

THE REASONABLE ROBOT – INTRODUÇÃO: A INTELIGÊNCIA ARTIFICIAL E A LEI¹

The reasonable robot – introduction: Artificial Intelligence and the law

Ryan Abbott²

RESUMO:

A inteligência artificial (IA) e as pessoas não competem em um campo de igualdade. Veículos que dirigem por conta própria podem ser mais seguros que motoristas humanos, mas as leis muitas vezes penalizam tal tecnologia. As pessoas podem fornecer um serviço superior ao cliente, mas as empresas estão se automatizando para reduzir seus impostos. A IA pode inovar mais efetivamente, mas uma estrutura legal antiquada restringe a IA inventiva. Em *The Reasonable Robot*, Ryan Abbott argumenta que a lei não deve discriminar entre a IA e o comportamento humano e propõe um novo princípio legal que, em última instância, melhorará o bem-estar humano. Este trabalho deve ser lido por qualquer pessoa interessada na relação em rápida evolução entre a IA e a lei.

Palavras-chave: Inteligência Artificial; Neutralidade Legal; Responsabilidade Civil; Propriedade Intelectual

ABSTRACT:

*Artificial intelligence (AI) and people do not compete on a level playing field. Self-driving vehicles may be safer than human drivers, but laws often penalize such technology. People can provide superior customer service, but companies are automating to reduce their taxes. AI can innovate more effectively, but an antiquated legal framework restricts inventive AI. In *The Reasonable Robot*, Ryan Abbott argues that the law should not discriminate between AI and human behavior and proposes a new legal principle that will ultimately improve human welfare. This work should be read by anyone interested in the rapidly evolving relationship between AI and the law.*

Keywords: Artificial Intelligence; Legal Neutrality; Tort Law; Intellectual Property

¹ Tradução da introdução do livro **The Reasonable Robot**: Artificial Intelligence and the Law, publicado pela Cambridge University Press em 2020, realizada por Lukas Ruthes Gonçalves, mestre e doutorando em direito pela UFPR. A versão original do livro, bem como suas referências, pode ser acessada através do link: <https://doi.org/10.1017/9781108631761>.

² Ryan Abbott, MD, JD, MTOM, PhD, é Professor de Direito e Ciências da Saúde na Faculdade de Direito da Universidade de Surrey, e Professor Assistente Adjunto de Medicina na UCLA. Médico e advogado de patentes, a pesquisa de Abbott sobre direito e tecnologia ajudou a moldar o diálogo internacional sobre estes tópicos. Ele serviu como especialista para a Organização Mundial da Saúde, a Organização Mundial da Propriedade Intelectual, a Comissão Europeia e o Parlamento do Reino Unido. Abbott também liderou os primeiros pedidos de patente de invenções feitas de forma autônoma por uma IA. Em 2019, ele foi nomeado um dos 50 melhores profissionais em Propriedade Intelectual pela revista **Managing IP**. E-mail: rryanabbott@gmail.com.

A ascensão da poderosa IA será a melhor ou a pior coisa que já aconteceu com a humanidade. Ainda não sabemos qual.

- Stephen Hawking

SUMÁRIO

1. Neutralidade legal da IA. **2.** Área tributária. **3.** Responsabilidade civil. **4.** Propriedade intelectual. **5.** Área criminal. **6.** O futuro da IA. Referências.

A inteligência artificial (IA) está fazendo mais do que nunca, e muitas vezes o está fazendo mais barato, mais rápido e melhor do que as pessoas. Em 2017, a empresa DeepMind desenvolveu uma IA, AlphaGo Master, que venceu o campeão mundial do jogo de tabuleiro Go. Muitos especialistas haviam previsto que o domínio da IA Go levaria mais uma década, dada a complexidade do jogo. Há mais configurações de tabuleiro possíveis no Go do que há átomos no universo. Mais tarde, em 2017, uma versão revisada do AlphaGo, AlphaGo Zero, venceu cem jogos seguidos contra o AlphaGo Master, sem perder nenhum. Isto foi possível após treinamento de apenas três dias jogando contra si mesmo. Ao contrário de seus predecessores, AlphaGo Zero nunca aprendeu com os exemplos humanos.

Go foi o último jogo de tabuleiro tradicional no qual as pessoas podiam superar as máquinas. Agora existe todo um campo de atividade no qual a IA ultrapassa dramaticamente as pessoas. Embora a vitória da AlphaGo tenha sido um marco técnico emocionante, ela teve um impacto social limitado porque jogar jogos de tabuleiro não é o empreendimento mais prático. Porém, eles são uma das medidas mais antigas de se medir a inteligência de uma máquina, e a ascendência da IA sugere que ela pode em breve automatizar uma gama mais ampla de tarefas, talvez mais cedo do que muitos antecipam, e pode fazê-lo de forma espetacular.

A Alphabet, empresa dona do DeepMind, não está investindo em IA para dominar o campo para jogos de tabuleiro competitivos. Em princípio, se uma IA pode ser treinada para reconhecer padrões em Go, então ela pode ser treinada para reconhecer uma pneumonia em um raio-X ou pedestres em uma estrada. De fato, o DeepMind já está sendo aplicado para resolver desafios práticos. Em 2018, a IA AlphaFold do DeepMind superou todos os seus noventa e oito concorrentes em um desafio que visava prever a estrutura tridimensional das proteínas - uma tarefa crítica para a descoberta de medicamentos. Ao contrário de jogar Go, prever o dobramento das proteínas é um problema científico importante, comum e da vida real. Da mesma forma, novamente em 2018, os pesquisadores descobriram que outra IA do DeepMind encaminhou corretamente pacientes com mais de cinquenta doenças oculares distintas para atendimento especializado em 94% dos casos, correspondendo à performance de médicos especialistas. Em 2019, o IA do DeepMind foi capaz de prever consistentemente o desenvolvimento de insuficiência renal aguda quarenta e oito horas antes do que os médicos humanos, o que poderia fazer com que cerca de 30% dos casos jamais ocorressem.

O impacto social futuro desses avanços será tremendo. Já estão, inclusive, sendo criados títulos impressionantes para descrever a disrupção futura dessa era como a Quarta Revolução Industrial, a Segunda Idade da Máquina e a Revolução da Automação. Entre outras coisas, prevê-se que a IA irá gerar uma enorme quantidade de riqueza ao mudar o futuro do trabalho. Esta tem sido há muito tempo a experiência com a IA automatizando o trabalho físico, como na fabricação de automóveis, mas agora ela está passando a automatizar o trabalho mental, e não apenas atividades de serviço relativamente simples como operar uma caixa registradora no McDonald's. A IA está completando tarefas executadas por médicos, advogados e cientistas.

O carro-chefe da IBM, a IA Watson, conhecida por ter ganhado um jogo de *Jeopardy!* em 2011, trabalha em uma série de áreas de especialização. Na área da saúde, Watson (agora uma marca composta por uma variedade de sistemas de IA) analisa a genética dos pacientes com câncer para ajudar a selecionar medicamentos apropriados, uma tarefa que um

grupo de especialistas humanos também pode fazer, mas que continua sendo uma atividade complexa e exigente. Para alguns pacientes, pode exigir cerca de 160 horas de trabalho coletivo por uma equipe de prestadores de serviços de saúde altamente treinados. Em contraste, um estudo de 2017 relata que Watson pode superar a prática padrão e que requer apenas cerca de 10 minutos para fazê-lo¹, embora seu desempenho tenha se mostrado controverso².

