunstanciais dos casos, linguagem empregada e organização dos tópicos contidos nas sentenças; e (iii) *Legal adviser support*: um consultor jurídico baseado no sistema *Watson*³, desenvolvido pela IBM, cuja habilidade principal é oferecer pareceres e apontar resultados mais precisos para processos judiciais.

No Brasil, um bom exemplo de aplicabilidade de sistemas inteligentes no Direito é a elaboração de peças processuais pelo programa *Dra. Luzia*. Desenvolvido pela Legal Labs, seu objetivo é auxiliar procuradorias da Fazenda Pública ligadas ao ajuizamento de execuções fiscais. Ele extrai dados públicos que apoiam o peticionamento individual ou em bloco e efetua uma melhor gestão de tais processos em seu acompanhamento.

Além disso, deve-se mencionar o importante projeto de pesquisa e desenvolvimento de aprendizado de máquina (*machine learning*) sobre dados judiciais das repercussões gerais do Supremo Tribunal Federal (STF). Chamado

³“A reputação do Watson vem da sua capacidade de entender e resolver perguntas em língua natural (QA, ou *question answering*), bem como de buscar no *big data* as informações para as perguntas que lhe são endereçadas (IE, ou *information extraction*)” (WOLKART, 2019, p. 753, grifos do autor).

Victor, o programa foi desenvolvido pelo próprio STF em convênio com a Universidade de Brasília. Apesar de estar em fase inicial de desenvolvimento, tem pretensões bastante inovadoras; um dos objetivos básicos desse projeto é a aplicação das técnicas do *machine learning* na busca e no reconhecimento de padrões nos processos jurídicos relativos a julgamentos de repercussão geral do STF.

Dados os esforços multidisciplinares envolvidos, esse projeto tem como principal objetivo desenvolver um sistema composto de algoritmos baseado no *deep learning* que propicie a automação de análises textuais desses processos, a fim de identificar os temas mais recorrentes de repercussão geral. De acordo com os pesquisadores envolvidos nesse projeto⁴, o algoritmo do programa *Victor* não tomará a decisão final acerca da repercussão geral, sendo sua atuação restrita à organização de tais processos, a fim de identificar de forma mais clara e consistente os temas mais recorrentes veiculados nos recursos extraordinários sob o regime de repercussão geral. Com isso, espera-se gerar “mais qualidade e velocidade ao trabalho de avaliação judicial, com a redução de tarefas de classificação, organização e digitalização de processos” (MAIA FILHO; JUNQUILHO, 2018, p. 226).

Sem desmerecer as inovações proporcionadas pela inserção das tecnologias de IA no universo jurídico, deve-se destacar que a atividade jurista-

⁴Em 30/8/2018, segundo a ministra Cármen Lúcia, ex-presidente do STF, os testes realizados naquela data, iniciados com 27 temas mais recorrentes no Tribunal, que representavam 60% do total de temas regularmente identificados, obteve um nível de precisão na triagem de 84%, podendo atingir 95% no mês seguinte. “A ministra ressaltou que o trabalho de conversão de imagens em texto, por exemplo, que um servidor executa em três horas, será feito em cinco segundos com a nova ferramenta. Salientou que a ferramenta possibilitará um melhor aproveitamento de recursos materiais e humanos do Tribunal, acelerando a análise dos processos e reduzindo o congestionamento na admissibilidade dos recursos nos tribunais de origem, auxiliando o Poder Judiciário a cumprir sua missão em diversas instâncias” (MINISTRA..., 2018).

dicional não se restringe apenas ao processo decisório, estando nela incorporados outros elementos tão relevantes quanto ele. Nesse sentido, uma das grandes dificuldades a ser enfrentada por esse processo de automação de textos legais para códigos tecnológicos é a necessidade de constante atualização, que acaba ficando sob responsabilidade de programadores ou outros profissionais de Tecnologia da Informação, os quais, na maioria das vezes, não têm muita familiaridade e *expertise* no trato não só das fontes legais do sistema jurídico, como também das questões éticas, sociais, políticas e filosóficas que as permeiam.

Ainda que a atividade jurisdicional envolva, em algumas ocasiões, julgamentos discricionários baseados em premissas e condições muito subjetivas dos magistrados, o simples emprego de instrumentos automatizados não é capaz de evitar enviesamentos⁵, já que estes costumam ser alimentados por base de dados que refletem tendências igualmente inclináveis a certas conclusões.

Por fim, com base nos estudos de Sourdin (2018, p. 1.130), apesar de a linguagem binária (sintática e semântica) ser perfeitamente conhecida dos sistemas computacionais, permitindo

⁵De acordo com Salas (2018), em reportagem do caderno de Tecnologia do jornal *El País*, um trágico exemplo disso foi o problema enfrentado pelo mais utilizado sítio de buscas da internet, o Google. “Em junho de 2015, um usuário do Google Photos descobriu que o programa etiquetava seus amigos negros como gorilas. A inteligência artificial do Google não era capaz de distinguir a pele de um ser humano da dos macacos, como gorilas e chimpanzés. Esse viés racista da máquina forçou o Google a pedir desculpas e prometeu encontrar uma solução para o erro. Dois anos depois, a solução é clara: para que o programa não confunda seres humanos com gorilas, tirou os gorilas do buscador. E os chimpanzés. E os macacos” (SALAS, 2018). Apesar de o Google ter tomado providências para resolver o problema – ainda que de maneira não muito satisfatória –, partindo da suposição de que um caso semelhante poderia ser repetido num processo judicial, isto é, um julgamento feito pela máquina com um algoritmo enviesado, nos casos em que a parte prejudicada tivesse esbarrado na certificação do trânsito em julgado, tal dano poderia ser irreparável em algumas ocasiões.

que realizem o perfeito processamento de informações, os sistemas ainda não estão aptos a incorporar os significados intrínsecos atrelados a expressões e situações particulares.

