

**INTELIGÊNCIA ARTIFICIAL: PERSPECTIVAS A PARTIR DA FILOSOFIA
DO DIREITO**

Marco Almada¹

Resumo: Tecnologias baseadas em aprendizado de máquina e inteligência artificial ocupam cada vez mais espaço nas sociedades contemporâneas, alterando dinâmicas sociais já existentes ou mesmo introduzindo novos modos de interação interpessoal. Estas mudanças se tornam relevantes para a filosofia jurídica conforme afetam as relações sociais tuteladas pelo Direito ou o funcionamento dos sistemas jurídicos, o que pode ocorrer com a automação de atividades hoje desempenhadas por juízes, advogados e outros atores jurídicos. Mas, ao mesmo tempo que a inteligência artificial traz novas questões para a filosofia jurídica, também pode contribuir com novas abordagens no estudo de relações sociais complexas e de seus desdobramentos jusfilosóficos. Neste artigo, proponho questões e abordagens para que o diálogo entre a Filosofia do Direito e a pesquisa em Inteligência Artificial ocorra de forma produtiva para ambos os campos e para as funções sociais do Direito.

Palavras-chave: Filosofia do Direito; Direito da Inteligência Artificial; Inteligência Artificial no Direito.

Introdução

O que a filosofia jurídica pode nos dizer sobre a inteligência artificial? Em um primeiro momento, podemos ter a impressão de que não há motivos para estabelecer um diálogo entre os dois campos, uma vez que eles lidam com objetos bem diferentes entre si: de um lado, temos a filosofia jurídica como o campo dedicado a investigar as questões filosóficas relacionadas ao direito², objeto que é tido como uma das mais complexas criações³ das sociedades humanas; do outro, temos um campo multidisciplinar que, embora conectado com a filosofia já desde seus momentos iniciais (TURING, 1950), tem como objeto central não um sistema social, mas sim sistemas

¹Mestre em Engenharia de Computação pela Universidade Estadual de Campinas (Unicamp). Graduando em Direito pela Universidade de São Paulo (USP). Pesquisador no Lawgorithm. E-mail: marco.almada@usp.br.

²Não é meu intuito, aqui, delimitar o objeto próprio da filosofia jurídica, então adoto uma visão ampla, que engloba tanto as questões a respeito da natureza do direito quanto questões filosóficas relativas aos ramos do sistema jurídico, dentre outras reflexões que se façam relevantes.

³ Como salienta MARMOR (2012 xix): “Law is one of the most complex, intricate and sophisticated creations of human societies”. Ainda que se possa questionar a ideia de que toda forma de direito é um artefato cuja criação é resultado de um projeto consciente, o funcionamento dos sistemas jurídicos modernos ainda pode ser visto como uma forma de tradição (PRIEL, 2018), que também envolve um processo criativo, mesmo que mais difuso.

computacionais capazes de tomar decisões adequadas diante de situações concretas (RUSSELL; NORVIG, 2010, p. 30).

Estes sistemas computacionais, contudo, passam a ser relevantes para o direito na medida em que carros autônomos (LOPES; ROBERTO, 2018), assistentes pessoais digitais, classificadores de resultados de acórdãos de tribunais (OLIVEIRA; WASSERMANN, 2017), dentre outras soluções, ganham espaço na sociedade. Com isso, a inteligência artificial cada vez mais se torna relevante tanto para o entendimento de relações sociais tuteladas pelo direito quanto para o próprio funcionamento de um sistema jurídico, através de automação de tarefas mecânicas — como a construção de peças repetitivas — ou, no futuro, mais complexas.

Neste artigo, sustento que é do interesse dos estudiosos da filosofia jurídica tomar contato com os desenvolvimentos da inteligência artificial (IA). Para isso, após uma breve discussão a respeito dos objetivos dos usos da IA, discuto duas versões desta tese: uma mais fraca e uma mais forte. A versão fraca de minha tese é a de que o contato com a IA pode trazer mudanças *metodológicas* na filosofia do direito, ao permitir uma expansão do alcance das reflexões jusfilosóficas. Neste caso, a mudança em função da IA não seria obrigatória, ao contrário do que vemos na versão mais forte, segundo a qual a inteligência artificial pode trazer mudanças nos *objetos* típicos da filosofia do direito, conforme as características dos sistemas inteligentes se tornam elementos relevantes para a confecção e aplicação das normas jurídicas, processos esses que, por sua vez, também podem estar sujeitos à automação.