Várias empresas afirmam que sua IA já pode superar os médicos humanos em certas áreas da prática médica. Isto não é surpreendente. As máquinas são capazes de memorizar cada pedaço da literatura médica já criada e processar a experiência da prática médica a partir de incontáveis experiências humanas. Além disso, elas nunca precisam de uma pausa para descansar. Em 2017, uma empresa chinesa relatou que seu robô Xiao Yi fez e passou, com folga, o Exame Nacional de Licenciamento Médico, o teste necessário para se tornar um médico na China. Xiao Yi conhece o conteúdo de dezenas de livros médicos, milhões de registros médicos e centenas de milhares de artigos, mas para passar no teste também teve que aprender, raciocinar e fazer julgamentos por si só. Pesquisadores da IBM relataram até que Watson passou tranquilamente pelo exame equivalente nos Estados Unidos depois de ter sido proibido de formalmente fazê-lo.

Naturalmente, apenas passar nos testes não faz de alguém, ou alguma coisa, um médico. Ainda que a IA seja consistentemente melhor que um médico no diagnóstico de certas doenças, no gerenciamento de prescrições ou na realização de cirurgias, é pouco provável que a ela automatize completamente o atendimento médico. Mas estes avanços já sugerem que existem aspectos da assistência médica suscetíveis à automatização, e que menos médicos serão necessários uma vez que haja médicos mais eficientes auxiliados pela IA.

1 NEUTRALIDADE LEGAL DA IA

A lei desempenha um papel crítico no uso e desenvolvimento da IA. As leis estabelecem regras e padrões de comportamento obrigatórios para

garantir o bem-estar social e proteger direitos individuais, e elas podem nos ajudar a perceber os benefícios da IA ao mesmo tempo em que minimizam seus riscos - que são significativos. A IA tem estado envolvida em quedas rápidas no mercado de ações, crimes cibernéticos e manipulação social e política. Tecnólogos famosos como Elon Musk e acadêmicos como Stephen Hawking até argumentaram que a IA pode condenar a raça humana. A maioria das preocupações, no entanto, se concentram em problemas mais práticos e de menor prazo, tais como desemprego, discriminação e segurança relacionados a tecnologia.

Embora os riscos e benefícios da IA sejam amplamente reconhecidos, há pouco consenso sobre a melhor forma de regular a IA e as jurisdições em todo o mundo estão definindo quais seriam as ações a serem tomadas. Já existe uma divisão internacional significativa em relação à extensão na qual a IA pode ser usada na vigilância estatal de seus residentes, se empresas ou consumidores seriam os “donos” dos dados pessoais vitais para o desenvolvimento da IA, e quando os indivíduos teriam direito a uma explicação para as decisões tomadas pela IA (desde aprovações de crédito até sentenças criminais).

É tentador esperar que a IA se encaixe perfeitamente nas regras existentes, mas as leis destinadas a regular o comportamento dos atores humanos muitas vezes têm consequências involuntárias e negativas quando as máquinas começam a agir como pessoas. Apesar disso, as leis centradas na IA têm passado por um desenvolvimento lento, devido em parte à preocupação de que um ambiente regulatório excessivamente oneroso possa deter a inovação. No entanto, a IA já está sujeita a regulamentações que podem ter sido criadas há décadas para lidar com questões como privacidade, segurança e concorrência desleal. O que é necessário não é necessariamente mais ou menos lei, mas a lei certa.

Em 1925, o juiz Benjamin Cardozo advertiu uma turma de graduação em Direito que “as novas gerações trazem consigo seus novos problemas que exigem novas regras, a serem padronizadas, de fato, após as regras do passado, e ao mesmo tempo adaptadas às necessidades e à justiça de outro dia e hora”³. Este é o caso da IA, mesmo que só se diferencie em grau de outras tecnologias disruptivas como computadores pessoais e

a Internet. Um regime jurídico otimizado para a IA é ainda mais importante se a IA se revelar diferente em espécie.

Não é provável que haja uma única mudança legal, como a concessão de personalidade jurídica à IA semelhante a uma empresa, que resolverá questões em todas as áreas da lei, e é por isso que é necessário fazer o difícil trabalho de pensar sobre as implicações da IA em diferentes ambientes. A este respeito, é promissor que tenha havido nos últimos anos esforços para articular padrões políticos ou os melhores princípios, tais como confiabilidade e sustentabilidade especificamente para a regulamentação da IA pelos governos, grupos de reflexão e indústria. Por exemplo, a Organização para Cooperação e Desenvolvimento Econômico (OECD) adotou seus Princípios sobre Inteligência Artificial em maio de 2019⁴, e um mês depois o G20 adotou princípios de IA centrados no ser humano, guiados pelos princípios delineados pela OECD⁵.

A tese central deste livro sustenta que precisa haver um novo princípio orientador para a regulamentação da IA, um princípio de *neutralidade legal da IA*, afirmando que a lei não deve discriminar entre o comportamento humano e o da IA. Atualmente, o sistema legal não é neutro. Uma IA que seja significativamente mais segura que uma pessoa pode ser a melhor escolha para dirigir um veículo, mas as leis existentes podem proibir veículos sem condutor. Uma pessoa pode ser a melhor escolha para a fabricação de bens, mas uma empresa pode automatizar porque economiza em impostos. A IA pode ser melhor para gerar certos tipos de inovação, mas as empresas podem não querer usar a IA se isso restringir a propriedade dos direitos de propriedade intelectual. Em todos esses casos, um tratamento jurídico neutro acabaria beneficiando o bem-estar humano, ajudando a lei a atingir melhor seus objetivos políticos subjacentes.

A IA pode se comportar como uma pessoa, mas não é como uma pessoa. As diferenças entre a IA e as pessoas ocasionalmente exigirão regras diferenciais. A diferença mais importante é que a IA, que carece de consciência e interesses humanos, não merece moralmente direitos, portanto, tratar a IA como se o fizesse só deveria ser justificado se isso be-

neficiasse as pessoas. Um exemplo disso seria se os próprios veículos autônomos precisassem ser titulares de apólices de seguro ou outras formas de garantia para cobrir potenciais danos a pedestres. Esta é essencialmente a razão pela qual se permite que empresas sejam partes em contratos e possuam propriedade própria. Seus direitos legais existem apenas para melhorar a eficiência das atividades humanas, como o comércio e o empreendedorismo e, como a IA, as empresas não merecem moralmente direitos. Elas são um membro de nossa comunidade jurídica, mas não de nossa comunidade moral.

Consequentemente, este livro não advoga que a IA tenha direitos ou que seja qualificada como uma pessoa jurídica. Nem que o princípio de neutralidade legal da IA seja um princípio moral de não-discriminação na forma como esse termo é tradicionalmente usado. As leis antidiscriminação têm ajudado a melhorar as condições para grupos historicamente marginalizados, principalmente por uma questão de justiça. Entretanto, as leis antidiscriminação também podem promover a concorrência e a eficiência.

Certamente, a neutralidade legal da IA não deve ser a força motriz por trás de cada decisão. Ela não deve vir em detrimento de outros princípios, como transparência e responsabilidade. Uma pessoa pode ser mais eficiente na mineração de minerais em condições perigosas, mas a automação pode ser preferível com base em considerações de segurança. Uma IA pode ser mais eficiente na identificação e eliminação de alvos militares, mas pode haver outras razões para não delegar decisões de vida ou morte a uma IA.

Ao invés de um princípio de política dispositiva, a neutralidade legal da IA um padrão apropriado que pode ser afastado quando há boas razões para isso. Este livro examina como tal princípio afetaria quatro áreas da lei - tributária, responsabilidade civil, propriedade intelectual e criminal - e argumenta que como a IA ocupa cada vez mais papéis uma vez reservados às pessoas, a IA precisará ser tratada mais como uma, e às vezes as pessoas precisarão ser tratadas mais como a IA.

