

INTELIGÊNCIA ARTIFICIAL: DESAFIOS E RISCOS ÉTICO-JURÍDICOS**ARTIFICIAL INTELLIGENCE: CHALLENGES AND ETHICAL-LEGAL RISKS**

Artigo recebido em 09/06/2020

Revisado em 24/07/2020

Aceito para publicação em 21/08/2020

Mateus de Oliveira Fornasier

Professor do Programa de Pós-Graduação Stricto Sensu (Mestrado e Doutorado) em Direito da Universidade Regional do Noroeste do Estado do Rio Grande do Sul (UNIJUI). Doutor em Direito Público pela Universidade do Vale do Rio dos Sinos (UNISINOS), com Pós-Doutorado em Direito e Teoria (Law and Theory) pela University of Westminster (Reino Unido)

Norberto Knebel

Doutorando em Direito pela Universidade Regional do Noroeste do Estado do Rio Grande do Sul (UNIJUI). Mestre em Direito pelo Centro Universitário La Salle (UNILASALLE). Bolsista CAPES.

RESUMO: Este artigo objetiva compreender o que é a IA e a sua aplicabilidade social para o aumento da efetividade dos mais variados processos de produção, de gerenciamento e de comunicação na sociedade. Resultados: i) ainda é muito difícil prever o momento em que a humanidade obterá uma IA que se comporte como a inteligência humana, mas o atual grau de desenvolvimento dessa tecnologia já representa riscos para a humanidade; ii) iniciativas de aproximação (e não apenas de informação) do público em relação às decisões importantes atinentes à IA é um meio de se fazer com que o desenvolvimento da IA se dê para as pessoas, e não às suas custas. Assim, instituições públicas e da sociedade civil devem ser reprojctadas para superar o déficit democrático e o medo em relação à IA. Metodologia: método de procedimento hipotético-dedutivo, com abordagem qualitativa e técnica de pesquisa bibliográfica e documental.

PALAVRAS-CHAVE: Inteligência artificial; regulação; ética.

ABSTRACT: This article aims to understand what AI is and its social applicability to increase the effectiveness of the most varied production, management and communication processes in society (both in public and private sectors). Results: i) it is still very difficult to predict the moment when humanity will obtain an AI that behaves like human intelligence, but the current degree of development of such technology already represents risks for mankind; ii) initiatives to bring the public closer (and not just information) to the important

decisions related to AI is a means of making AI development happen to people, and not at their expense. Thus, public and civil society institutions must be redesigned to overcome the democratic deficit and fear in relation to AI. Methodology: hypothetical-deductive procedure method, qualitative approach and bibliographic and documentary research technique.

KEYWORDS: Artificial intelligence; regulation; ethics.

SUMÁRIO: Introdução. 1 Inteligência artificial: evolução histórica e definições. 2 Desafios e riscos éticos da aplicação da IA. 3 Expectativa para a regulação jurídica no Brasil. Conclusão. Referências.

INTRODUÇÃO

Mão de obra e processos produtivos mais baratos e eficientes. Uma Administração Pública mais eficiente, mais confiável, menos custosa e menos burocratizada. Veículos autônomos que conduzem pessoas sem causar acidentes. Cirurgias que não padecerão de erros da falibilidade humana. Sistemas que aprenderão profundamente a partir de dados registrados, melhorando a saúde da população. Eis algumas das promessas da aplicação da IA propagadas nos meios de divulgação científica e tecnológica da atualidade. Mas o lado reverso da IA é repleto de previsões sobre o fim da humanidade (ou da civilização): máquinas que perpetrarão discriminações e abusos, causando maior exclusão social; que tomarão empregos; que causarão destruição em massa em conflitos; que não respeitarão os direitos civis daqueles que têm seus dados coletados para melhorar os sistemas, conforme os interesses das empresas contratadas para desenvolvê-los.

Trata-se de questão de grande relevância social, portanto, estudar a aplicação da IA na sociedade. Muito se tem apresentado preocupações éticas relacionadas a essa tecnologia na literatura acadêmica, pois ela permite um desenvolvimento com profundas implicações para a humanidade. A continuidade da sociedade como conhecida, da produção e da própria dignidade humana — já que há quem levante a questão de que o ser humano, dentro em breve, talvez se torne desnecessário para a produção material e de conhecimento, com a IA — representam sérias preocupações, as quais redundam em diversas questões jurídicas relevantes relacionadas à regulação de tal tecnologia.

Nesse sentido, quais são os principais desafios e riscos que a aplicação das tecnologias atinentes à IA acarreta? Como hipótese a essa questão, apresenta-se que preocupações éticas (perpetração de preconceitos sociais e a ampla possibilidade de cometimento de fraudes) são

tão importantes quanto questões técnicas (complexidade do desenvolvimento de tais sistemas), econômicas (custos para o desenvolvimento do conhecimento relacionado à IA e para sua implementação) e político-burocráticas (desenvolvimento de lideranças e de políticas capazes de tornarem efetiva a aplicação da IA na sociedade).

O objetivo geral dessa pesquisa, desenvolvida pelo método de procedimento hipotético-dedutivo, com abordagem qualitativa e técnica de pesquisa bibliográfica e documental, é compreender o que é a IA e a sua aplicabilidade social para o aumento da efetividade dos mais variados processos de produção, de gerenciamento e de comunicação na sociedade, fornecendo-se assim subsídios para o desenvolvimento de políticas públicas, de regulação e de legislação relacionadas a esse ramo tecnológico. Para a sua consecução, dividiu-se seu desenvolvimento em três seções. A primeira compreende as principais definições e terminologias da IA. Já a segunda apresenta sérios riscos e desafios acarretados pela aplicação da IA. E a terceira debate os incipientes movimentos legislativos no Brasil para regulação da IA, tendo em vista perceber-se como esses projetos recebem esse cenário de desafios, riscos e oportunidades ligados ao tema.

1 INTELIGÊNCIA ARTIFICIAL: EVOLUÇÃO HISTÓRICA E DEFINIÇÕES

A IA não é um conceito novo — grande parte de seu apoio teórico e tecnológico foi avançado nas últimas seis décadas. O início oficial da AI é considerado a “conferência de Dartmouth” (1956) (LI; DU, 2017) e, em certa medida, o teste de Turing, a ela anterior, ofereceu pensamentos sobre como reconhecer uma “máquina inteligente”. A IA encontra suas raízes históricas profundas na construção de autômatos, na reflexão sobre a lógica e suas consequências, e no desenvolvimento de máquinas de calcular (LEXCELLENT, 2019, p. 13 e ss.). Mas o fundador da ideia de IA foi o matemático britânico, que em 1950 lançou o seu conceito ao descrever o “jogo da imitação” em um famoso artigo, cujo principal questionamento pode ser assim descrito: pode um homem, conectado por um teleprinter, ao que ele não sabe ser uma máquina em uma sala vizinha, ser enganado e manipulado pela máquina com uma eficiência comparável à de um ser humano? Para Turing, a IA era a mais completa farsa da psicologia humana.