Diante destas possibilidades, torna-se relevante entender os impactos jurídicos da inteligência artificial, seja para coibir seus excessos, seja para aproveitar o potencial de transformação por meio destas tecnologias, seja em uma dimensão puramente descritiva das mudanças legais. Estabelecida a pertinência dos estudos sobre a inteligência artificial, concluo com algumas sugestões para que os filósofos jurídicos, sozinhos ou em equipes multidisciplinares, abordem os temas ligados à inteligência artificial que sejam relevantes para seus tópicos de interesse.

Inteligência artificial e o direito

Antes de começarmos a discussão a respeito da inteligência artificial, é necessário entender a que nos referimos ao usar o termo. Uma definição possível —

presente tanto no imaginário popular quanto em tratamentos acadêmicos como em Turing (1950), artigo tido como um dos momentos fundadores do campo — é a de que a inteligência artificial é uma disciplina que pretende construir sistemas capazes de replicar o pensamento humano ou condutas humanas.

Esta visão a respeito do objetivo da inteligência artificial se faz presente em vários dos debates atuais a respeito do tema, como a busca por meios de evitar que uma inteligência que exceda as capacidades humanas fuja ao controle (TRAZZI; YAMPOLSKIY, 2018), ou as propostas de atribuição de direitos e obrigações jurídicas a robôs (BRYSON; DIAMANTIS; GRANT, 2017).

Contudo, muitas das tecnologias hoje associadas à inteligência artificial não são desenvolvidas com tal intuito mimético. Em vez disso, desenvolvedores de sistemas como carros automáticos e ferramentas de tradução (RUSSELL; NORVIG, 2010, p. 28–29) têm como objetivo criar sistemas que resolvem um determinado problema, mesmo que isso não seja feito a partir de um modelo da forma como um humano resolveria o mesmo problema. Assim, para os fins da presente discussão, considerarei que a inteligência artificial é um campo cujo objetivo é a construção de *agentes racionais*, isto é, sistemas computacionais que tomam ações de maneira racional para a realização de um objetivo.

A racionalidade destes agentes é algo essencialmente instrumental, uma vez que é voltada a permitir que a IA gere o melhor resultado esperado para uma ação diante das circunstâncias concretas, que podem envolver incertezas ou limitações à sua capacidade computacional (RUSSELL; NORVIG, 2010, p. 4–5). Por exemplo, um carro autônomo terá acesso a informações sobre o destino do passageiro, sobre seus arredores — através de sensores — e sobre o trânsito da região, e com base nisso determinará o melhor caminho para chegar a tal destino, dadas condições de contorno (como as leis de trânsito). Uma vez que as ações do sistema podem afetar de maneira direta ou indireta os direitos ou interesses de pessoas naturais ou jurídicas — por exemplo, no caso de um atropelamento (LOPES; ROBERTO, 2018) —, o uso da IA pode surgir como um fator na aplicação das normas jurídicas existentes, o que leva a uma preocupação com a regulação destes sistemas.

A perspectiva regulatória, que Maranhão (2017b) denomina Direito da IA, é complementada pela possibilidade de aplicação de sistemas inteligentes a domínios jurídicos. Como exemplo, temos o uso de técnicas de processamento computacional de

linguagens naturais em apoio a leituras ditas *originalistas* da constituição americana (STRANG, 2016). Uma vez que as versões contemporâneas do originalismo buscam sua justificação no sentido convencional originalmente atribuído ao texto constitucional (STRANG, 2016, p. 1197), os defensores destas correntes buscam meios para identificar acuradamente tal sentido, e, para isso, *corpora* de textos da época da elaboração da constituição americana (STRANG, 2016, p. 1210) são processados por algoritmos em busca do sentido corrente dos termos juridicamente relevantes.