2 ÁREA TRIBUTÁRIA

A automação envolve muito mais do que deixar as pessoas sem emprego, o que o economista John Maynard Keynes chama de “desemprego tecnológico”, mas é uma das coisas que mais preocupa as pessoas. Hoje, este é um frequente tema de pesquisas sobre os mercados de trabalho, algumas que prevêm o desemprego tecnológico de longo prazo e outras que não o fazem. É também uma preocupação antiga. Os Luditas, um grupo de trabalhadores ingleses, se opunham à eliminação de empregos pela automação, destruindo periodicamente máquinas em atos de protesto durante a Primeira Revolução Industrial. A história tem mostrado que seus receios eram descabidos, pelo menos no que diz respeito às preocupações com o desemprego de longa duração. No final, as máquinas resultaram não apenas em grandes ganhos de produtividade, mas também em mais empregos para todos, e desde então as novas tecnologias têm consistentemente resultado na criação geral de empregos. Motores a vapor, energia elétrica e computadores pessoais, todos eliminaram empregos, mas eles criaram mais empregos do que eliminaram.

Na virada do século XX, cerca de 40 por cento da força de trabalho dos EUA estava empregada na agricultura. Agora, menos de 2% da força de trabalho trabalha na agricultura, o que não se traduziu em um aumento de 38% no desemprego. De fato, mesmo quando o emprego na agricultura e a contribuição relativa da agricultura para a economia americana diminuiu, a produtividade dos trabalhadores agrícolas disparou e a contribuição absoluta da agricultura para a economia aumentou.

Para a Quarta Revolução Industrial, o fato de a história se repetir pode não ser tão ruim. Apesar de algumas reticências, os riscos da automação podem ser exagerados e novamente resultar em ganhos de emprego a longo prazo. De todo modo, a Primeira Revolução Industrial foi acompanhada por décadas de agitação social generalizada, aumentando as disparidades de renda e sofrimento individual. Uma abordagem regulatória proativa deveria nos permitir aproveitar ao máximo a automação e, ao mesmo tempo, limitar alguns de seus efeitos nocivos - ainda mais

importante se a IA resultar em um novo tipo de revolução industrial com um aumento permanente do desemprego de longo prazo.

Mas para todo o debate sobre a IA estar deixando as pessoas desempregadas, isto pode ocorrer por uma razão muito surpreendente. As leis tributárias tratam as pessoas e a tecnologia de automação de forma diferente, mesmo quando elas estão realizando as mesmas tarefas. Por exemplo, a automação permite que as empresas evitem os impostos salariais de funcionários e empregadores. Portanto, se um chatbot custa à Sephora o mesmo ou até um pouco mais antes dos impostos do que um empregado que faz o mesmo trabalho, na verdade custa menos para a empresa automatizar depois dos impostos.

Além de evitar impostos salariais, as empresas podem acelerar as deduções de impostos para alguma IA quando esta possui um componente físico ou cumpre algumas exceções aplicáveis para software - mas não para trabalhadores humanos. Em outras palavras, os empregadores podem recuperar grande parte do custo de uma IA antecipadamente como dedução de impostos, o que pode ser mais valioso para algumas grandes empresas do que atrasar as despesas salariais ao longo do tempo. Finalmente, os empregadores também recebem uma variedade de incentivos fiscais indiretos para automatizar. Em resumo, nossas leis fiscais impedem as pessoas e a IA de competir por seus méritos. Embora o sistema não tenha sido projetado para fazer isso, ele tributa principalmente o trabalho e não o capital. Isto tem tido o efeito não intencional de incentivar ineficientemente a automação, já que a IA tem assumido o papel tanto do capital quanto do trabalho.

O que é mais preocupante é que a IA não paga impostos! Isto soa ridículo, mas os impostos de renda e emprego são as maiores fontes de receita para o governo, juntos representando quase 90% do total da receita tributária federal. Pelo contrário, os impostos de renda das empresas geram menos de 8% da receita federal. Sob a Lei de Cortes de Impostos e Empregos de 2017, a alíquota estatutária do imposto de sociedades foi reduzida para 21% e a receita do imposto de sociedades tem tido uma tendência acentuada de queda. Qualquer que seja a alíquota legal, a alíquota

efetiva do imposto corporativo - o que as empresas pagam após levar em conta as isenções fiscais - é substancialmente menor.

Em 2018, a taxa anual de imposto do S&P 500, que se refere a 500 grandes empresas que têm ações ordinárias listadas em uma das três principais bolsas de valores dos EUA, era inferior a 18%⁶. No entanto, isto inclui todos os impostos dos níveis federal e estadual, bem como das autoridades estrangeiras. A Amazon chamou indesejada atenção naquele ano ao relatar um lucro nos EUA de US\$ 11,2 bilhões antes da tributação juntamente com uma conta de impostos negativa de US\$ 129 milhões⁷. A taxa total efetiva de impostos da Amazon para 2018 foi de 11%, incluindo impostos estrangeiros, estaduais e diferidos.

Portanto, a IA não paga imposto de renda nem gera impostos sobre o emprego. Não compra bens e serviços, portanto, não é cobrado imposto sobre vendas. Não compra ou possui propriedade, portanto, não paga impostos sobre a propriedade. A IA simplesmente não é um contribuinte, pelo menos não na mesma medida que um trabalhador humano. Se todo o trabalho fosse automatizado amanhã, a maior parte da base tributária desapareceria imediatamente. O que acontece é que quando as empresas se automatizam o governo perde receitas - potencialmente centenas de bilhões de dólares no total. Isto pode ser suficiente para restringir significativamente a capacidade do governo de pagar por coisas como seguridade social, defesa nacional e assistência médica. A longo prazo, a perda de receita deveria se reequilibrar se as pessoas que ficaram desempregadas eventualmente retornassem a tipos de trabalho semelhantes, e deveria haver ganhos de receita se a automação tornar as empresas mais produtivas e se as pessoas continuarem a encontrar tipos de trabalho mais bem remunerados. Este não será o caso se nos encaminharmos para um futuro de trabalho com taxas de desemprego mais elevadas, a menos que haja um aumento da produtividade que supere drasticamente o desemprego.

Só recentemente surgiu o debate público sobre a tributação da IA, e tem sido principalmente em relação ao abrandamento da taxa de automação nas empresas, não como uma tentativa de elaborar políticas neutras do ponto de vista fiscal ou garantir a receita do governo. A questão de como a lei deve responder permanece. A automação não deve ser desencorajada por princípio; na verdade, ela deve ser bem-vinda quando

melhora a eficiência. Porém, a automação para fins de economia fiscal pode não tornar as empresas mais produtivas ou resultar em quaisquer benefícios para o consumidor, e pode resultar em quedas de produtividade para reduzir a carga tributária. Isto não é socialmente benéfico.

As opções, uma vez que os legisladores concordem que eles não querem beneficiar a IA em detrimento dos trabalhadores humanos, poderia ser reduzir os benefícios fiscais que a IA recebe sobre as pessoas ou reduzir os impostos existentes que só se aplicam aos trabalhadores humanos. Por exemplo, os impostos sobre a folha de pagamento poderiam ser eliminados, o que pode ser uma maneira melhor de alcançar a neutralidade, uma vez que reduz a complexidade fiscal e acaba com a tributação de algo de valor social - isto é, o trabalho humano. No entanto, isto eliminaria cerca de 35% da receita tributária atual do governo federal dos EUA.

Há muitas maneiras de garantir uma receita tributária adequada, como por exemplo, aumentando os impostos sobre propriedade ou sobre vendas, o que pode ser uma forma mais progressiva de coletar fundos, pois tributaria a renda independentemente de sua fonte - trabalho ou capital. Poderia certamente ser progressivamente concebida através da aplicação de impostos de propriedade relativamente mais altos para propriedades de maior valor e impostos de venda mais altos para, digamos, bens de luxo. O imposto sobre a renda também poderia ser elevado através do aumento das alíquotas estatutárias marginais de impostos para pessoas com altos rendimentos ou das alíquotas efetivas de impostos através da eliminação de coisas como a regra *step-up in basis*, a qual reduz a responsabilidade fiscal por bens herdados.