A antiga receita de redes neurais artificiais, desenvolvida já na década de 1950 e renomeada como “aprendizado profundo” (*deep learning*), voltou a ser tendência em meados da década de 2010. Em outubro de 2012 foi organizado um concurso de reconhecimento de imagens chamado ImagNet, cujo objetivo era reconhecer milhares de imagens tiradas do

Flickr, atribuindo a eles o rótulo correto entre cerca de 1.000 categorias (gato persa, husky, robalo, bisonte, acordeão, berço, trator, etc.). A taxa de erro foi de apenas 16% no *deep learning* (equipe de Geoffrey Hinton), contra uma taxa de 25% no SMS, método de triagem concorrente. Outro setor completamente dominado por esses novos cérebros artificiais foi a tradução automática. O software desenvolvido por Alex Graves foi o melhor a reconhecer escritos em árabe, farsi e francês.

Mas a maturidade de um campo pode ser medida tanto pelos sucessos quanto pelos primeiros fracassos, e o *deep learning* forneceu vários deles: por exemplo, pessoas negras eram confundidas com gorilas, e em março de 2016, uma conta no Twitter criada por um robô Microsoft Tay tornou-se “racista” ao se inspirar um pouco demais os tweets encontrados. Em 2017, um carro do Uber matou um pedestre, e em setembro daquele mesmo ano, um algoritmo prometeu adivinhar a orientação sexual das pessoas. Assim, logo emergiu o medo de desenvolver uma IA que escaparia ao controle dos seres humanos.

Existem dois grandes tipos de IA: a IA fraca, em que os computadores simulam um comportamento inteligente, mas não necessariamente têm uma mente (ou seja, não têm consciência, não compreendem o que estão realizando); e a IA geral (ou forte), segundo a qual uma máquina realmente tem uma mente (ou seja, compreende o que está realizando, não apenas simulando comportamento) ou apenas a simular uma mente (NEAPOLITAN; JIANG, 2018, p. 2-3).

Quando considerada como sendo o desenvolvimento de uma entidade artificial, capaz de aprender e tomar decisões em um ambiente complexo e mutável, afetando esse ambiente e comunicando seus conhecimentos e decisões aos humanos, a IA (IA) ainda nem está próxima de ser desenvolvida pela humanidade. Mas, em seus esforços para desenvolver a IA, os pesquisadores analisaram o comportamento e o raciocínio de entidades inteligentes (como seres humanos, por exemplo) e desenvolveram algoritmos baseados nesses estudos. Tais algoritmos foram usados para resolver muitos problemas interessantes, incluindo o desenvolvimento de sistemas que se comportam de forma inteligente em domínios limitados. Dentre tais algoritmos destacam-se sistemas que realizam diagnósticos médicos, identificam problemas com software, tomam decisões financeiras, navegam em terrenos difíceis, monitoram falhas de veículos espaciais, reconhecem fala e rostos, entendem textos, planejam itinerários, rastreiam alvos, aprendem acerca de preferências individuais, aprendem as relações causais entre os genes (bem como quais genes afetam um fenótipo).

Embora seja difícil imaginar a inteligência geral artificial do nível humano (geral) sem alguns dos atributos funcionais que associamos à consciência nos seres humanos (i.e.

consciência do mundo ou integração cognitiva), é possível imaginar uma IA geral do tipo não humano, totalmente desprovida de emoções, funcional e fenomenologicamente falando (SHANAHAN, 2019, p. 99). Muitos filósofos, contudo, concordariam que a emoção, embora talvez não seja necessária, é possível na IA do nível humano, funcional e/ou fenomenologicamente, principalmente se a IA estiver em conformidade com o plano do cérebro biológico. No futuro, o lado dessa divisão em que um determinado sistema de IA cai terá claramente implicações morais significativas.

A eventual criação de uma IA no nível humano traria implicações consideráveis para a humanidade em geral, especialmente porque o advento da IA no nível humano muito provavelmente seria seguido rapidamente pela IA com inteligência sobre-humana — e esse processo continuaria, exponencialmente.

Mas também é saliente que mesmo a IA do estado atual tem um rico potencial para transformar a sociedade da substituição “trivial” do trabalho humano tedioso (por exemplo, controlando robôs para limpar o chão e cortar a grama) para um novo papel mais complexo como facilitador social internacional (ajudando as pessoas a se comunicarem com mais facilidade, com tradução simultânea entre idiomas, por exemplo), ajudando o Estado a fazer um uso substancialmente melhor de recursos públicos escassos (por exemplo, sistemas que ajudem a melhor organizar sistemas públicos de saúde, identificando modos de realizar economia no orçamento de compras).

Assim, tem-se que o espectro de possibilidades expostas entre os estudiosos vai desde aqueles que argumentam que a IA mudará para sempre a vida humana em um futuro próximo até aqueles que argumentam que um apocalipse da IA é apenas ficção científica; mas a verdade é que nenhum dos campos entende completamente a IA geral os impactos que ela poderia ter no nosso mundo (HANEY, 2018, p. 168-169). Até que a humanidade crie a IA geral, ela estará além da compreensão da humanidade. Essa realidade representa um destino irônico para a humanidade: primeiro, deve entender a IA geral para controlá-la, mas não pode entendê-la até que seja criada. Ademais, o resultado positivo da criação da IA geral não pode ser tido como certo. O consenso geral no campo da IA é que nenhum conjunto de regras é capaz de controlar a totalidade do que a regulamentação para a IA geral. Mas não importa quantos padrões possam ser reconhecidos e tendências possam ser rastreadas, o futuro da IA não acontecerá por si só.

2 DESAFIOS E RISCOS ÉTICOS DA APLICAÇÃO DA IA

As tendências futuras próximas relacionadas aos possíveis desenvolvimentos da IA, segundo Luciano Floridi (2020, p. 140), são: i) mudança de dados históricos para dados sintéticos; e ii) tradução de tarefas difíceis (em termos de habilidades) em tarefas complexas (em termos de computação). Tais tendências são complementares e previsíveis. Isso significa dizer que, cada vez mais, se procurará desenvolver a IA usando dados, tanto quanto possível, híbridos e preferencialmente sintéticos, através de um processo de ludificação (projeto/redesenho das realidades com as quais se lida — o que vem a demonstrar que o futuro da IA depende das habilidades de design e engenhosidade) de interações e tarefas, afastando-se do uso de dados puramente históricos. E isso será feito mediante a tradução, tanto quanto possível, dos problemas difíceis em problemas complexos, envolvendo as realidades em torno das habilidades de nossos artefatos. Ou seja: procuraremos criar dados híbridos ou sintéticos para lidar com problemas complexos, ludificando tarefas e interações em ambientes envelopados. Quanto mais isso for possível, mais bem-sucedida será a IA.

Essa tendência também dependerá da capacidade de negociar os resultantes (e graves) problemas éticos, legais e sociais (ELSI), desde novas formas de privacidade (preditiva ou baseada em grupo) até estímulos de comportamentos e autodeterminação. A própria ideia de que estamos cada vez mais moldando nossos ambientes para torná-los compatíveis com a IA deve fazer qualquer um refletir. Antecipar essas questões, para facilitar o ELSI positivo e evitar ou mitigar as negativas, é o valor real de qualquer análise prospectiva. É interessante tentar entender quais podem ser os caminhos de menor resistência na evolução da IA. Mas seria bastante estéril tentar prever “quais grãos crescerão e quais não crescerão” e depois não fazer nada para garantir que os bons grãos cresçam e os ruins não. Assim, o desafio à frente em relação à IA não será tão relacionado à inovação digital em si, mas à governança das tecnologias digitais, incluindo-se a IA.