Com esse breve quadro, não se quer aqui refutar os ganhos do Poder Judiciário (e da sociedade) com a inserção dos recursos tecnológicos inteligentes no exercício de sua atividade. Entretanto, o objeto deste estudo, que será mais especificamente discutido nos próximos tópicos, é abordar o fato de essas novas ferramentas inteligentes deverem passar por um apurado filtro de controle ético, filosófico e social, sob pena de uma indevida e mecânica emulação da capacidade do raciocínio jurídico.

4 Uma necessária *accountability*⁶ para os sistemas de inteligência artificial

Sem sombras de dúvidas, os sistemas baseados em IA têm-se popularizado cada vez mais. Aos poucos, tal tecnologia deixa de ser privilégio de alguns para estar presente no cotidiano de pessoas, de empresas e do poder público. Sistemas de reconhecimento facial, *drones* inteligentes voltados ao monitoramento de grandes áreas rurais e urbanas, câmeras de segurança espalhadas em pontos estratégicos das cidades para o incremento da proteção da população,

⁶*Accountability*, termo da língua inglesa que ainda não tem uma tradução mais precisa na língua portuguesa, é conceito que denota práticas a serem observadas por aqueles que exercem relevantes funções em dada sociedade, a exemplo dos Poderes Públicos e das grandes corporações empresariais. Nesse sentido, de forma simplificada, *accountability* é o agir pautado por responsabilidade ética, transparência das ações, com uma devida e adequada prestação de contas de tais atos. Tal concepção liga-se à ideia de governança e também à de responsabilidade civil. Essas são apenas umas das razões pelas quais na atualidade esse assunto, combinado ao campo da IA, tem atraído cada vez mais a atenção de governos, corporações empresariais e organizações nacionais e internacionais.

uso militar voltado ao campo da espionagem, entre outros, são apenas alguns exemplos da utilização do novo avanço tecnológico baseado no mapeamento da forma cognitiva dos seres humanos.

Por conta disso, as decisões têm sofrido um intenso processo de automatização baseada em critérios na maioria das vezes não conhecidos ou bem explicados por seus criadores, de modo que passam a ter grande influência no dia a dia das pessoas sem que elas necessariamente percebam. De acordo com Gutierrez (2019, p. 86-87), algumas questões devem ser objetivamente respondidas, tais como:

– Como garantir que os sistemas de decisões automatizadas não discriminem (e, assim, respeitem o direito constitucional à não discriminação) ou não firam o direito à privacidade? – Quais são os critérios que estão embasando ou podem definir possíveis decisões de sistemas automatizados e que porventura podem ter como efeito a discriminação, ameaça à vida, à democracia ou ao cumprimento das leis vigentes? – É possível assegurar que um sistema de decisões automatizadas de determinada empresa está cumprindo as regras contratuais, as legítimas expectativas dos seus clientes e as leis vigentes?

O fato é que a tecnologia intelectual de hoje funciona como modalidade de cálculo racional da vida, o que faz com que essa extraordinária forma de pensar tenha um aumento exponencial de produtividade e eficiência. Isso não implica dizer que, paralelamente a todas essas relações entre pessoas e máquina, um de seus elementos mais caros, como a ética do comportamento humano, deva ser reduzido em sua complexidade – principalmente porque cada vez mais as relações humanas têm sido intermediadas por alguma modalidade de tecnologia da informação, seja ela de cunho inteligente ou não.

Apesar de se viver na Era do Pensamento Pragmático de alto poder de encantamento, é possível dizer que boa parte dessa sedução se liga ao fato de que, até agora, poucos ainda não perceberam que foram e continuam sendo encantados por esse pensamento. O problema não é reduzir a importância do pensamento pragmático, mas acreditar que todas as soluções para os efeitos colaterais do uso excessivo e irrefletido da tecnologia de hoje e do futuro só serão dadas também por essa mesma corrente do pensamento. Se levado ao seu extremo, o pragmatismo conduz à indolência do pensamento, “reduzindo-o a uma tendência hegemônica de se pensar e que, apesar de sua efetividade, gera o empobrecimento da contínua tarefa que todos os seres humanos herdaram que é a de construir a sua própria humanidade” (MARIZ, 2015, p. 71-72).

Na discussão das questionáveis condutas no manejo de sistemas de IA, já existem pessoas e entidades que voltam suas preocupações e estudos

aos sujeitos que desenvolvem e utilizam essa tecnologia, com o intuito de criar e de aprimorar importantes mecanismos de responsabilização. Entre as precursoras a se pronunciarem sobre o assunto, cita-se a IBM ([2017]), que lançou o manifesto *Data Responsibility @IBM*⁷ no final de 2017. Por ser a IBM detentora da maior gama de conhecimentos a respeito da concepção da IA, os trabalhos e palestras de sua CEO, Ginni Rometty, ao apresentar a posição da empresa a respeito do tema, têm revelado importantes contribuições ultimamente.