Mais ambiciosamente, a neutralidade legal da IA pode provocar uma mudança mais fundamental na forma como o trabalho *versus* capital e os trabalhadores *versus* empresas são tributados. Novos regimes fiscais poderiam visar a IA, bem como outras tecnologias de automação às quais se aplicam considerações semelhantes, mas isso provavelmente aumentaria os custos de conformidade e a complexidade fiscal. Também “taxaria a inovação” no sentido de que poderia penalizar os modelos de negócios que são legitimamente mais produtivos com menos trabalho humano. Uma solução melhor seria aumentar os impostos sobre ganhos de capital e as taxas de impostos corporativos para reduzir a dependência

nos impostos sobre o trabalho. Antes de a IA entrar em cena, havia críticas de longa data sobre a medida em que o capital é favorecido em relação ao trabalho na política fiscal. A Quarta Revolução Industrial pode dar o impulso necessário para finalmente abordar esta questão.

O lado negativo do aumento da tributação do capital é, em grande parte, uma preocupação com a concorrência tributária internacional. Há uma crença histórica de que a mão-de-obra deve ser mais tributada do que o capital, porque o capital é mais móvel e deixará as jurisdições com taxas tributárias maiores. Estas preocupações podem ser exageradas, particularmente em grandes mercados desenvolvidos, como os Estados Unidos. Historicamente, taxas relativamente altas de impostos corporativos não têm sido uma barreira para investimentos baseados nos EUA.

Os Estados Unidos, que têm a maior economia do mundo, não têm um sistema tributário relativamente progressivo - ou seja, um sistema baseado na capacidade de pagamento de uma pessoa. Existem maiores disparidades de riqueza nos Estados Unidos do que em qualquer outro país desenvolvido. Com a IA provavelmente resultando em ganhos financeiros maciços, mas mal distribuídos, ela exigirá e nos permitirá repensar como alocar recursos e redistribuir riqueza. Se optarmos por reduzir a desigualdade de renda, isto deve ser feito principalmente através da tributação. Novas leis para garantir que a IA contribua com sua parcela justa para a receita do governo poderiam financiar programas de reciclagem para os trabalhadores e aumentar os benefícios sociais. Se a IA causar um aumento do desemprego a longo prazo, a receita tributária subsequente poderia até mesmo apoiar uma renda básica universal que permitiria aos governos pagar a cada cidadão independentemente de seu emprego.

3 RESPONSABILIDADE CIVIL

A IA fará todo tipo de coisas que só uma pessoa costumava fazer como dirigir. É difícil dizer exatamente quando isso vai acontecer. Empresas como Uber e Tesla afirmaram que estariam usando ou vendendo veículos totalmente autônomos (AVs) antes de 2020. Outros fabricantes de

automóveis agora atestam que estarão vendendo AVs no início da década de 2020. Em contraste, um grupo de especialistas da Comissão Europeia prevê que veículos totalmente autônomos não estarão disponíveis comercialmente antes de 2030⁸. Independentemente de quando os veículos autônomos se tornarão comuns, pesquisas frequentemente relatam atitudes públicas negativas sobre eles. A maioria das pessoas diz que se sentiriam inseguras sendo levadas pelo seu carro, mas os AVs já podem ser mais seguros do que as pessoas. Os motoristas humanos são perigosos - cerca de 94% dos acidentes envolvem erros humanos. No mundo inteiro, mais de um milhão de pessoas morrem a cada ano em acidentes com veículos motorizados e dezenas de milhões são feridos. Isto é quase exclusivamente o resultado do fato de as pessoas serem péssimos motoristas.

Houve pelo menos uma fatalidade causada por um AV. Operado por Uber, o AV colidiu com um pedestre no Arizona porque não conseguiu detectá-lo a tempo de parar. Mais recentemente, os reguladores relataram que um sistema “Autopilot” da Tesla pode ter sido o culpado por um acidente fatal em março de 2019. Um porta-voz da empresa observou em resposta que os motoristas da Tesla fizeram mais de um bilhão de milhas com o sistema de piloto automático de seus carros ativados, e que ele tende a tornar os motoristas mais seguros⁹. As fatalidades relatadas anteriormente envolvendo o sistema Autopilot da Tesla foram finalmente determinadas pelas entidades reguladoras como não sendo culpa da AV. Entretanto, esses incidentes abordam os desafios da interação homem-máquina - os motoristas humanos devem estar preparados para retomar o controle do veículo em cima da hora, mas é difícil para as pessoas permanecerem alertas e engajadas enquanto um AV está dirigindo.

Inevitavelmente, os carros que dirigem por conta própria causarão fatalidades. Mas o perfeito não deve ser o inimigo do bom. Os AVs não precisam ser inofensivos para tornar as pessoas mais seguras; eles só precisam ser melhores motoristas do que as pessoas. Seja em 2025, 2035 ou 2045, os AVs não serão apenas condutores mais seguros do que as pessoas, mas muito mais seguros do que elas. Os AVs estão melhorando rapidamente e os motoristas humanos não, o que é importante no que diz respeito à responsabilidade legal por atos ilícitos.

A lei de responsabilidade civil define o que constitui um dano legal e determina os casos em que alguém pode ser responsabilizado civilmente em oposição a criminalmente. Os acidentes causados por pessoas geralmente exigem negligência para a responsabilização. Negligência significa que as ações de alguém estão abaixo do padrão de uma pessoa razoável. Para aplicar este teste, ao longo dos séculos, os tribunais desenvolveram o conceito de uma hipotética pessoa razoável que estabelece o padrão de comportamento humano. Assim, se um motorista é processado por colidir com uma criança que fugiu para a rua, e uma pessoa razoável teria evitado a criança, então o motorista seria considerado responsável. Se uma pessoa razoável não tivesse evitado a criança, então o motorista não seria responsabilizado. O paradigma da “pessoa razoável” não é exclusivo da responsabilidade civil; é uma norma que se aplica em muitas áreas, inclusive na legislação criminal e contratual.

As pessoas mencionam a responsabilidade civil quando falam sobre a lei e a condução de automóveis, muitas vezes no contexto da preocupação de que ninguém - ou nada - será responsabilizado por acidentes causados por veículos automotores. Essas preocupações são descabidas. Os AVs são produtos, e já existe um regime legal construído em torno de lesões causadas por produtos. A lei de responsabilidade por defeitos do produto poderia simplesmente ser aplicada aos acidentes automobilísticos. Simplificando, há uma regulação diferente para acidentes causados por pessoas e para acidentes causados por produtos. A lei responsabiliza estritamente os fabricantes e revendedores de produtos pelos danos causados quando uma máquina é defeituosa, ou quando suas propriedades estão desvirtuadas. A responsabilidade estrita se refere ao fato de que a responsabilidade se baseia na causalidade, sem que haja negligência: a IA causou um dano independentemente de a conduta de um fabricante ser socialmente censurável? A responsabilidade estrita é uma barra mais fácil de ser superada para a responsabilização, o que é uma boa regra para a maioria dos produtos. Há mais responsabilidade para os fabricantes, incentivando-os assim a fabricar produtos mais seguros, já que eles estão na melhor posição para melhorar sua segurança e lucrar com a diminuição de acidentes. No entanto, uma maior responsabilidade para os fabri-

cantes não se traduz necessariamente em menos acidentes se um produto for mais seguro do que a norma existente. Nesse caso, a lei de responsabilidade por fatos ou vícios de produto tornaria as pessoas menos seguras. Quando a IA tem mais responsabilidade do que uma pessoa, ela torna o custo da automação mais alto. Este não é um resultado desejável. A automação deve ser encorajada, ou pelo menos não desencorajada, através da responsabilidade civil em situações em que ela melhoraria a segurança.