Além do uso do processamento de linguagem natural, há também o emprego de técnicas de aprendizado de máquina na tentativa de *predição* de decisões judiciais, em uma releitura da visão preditiva de direito defendida por Oliver Wendell Holmes (FELIN *et al.*, 2017), ou para a identificação automática de argumentos presentes em peças judiciais, dentre outros usos possíveis (MARANHÃO, 2017b). A seguir, destaco uma aplicação de técnicas computacionais que pode se beneficiar de técnicas de inteligência artificial: a simulação de fenômenos jurídicos.

Simulações como uma ferramenta da filosofia jurídica

A delimitação precisa do que é uma simulação computacional ainda é uma questão em aberto na filosofia da ciência (SAAM, 2017). Porém, de um ponto de vista operacional, podemos entender uma simulação como um sistema computacional que representa o comportamento de um sistema-alvo ao longo do tempo (HUMPHREYS, 2004). Nas ciências naturais e engenharias, estas simulações são usadas para testar hipóteses sem que seja necessário construir um experimento físico (WEISBERG, 2012), e este uso, por sua vez, serviu de ponto de partida para aplicações em ciências sociais como a economia (ALMADA; ATTUX, 2018).

Os usos de simulações em ciências sociais podem fornecer apoios a argumentos jusfilosóficos, da mesma forma que outras pesquisas nestas ciências. Há, no entanto, um interesse mais direto: simulações podem servir como uma forma particular de *experimento mental* (SAAM, 2017), um método já adotado na teoria política analítica (BROWNLEE; STEMPOWSKA, 2017) e também na filosofia do direito — por exemplo, no argumento do “homem mau” proposto por Holmes (1997) como parte de sua justificativa para definir a previsão das decisões judiciais como objetivo do direito.

Um *experimento mental*⁴ é uma narrativa de um caso excepcional de uma teoria (GENDLER, 2000, xi), a partir da qual é possível identificar contradições ou consequências não-intuitivas de uma teoria, dentre outros usos (BROWNLEE; STEMPOWSKA, 2017, p. 21). Para um filósofo do direito que considere tais narrativas como uma técnica válida, as simulações computacionais podem ser também uma abordagem interessante, uma vez que elas podem ser entendidas como um tipo particular de experimento mental formalizado (SAAM, 2017). Mas, nesse caso, por que adotar as simulações, dado que seu uso adequado exige uma certa familiaridade com ferramentas computacionais que não são parte do repertório típico do pesquisador interessado em questões jurídicas?

A primeira vantagem das simulações estaria na *escala* dos modelos que podemos construir. Experimentos mentais são dependentes das operações que o cérebro humano pode realizar, e as simulações computacionais podem atuar como ferramentas de extensão (HUMPHREYS, 2004) destes limites. Como um exemplo, a literatura de neurociências afirma que um ser humano consegue acompanhar simultaneamente em torno de 7 objetos (TRAZZI; YAMPOLSKIY, 2018), mas um modelo baseado em agentes pode ter dezenas de agentes (ou muito mais, como ocorre em jogos online massivos) interagindo ao mesmo tempo, o que pode ser útil, por exemplo, em estudos sobre tolerância e diversidade como o proposto por Muldoon, Borgida e Cuffaro (2012). Nessa mesma linha, uma simulação pode ser utilizada para reproduzir os acontecimentos de séculos em um tempo fictício que transcorre de maneira mais rápida e mais inteligível que o tempo de parede, facilitando assim a investigação de efeitos de longo prazo.

Além disso, a simulação também pode contribuir para o exame das premissas e intuições envolvidas na teoria simulada, permitindo que estas sejam substituídas ou removidas; com isso, seria possível verificar se uma explicação proposta tem o alcance desejado ou se, como muitas vezes apontam críticos da abordagem analítica (MONTEIRO BRITO JÚNIOR, 2018), ela é fruto de intuições culturalmente enraizadas que não seriam generalizáveis.

A construção de simulações computacionais pode ser um processo demorado e complexo, o que torna seu uso pouco interessante para muitos dos cenários em que hoje

⁴ Também conhecido como *experimento de pensamento*, a partir do termo alemão *Gedankenexperiment*, cunhado por Ernst Mach (1838–1916) (SORENSEN, 1992, p. 51).

os experimentos mentais são empregados. Porém, diante das vantagens acima postas, o uso das simulações pode ser interessante para filósofos que estejam interessados, por exemplo, no estudo de teorias da justiça em populações marcadas por grande diversidade (MULDOON, 2016).