Em 2018, durante a reunião anual do Fórum Econômico Mundial em Davos, Rometty (2018, tradução nossa) fez um discurso sobre a posição da IBM em torno do assunto:

Quando se fala a respeito de novas capacidades de Inteligência Artificial, devemos ser claros sobre quando e como está sendo aplicada e sobre quem a treinou, com quais dados e como. Está refletindo a nossa *expertise* profissional? Existem preconceitos não intencionalmente inseridos nela? Nós devemos explicar por que seus algoritmos tomam as decisões que tomam. E se uma empresa não consegue fazê-lo, seus produtos não deveriam estar no mercado. Por último, nós ainda temos a responsabilidade de nos certificarmos de que as novas ondas tecnológicas não podem deixar ninguém à margem de seu processo. Isso implica que devemos investir em modernas formas de desenvolvimento de nossas habilidades com o intuito de garantir que a força de trabalho no mundo tenha o conhecimento e a experiência necessária para trabalhar com tecnologias

⁷Um dos seus princípios mais intrigantes está no seu art. 4^o, ao afirmar que a IBM acredita veementemente que a IA não pode substituir nem substituir as pessoas na tomada de decisões, em seus julgamentos, em suas intuições ou nas suas escolhas éticas. As companhias devem ser capazes de explicar quais são, de fato, as recomendações contidas em seus algoritmos. No original: “We firmly believe that artificial intelligence cannot and will not replace human decision-making, judgment, intuition or ethical choices. Companies must be able to explain what went into their algorithm’s recommendations” (IBM, [2017]).

como a IA ou o *blockchain*, preparando-se para os novos empregos que serão criados por essa era.⁸

No papel parece ser um bom discurso; no entanto, apenas na prática a sociedade global poderá descobrir o nível de comprometimento que a IBM de fato terá com os adquirentes de seus serviços e os afetados por eles. Apesar de, no discurso, haver uma suposta preocupação com os impactos que os sistemas de IA estão causando no mundo de hoje, fica patente também o argumento de solidificação do pensamento pragmático, que não está isento de críticas.

Além de esse tema ter sido debatido em 2018 no Fórum Econômico Mundial em Davos, foram realizados vários fóruns multilaterais, cuja temática principal girou em torno de práticas mais éticas na utilização dos sistemas de tecnologia de IA, com a criação de recomendações para agentes públicos e privados voltadas à criação de ações mais transparentes e passíveis de controle por parte de todos os setores da sociedade. De acordo com Gutierrez (2019, p. 88), uma das questões que assume o protagonismo nesse debate “é a definição universal sobre o que consiste um sistema de IA, qual é o seu ciclo de formação e quais são os seus principais atores”.

Simplificadamente, não há mais margens para dúvidas de que o tema da AI tem ocupado cada vez mais as discussões de todos os setores da academia, de governos e dos setores empresariais.

⁸No original: “When it comes to the new capabilities of artificial intelligence, we must be transparent about when and how it is being applied and about who trained it, with what data, and how. Does it reflect professional expertise? Are unintended biases built in? We must explain why its algorithms make the decisions they do. If a company can’t do that, their products shouldn’t be on the market. Lastly, we also have a responsibility to make sure that new waves of technology don’t leave anyone behind. That means investing in modern skills training to ensure the global workforce has the knowledge and experience to work in partnership with technologies like AI or blockchain, and is prepared for the ‘new collar’ jobs this era will create” (ROMETTY, 2018).

5 Algumas dificuldades enfrentadas pela inteligência artificial no trabalho de decodificação do processo cognitivo humano

As considerações a seguir configuram uma síntese do capítulo 2 (Os caminhos vitais) do livro de Carr (2011, p. 26), intitulado *A geração superficial: o que a internet está fazendo com os nossos cérebros*. Na obra, o autor descreve um breve roteiro da forma como opera o cérebro humano em seu processo de cognição desenvolvido por neurocientistas, biólogos e psicólogos.

A princípio, pode parecer que esta seção esteja descontextualizada do objetivo deste estudo – isto é, demonstrar a necessidade de aprimoramento dos mecanismos éticos e sociais com a utilização da IA no universo jurídico. Todavia, justifica-se tendo em vista que os sistemas de IA que atuam no mundo jurídico trabalham com o mecanismo do *machine learning*, com destaque para o *deep learning*, por operarem com mais autonomia do que o sistema analítico; nesse sentido, várias questões imprescindíveis devem ser apontadas, sob pena de o processo de alimentação de dados da máquina na criação do seu algoritmo não refletir necessariamente o ambiente cultural que ela pretende regular.

5.1 Os meandros pelos quais transitam o conhecimento

Mesmo com os avanços no conhecimento sobre o funcionamento físico do cérebro humano durante o século XX, uma antiga premissa entre neurologistas e biólogos permanecia intacta: a de que a estrutura do cérebro adulto jamais se alteraria. Apesar de a crença na imutabilidade do cérebro adulto estar amplamente difundida, havia uns poucos heréticos que a refutavam.

Com um novo enfoque, alguns psicólogos e biólogos começaram a constatar, no rápido e

crescente fluxo da pesquisa cerebral, algumas indicações de que até mesmo o cérebro de um adulto era maleável, dotado de plasticidade. Um dos grandes defensores dessa nova possibilidade era o biólogo britânico J. Z. Young, para quem o cérebro de fato poderia estar em constante fluidez, podendo ser adaptado a qualquer tarefa que lhe fosse requerida. Para Young (1951, p. 36, tradução nossa), “toda ação pode deixar alguma impressão no sistema nervoso”.