Conforme enuncia Cowls (2020, p. 101; p. 11-112), a IA está mudando rapidamente o relacionamento entre as pessoas e o Estado com o seu poder de racionalização. Mas o poder "reduzidor" dos sistemas de IA parece representar uma ameaça à democracia (a não ser que tais sistemas venham a ser desenvolvidos com preferências, perspectivas e prioridades públicas em mente). Ou seja, é necessário superar o pensamento de que se deve ater apenas à conformidade legal mínima e à fé no livre mercado, passando-se a considerar a opinião pública como sendo elemento constitutivo da legitimação do uso da IA na sociedade. Embora a bagunça da política possa impedir respostas claras sobre o uso da IA, isso é preferível aos

sistemas de IA "friamente racionais", mas democraticamente deficientes que até então se tem desenvolvido. Ou seja, os sistemas de IA não devem ser apenas treinados com os dados humanos acumulados com o mínimo consentimento e implantados apenas com poder e lucro em mente, mas sim, que sejam projetados e implantados de maneiras que respeitem as preferências, perspectivas e prioridades das pessoas.

Isso nos obriga a tornar mais concretas as visões e valores que devem moldar nossas sociedades. Com o surgimento da internet, muito se falou da ideia de que agora revelamos nossos pensamentos, sentimentos e preocupações mais íntimos on-line, tornando-os digitais. Mas treinar os sistemas de IA para operar de forma equitativa pode ter um efeito reverso. Ao manifestar pontos de vista e valores das pessoas, das diversas maneiras discutidas acima, talvez a tecnologia digital não revele quem as pessoas realmente são, mas sim, que fazemos com que a tecnologia aja mais como nós. Assim, não apenas se deve aspirar a uma IA que raciocine de modo transparente e explicável, mas também de maneira pública, refletindo não a sociedade como ela é, mas a sociedade como deveria ser. Pedir às pessoas suas perspectivas, preferências e prioridades para o que a IA deve fazer é o primeiro passo essencial nesse processo.

É claro que a aspiração de se perguntar às pessoas o que desejam com a IA também é arriscado, pois atribui-se um enorme poder a quem faz a pergunta. Assim, pragmaticamente, deve-se investir a responsabilidade para gerenciar esse processo em instituições públicas equitativas, organizações da sociedade civil e outras organizações sem fins lucrativos que, juntas, podem representar a sociedade como um todo e/ou assembleias dentro dele, especialmente aqueles que mais perdem com a IA projetada e implantada negligente ou perniciosamente.

Conforme cada vez mais cada vez mais a tomada de decisões vai sendo delegada a algoritmos, surgem questões importantes em circunstâncias em que essas decisões têm consequências diretas para os direitos individuais, às oportunidades pessoais e ao bem coletivo. Em certas formas de tomada de decisão no sistema de justiça criminal em relação à liberdade pessoal (i.e. determinações de custódia, sentença e liberdade condicional), isso é muito mais evidente, mas essa hipótese também se aplica ao acesso à educação, moradia, assistência social, crédito e seguro. Assim, o fato de as implicações sociais desses novos métodos só poderem ser compreendidas com a compreensão adequada de seus fundamentos técnicos gerais se algo crucial para os formuladores de políticas (SCANTAMBURLO; CHARLESWORTH; CRISTIANINI, 2019, p. 49 e ss.). Mesmo que se parta da análise da situação de sistemas algorítmicos decisórios para a execução no sistema de justiça criminal

(i.e. análise dos antecedentes, do comportamento violento, da região onde o réu vive e outras circunstâncias socioeconômicas que podem ser importantes para o sistema penal e sopesados/classificados por ferramentas de IA), situações semelhantes emergem de outros algoritmos utilizados no controle do acesso a oportunidades, para explicar como o aprendizado de máquina funciona e, como resultado, como as decisões são tomadas por algoritmos inteligentes modernos ou “classificadores”.

Embora as decisões sobre máquinas sejam frequentemente julgadas com base apenas em sua precisão, em tais circunstâncias, há considerações normativas mais amplas. Desse modo, a decisão baseada em aprendizado de máquina nas decisões de justiça criminal sobre a apresentação de acusações criminais, duração da pena de prisão, elegibilidade para liberdade condicional, ou ainda, decisões administrativas estatais sobre a proteção da criança e a elegibilidade para benefícios sociais diferem fundamentalmente de seu uso em outras esferas, mesmo aquelas que também podem ter um impacto positivo ou negativo nos indivíduos (i.e. decisões de diagnóstico médico). Assim, os sistemas que utilizam algoritmos de tomada/apoio à decisão com potencial prejudicial aos direitos humanos individuais ou coletivos merecem uma regulação significativamente mais qualificada do que os sistemas que usam algoritmos de tomada de decisão para processar objetos, dotada de princípios normativos próprios. Ou seja, a contínua expansão da tomada de decisões sobre máquinas nos setores público e privado, muitas vezes justificada com base na eficiência e economia de custos, ainda precisa desenvolver o grau de análise e avaliação regulatórias informadas que seu mérito problemático demanda.

Os sistemas de saúde em praticamente todos os países são desafiados pelo aumento de custos e pela piora de resultados. Cada vez mais, formuladores de políticas, políticos, empreendedores clínicos e cientistas de computadores e dados argumentam que sistemas de IA trarão uma parte essencial da solução, particularmente mediante o aprendizado de máquina (*machine learning*, ou ML). Esse argumento não decorre da crença de que todas as necessidades de cuidados de saúde serão atendidas em breve por médicos autômatos, mas sim, da definição contrafactual clássica de IA como uma terminologia genérica para várias técnicas que podem ser usadas para fazer com que as máquinas concluam tarefas. A automação dessa natureza pode oferecer grandes oportunidades para a melhoria dos serviços de saúde e da saúde dos pacientes, melhorando significativamente as capacidades clínicas humanas em diagnóstico, descoberta de medicamentos, epidemiologia, medicina personalizada e eficiência operacional. Mas a eventual incorporação de soluções de IA na prática clínica acarreta pelo menos três questões: as possibilidades e limitações técnicas; o quadro ético, regulatório e

jurídico; e a estrutura de governança (MORLEY et al., 2019). Ademais, é muito importante que os desafios éticos levantados pela implementação da IA em ambientes de saúde sejam enfrentados proativa, e não reativamente.

Blasimme e Vayena (2020) argumentam que os desafios éticos que a IA na medicina suscita demanda novos mecanismos éticos de regulação para além dos altamente proliferados códigos de conduta e diretrizes atuais. O amplo escopo de possíveis aplicações para pesquisa, uso clínico e saúde pública pode fazer com que alguns usos específicos da IA não sejam cobertos pelos mecanismos de supervisão existentes. E a dependência apenas das ferramentas reguladoras existentes provavelmente falhará em garantir níveis adequados de confiança e responsabilidade pública. Assim, sugerem que, no domínio da pesquisa, os comitês de ética terão de incorporar avaliações reflexivas dos méritos científicos e sociais da pesquisa orientada pela IA e, com isso, terão de abrir espaço para novas figuras profissionais (cientistas sociais, por exemplo). Financiadores de pesquisa, poderão exigir que os mecanismos de monitoramento e capacidade de resposta façam parte dos planos de pesquisa, para isso criando comitês multidisciplinares de avaliação periódica dos dados dessas atividades, a fim de ajustar suas políticas de financiamento no futuro. Ademais, ao atingir escalas que reivindicam dados de comunidades ou populações inteiras, faz-se necessário que a pesquisa orientada pela IA assuma formas adequadas de inclusão para garantir o aprendizado social em diferentes comunidades epistêmicas (incluindo públicos leigos e atores não acadêmicos).