Inteligência Artificial e as relações jurídicas

Enquanto ferramenta auxiliar, a IA desempenha um papel instrumental que é similar àquele desempenhado por disciplinas como a economia ou a sociologia, ao fornecer substratos para a reflexão jusfilosófica. Tais substratos podem auxiliar o trabalho do filósofo jurídico ou sugerir novas possibilidades de investigação, mas não seriam em princípio essenciais para o tratamento dos temas da filosofia do direito⁵. Mas, se é possível, com menor ou maior prejuízo, abrir mão destes usos da IA no direito, há uma outra interface que mais difícil de afastar: a inteligência artificial enquanto *objeto* de relações jurídicas. Um exemplo é a questão da personalidade jurídica robótica, discutida em 2017 pelo Parlamento Europeu.

Uma vez que o conceito de personalidade jurídica está dissociado da noção de *Homo sapiens* — o que se reflete, por exemplo, na existência da pessoa jurídica — parte da literatura sustenta que os robôs — que são sistemas inteligentes programáveis, dotados de flexibilidade para interagir com o ambiente externo e de capacidade mecânica para atuar neste — poderiam, no futuro, ser considerados capazes de, por conta própria, adquirir direitos e contrair obrigações jurídicas⁶.

Esta ideia foi criticada com base na sua incompatibilidade com a regulação europeia vigente, mas, de uma forma mais relevante para a presente discussão, ela também pode ser criticada de um ponto de vista conceitual. Um primeiro problema é salientado por Nevejans (2016): mesmo que os problemas de técnica jurídica na criação de uma personalidade jurídica robótica fossem sanados, isso pode trazer resultados socialmente indesejáveis, uma vez que muitos dos casos de uso em que estes robôs são críticos — como a limpeza de lixo tóxico ou o uso contínuo em produção fabril —

⁵Ainda que se possa, por exemplo, criticar a filosofia analítica do direito por não levar em conta diferenças etnográficas (MONTEIRO BRITO JÚNIOR, 2018), este tipo de crítica externa em geral é voltado à validade da abordagem teórica em si e não aos critérios internos que dão origem a um exemplar válido daquela abordagem.

⁶No contexto do direito brasileiro, uma primeira discussão nesse sentido foi proposta por Castro Júnior (2009).

expõem o robô a uma vida útil reduzida. Desta forma, o uso destes robôs seria incompatível com uma plena atribuição de direitos a eles, já que atentaria contra algo tipicamente visto como um direito básico: a continuidade da existência da personalidade (NEVEJANS, 2016, p. 15).

Mesmo uma atribuição mais limitada de direitos, ainda que possível (BRYSON; DIAMANTIS; GRANT, 2017, p. 281), pode se mostrar problemática. Na ausência de autonomia plena dos robôs, a atribuição de direitos ou de obrigações a eles pode ser, no fundo, uma forma de mascarar responsabilidades jurídicas que caberiam a seus controladores (BRYSON; DIAMANTIS; GRANT, 2017, p. 285–287): por exemplo, se um robô detém a propriedade jurídica de um bem, pode ser que o uso, a fruição e a disposição deste bem acabem sendo exercidos não pela vontade do titular da propriedade, mas do controlador humano. Assim, a atribuição de direitos — mesmo sem atingir uma personalidade jurídica plena — não seria mais que uma ficção jurídica, que pode ou não ser socialmente interessante.

Tal ficção pode, ao menos em princípio, servir como uma forma de proteger interesses válidos, como o desenvolvimento tecnológico. Por exemplo, atribuir responsabilidade civil ao robô e não ao seu desenvolvedor pode estimular a adoção de tecnologias robóticas ao diminuir o risco jurídico envolvido nesse processo. Contudo, Bryson, Diamantis e Grant (2017, p. 288) sustentam que as dificuldades envolvidas na operacionalização da personalidade jurídica robótica podem dar margem a cenários em que as pessoas naturais eventualmente lesadas por um robô não tenham meios adequados para buscar reparação por vias jurídicas. Com base nisso, Bryson, Diamantis e Grant (2017) sustentam que a atribuição de personalidade jurídica a robôs não seria desejável, ainda que reconheçam a possibilidade em abstrato de que os benefícios dessa escolha possam vir a superar os seus eventuais impactos negativos.