Ao invés vez de aplicar a legislação padrão de responsabilidade por produtos à IA, ela deveria avaliar os acidentes causados pela IA sob um padrão de negligência. Em certo sentido, isto trataria a IA como uma pessoa e se concentraria no ato da IA ao invés de seu design. A lei perguntaria se a IA se comportou de tal forma que, se uma pessoa tivesse feito o mesmo, o ato teria ficado abaixo do padrão de uma pessoa razoável. Se assim fosse, então o fabricante da IA seria financeiramente responsável. Como com os réus humanos, a lei normalmente não se preocupa com o que uma pessoa estava pensando ou se ela achava que o que estava fazendo parecia razoável. A lei analisa objetivamente se uma pessoa razoável teria cometido o ato.

Aqui, como na legislação tributária, IA e pessoas competem nos mesmos tipos de atividades e exibem comportamentos semelhantes. Atualmente, a lei tributária incentiva a automação, enquanto a lei de responsabilidade civil desestimula a automação. Um princípio de neutralidade legal da IA que julga seu comportamento por um padrão de negligência encorajaria o desenvolvimento e a adoção de tecnologias mais seguras. Novamente, isto não trataria a IA e as pessoas diretamente da mesma forma legal, já que a IA não seria pessoalmente responsável pelos danos causados por ela. A IA é possuída como uma propriedade, não tem recursos financeiros e não é influenciada pelo espectro da responsabilidade como as pessoas são. A responsabilidade baseada em negligência para a IA funcionaria como um mecanismo baseado no mercado para incentivar a introdução de tecnologias que melhoram a segurança com o benefício de não exigir financiamento governamental, encargos regulamentares adicionais para a indústria, ou novas responsabilidades administrativas.

Aplicar uma estrutura de negligência à IA é a parte menos importante da história da responsabilidade civil. Chegará o momento em que o desempenho da IA não será apenas mais seguro, mas substancialmente mais seguro do que o de uma pessoa - a ponto de que carros que se dirigem sozinhos quase nunca causarão acidentes, e quase qualquer acidente causado por uma pessoa seria considerado negligente em comparação com uma IA. Quando isto acontecer, não importará qual regime de responsabilidade a lei se aplica à IA; importará qual regime de responsabilidade que aplicamos a nós mesmos porque até lá nós - pessoas - seremos o maior perigo na estrada. Os tribunais, neste ponto, deveriam manter os motoristas humanos no padrão dos carros que dirigem sozinhos - o *Reasonable Robot standard* (padrão razoável de robô), embora possa ser mais preciso, se menos cativante, chamá-lo de padrão razoável de IA. Hoje em dia, se uma criança corre na frente do carro de uma pessoa enquanto ela está dirigindo à noite e não consegue parar, a pessoa provavelmente não seria responsável. Entretanto, em um futuro em que o padrão razoável da pessoa é representado por uma IA que poderia ter parado, aí ela seria responsável.

Carros autônomos são apenas um exemplo de como a IA irá interferir com a lei de responsabilidade civil. Com o aumento dos custos de saúde e as evidências sugerindo que a IA pode superar as pessoas em alguns aspectos do cuidado com a saúde, as pessoas poderão em breve consultar o Dr. Watson para obterem cuidados. Atualmente, a IA só pode prevalecer sobre pessoas em aspectos muito restritos da prática médica, mas está melhorando rapidamente, e os médicos humanos não estão. O que deve ser lembrado é que o Watson não precisa ser perfeito para melhorar a segurança - apenas um pouco melhor do que os médicos humanos, e essa barra é baixa. Não se enganem, os médicos humanos são absolutamente perigosos. As pessoas não devem deixar de consultar médicos, mas o erro médico mata muito mais pessoas do que os acidentes de carro. Na verdade, para os médicos que fazem um juramento de em primeiro lugar nunca causar o mal a alguém, não será ético permitir que eles concorram com uma IA muito mais segura.

4 PROPRIEDADE INTELECTUAL

Não deve causar surpresa o fato de que a IA tenha gerado invenções científicas de forma autônoma por décadas. Mas enquanto a lei fornece direitos de propriedade intelectual - direitos a certas criações intangíveis, como direitos autorais para livros e música ou patentes para certos tipos de descobertas - para a produção humana, ela permanece atrasada quando a IA cria “produtos da mente”. Legalmente, não está claro se as invenções geradas pela IA, aquelas feitas sem os inventores tradicionais, são elegíveis para proteção de patentes.

Na maioria dos casos, a existência de uma patente exige que o inventor seja uma pessoa física (um indivíduo humano) e o titular inicial desse direito. Os inventores têm a capacidade de transferir seus direitos para outros; isto pode acontecer automaticamente quando empregados criam algo dentro do escopo do emprego. De fato, a maioria das patentes são de propriedade de pessoas artificiais na forma de empresas. Ainda assim, a exigência de que um inventor seja uma pessoa física garante o direito dos criadores humanos de serem reconhecidos mesmo quando as empresas possuem direitos relacionados à propriedade intelectual.

Estas leis não foram projetadas com a IA em mente e até o ano de 2019 não há qualquer lei especificamente sobre invenções geradas por IA em nenhuma jurisdição. Como resultado, não está claro se uma invenção gerada por IA poderia ser patenteável, quem poderia ser o proprietário de tal patente e quem - ou o que - poderia ser considerado um inventor. Estas não são apenas questões acadêmicas. Em 2019, a Siemens relatou que não tinha sido capaz de requerer proteção para múltiplas invenções geradas por IA porque não conseguia identificar uma pessoa física qualificada como inventor¹⁰. Enquanto isso, os escritórios de patentes provavelmente têm concedido patentes sobre invenções geradas por IA por décadas - mas apenas porque ninguém está revelando o envolvimento da IA.

A lei deve permitir patentes para invenções geradas por IA e até mesmo reconhecer a IA como inventora quando a ela atende aos critérios de invenção. A principal razão se baseia no motivo pelo qual a lei concede direitos de propriedade intelectual em primeiro lugar: para incentivar certas

atividades de valor social. Se não houvesse uma lei que impedisse as pessoas de copiar ou usar produtos intangíveis como um novo medicamento ou um processo industrial, este tipo de coisa estaria sendo subproduzido. Isto é referido como o problema do *free rider*, para o qual as patentes são uma solução. Elas fornecem a um inventor um monopólio temporário sobre uma invenção, impedindo que terceiros a utilizem ou a copiem sem permissão. A perspectiva de uma patente proporciona assim uma motivação financeira adicional para os inventores.

As máquinas não têm uso para patentes, mas as pessoas que constroem, possuem e usam IA têm. Permitir patentes para invenções geradas por uma IA inventiva a tornaria mais valiosa e incentivaria o seu desenvolvimento, o que se traduziria em recompensas pelo esforço a montante do estágio de invenção e, em última instância, resultaria em mais inovação. Em contraste, não permitir patentes para invenções geradas por IA desencorajaria as empresas de usá-las para gerar nova propriedade intelectual, mesmo nos casos em que ela seria mais eficaz do que uma pessoa.

Além disso, o reconhecimento da IA como inventora protegeria os direitos morais humanos porque impediria as pessoas de receberem um reconhecimento imerecido. Assumir o crédito pelo trabalho de uma IA não importaria para uma máquina, mas diminuiria as realizações das pessoas que legitimamente criaram obras patenteáveis. Além disso, o reconhecimento da IA como inventora daria valor aos seus desenvolvedores e reduziria o trabalho dos escritórios de propriedade intelectual. Assim como no caso de carros autônomos, a IA não possuiria direitos de propriedade intelectual - esses seriam de propriedade do proprietário da IA.