No tocante à atenção ao paciente, a validação clínica é crucial. Padrões de evidência *ad hoc* são necessários para a inovação clínica responsável, mas insuficientes para regular possíveis problemas éticos. Hospitais e clínicas, assim, devem organizar órgãos de supervisão clínica da IA encarregados do aconselhamento dos seus administradores acerca da adoção de uma determinada tecnologia da IA e o monitoramento periódico de seus efeitos nas jornadas dos pacientes e no engajamento dos pacientes durante todo o período de cuidados. Além disso, os requisitos de consentimento terão de ser adaptados à presença do altamente automatizado processamento de dados altamente automatizado. E quanto à saúde pública, o novo nível de precisão possibilitado pela IA na promoção da saúde ou na vigilância de doenças terá que ser negociado no nível das comunidades-alvo, ou resultará na falta de confiança pública. Os limites aceitáveis de coleta de dados e análise algorítmica terão de ser resultado de deliberações inclusivas em toda a comunidade, especialmente sobre aquelas que se vier a coletar e processar dados.

À medida que a IA é cada vez mais integrada à prática clínica, vários desafios cruciais relacionados à privacidade dos pacientes persistirão, especialmente no que diz respeito à

aquisição de dados, relatórios e potencial re-identificação dos dados do paciente. Assim, Rohringer, Budhkar e Rudicz (2019) sopesam o interesse de empresas — que ultrapassam limites de direitos fundamentais na busca de dados de pacientes para treinar seus sistemas (sendo um bom exemplo disso a declaração de ilegalidade e encerramento, pelo órgão britânico regulador de proteção de dados, da parceria entre a Royal Free NHS Foundation Trust e a subsidiária do Google DeepMind, devido à insuficiência do consentimento informado relacionado ao uso secundário de dados privados) — e o interesse público através de novas regulamentações sobre a privacidade desses dados (HERN, 2017).

Conforme descrito pelos autores, os usos avançados da IA podem servir para mudar a responsabilidade pelos cuidados e monitoramento dos profissionais de saúde para os próprios pacientes. A IA avançada está sendo usada atualmente em “relógios inteligentes” de consumo para monitorar intensivamente pacientes com doença pulmonar obstrutiva crônica de uma maneira que não é possível *in situ*, mas que também exige que os pacientes cuidem do equipamento e sigam o protocolo. Isso levanta questionamentos ético-jurídicos (por exemplo, acerca do tipo de paciente que será favorecido com essa tecnologia, ou se os pacientes despreparados para gerenciar e manter seus próprios dados recebem cuidados dentro de padrões de qualidade aceitáveis, os modos pelos quais se estabelece se o grau em que o consentimento é informado e não coagido, etc.). Diretrizes jurídicas sólidas atinentes a esses questionamentos exigirão normas que substituam as normativas fragmentadas que atualmente caracterizam as mais diversas jurisdições competentes para tal. Assim, uma ênfase maior em normas técnicas internacionais (aí incluindo-se normas para gerenciamento de dados) poderia ser um passo crucial para a segurança do paciente.

Ademais, um “cabo-de-guerra” entre o interesse em se alcançar a precisão máxima nos modelos de aprendizado por máquina (ML) e o interesse em preservar a privacidade completa do paciente. Buscar métodos que possam filtrar informações demográficas privadas e comprometer minimamente o desempenho do modelo é uma possível solução sugerida. A demência, por exemplo, pode ser detectada por classificadores de ML treinados em identificar padrões linguísticos naturais da fala. Apesar do forte condicionamento que a idade avançada possa causar à probabilidade de desenvolver demência, a revelação da idade precisa de uma pessoa nem sempre pode ser permitida. Assim, redes neurais (computadorizadas) avançadas foram desenvolvidas para isolar e filtrar informações confusas desse tipo, enquanto reduz modestamente a precisão.

Chesterman (2020) aponta para a disparidade de velocidades com que, por um lado, aumenta a capacidade de processamento e a conectividade ocorrem e, por outro, a velocidade

com a qual do Direito é capaz de fornecer soluções de regulação para o uso de aparatos tecnológicos. Desde que os computadores passaram a representar um dos elementos tecnológicos com maior inovação desde a década de 1960, a eficiência com a qual os dados são processados vem levantando questões regulatórias — o que pode ser bem compreendido no que tange à privacidade. Para dados públicos protegidos pela “obscuridade prática” dos registros em papel disponíveis apenas em órgãos governamentais, as chances de um terceiro descobrirem informações privadas eram baixas. Mas a Internet praticamente acabou com essa obscuridade prática.

Há muito já é claro que a velocidade representa um desafio à regulamentação, principalmente no o comércio de alta frequência (que aponta para o perigo de que a velocidade da tomada de decisões possa frustrar as tentativas humanas de impedi-lo quando fugir do controle) e Direito da Concorrência (em que o conluio mediante algoritmos pode ser impossível de detectar). Esses problemas se aplicam a muitas das outras áreas de atividade desafiadas por novas formas de tecnologia. Um modo de enfrentá-los seria a diminuição da velocidade, localizando-se e compartimentando-se dados, e introduzindo-se latência artificial nos algoritmos de negociação, tornando mais lentas as estruturas de funcionamento do mercado digital. Essa abordagem pode ser a única maneira de continuar a confiar em ferramentas reguladoras projetadas para humanos e operando em uma escala de tempo humana, mas põe sob risco aquilo que torna esses sistemas valiosos, que é a sua inovação e efetividade.

Danaher (2019, p. 135-136) adiciona outro argumento, que pretende ter um amplo significado no nível da civilização, mas envolver especulações menos fantasiosas sobre a provável inteligência futura das máquinas, à crescente onda de pânico sobre robôs e IA: a ascensão dos robôs criará uma crise de paciência moral, sendo reduzida a capacidade e a disposição dos humanos de agirem moral e responsabilmente no mundo, reduzindo-nos, assim, a pacientes morais. Dada a centralidade que caracteriza essa capacidade e essa vontade no sistema de valores dos modernos Estados democráticos liberais, a crise da paciência moral tem um amplo significado civilizatório, pois ameaça algo que é fundamental e pressupõe em muitos discursos morais e políticos contemporâneos. Essa ameaça não deriva de nenhuma especulação fantasiosa sobre as futuras e poderosas e superinteligentes formas futuras de robôs e IA, mas sim, de extrapolações plausíveis das tendências existentes. Isso não levaria a humanidade à extinção ou à aniquilação, mas minará um valor social fundamental.

É claro que esse argumento poderia ser objetado em razão do fato de que a humanidade tem o poder de moldar a tecnologia e usá-la para impedir que ela se desenvolva

de maneiras que prejudicam a agência humana, mas isso seria deveras difícil, pois: i) o desenvolvimento de robôs e IA na indústria e nos serviços é impulsionado pelo apelo sedutor do sucesso econômico; ii) o crescimento de robôs e IA no setor público decorre do compromisso com valores geralmente aceitos (eficiência, precisão, relação custo-benefício) e da necessidade de responder à complexidade do mundo exterior; e iii) a crescente disposição do ser humano em terceirizar o desenvolvimento moral pessoal para robôs e IA é facilitada por vieses psicológicos e heurísticos. Ou seja, as tendências decorrem de forças que são maiores e mais poderosas do que a própria agência humana.