Young não foi o primeiro a discutir essa concepção. Setenta anos antes, o psicólogo americano William James já havia escrito algo semelhante a respeito da adaptabilidade cerebral. Para James (1890, p. 104-106), o cérebro é dotado de extraordinário grau de plasticidade, e, por isso, tanto forças externas quanto internas podem, de uma hora para outra, transformar a estrutura cerebral em algo bem diferente do que era antes. Essa mesma ideia foi adotada posteriormente por Sigmund Freud com algumas variações, que foram por ele descritas num manuscrito de 1895, infelizmente não publicado. Com o título *Project for a Scientific Psychology*, o manuscrito defendia que especialmente as barreiras de contato entre os neurônios poderiam alterar-se em resposta às experiências de uma pessoa.

Tais estudos foram frequentemente contestados pela maioria dos médicos e cientistas que estudavam os mecanismos de conhecimento do cérebro, pois todos estavam convictos de que a plasticidade do cérebro terminava na infância e que os caminhos vitais, sendo estabelecidos, não poderiam ser alargados ou estreitados, muito menos modificados⁹.

⁹De acordo com Carr (2011, p. 39), “a concepção do cérebro adulto como um aparato físico e imutável se desenvolveu a partir de, e foi ancorada por, uma metáfora da Idade Industrial que representava o cérebro como um maquinismo mecânico. Como uma máquina a vapor ou um dínamo elétrico, o sistema nervoso era composto por muitas partes, e cada uma tinha um propósito específico e predeterminado que contribuía de um modo essencial à operação bem-sucedida do todo. As partes não poderiam

Por conta da percepção de que o cérebro adulto é provido de uma estrutura estática, houve uma compreensão de que alguns problemas mentais não poderiam ser tratados, porque as intervenções seriam ineficazes ou injustificadas. Assim, como o cérebro não poderia mudar, a natureza humana por ele emanada parecia, necessariamente, fixa e inalterável. Não haveria regeneração, apenas degradação.

No final da década de 1960, Michael Merzenich – ao fazer seu trabalho de pós-doutoramento na Universidade de Wisconsin, em que ele mapeou o sistema cerebral por meio de sondas com a espessura de um fio de cabelo – começou a criar mapas muito detalhados no intuito de buscar novos *insights* a respeito da estrutura do cérebro. Após remover um pedaço do cérebro de um macaco, deixando uma pequena porção exposta, ele introduziu nela um micro eletrodo para registrar as sensações das mãos daquele animal. O resultado foi surpreendente: ao fazer pequenas incisões nos nervos sensoriais das mãos do macaco, ele percebeu que os nervos do animal voltavam a crescer a esmo, causando uma confusão no sistema cerebral. Isto é, as vias neurais do macaco haviam sido tecidas em um novo mapa mental, gerando aí um novo arranjo de nervos naquelas mãos (SYKA; MERZENICH, 2005).

Apesar de o trabalho de Merzenich, num primeiro momento, não ter despertado o debate na comunidade acadêmica dominante, voltada às concepções estáticas do cérebro, a partir da década de 1980, o seu trabalho começou a ser discutido mais seriamente. A neuroplasticidade

mudar, nem em sua forma nem em sua função, porque isso levaria, imediata e inexoravelmente, à avaria da máquina. Diferentes regiões do cérebro, e mesmo circuitos individuais, desempenhavam, papéis perfeitamente definidos no processamento de insumos sensoriais, em dirigir os movimentos dos músculos e em formar memórias e pensamentos; e esses papéis, estabelecidos na infância, não eram suscetíveis de alteração. Quando se tratava de cérebro, a criança era, de fato, como Wordsworth escrevera, o pai do homem”.

do cérebro era uma realidade para a qual não se podiam mais fechar os olhos.

Incentivados pelas descobertas de Merzenich, outros cientistas, após a realização de muitas outras experiências, chegaram à conclusão de que o cérebro não é exatamente plástico, mas, sim, maciçamente plástico. Apesar de a plasticidade cerebral ir diminuindo com o passar dos anos, ela nunca desaparece, pois os neurônios de uma pessoa sempre estão rompendo antigas conexões para formar novas, e, enquanto isso, várias células nervosas estão sendo constantemente criadas.

A plasticidade cerebral faz a harmonização entre duas filosofias que sempre estiveram em conflito: o empirismo, de John Locke, e o racionalismo, de Immanuel Kant. Para os empiristas, no nascimento a mente está em branco (“*tabula rasa*”) e o que se conhece advém das experiências acumuladas ao longo da vida; logo, o ser humano é produto da cultura, e não da natureza. Por sua vez, segundo os racionalistas, o ser humano nasce com modelos mentais inatos que determinam o modo como o mundo é percebido e compreendido, de forma que as experiências são submetidas a tais modelos mentais para serem filtradas; logo, há o predomínio da natureza, e não da cultura.

Nesse sentido, os genes humanos especificam muitas das conexões entre neurônios. Essas conexões geneticamente determinadas formam o modelo inato de Kant, que é a arquitetura básica do cérebro. No entanto, as experiências humanas regulam a intensidade, ou a efetividade de longo prazo, dessas conexões e, de acordo com o que sustenta Locke, configuram o remodelamento continuado da mente e a expressão de novos padrões de comportamento. De acordo com a neurociência, as filosofias em oposição dos empiristas e dos racionalistas encontraram seu fundamento comum na plasticidade cerebral.