Alguns críticos têm argumentado a favor da proibição dos direitos de propriedade intelectual sobre obras geradas por IA com base no fato de que a lei deveria proteger apenas os resultados da atividade mental e as máquinas não podem pensar. Este não é o foco correto: Se e como uma IA pensa não deve ser relevante. O Congresso percebeu a necessidade de um padrão funcional para invenções nos anos 50. Antes disso, os tribunais utilizavam um teste denominado de "*Flash of Genius*", que exigia que a faísca inventiva chegasse a uma pessoa em um momento de clareza e

não como resultado de uma pesquisa metódica e laboriosa. A natureza do teste nunca foi totalmente clara, mas o processo envolvia o raciocínio subjetivo dos juízes sobre o que um aspirante titular de patente poderia estar pensando.

Eventualmente, as pessoas perceberam que o teste era uma ideia terrível. Ele é difícil de aplicar, mas mais importante, não deveria importar para a lei se a invenção vem de Einstein ou de uma sala cheia de macacos. O que interessa à sociedade é gerar inovação socialmente benéfica, não como uma IA é projetada ou se ela pensa em um sentido filosófico. A inventividade da produção de uma IA ao invés de um antropomorfismo desajeitado deve guiar a lei de propriedade intelectual.

Em 2019, advogados de patentes liderados pelo autor deste livro anunciaram que haviam depositado os primeiros pedidos de patente para divulgar que eles se baseavam em invenções geradas por IA como parte do *Artificial Inventor Project* (Projeto Inventor Artificial). As aplicações nomeiam uma IA, “DABUS”, como inventora de um projeto de recipiente funcional e de um tipo de sinal de emergência, com o proprietário da IA listado como o requerente da patente e futuro proprietário de quaisquer Direitos resultantes. Antes de o *Artificial Inventor Project* anunciar os pedidos feitos, o Escritório de Propriedade Intelectual do Reino Unido (UKIPO) havia examinado os arquivos e essencialmente determinado que eles eram patenteáveis, exceto que um inventor humano não havia sido nomeado. Após a divulgação do inventor IA, o UKIPO rejeitou os pedidos, que agora estão sob apelação nas Cortes Superiores do Reino Unido. Em seu aviso de rejeição, o auditor da UKIPO escreveu que

As invenções criadas pelas máquinas de IA provavelmente se tornarão mais prevalentes no futuro e há uma questão legítima sobre como ou se o sistema de patentes deve lidar com tais invenções. Descobri que o sistema atual não atende a tais invenções e nunca foi previsto que o faria, mas os tempos mudaram e a tecnologia avançou. É correto que isto seja debatido mais amplamente e que quaisquer mudanças na lei sejam consideradas no contexto de tal debate, e não arbitrariamente inseridas na legislação existente¹¹.

Esses mesmos pedidos de patentes estão atualmente sendo considerados em várias jurisdições ao redor do mundo, incluindo os Estados Unidos e o Escritório Europeu de Patentes¹².

Enquanto isso, a IA está melhorando exponencialmente e os pesquisadores humanos não. Isto é excitante porque significa que a sociedade provavelmente testemunhará o mesmo tipo de fenômeno com a IA inventiva e com os carros autônomos: grandes melhorias em relação ao desempenho humano. Quando a IA superar o desempenho das pessoas, ela se tornará a forma padrão de realização da pesquisa. Ao invés de a Pfizer perguntar a seus cientistas se um medicamento que trata uma condição imunológica pode tratar outra, ou a Exxon seus químicos para projetar melhores catalisadores, ambas as empresas usarão, por exemplo, a IA do DeepMind para completar a tarefa. Na lei de patentes, o padrão humano frequentemente comparado ao padrão de pessoa razoável é a “pessoa com habilidade comum na arte”, ou a pessoa hábil. Esta pessoa hipotética representa o trabalhador médio em um campo e serve como referência para o comportamento humano. A ideia é que para se obter um certificado de patente, uma pessoa deve realizar algo a mais do que o trabalhador médio de uma área faria. Portanto, se uma pessoa inventa algo que seria óbvio para a pessoa hábil, não lhe pode ser concedida uma patente. Se a invenção não for óbvia, então lhe pode ser.

A IA deve mudar o padrão da pessoa qualificada. Como a pessoa hábil reflete o trabalhador médio em um campo de invenção, o conceito deve mudar uma vez que o trabalhador médio seja complementado pela IA. Neste momento, talvez já na atualidade em alguns campos, a pessoa hábil deve se tornar a pessoa hábil que usa a IA. Isto deve elevar o nível de patenteabilidade porque o auxílio da IA tornará os trabalhadores médios mais sofisticados e instruídos - tornando mais óbvias as invenções. Uma vez que a IA passe de ser um auxiliar rotineiro para automatizar o trabalho inventivo, a pessoa hábil deve se tornar uma IA inventiva. Isto deve elevar ainda mais o requisito da patenteabilidade porque a IA inventiva do futuro encontrará mais facilmente as invenções óbvias, e esse requisito continuará a subir à medida que as máquinas continuem a melhorar.

Sem um limite claro para a sofisticação da IA, será difícil para uma pessoa sozinha encontrar qualquer coisa não óbvia. Eventualmente, tudo será óbvio para uma IA superinteligente. Isto pode significar o fim do sistema de patentes, mas isto não deve ser motivo de preocupação. Uma vez que a IA inventiva superinteligente seja de uso corrente, os custos financeiros da inovação serão triviais, o impulso para incentivar será desnecessário e a inovação futura será autossustentável.

5 ÁREA CRIMINAL

Semelhante ao seu papel em invenções, a IA já está se engajando de forma autônoma em atividades que seriam criminosas para uma pessoa física. Além disso, a IA pode fazê-lo de forma indetectável, ou irredutível, ao ato ilícito de uma pessoa. Em outras palavras, da mesma forma que algumas obras geradas por IA carecem de um ator que tradicionalmente se qualificaria como inventor, há casos de crimes gerados por IA em que nenhuma pessoa física pode ser responsabilizada criminalmente. Hoje, pelo menos, quase todos os crimes envolvendo IA são passíveis de serem redutíveis ao crime humano. Dito de forma simples, se uma pessoa ataca alguém com um computador, ela cometeu assalto, não o computador.

Entretanto, pode haver momentos em que não é possível reduzir o crime da IA a um indivíduo devido à autonomia, complexidade ou explicabilidade limitada da IA. Tal caso pode envolver vários indivíduos que contribuem para o desenvolvimento de uma IA durante um longo período de tempo, como com software de código aberto, onde milhares de pessoas podem colaborar informalmente para criar uma IA. Outro caso nesta categoria pode apresentar uma IA que se desenvolve em resposta ao treinamento com dados. Atribuir reponsabilidade para uma operação de IA aonde a máquina aprendeu a se comportar com base no acesso a milhões ou bilhões de pontos de dados de fontes heterogêneas pode ser praticamente impossível.

O direito penal fica aquém do esperado dada a possibilidade de uma entidade autônoma poder se envolver em atividades criminosas sem ser responsabilizada. Isto provavelmente se tornará um problema mais signi-

ficativo à medida que a IA se tornar mais avançada, comum e independente. Uma solução para crimes gerados pela IA é responsabilizar a própria IA e condená-la por um crime. Um pequeno, mas crescente número de acadêmicos está desenvolvendo tais argumentos, e a punição criminal da IA pode parecer seguir o princípio da neutralidade legal da IA, que adverte contra o tratamento legal diferenciado entre uma ação da IA e o comportamento humano.