3 EXPECTATIVA PARA A REGULAÇÃO JURÍDICA NO BRASIL

Dentro desse complexo cenário que implica questionamentos éticos ao desenvolvimento tecnológico das IAs — entre oportunidades, desafios e riscos —, a regulação precisa encontrar estratégias para conter suas consequências negativas e incentivar a inovação, criando caminhos juridicamente estáveis para produção de tecnologia. No Brasil, tramitam em fase inicial três projetos de lei: um de iniciativa da Câmara dos Deputados — o Projeto de Lei 21/2020¹ — que, conforme seu histórico de tramitação, foi apresentado em fevereiro de 2020 propondo, na ementa: “estabelece princípios, direitos e deveres para o uso de IA no Brasil”; e dois de iniciativa do Senado Federal — desde 2019, o PL 5091² (que “estabelece os princípios para o uso da IA no Brasil”) e o PL 5691³ (que “institui a Política Nacional de Inteligência Artificial”).

Na Câmara, a apresentação do projeto de lei define os conceitos de (I) “sistema de inteligência artificial” — sistema desenvolvido por humanos que é capaz de prever, recomendar ou tomar decisões; (II) “ciclo de vida do sistema de inteligência artificial” — o tempo de operação do sistema determinado por sua programação; (III) “conhecimento em inteligência artificial” — as habilidades e recursos adquiridos e utilizados no ciclo de cada sistema; (IV) “agentes de inteligência artificial” — pessoa física ou jurídica, podem ser o agente de desenvolvimento do sistema ou de operação; (V) “partes interessadas” — todos os afetados pelo funcionamento do sistema; (VI) “relatório de impacto de inteligência artificial” — contém a descrição do funcionamento do sistema e os mecanismos de gestão de risco.

¹ De autoria do Deputado Eduardo Bismarck (PDT/CE)

² De autoria do Senador Styvenson Valentim (PODEMOS/RN)

³ De autoria do Senador Styvenson Valentim (PODEMOS/RN)

Já em relação aos projetos tramitando no Senado Federal, o PL 5051/2019 busca estabelecer os princípios para o uso da IA no Brasil, tendo em vista o reconhecimento do bem-estar humano, no art. 2 do texto inicial: (I) o respeito à dignidade humana, à liberdade, à democracia e à igualdade, (ii) o respeito aos direitos humanos, à pluralidade e à diversidade; (iii) a garantia da proteção da privacidade e dos dados pessoais; (iv) transparência, a confiabilidade e a possibilidade de auditoria dos sistemas; (v) a supervisão humana. E a disciplina do uso tendo por objetivo a promoção e a harmonização da valorização do trabalho humano e o desenvolvimento econômico (art. 3º). Também, define a responsabilidade civil por danos decorrentes do uso de IA na figura do seu supervisor (art. 4, §2º), tendo em vista que a lei define a tomada de decisões por uma IA como auxiliar da decisão humana (art. 4º), ou seja, é necessária e fundamental a supervisão humana desses instrumentos, reconhecendo suas implicações. Importante nesse projeto, também, é a formação de diretrizes para a atuação da Administração Pública em (I) promover educação sobre a IA; (II) políticas de proteção e qualificação do trabalho humano; (III) a garantia da adoção gradual da IA; (IV) a ação proativa na regulação das aplicações de IA.

Também no Senado Federal, o PL 5691/2019, em seu texto inicial, da mesma maneira que o outro PL da Casa, não desenvolve sobre os conceitos como o PL da Câmara, restringe-se aos princípios⁴ e diretrizes⁵ da pretendida Política Nacional de Inteligência Artificial. Essa política busca o desenvolvimento tecnológico aliado à transparência democrática das decisões baseadas em IA, até mesmo às autônomas. Também, integra um processo de transição para os mecanismos de automação, mitigando prejuízos ao emprego (art. 3, XI). É uma política que se relaciona diretamente com a responsabilidade humana e das organizações que desenvolvem e operam IA, obrigando-as a estabelecer relações abertas, inteligíveis e debatidas sobre o seu uso, sustentando formas de decisões rastreáveis e princípios de governança ligados à proteção dos riscos ligados à IA e a proteção de dados pessoais — tendo em vista o cumprimento da LGPD.

4 Conforme art. 2 do texto inicial: (i) desenvolvimento inclusivo e sustentável, (ii) respeito à ética, aos direitos humanos, aos valores democráticos e à diversidade, (iii) proteção da privacidade e dos dados pessoais, (iv) transparência, segurança e confiabilidade -

5 Conforme o art. 3 do texto inicial: (i) estabelecimento de padrões éticos para o uso da inteligência artificial, (ii) promoção de crescimento inclusivo e sustentável; (iii) melhoria da qualidade e da eficiência dos serviços oferecidos à população; (iv) estímulo a investimentos públicos e privados em pesquisa e desenvolvimento; (v) promoção da cooperação e interação entre entes públicos, entre os setores público e privado e entre empresas; (vi) desenvolvimento de estratégias para incrementar o intercâmbio de informações e a colaboração entre especialistas e instituições nacionais e estrangeiras; (vii) estímulo às atividades de pesquisa e inovação; (viii) desenvolvimento de mecanismo de fomento à inovação; (ix) capacitação de profissionais da área de tecnologia; (x) valorização do trabalho humano; (xi) promoção de uma transição digital justa com a mitigação das consequências adversas da IA para o mercado de trabalho e para as relações trabalhistas

A responsabilidade e a caracterização dos agentes responsáveis pelas IA suscita o debate sobre a personalização jurídica dos robôs, que sustenta a possibilidade de os robôs autônomos, justamente por possuírem inteligência, tenham reconhecido *status* de personalidade. A tese de Castro Junior (2009, p. 190) reflete sobre a abrangência do conceito de “personalidade jurídica” em um contexto pós-humanista, tendo em vista que o conceito de personalidade não é mero reflexo do conceito de “ser humano” – diferentemente do paradigma antropocentrismo do Direito brasileiro. A cidadania robótica corresponderia a uma disciplina jurídica própria de entes inteligentes que tomam decisões autônomas, relacionando o exercício da cidadania com a constatação de uma personalidade jurídica — sendo a evolução tecnológica desses comportamentos inteligentes o caminho para a aquisição dessa personalidade.