Embora ofereça uma alternativa ao determinismo genético, a neuroplasticidade também

impõe a sua própria forma de determinismo ao comportamento humano. De acordo com estudos mais recentes da neurociência, circuitos particulares do cérebro começam a transformar essas atividades em hábito quando se fortalecem por meio da repetição de atividades físicas ou mentais. Exatamente nesse ponto reside o paradoxo da neuroplasticidade, pois ela acaba prendendo o cérebro, em tais ocasiões, a comportamentos rígidos.

Apesar de a plasticidade cerebral ter sido uma grande descoberta da neurociência nas últimas décadas, pode-se dizer que plástico não significa elástico, visto que o sistema cognitivo dos seres humanos não consegue retornar repentinamente ao seu estado anterior, a exemplo de um elástico de borracha. Assim, nada garante que o novo estado mental seja realmente desejável. Infelizmente, maus hábitos ficam arraigados no cérebro, assim como os bons hábitos cultivados. Da mesma forma que o cérebro pode construir circuitos novos e mais fortes por meio de práticas físicas ou mentais, tais circuitos também podem acabar sendo dissolvidos pela negligência. A possibilidade de decadência intelectual é inerente à maleabilidade do cérebro – o que, contudo, não significa que, diante de um esforço concentrado, não se possa mais redirecionar os sinais neurais para se reconstruírem algumas habilidades perdidas.

Nesta seção, almejou-se dizer e demonstrar que os ensinamentos da neurociência estão totalmente conectados aos estudos e às pesquisas que envolvem a IA. Na tentativa de transcrever o que pode ser aprendido pela máquina a respeito da forma como opera o processo de cognição no cérebro humano, a tarefa pode não ser tão perfeita assim – daí a necessidade de criar, discutir e aprimorar, a cada passo dado por essa nova tecnologia, instrumentos capazes de efetuar um forte filtro ético a respeito das formas pelas quais os sistemas de IA se têm

manifestado nos dias de hoje e quais podem ser as suas consequências no futuro.

5.2 Podem as máquinas pensar?

Um dos textos mais estudados a respeito da AI, desde seus primórdios, é o *Computing Machinery and Intelligence*, de Turing (1950). Nesse estudo, Turing procura demonstrar que, diante de determinadas condições, uma máquina, nos moldes de um computador digital, pode ocupar o lugar de um ser humano num diálogo. O objetivo maior do matemático britânico foi responder à seguinte questão: podem as máquinas pensar?

Para tanto, ele propõe uma espécie de jogo em que seus participantes, um homem A e uma mulher B, tentam dificultar o trabalho de um terceiro participante C, cuja função é interrogá-los, no intuito de determinar, por meio de suas produções linguísticas – às quais tem acesso apenas indiretamente –, quem é o homem e quem é a mulher. O objetivo do participante A é tentar fazer com que o interrogador C faça a identificação de modo equivocado ao lhe oferecer respostas falsas. A participante B, por sua vez, tem o objetivo de ajudar o interrogador, fornecendo-lhe respostas corretas e tentando convencê-lo de que seu adversário não está falando a verdade. Nessas condições, Turing (1950) propõe uma pequena alteração no jogo: a substituição do participante A por uma máquina, cuja função passaria a ser a de enganar o interrogador. Essa nova configuração do jogo é o que, para Turing, pode oferecer o contexto ideal, ou as condições necessárias, para que a máquina desempenhe aquilo que ele chama de “pensar”.

Desde o início de seu estudo, Turing deixa bem claro que seu propósito maior é unicamente fazer com que a máquina seja bem-sucedida no seu objetivo de não deixar o interrogador

saber que seu interlocutor não é humano. Assim, independentemente do que já foi discutido até hoje a respeito dos acertos e erros cometidos por Turing no seu importante estudo, o fato é que as máquinas hoje em dia conseguem, sim, em alguma medida, pensar com base em dados¹⁰ coletados que alimentam o seu sistema operacional, convertidos em linguagem algorítmica.

Contudo, uma questão primordial deve ser levantada a respeito daquilo que pensam as máquinas: elas são capazes de desempenhar tão bem quanto os seres humanos a função, por exemplo, de um falante inserido num contexto conversacional? Em qualquer ambiência na qual a linguagem é empregada ao conduzir a comunicação entre duas pessoas, sempre devem ser levadas em consideração algumas questões que vão muito além de aspectos puramente gramaticais. Nesse sentido, pode-se perguntar também de que maneira fatores subjetivos e extralinguísticos – como a intencionalidade, o eufemismo, a intuição, a ironia etc. – poderiam ser equacionados pelas máquinas; de que forma a criatividade do interlocutor poderia ser imitada; como o sistema operacional poderia imitar a criatividade do comunicador; de que maneira operaria a deformação da regra feita pela interpretação do hermenauta, entre outros questionamentos.

Tudo isso sem olvidar outra questão não menos importante levantada por Sourdin (2018) a respeito da dificuldade encontrada pelos sistemas de IA na diferenciação entre a semântica e a sintaxe, dado que a linguagem binária, natural em sistemas computacionais, não está apta a incorporar os significados intrínsecos atrelados a expressões e situações particulares.