A aplicação da neutralidade legal da AI é mais simples em áreas onde a lei se preocupa principalmente em promover certos tipos de comportamento, tais como incentivar a inovação no caso da lei de patentes ou melhorar a segurança na responsabilidade civil. Mas algumas áreas da lei estão mais explicitamente preocupadas com as motivações intrínsecas de alguém para agir de uma certa forma. Este é particularmente o caso no direito penal, que geralmente exige não apenas que alguém tenha se envolvido em algum tipo de conduta proibida, mas também que ele o tenha feito por más razões. Por exemplo, para cometer um crime uma pessoa pode precisar causar dano, talvez até mesmo um dano que constituiria um delito, mas ela também deve ter a intenção de causar esse dano - ou pelo menos ter algum tipo de estado mental injusto. Culpa e retribuição são motivos fundacionais para a punição criminal. A primeira se refere ao quanto as pessoas são moralmente culpadas por suas atividades, e a segunda se preocupa em punir as pessoas porque elas merecem, e não porque isso resulta em um bom resultado social. Uma IA pode se envolver nos mesmos tipos de comportamentos antissociais que a lei criminal condena, mas não está claro que uma IA possa alguma vez ser culpada porque apenas executa a programação. Uma preocupação ainda mais fundamental, de natureza semelhante, é a exigência de que um crime envolva um ato voluntário. Como a IA não é consciente, não está claro que ela seja capaz de realizar um ato - apenas de causar danos físicos. Um furacão não pode realizar um ato, mas não a extensão dos danos que pode causar é imensa.

Intuitivamente, isto faz com que a ideia de punir a IA pareça incoerente. Mas, como no caso dos impostos sobre robôs, a punição criminal da IA não é tão ridícula quanto possa parecer. A lei já pune criminalmente

peças artificiais sob a forma de empresas. Mesmo que não possuam literalmente estados mentais, as corporações podem enfrentar acusações diretamente quando seus procedimentos defeituosos causam danos, particularmente quando problemas estruturais em sistemas e processos corporativos são difíceis de se reduzir às ações errôneas de indivíduos. A lei pune criminalmente as ofensas de responsabilidade estrita, atos que não exigem qualquer estado mental injusto, como a intenção de causar danos. A punição pode até mesmo ser imposta em caso de ausência de ação. Em suma, punir uma pessoa artificial por não agir, mesmo sem evidência de intenção danosa, não é algo que possa ser descartado. O direito penal pode - e, quando empresas estão envolvidas, já o faz - apelar para a elaboração de ficções legais para fornecer uma base para punir algumas entidades artificiais.

Uma IA não é uma empresa, e sob as atuais estruturas legais ela não pode ser responsabilizada criminalmente, mas as leis podem ser alteradas. Em 2017, o Reino da Arábia Saudita anunciou que havia dado a cidadania a um robô chamado Sophia, fabricado pela Hanson Robotics. Embora esta tenha sido provavelmente mais uma manobra publicitária do que um ato de boa-fé, não há nenhum princípio legal imutável que proíba a cidadania a um robô. No Reino Unido, por exemplo, o Parlamento tem a soberania para aprovar qualquer legislação que queira, e não está vinculado a uma constituição escrita. De fato, a lei inglesa costumava incorporar a punição a objetos inanimados. No século XI, se bens pessoais causavam a morte de uma pessoa, a propriedade era dada como *deodand*, confiscada como uma coisa maldita e dada a Deus. O remédio da *deodand* não foi formalmente abolido até um ato do parlamento em 1846.

Outras jurisdições têm se mostrado ainda mais flexíveis. A lei indiana reconhece que animais, rios e até entidades divinas podem ter personalidade jurídica, e a Bolívia concedeu direitos à “Mãe Terra” em 2010¹³. Mas só porque as leis podem mudar, não significa que devam mudar. Mudanças legais podem acarretar custos significativos, e se feitas de maneira inadequada podem minar o Estado de Direito e a confiança no sistema jurídico. Para responder à questão de se a IA deve ser criminalmente responsável, é necessário um exame sério dos custos e benefícios

da punição da IA, incluindo se os compromissos doutrinários e teóricos do direito penal são consistentes com a imposição de condenações criminais à AI. Punir a IA poderia gerar uma dissuasão geral: desencorajar outros potenciais infratores de cometer crimes. A perspectiva de punição por crimes gerados por IA não deteria diretamente uma IA assim como os direitos de propriedade intelectual para obras geradas por IA não motivariam uma IA. O objetivo seria impactar o comportamento dos desenvolvedores, proprietários ou usuários de IA. Isto poderia ocorrer se a punição envolvesse confisco ou destruição de uma valiosa IA ou penalidades financeiras dirigidas aos seus proprietários. A punição da IA também poderia beneficiar psicologicamente as vítimas de crimes gerados por IA que vissem o Estado afirmar seus direitos e punir a entidade que lhes causou danos. Isso garantiria aos cidadãos que a atividade criminosa, mesmo por IA, não seria tolerada na sociedade.

Entretanto, a punição da IA não deve violar os princípios profundamente defendidos do direito penal, tais como a exigência de um ato voluntário e a capacidade de culpabilidade. Uma solução é a abordagem que tem sido aplicada para permitir a punição corporativa. As empresas não têm estados mentais, mas a lei permite que estados mentais possuídos por agentes humanos de uma empresa sejam imputados à ela. Assim, se os diretores da empresa optarem por um esquema ilegal de fixação de preços, a empresa é considerada como tendo a intenção de se engajar em uma ação ilegal. No caso da IA, a lei poderia de forma semelhante imputar estados mentais dos proprietários, usuários ou desenvolvedores de IA, embora possa ser mais difícil fazer isso para a IA do que para uma empresa, particularmente no caso de um crime gerado pela IA onde não há um suprimento pronto de indivíduos culpáveis. Uma empresa, ao contrário da IA, é composta de pessoas.

Uma opção diferente seria notar que o direito penal nem sempre exige culpabilidade. Por exemplo, os crimes de responsabilidade estrita não exigem um estado mental particular. Dependendo da jurisdição, vender álcool a um menor pode ser um crime, independentemente de alguém ter acreditado razoavelmente que era um adulto. Os crimes de responsabilidade estrita são desaprovados porque a sociedade não quer

punir pessoas que tenham agido sem culpa moral; isto os trataria como um meio para atingir um fim sem respeitá-los como indivíduos. A mesma restrição não se aplica à IA porque ela não encara a punição de uma maneira negativa ou possui direitos humanos. Entretanto, isto não supera o fato de a IA não cumprir com a exigência de um ato voluntário.

Mais ambiciosamente, a lei poderia permitir que a IA seja de fato capaz de agir e que sua tomada de decisão envolva algo suficientemente análogo a um estado mental humano. Funcionalmente, a IA pode adquirir e processar informações, engajar-se na lógica e na razão para determinar o melhor meio de atingir um objetivo e agir sobre o mundo de forma a aumentar a probabilidade de que esse objetivo ocorra. Os filósofos podem debater se esses comportamentos contam como atos genuínos ou estados mentais, mas de qualquer forma a lei poderia tratá-los como tal.

Como uma questão prática, pode ser difícil raciocinar sobre o que uma máquina estava pensando, mas os jurados muitas vezes não conhecem diretamente o estado mental de um réu humano e inferem o que ele estava pensando com base em seu comportamento. Os jurados poderiam fazer inferências semelhantes sobre o conhecimento, intenção e objetivos de uma IA com base em seu comportamento. Por exemplo, se um carro que dirige sozinho atropela alguém, pode ser considerado como tendo a intenção de causar danos se ele mudar repetidamente seu rumo para atingir um pedestre em movimento.