As iniciativas regulatórias europeias apontam para a instituição da figura jurídica da personalidade dos robôs. A Resolução do Parlamento Europeu n. 2015/2103, e as Recomendações Quanto ao Conteúdo da Proposta Requerida anexo a essa, apontam para a tendência de personalização dos robôs a longo prazo, devido a ampliação das atividades autônomas e desenvolvidas de forma inteligente (via aprendizado de máquina), também, o debate acerca da responsabilidade civil, atualmente restrito no âmbito dos engenheiros, evoluiu para a responsabilização das personalidades eletrônicas – por meio de fundos específicos.⁶ É uma especulação regulatória que tem em vista o desenvolvimento da tecnologia e já leva em conta realidades de personalização de inteligências artificiais, como o caso Sophia.⁷

Entretanto, os projetos de lei brasileiros caminham noutro sentido – ligado ao princípio da precaução do Direito Ambiental -, conforme apontam Andrade e Faccio (2019, p. 173), na qual os possíveis danos desconhecidos devem ser levados em conta pelo agente. É um dever de cuidado com os riscos, mesmo que não plenamente conhecidos. Também, frente aos direitos civis da propriedade privada, o enquadramento jurídico razoável das inteligências artificiais é no conceito de coisa, atestam Ehrhardt Júnior e Silva (2020). Os projetos em tramitação tratam, de forma unânime, da noção de centralidade no ser humano — o PL

⁶ “Criar um estatuto jurídico específico para os robôs a longo prazo, de modo a que, pelo menos, os robôs autônomos mais sofisticados possam ser determinados como detentores do estatuto de pessoas eletrônicas responsáveis por sanar quaisquer danos que possam causar e, eventualmente, aplicar a personalidade eletrônica a casos em que os robôs tomam decisões autônomas ou em que interagem por qualquer outro modo com terceiros de forma independente” (UNIÃO EUROPEIA, 2017, p. 17).

⁷ Trata-se do reconhecimento de personalidade jurídica da robô Sophia pela Arábia Saudita, ver em: <https://revistagalileu.globo.com/Tecnologia/noticia/2017/10/arabia-saudita-torna-se-primeiro-pais-conceder-cidadania-para-um-robo.html>.

21/2020 define a condição de agente humano (pessoa física ou jurídica) desenvolvedor e operador das IA, o PL 5051/2019 apregoa o conceito de “supervisão humana” (art. 2º, V) e o PL 5691/2019 obriga a solução necessária às IA da intervenção humana absoluta sobre as máquinas (art. 4º, VII). Enfim, a expectativa legislativa sobre o tema no Brasil aponta para o esclarecimento da relação humano e coisa (máquina/ propriedade).

Esse poder de intervenção sobre as máquinas sugere um conteúdo de escrutínio democrático e decisões humanas sobre a tomada de decisões das IA. O texto de apresentação do PL 21/2020 além de apregoar o respeito aos direitos humanos e aos valores democráticos (art. 4, III) aponta para a transparência e explicabilidade dos sistemas de inteligências artificial (art. 6, IV) e a governança multi participativa entre os diferentes agentes envolvidos, públicos ou privados (art. 10, VI), no PL 5051/2019 consta que a disciplina do uso das IA é da transparência, confiabilidade e auditoria (art. 2, IV) e a ação proativa na regulação das aplicações (art. 5, IV) e no PL 5691/2019 caracterizam-se os mesmos princípios como, também, os da participação social da população (art. 4, V), de que as decisões das máquinas sejam rastreáveis (art. 4, VIII) e de seguir padrões de governança que garantam mitigação de riscos potenciais (art. 4, IX).

O cenário apontado por esses projetos legislativos compactua com a tendência à autorregulação e a auditabilidade social das tecnologias, reconhecendo a impossibilidade de uma regulação puramente estatal e dogmática — buscando instrumentos adaptados a flexibilidade e capacidade de resposta das novas tecnologias (SANTOS; MARCO, MULLER, 2019, p. 3079-3081). Vislumbra-se o incentivo a dois conjuntos de práticas: (a) da autorregulação ligada a responsabilidade social empresarial; e (b) da auditabilidade social dos algoritmos. A primeira institui-se a partir da realidade das grandes empresas transnacionais, operadoras majoritárias dessas tecnologias, que constituem em suas atividades ordens jurídicas para além das fronteiras por meio de seus *compliance* e suas práticas de auto regulação — formando expectativas sobre elas para uma determinação própria que respeite Direitos Humanos, ou seja, não como opostos à ordem estatal e seus procedimentos democráticos, mas como mecanismos internos capazes de aproximar essas ordens privadas dos interesses de ordem positiva.

Essa tendência à autorregulação consiste em um sistema de correção, ou regulação regulada pelas empresas que estão presentes, já no ordenamento jurídico brasileiro, na Lei Geral de Proteção de Dados (LGPD) — Lei nº 13.709/2018 — como “boas práticas e da governança” (art. 50) na qual os controladores e operadores de dados pessoais podem e devem formular práticas de governança que apreciem os princípios norteadores dessa lei, sendo a

LGPD um marco para o fomento de uma cultura corporativa de proteção dos dados pessoais (BIONI, 2019, p. 32-33). No cenário de um possível marco normativo das Inteligências Artificiais, os agentes desenvolvedores e operadores se inserem nesse mesmo contexto de ente regulador, ao nível interno, tanto no ato de criação dos sistemas como de suas atividades, consagra a realidade de regulação *by design*, ou seja, que os próprios desenvolvedores também são responsáveis pela preservação de princípios de ordem pública — algo que Magrani (2019, p. 261) chama de direito como metatecnologia, uma regulação ao nível do desenvolvimento.

Além da autorregulação, a auditabilidade social,⁸ ou *accountability*, dos algoritmos é a tendência que não há como esperar que os algoritmos justos partam somente do cumprimento da lei ou do Direito estatal, mas que possam ser fundamentados em princípios abertos e verificáveis (COURTLAND, 2018). Os três projetos de lei estudados apontam para a transparência e verificabilidade das decisões tomadas pelas inteligências artificiais, ou seja, como e por que elas tomam decisões deve passar por escrutínio público – tamanho é o campo de pessoas e organizações afetadas por essas decisões -, sendo fundamental a criação de mecanismos de esclarecimento de forma clara e com linguagem acessível (VEDDER; NAUDTS, 2017, p. 214). É essencial para a manutenção da livre expressão, tendo em vista evitar decisões enviesadas pelas IA que contenham conteúdo discriminatório, contemplando uma realidade de (re) interpretação dos algoritmos complexos, para além da ingenuidade técnica, mas como algo que implica toda a sociedade e exige abordagem transdisciplinar (KITCHIN, 2017, p. 26-27).

Essa abordagem social das tecnologias só é possível por meio do acesso a dados transparentes e abertos sobre as tecnologias, por isso verifica-se importante a tendência dos projetos de lei brasileiros em contemplar a educação sobre as IA - art. 14 do PL 21/2020 e art. 5, I do PL 5051/2019. O marco regulatório sobre as IA no Brasil precisa atender a esse dilema ético sobre a capacidade dos cidadãos em participar de decisões de complexos sistemas tecnológicos, preservando práticas democráticas em interferir no bom desenvolvimento de sistemas que possibilitam uma série de oportunidades – para os serviços públicos ou para a produção privada. As ferramentas educativas sobre IA ultrapassam a esfera da educação pois se transformam em condição para o exercício de direitos acerca dessa tecnologia, é construção

⁸ Conforme definição disponível em Sérgio Amadeu Silveira (2019, p. 50) como: “permitir que terceiros interessados analisem, compreendam e revisem o comportamento do algoritmo por meio da divulgação de informações que permitam o monitoramento, verificação ou crítica, inclusive por meio do fornecimento de documentação detalhada, APIs tecnicamente adequadas e termos de uso permissivos.”

de um discernimento público sobre o tema, superando a visão alienada e distante em relação à tecnologia.