Nesse contexto, um dos problemas a serem enfrentados pelo *machine learning* é exatamente estabelecer em que extensão e medida se poderá dimensionar a mais adequada relação entre a linguagem e o pensamento. Se se parte da premissa de que os limites da linguagem humana são os limites do mundo em que se vive, como poderá a máquina ser capaz de

¹⁰De acordo com Barocas e Selbst (2016), a mineração de dados, desde a sua coleta até a apresentação do resultado, costuma utilizar cinco mecanismos sobre os quais podem recair distorções: a) definição do problema; b) treinamento de dados; c) seleção de dados; d) utilização de *proxies*; e e) utilização de *masking*. Em tais hipóteses, “ao se definir o alvo variável, nomeando e coletando os dados de treinamento, usando a seleção de dados e tomando decisões com base nos modelos resultantes, cada um desses passos pode criar a possibilidade de um resultado final que tenha um impacto desproporcional em grupos protegidos, seja por especificar o problema a ser resolvido de modo a afetar grupos de forma distinta, por cometer falhas em reconhecer ou endereçar distorções estatísticas, por reproduzir preconceitos históricos, por estarem baseados em uma gama insuficiente de fatores. Mesmo nas situações em que os mineradores de dados tenham agido com bastante cautela, ainda existe a possibilidade de surgir resultados discriminatórios com os modelos que, ainda que de forma não intencional, utilizaram-se de *proxies* como variáveis em casos de grupos protegidos” (BAROCAS; SELBST, 2016, p. 675, tradução nossa).

reproduzir um modelo de realidade por ela não dominado, já que ainda não existe um modelo de linguagem que envolva todas as possibilidades de interpretação da realidade?

Portanto, propõe-se aqui que, no caso do desenvolvimento de sistemas de IA, tão válido quanto o processo de sua criação é o processo de sua ética aplicada diante do componente humano que a utiliza.

6 As dificuldades na auditoria de sistemas de inteligência artificial

Por conta da multiplicidade de sistemas de IA e suas múltiplas possibilidades de combinações, uma auditoria condizente torna-se bastante complexa. Não se quer dizer que um monitoramento viável e adequado seja algo impossível; o fato é que, na atualidade, muito é dito e desenvolvido a respeito da eficiência da aplicação de tais sistemas nas mais variadas áreas profissionais e do conhecimento, mas até agora pouco se fala a respeito de um *software* inteligente capaz de compreender as múltiplas linguagens traduzidas para esses sistemas, voltadas ao exame de eventuais vieses que podem ter sido captados pelas máquinas que acabam repetindo os preconceitos que são percebidos na era da modernidade.

Com o advento da tecnologia inteligente, o conjunto de variáveis envolvidas aumentou acentuadamente, e os limites de tempo e espaço para guardar e propagar as informações também se expandiram. Para Bell (1977, p. 45), a característica distintiva da tecnologia inteligente é o seu “esforço no sentido de definir a ação racional e identificar os meios para se chegar a tanto”.

Se, durante algum tempo, era suficiente a abertura para a auditoria dos códigos-fonte dos algoritmos para demonstrar a sua conformidade com leis ou padrões exigidos por governos, empresas, sociedade civil organizada e organizações internacionais, diante da atual diversidade de novos sistemas de IA, percebe-se que o grau de eficiência do seu monitoramento pode ser de eficácia reduzida. Um sistema de IA baseado num sistema analítico, por exemplo, pode ser de fácil controle; todavia, em sistemas inteligentes cujo modo operacional se dá por *machine learning*, seus mecanismos de monitoramento podem apresentar resultados totalmente insatisfatórios.

Como se comentou anteriormente, diante do fato de os algoritmos do *machine learning* serem desenvolvidos para interagir com ambientes externos e dinâmicos, seus códigos não determinam previamente as correlações e os parâmetros de análise, mas sim um padrão de probabilidades a ser preenchido com base na interação com o ambiente externo.

Para Gutierrez (2019, p. 90), “auditar tais códigos-fonte desses algoritmos tenderia a resultados inócuos”.

Todavia, ainda de acordo com Gutierrez (2019, p. 90),

[i]sso não significa, [...], a impossibilidade de auditorias e processos de *accountability* sobre sistemas de IA baseados em *machine learning* supervisionado. Nesse caso específico, é possível que se faça um registro dos *logs* de treinamento e calibragem dos sistemas de IA. A auditoria seria focada não no código-fonte, mas nesses *logs* que não são os *inputs* paramétricos desse tipo de sistemas de IA. Aliás, a construção e revisão desses parâmetros por equipes interdisciplinares e baseadas em amplo espectro de diversidade têm sido um mecanismo alternativo por empresas para evitar *by default* que esses sistemas tenham vícios de origem ou incorram em decisões éticas ou legalmente condenáveis. Embora ainda não sejam um requisito regulatório, os registros desses *logs* podem ser um importante recurso para empresas preocupadas em demonstrar seu compromisso e transparência [...]. Isso porque os parâmetros de correlações são formulados de maneira independente pelos sistemas a partir da interação com o ambiente dinâmico. E como foram formulados a partir de lógicas incomuns ao raciocínio humano, há grande dificuldade para se explicar de forma humanamente inteligível como esses sistemas chegaram a determinadas correlações ou resultados. E aqui, talvez, tenhamos de reconhecer que somos mesmo humanamente incapazes de fazê-lo e que necessitamos de outros ferramentais.