Punir a IA vem com custos, mesmo que não viole nenhum princípio fundamental do direito penal. Os custos viriam na forma de mudanças legais significativas e um nível de disrupção que deveria ser evitado sem uma boa causa, e punir a IA poderia enviar uma mensagem preocupante de que a ela está moralmente em pé de igualdade com as pessoas. Direitos e obrigações muitas vezes andam de mãos dadas, e a punição poderia reforçar o ponto de vista de que a IA merece direitos. A sociedade já viu isso antes com os direitos concedidos às empresas nos Estados Unidos aumentando gradualmente ao longo do tempo: Uma IA elegível para punição hoje poderia ser uma IA elegível para votar amanhã.

Estão disponíveis melhores respostas para a perspectiva de crimes gerados por IA. Uma opção seria expandir as penalidades criminais ou civis

dirigidas às pessoas. Novos deveres legais poderiam ser criados para desenvolver responsabilmente, supervisionar ou permanecer imputável por uma IA, com a responsabilidade sendo atribuída pelo não cumprimento de tais deveres. Isso se qualificaria como responsabilidade baseada na conduta humana e não na responsabilidade pela conduta prejudicial da própria IA. Punir diretamente proprietários, usuários e desenvolvedores de IA seria provavelmente uma maneira mais eficaz de influenciar seu comportamento do que indiretamente através da punição da IA. Expandir a responsabilidade civil ao invés da criminal pode ser uma resposta melhor, porque fazer o segundo poderia desestimular demais atividades como o desenvolvimento da IA, o que gera benefícios sociais, especialmente porque o crime gerado pela IA ainda não tem sido um problema significativo. Portanto, a punição da IA deve ser evitada - não porque seja incompatível com a lei criminal, mas simplesmente porque é uma má ideia.

6 O FUTURO DA IA

A IA considerada no contexto do direito tributário, responsabilidade civil, propriedade intelectual e direito penal fornece insights sobre como a IA afetará os padrões legais existentes e como os padrões legais moldarão seu desenvolvimento. A IA promete ser altamente disruptiva - e se a história for um guia - de formas inesperadas. Talvez, em retrospectiva, a IA provará ter sido meramente parte de outra revolução industrial. Entretanto, nosso sistema jurídico não tem feito historicamente o melhor trabalho de limitar os danos causados pela disrupção tecnológica. Uma abordagem diferente de nossas estruturas legais poderia ajudar a otimizar os benefícios sociais da IA. Os princípios usados para desenvolver um sistema legal precisam ser, se não repensados, redesignados com respeito à IA antes que os eventos ultrapassem a sociedade.

Tais eventos estão bem encaminhados. Considere mais uma vez a IA e os jogos de tabuleiro, um tópico com lições importantes para pensar sobre o futuro da IA. Em primeiro lugar, ele nos ensina que a IA é uma tecnologia excepcional - algo que pode agir como uma pessoa independentemente de seus criadores. Em 1997, foi o Deep Blue da IBM que derrotou o campeão mundial Garry Kasparov no xadrez, não os progra-

madores do Deep Blue que, mesmo que tivessem jogado contra ele como uma equipe, não teriam tido chance de vencer. Ao invés disso, os programadores criaram uma entidade autônoma que se engajava em uma atividade além de suas próprias capacidades. Naquela época, o Deep Blue era um dos mais poderosos supercomputadores já construídos, capaz de avaliar 200 milhões de posições de xadrez por segundo. Hoje em dia, os programas de xadrez executados em smartphones podem vencer os melhores jogadores humanos do mundo.

A IA e o xadrez também apontam para um futuro próximo no qual o melhoramento humano através da IA é mais importante do que a automação. Após sua partida com o Deep Blue, Kasparov teve uma realização: Uma pessoa e uma IA podem jogar xadrez em colaboração e complementar-se mutuamente. Em 1998, ele venceu o primeiro torneio de xadrez “centauro”, onde um jogador humano e uma inteligência artificial jogam em equipe. Não é surpreendente que uma pessoa auxiliada por uma IA tenha se mostrado melhor do que alguém jogando sem assistência. Mas uma pessoa e uma IA também superaram uma IA jogando sozinha. Os grandes mestres são bons em estratégia de xadrez de longo prazo, mas pobres em calcular rapidamente milhões de jogadas possíveis. O inverso é verdadeiro para a IA jogando xadrez. Como as pessoas e a IA são fortes em diferentes dimensões, elas podem se sair melhor trabalhando juntas do que independentemente.

A lição final é a seguinte: A automação é provavelmente inevitável. Em 2017, no mesmo ano em que o AlphaGo venceu o melhor jogador humano do mundo, o programa de xadrez Cryptic venceu a melhor equipe composta de uma IA e um humano¹⁴. Eventualmente, as pessoas podem simplesmente atrapalhar.

REFERÊNCIAS

CARDOZO, Benjamin Nathan. **Selected Writings of Benjamin Nathan Cardozo** 417 (Margaret E. Hall ed., 1947).

Gear 2030, Final Report of the High-Level Group on the Competitiveness and Sustainable Growth of the Automotive Industry in the European Union, at 40 (July 2017).

Law No. 071, **Ley de Derechos de la Madre Tierra**, Diciembre 21, 2010, LA ASAMBLEA LEGISLATIVA PLURINACIONAL (Bolivia), www.planificacion.gob.bo/uploads/marco-legal/Ley%20N%C2%B0%20071%20DERECHOS%20DE%20LA%20MADRE%20TIERRA.pdf.

LEE, Timothy B. Autopilot Was Active When a Tesla Crashed into a Truck, Killing Driver, **Ars Technica**, May 16, 2019, <https://arstechnica.com/cars/2019/05/feds-autopilot-was-active-during-deadly-march-tesla-crash/?comments=1&post=37374819>.

Ministry of Foreign Affairs of Japan, **G20 Ministerial Statement on Trade and Digital Economy** (2019), www.mofa.go.jp/files/000486596.pdf.

Organisation for Economic Co-Operation and Development, **Principles on Artificial Intelligence** (2019), www.oecd.org/going-digital/ai/principles/.

KRANTZ, Matt. PepsiCo Paid No Tax? Neither Did These Other 33 Profitable S&P 500 Companies, **Investor's Business Daily**, July 18, 2019, www.investors.com/etfs-and-funds/personal-finance/corporate-tax-rate-zero-profitable-us-companies-sp500/.

RUBIN, Richard. Does Amazon Really Pay No Taxes? Here's the Complicated Answer, **The Wall Street Journal**, June 14, 2019, www.wsj.com/articles/does-amazon-really-pay-no-taxes-heres-the-complicated-answer-11560504602.

STRICKLAND, Eliza. How IBM Watson Overpromised and Underdelivered on AI Health Care, **IEEE SPECTRUM**, Apr. 2, 2019, <https://spectrum.ieee.org/bio-medical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care>.

WEIBEL Beat. AI Created Inventions – Digital Inventor Computer-Implemented Simulations – Digital Twin, **WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI)**, Sept. 30, 2019, www.wipo.int/meetings/en/doc_details.jsp?doc_id=454861.

WRZESZCZYNSKI, Kazimierz O. et al. Comparing Sequencing Assays and Human-Machine Analyses in Actionable Genomics for Glioblastoma, 3 **NEUROL. GENET.** (2017), DOI: 10.1212/NXG.000000000000164.

Ex parte Stephen L. Thaler, No. BL O/741/19 (U.K. I.P.O. Dec. 4, 2019).

Veja www.artificialinventor.com para o status atual desses pedidos.

E-mails com Arno Nickel, Gerente Geral do InfinityChess (Nov. 7 & 10, 2019) (de posse do autor).

Recebido: 12/11/2020

Aprovado: 05/01/2021