Os projetos de lei mencionados apontam para uma opção legislativa correspondente a responsabilização humana dos desenvolvedores das inteligências artificiais – e a coisificação das decisões autônomas –, ao mesmo tempo que reconhece as oportunidades nos diversos campos da sociedade de sua aplicação, oferecendo mecanismos de prestação eficiente de serviços. Nesse cenário, entre as possíveis lacunas dos projetos de lei em andamento no Brasil, é preciso elencar um aspecto que não consta em todos os textos iniciais que deve ser atendido: (a) a definição dos agentes envolvidos no processo de desenvolvimento e operação da IA, principalmente ao que se refere o esclarecimento da responsabilidade. E um aspecto que não consta em qualquer projeto: (b) a criação de uma agência reguladora, independente e ligada a formas de regulação cooperativa, facilitadora das formas de autorregulação e auditabilidade plena dos algoritmos.

Quanto à definição dos agentes envolvidos, é importante definir os responsáveis e a definição de sistema, tanto na perspectiva da responsabilidade como esclarecimento necessário à efetiva participação e auditabilidade. A LGPD fez essa classificação em seu art. 5º, algo que pode ser considerado virtuoso, quando se analisa tal dispositivo a partir da perspectiva da efetiva caracterização dos participantes, mesmo em um cenário de complexidade, tendo em vista que os conceitos não são restritivos — o PL 21/2020 adota essa medida em seu texto inicial, fazendo a importante distinção entre sistema e agente, além disso, apontando a diferença entre agente desenvolvedor da IA e o agente operador. Esse esclarecimento acerca dos envolvidos na produção e operação de inteligências autônomas é essencial para cumprimento do princípio da transparência essencial para o andamento de uma auditabilidade social dos algoritmos.

A proposta de uma agência reguladora de IA e robótica está ausente nos três projetos de lei estudados, todavia, considerando os méritos da regulação europeia — e o já aplicado na LGPD com a Autoridade Nacional de Proteção de Dados — uma “Agência Federal Brasileira para a Robótica e da Inteligência Artificial”, conforme apontado por Paulo Novais e Pedro Miguel Freitas (2018), é mais capaz de lidar com o tema de forma multidisciplinar e complexa em relação à dureza do direito positivo – sem abrir mão dos princípios democráticos positivados —, unindo políticos, acadêmicos e profissionais em práticas ativas de regulação das inteligências artificiais. A multiplicidade de ordens jurídicas e a intensa transformação das tecnologias exige um organismo mediador horizontal, compreendo o alinhamento entre ordens públicas, privadas e conjuntas, aproximando o público das tomadas

de decisões por meio de efetiva auditabilidade social, debatendo de forma mais transparente e informada os riscos e oportunidades envolvidas no entorno das IA.

CONCLUSÃO

Embora se observe, ao longo das últimas seis décadas, um grande desenvolvimento da computação e da IA, ainda é muito difícil prever a obtenção de uma IA que pense e se comporte como humanos. Mas isso não significa que o atual desenvolvimento das tecnologias relacionadas à IA não representem riscos para a humanidade: sistemas bélicos embasados em IA, nos preconceitos na programação e nos dados de aprendizado dos algoritmos, substituição da mão de obra humana na produção, desrespeito à privacidade e à autonomia humanas, já presentes na IA atual representam um campo de discussão ético-jurídico-política bastante desafiador.

Não se pode conceber um desenvolvimento de IA que parta apenas do setor público, sendo necessárias sinergia e colaboração entre público, privado e universidades para que a implementação da IA pelos Estados em seus serviços seja efetiva — mas isso não pode ocorrer sem um planejamento cuidadoso e detalhado da questão política e administrativa, o que envolve grande capacidade de liderança, gerenciamento e comunicação. Aliás, essa última característica — qualidade de comunicação — deve ser muito bem administrada em relação ao público no que tange à aplicação da IA em setores tais como educação, segurança, serviços e inteligência em geral, pois diversas preocupações éticas (principal, mas não unicamente, atinentes às possibilidades de fraudes e de viés prejudicial tanto dos dados com que são alimentados os sistemas quanto dos valores interpretativos que tais tecnologias venham a incorporar em sua programação), pois diversas preocupações éticas emanam de tais aplicações..

Os modos pelos quais a IA tem sido desenvolvida e aplicada até então têm demonstrado um déficit democrático, legal, social — e, em *ultima ratio*, ética. O grande potencial de racionalização (burocratização) faz com que a compreensão e a programação dos sistemas de IA (no que tange não apenas às informações com que as máquinas são alimentadas, mas também, em relação aos valores sociais com que operam) se afaste da opinião pública ponderada e debatida, e isso em grande medida aumenta o pânico com o qual se encara o futuro da sociedade em relação à IA. Assim, iniciativas de aproximação (e não apenas de informação) do público em relação às decisões importantes atinentes à IA é um meio de se fazer com que o desenvolvimento da IA se dê para as pessoas, e não às suas custas.

Assim, instituições públicas e da sociedade civil devem ser reprojatadas para superar o déficit democrático e o medo em relação à IA.

Aqueles que possuem a capacidade de regulação devem sopesar todos esses riscos, a fim de reinventar instituições, instrumentos e formas normativas, de modo que o dano existencial real que o uso da IA pelos setores público e privado representa à humanidade possa ser evitado, mitigado ou, ao menos, compreendido a tempo. Enquanto isso, a regulação brasileira acerca do tema é incipiente, tendo sido aqui estudados três projetos de lei que buscam estabelecer os princípios sobre o uso das IA. Foi identificada uma relação direta com os dilemas éticos trabalhados, estando as ideias de regulação alinhadas à perspectiva dúplice de oportunidades e riscos dessa tecnologia, apontando para a necessidade de regulação para uso adequado tanto para proporcionar o desenvolvimento dessas oportunidades como de reduzir os riscos, e a opção dos legisladores parece ser de cunho humanista — adequado à Constituição Federal de 1988 — separando as práticas humanas das decisões de máquinas, importando na figura necessária da responsabilização e supervisão humana das IA, mesmo quando autônomas. A efetiva regulação, nesse contexto, se dá no incentivo da autorregulação ao nível do desenvolvimento, contrária à regulação repressiva, mas dando condições aos desenvolvedores de adoção de princípios democráticos no âmbito dos engenheiros — sendo fiscalizada socialmente pelos mecanismos da auditabilidade social dos algoritmos, ou seja, por uma sociedade informada e educada sobre como e porquê as IA tomam decisões. Carece-se, entretanto, do estabelecimento de uma agência que intermedie práticas regulatórias, como facilitadora de uma cultura pública e privada de auditabilidade das IA.

REFERÊNCIAS

ANDRADE, Fabio Siebeneichler de; FACCIO, Lucas Girardello. Notas sobre a responsabilidade civil pela utilização da Inteligência Artificial. **Revista da AJURIS**, v. 46, n. 146, p. 153-181, 2019. Disponível em:

http://www.mpsp.mp.br/portal/page/portal/documentacao_e_divulgacao/doc_biblioteca/bibli_servicos_produtos/bibli_boletim/bibli_bol_2006/Rev-AJURIS_n.146.pdf#page=144. Acesso em: 08 jun 2020.

BIONI, Bruno Ricardo. Inovar pela Lei. **GV/Executivo**, v. 18, n. 4, p. 31-33, jul/ago 2019. Disponível em:

<http://bibliotecadigital.fgv.br/ojs/index.php/gvexecutivo/article/download/79978/76432>. Acesso em: 08 jun 2020.