Segundo o autor, várias são as dificuldades e os desafios a serem enfrentados na busca por uma auditoria viável em termos de sistemas de IA. Contudo, outra questão que deve ser enfrentada é o fato de que, normalmente, quem está na ponta do desenvolvimento da tecnologia de IA são entidades de alto poder econômico, e, em algumas ocasiões, não existe o interesse em desenvolver paralelamente ao trabalho primordial uma pesquisa voltada ao controle do que ela está produzindo. Normalmente as entidades que voltam seus esforços para a construção de uma viável *accountability* a respeito do modo operacional dos sistemas de IA são organizações que dependem de doações de recursos dos setores público e privado, o que de certa maneira as impossibilita de acompanhar os desdobramentos mais atuais dessa tecnologia. Isso ocorre até mesmo porque grande parte do que se produz em termos de tecnologias de sistemas de IA está protegida por leis voltadas à proteção não só da propriedade intelectual como também do sigilo empresarial, e, quando tais tecnologias se tornam disponíveis no mercado, ainda existe a proteção do sigilo e da privacidade dos que a adquirem.

Nesse sentido, se por trás da criação, do desenvolvimento e da implementação da IA há toda uma organização em termos de material humano, incentivos financeiros e industriais em nível global, o mesmo

não se pode dizer a respeito da estrutura que está por trás dos que voltam as suas atenções e pesquisas ao controle ético e antropológico do que é considerado por alguns como o maior benefício da humanidade.

Desse modo, abrem-se novas perspectivas ou para fortalecer esse argumento ou para poder refutá-lo; e, independentemente da posição que se assuma, o fato é que as sociedades do futuro correrão sério risco de se submeterem a um rápido e intenso processo na contramão do que é prometido por tal tecnologia, caso o debate ético e antropológico a respeito dos sistemas de IA não seja efetivamente realizado e aprofundado por governos, sociedade civil, academia, setores empresariais, organizações nacionais e internacionais.

7 Conclusão

Diante do avanço tecnológico experimentado na atualidade, percebeu-se que por meio de algoritmos é perfeitamente possível produzir conexões e levantamento de dados que vão muito além da capacidade de cognição humana na realização de tais atividades. A cada dia, de modo perceptível ou não, a vida das pessoas tem sido regida pelos sistemas de IA, nas mais diversas áreas do conhecimento. O campo jurídico como um todo, inclusive a atividade judicial e as atuações de membros do Ministério Público, dos advogados e de autoridades policiais, está em franco e aberto processo de transformação. Perceber essas mudanças é um primeiro passo para, com base nelas, os profissionais da área jurídica poderem readequar as suas estratégias, a fim de conseguirem dar respostas mais rápidas, flexíveis e eficientes.

Contudo, é necessário que todos os envolvidos na área jurídica não abdicuem de um substrato mínimo ético essencial do processo de tomada de decisões, sob pena da redução do valor da Justiça a meros números estatísticos, que não atendem à realidade social que o Direito pretende regular.

Com isso, defende-se aqui a abertura de um amplo debate a respeito de sérias questões éticas envolvidas no processo não só de criação como também de aplicabilidade dos sistemas de IA nas áreas das Ciências Sociais Aplicadas, como ocorre com o Direito.

Apesar de o processo de tomada de decisões das máquinas da tecnologia do *machine learning* ser extremamente vantajoso e sedutor, a possibilidade de seus efeitos colaterais, a longo prazo, está diretamente ligada a essas dimensões – daí a necessidade de um redobrado trabalho de todos os setores, público e privado, a fim de que a sociedade de hoje e de amanhã, na busca de desenvolvimento social, não caia em situação completamente oposta, em que as desigualdades sociais acabem sendo agravadas.

Sobre o autor

Henrique Alves Pinto é mestre em Direito Público e Políticas Públicas pelo Centro Universitário de Brasília (UniCEUB), Brasília, DF, Brasil; doutorando em Direito Público e Políticas Públicas pelo UniCEUB, Brasília, DF, Brasil; professor de Direito Público e Direito Constitucional das Faculdades CNEC e FACTU, Unai, MG, Brasil; advogado.
E-mail: henrikiobrien@hotmail.com

Como citar este artigo

(ABNT)

PINTO, Henrique Alves. A utilização da inteligência artificial no processo de tomada de decisões: por uma necessária *accountability*. *Revista de Informação Legislativa: RIL*, Brasília, DF, v. 57, n. 225, p. 43-60, jan./mar. 2020. Disponível em: http://www12.senado.leg.br/ril/edicoes/57/225/ril_v57_n225_p43

(APA)

Pinto, H. A. (2020). A utilização da inteligência artificial no processo de tomada de decisões: por uma necessária *accountability*. *Revista de Informação Legislativa: RIL*, 57(225), 43-60. Recuperado de http://www12.senado.leg.br/ril/edicoes/57/225/ril_v57_n225_p43

Referências

BAROCAS, Solon; SELBST, Andrew D. Big data's disparate impact. *California Law Review*, [Berkeley, CA], v. 104, n. 3, p. 671-732, 2016. DOI: <http://dx.doi.org/10.15779/Z38BG31>. Disponível em: <https://canvas.stanford.edu/courses/60360/files/1469413>. Acesso em: 3 dez. 2019.

BELL, Daniel. *O advento da sociedade pós-industrial: uma tentativa de previsão social*. Tradução de Heloysa de Lima Dantas. São Paulo: Cultrix, 1977.