BLASIMME, Alessandro; VAYENA, Effy. The Ethics of AI in Biomedical Research, Patient Care and Public Health. In: DUBBER, Markus D.; PASQUALE, Frank; DAS, Sunit (eds.).

The Oxford Handbook of Ethics of Artificial Intelligence (Forthcoming), 2020. DOI: <http://dx.doi.org/10.2139/ssrn.3368756>

BRASIL. Câmara dos deputados. **Projeto de Lei 21/2020**. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2236340>. Acesso em: 08 jun 2020.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 08 jun 2020.

BRASIL. Senado Federal. **Projeto de Lei 5091/2019**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/138790>. Acesso em: 08 jun 2020.

BRASIL. Senado Federal. **Projeto de Lei 5691/2019**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/139586>. Acesso em: 08 jun 2020.

CASTRO JUNIOR, Marco Aurélio. **Personalidade Jurídica do Robô e sua efetividade no Direito**. Tese (Doutorado). Faculdade de Direito. Universidade Federal da Bahia, 222f, 2009. Disponível em: <https://repositorio.ufba.br/ri/bitstream/ri/10719/3/Personalidade%20Juridica%20do%20Robo%20c3%b4%20e%20sua%20efetividade%20no%20Direito.pdf>. Acesso em: 08 jun 2020.

CHESTERMAN, Simon. “Move Fast and Break Things”: Law, Technology, and the Problem of Speed. **NUS Law Working Paper 2020**, v. 001, p. 1-25, 2020. DOI: <http://dx.doi.org/10.2139/ssrn.3516032>

COURTLAND, Rachel. Bias detectives: the researchers striving to make algorithms fair: As machine learning infiltrates society, scientists are trying to help ward off injustice. **Nature**, n. 558, p. 357-361, 2018. Disponível em: <https://www.nature.com/magazine-assets/d41586-018-05469-3/d41586-018-05469-3.pdf>. Acesso em: 08 jun 2020.

COWLS, Josh. Deciding How to Decide: Six Key Questions for Reducing AI’s democratic Deficit In: BURR, Christopher; MILANO, Silvia (eds.). **The 2019 Yearbook of the Digital Ethics Lab**. Cham: Springer, 2020, p. 101-116.

DANAHER, John. The rise of the robots and the crisis of moral patience. **AI & Society**, n. 34, p. 129–136, 2019. DOI: <https://doi.org/10.1007/s00146-017-0773-9>.

EHRHARDT JÚNIOR, Marcos; SILVA, Gabriela Buarque Pereira. Pessoa e sujeito de direito: reflexões sobre a proposta europeia de personalidade jurídica eletrônica. **Revista Brasileira de Direito Civil**, v. 23, n. 01, p. 57, 2020. DOI: DOI:10.33242/rbdc.2020.01.003.

FLORIDI, Luciano. What the near future of Artificial Intelligence Could Be In: BURR, Christopher; MILANO, Silvia (eds.). **The 2019 Yearbook of the Digital Ethics Lab**. Cham: Springer, 2020, p. 127-142.

HANEY, Brian S. The Perils and Promises of Artificial General Intelligence. **Journal of Legislation**, v. 45, n. 2, p. 151-170, 2018. Available at: <https://scholarship.law.nd.edu/jleg/vol45/iss2/1>. Accessed in: 27 fev 2020.

HERN, Alex. Royal Free breached UK data law in 1.6m patient deal with Google's DeepMind. **The Guardian**, 3 jul 2017. Available at: <https://www.theguardian.com/technology/2017/jul/03/google-deepmind-16m-patient-royal-free-deal-data-protection-act>. Accessed in: 28 fev 2020.

KITCHIN, Rob. Thinking critically about and researching algorithms. *Information, communication & Society*, v. 20, n.1, 2017. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/1369118X.2016.1154087?needAccess=true#> HR0cHM6Ly93d3cudGFuZGZvbmxpbmUuY29tL2RvaS9wZGYvMTAuMTA4MC8xMzY5MTE4WC4yMDE2LjExNTQwODc/bmVIZEFjY2Vzcz10cnVlQEBAMA==. Acesso em: 08 jun 2020.

LEXCELLENT, Christian. **Artificial Intelligence versus Human Intelligence: Are Humans Going to Be Hacked?** Cham: Springer, 2019.

LI, Deyi; DU, Yi. **Artificial intelligence with uncertainty**. Boca Raton: CRC Press, 2017.

MAGRANI, Eduardo. **Entre Dados e Robôs: A Ética Das “Coisas”**: da Ética do Discurso e Racionalidade Comunicativa ao Novo Materialismo de Sistemas Sociotécnicos. 2 ed. Porto Alegre: Arquipélagos, 2019.

MORLEY, Jessica; MACHADO, Caio C. V.; BURR, Christopher; COWLS, Josh; JOSHI, Indra; TADDEO, Mariarosaria; FLORIDI, Luciano. The Debate on the Ethics of AI in Health Care: a Reconstruction and Critical Review. **SSRN**, p. 1-35, 2019. DOI: <http://dx.doi.org/10.2139/ssrn.3486518>

NEAPOLITAN, Richard E.; JIANG, Xia. **Artificial Intelligence: With an Introduction to Machine Learning**. 2 ed. Boca Raton: CRC Press, 2018.

ROHRINGER, Taryn J.; BUDHKAR, Akshay; RUDZICZ, Frank. Privacy versus artificial intelligence in medicine. **University of Toronto Medical Journal**, v. 96, n. 1, p. 51-53, 2019.

SANTOS, Paulo Junior Trindade dos; MARCO, Cristhian Magnus de; MÖLLER, Gabriela Samrsla. Tecnologia Disruptiva e Direito Disruptivo: Compreensão do Direito em um Cenário de Novas Tecnologias. **Revista Direito e Práxis**, v. 10, n. 4, p. 3056-3091, dez. 2019. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/revistaceaju/article/view/45696/31176>. Acesso em: 08 jun 2020.

SCANTAMBURLO, Teresa; CHARLESWORTH, Andrew; CRISTIANINI, Nello. Machine Decisions and Human Consequences In: YEUNG, Karen; LODGE, Martin (eds.). **Algorithmic Regulation**. Oxford: University of Oxford Press, 2019, p. 49-81.

SHANAHAN, Murray. Artificial Intelligence In: SPREVAK, Mark; COLOMBO, Matteo (eds.). **The Routledge Handbook of the Computational Mind**. London; New York: Routledge, 2019, p. 91-100.

SILVEIRA, Sérgio Amadeu. **Democracia e os códigos invisíveis**: como os algoritmos estão modulando comportamentos e escolhas políticas. São Paulo: Edições SESC SP, 2019. [livro digital]

UNIÃO EUROPÉIA. Parlamento Europeu. Resolução do Parlamento Europeu (2015/2103), de 16 de fevereiro de 2017. Disposições de Direito Civil sobre Robótica. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.pdf. Acesso em: 08 jun 2020.

VEDDER, Anton; NAUDTS, Laurens. Accountability for the use of algorithms in a big data environment. **International Review of Law, Computers & Technology**, v. 31, n. 2, p. 206-224, 2017. Disponível em: <https://www.tandfonline.com/doi/pdf/10.1080/13600869.2017.1298547?needAccess=true>. Acesso em: 08 jun 2020.