BELLMAN, Richard. *An introduction to artificial intelligence: can computers think?* San Francisco: Boyd & Fraser Pub. Co., c1978.

CARR, Nicholas. *A geração superficial: o que a internet está fazendo com os nossos cérebros*. Tradução Mônica Gagliotti Fortunato Friaça. Rio de Janeiro: Agir, 2011.

CHITTENDEN, Tara. AI: artificial intelligence and the legal profession. [S. l.]: The Law Society, 2017. (Horizon Scanning: forward thinking). Disponível em: https://academia.edu/36920594/Artificial_Intelligence_and_Legal_Profession. Acesso em: 3 dez. 2019.

COPELAND, Michael. What's the difference between artificial intelligence, machine learning, and deep learning? *NVIDIA*, [s. l.], July 29, 2016. Disponível em: <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai>. Acesso em: 3 dez. 2019.

GUTIERREZ, Andriei. É possível confiar em um sistema de inteligência artificial?: práticas em torno da melhoria da sua confiança, segurança e evidências de *accountability*. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (coord.). *Inteligência artificial e direito: ética, regulação e responsabilidade*. São Paulo: Revista dos Tribunais, 2019. p. 83-97.

IBM. *Data Responsibility @IBM*. [S. l.]: IBM, [2017]. Disponível em: https://www.ibm.com/blogs/policy/wp-content/uploads/2017/10/IBM_DataResponsibility-USLetter_WEB.pdf. Acesso em: 4 dez. 2019.

JAMES, William. *The principles of psychology*. New York: H. Holt and Company, 1890.

KAPLAN, Andreas; HAENLEIN, Michael. Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implication of artificial intelligence. *Business Horizons*, [s. l.], v. 62, n. 1, p. 15-25, Jan./Feb. 2019. DOI: <https://doi.org/10.1016/j.bushor.2018.08.004>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0007681318301393>. Acesso em: 4 dez. 2019.

MAIA FILHO, Mamede Said; JUNQUILHO, Tainá Aguiar. Projeto Victor: perspectivas de aplicação da inteligência artificial ao direito. *Revista Direitos e Garantias Fundamentais*, Vitória, v. 19, n. 3, p. 219-238, set./dez. 2018. DOI: <http://dx.doi.org/10.18759/rdgf.v19i3.1587>. Disponível em: <http://sisbib.emnuvens.com.br/direitosegarantias/article/view/1587>. Acesso em: 4 dez. 2019.

MARIZ, Ricardo Spindola. A aprendizagem no passo e descompasso da sociedade. In: SOUSA, Carlos Ângelo de Meneses (org.). *Juventudes e tecnologias: sociabilidades e aprendizagens*. Brasília, DF: Liber Livro, 2015. p. 59-80.

MINISTRA Cármen Lúcia anuncia início de funcionamento do Projeto Victor, de inteligência artificial. *Notícias STF*, Brasília, DF, 30 ago. 2018. Disponível em: <http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=388443>. Acesso em: 3 dez. 2019.

ROMETTY, Ginni. We need a new era of data responsibility. *World Economic Forum*, [s. l.], 21 Jan. 2018. Disponível em: <https://www.weforum.org/agenda/2018/01/new-era-data-responsibility>. Acesso em: 4 dez. 2019.

RUSSELL, Stuart J.; NORVIG, Peter. *Artificial intelligence: a modern approach*. Englewood Cliffs, NJ: Prentice Hall, 1995.

SALAS, Javier. Google conserta seu algoritmo “racista” apagando os gorilas. *El País*, [s. l.], 16 jan. 2018. Disponível em: https://brasil.elpais.com/brasil/2018/01/14/tecnologia/1515955554_803955.html. Acesso em: 4 dez. 2019.

SARTOR, Giovanni; BRANTING, L. Karl. Introduction: judicial applications of artificial intelligence. *Artificial Intelligence and Law*, [s. l.], v. 6, n. 2-4, p. 105-110, June 1998. DOI: <https://doi.org/10.1023/A:1008223408127>.

SOURDIN, Tania. Judge v robot?: artificial intelligence and judicial decision-making. *University of New South Wales Law Journal*, [s. l.], v. 41, n. 4, p. 1.114-1.133, Nov. 2018. Disponível em: <http://www.unswlawjournal.unsw.edu.au/wp-content/uploads/2018/12/Sourdin.pdf>. Acesso em: 4 dez. 2019.

SYKA, Josef; MERZENICH, Michael M. *Plasticity and signal representation in the auditory system*. New York: Springer, c2005.

TURING, Alan M. Computing machinery and intelligence. *Mind*, [s. l.], v. 59, n. 236, p. 433-460, Oct. 1950. DOI: <https://doi.org/10.1093/mind/LIX.236.433>. Disponível em: <https://academic.oup.com/mind/article/LIX/236/433/986238>. Acesso em: 4 dez. 2019.

WOLKART, Erik Navarro. *Análise econômica do processo civil: como a economia, o direito e a psicologia podem vencer a tragédia da justiça*. São Paulo: Revista dos Tribunais, 2019. Apresentada originalmente como tese de doutorado, Universidade do Estado do Rio de Janeiro, 2018.

YOUNG, John Zachary. *Doubt and certainty in science: a biologist's reflections on the brain*. Oxford, UK: Clarendon Press, 1951.