A determinação dos ônus e dos benefícios envolvidos na personalidade jurídica robótica dependerá não só das capacidades que a tecnologia permite aos robôs, mas também dos arranjos institucionais e sociais que se estabeleçam em torno deles. Neste cenário, o filósofo jurídico pode contribuir com o debate por meio do reexame dos conceitos em disputa, o que exigirá um diálogo não só com a inteligência artificial, mas também com campos filosóficos como a ética, para determinar a pertinência ou não de alguma forma de personalidade robótica, bem como os fundamentos de uma eventual atribuição de direitos e obrigações.

Vinculação de agentes inteligentes a normas jurídicas

Outra dimensão da discussão a respeito da personalidade jurídica de robôs está ligada à *normatividade* do direito (MARMOR, 2012, p. 11): como podemos entender a ideia de que uma norma impõe um dever a um robô? Uma primeira tentativa de resposta a esta questão seria considerar que a obediência de um robô a uma norma é resultado necessário de sua programação. Sob esta perspectiva, o robô obedeceria às leis porque foi programado para isso, e eventual violação de normas seria fruto de um problema de construção.

A ideia de que o comportamento dos robôs é vinculado por normas *hard-coded* em sua programação encontrou uma expressão popular nas Leis da Robótica propostas por Isaac Asimov (NEVEJANS, 2016, p. 12–13), mas é insuficiente para descrever por que um agente inteligente obedece ou não a normas, em particular normas jurídicas. Estes agentes, como salienta a definição da Seção 2, são construídos para tomar decisões racionais para a realização de seu objetivo, mas é possível que, em uma situação concreta, as normas a que um agente está sujeito levem a obrigações conflitantes. Por exemplo, um carro autônomo deve ter entre seus comandos uma instrução para preservar a vida de seu motorista; mas, se a única opção para evitar o atropelamento de uma criança for forçar uma colisão com um muro que pode levar à morte do passageiro, este comando entra em conflito com a preservação da vida do pedestre.

Reconhecendo a possibilidade de conflitos normativos, tecnologias modernas de agentes inteligentes (ELSENBROICH; GILBERT, 2013, p. 72) adotam conjuntos *derrotáveis* de regras, em que a aplicação de uma ou outra regra pode ser afastada no caso concreto pela presença de outra. O agente inteligente considerará, pois, as circunstâncias concretas e os seus objetivos para determinar se uma norma jurídica que ele conheça deve ser seguida ou não, e, nesse sentido, dizer que um sistema inteligente S está obrigado a uma conduta ϕ nas circunstâncias C não significa algo essencialmente diferente do que significaria dizer que um humano reconhece a obrigação nas mesmas circunstâncias.

A aplicação de normas a sistemas inteligentes, contudo, traz novas questões em ao menos dois aspectos da normatividade. O primeiro destes é em relação à dimensão *psicológica* das normas: a coerção, muitas vezes tida como uma das características

centrais do direito, é um fenômeno ligado às vontades das partes envolvidas (EDMUNDSON, 2012, p. 464). Porém, ao menos no horizonte atual da tecnologia, não é possível dizer que um sistema inteligente tenha uma vida psicológica interior, o que prejudica a aplicação de coerção a sistemas inteligentes.

Tais sistemas, no entanto, não são inteiramente imunes às consequências jurídicas de seus atos, uma vez que a sanção jurídica pode afetá-los de outras formas. Um robô criado para obter lucros através de operações financeiras pode receber uma multa por ações indevidas, o que levaria sua programação a reconhecer que aquele curso de ação pode trazer efeitos indesejados. Não é impossível, portanto, desenhar efeitos jurídicos que possam afetar a conduta de agentes inteligentes, e o filósofo jurídico pode contribuir com a conceituação destes efeitos e a extensão da noção de coerção para além de suas bases antropocêntricas atuais.

Mas, se a possibilidade de avaliação dos efeitos de uma norma pode permitir que o agente aja de maneira compatível com as *regras* a que ele está sujeito⁷, a consideração de *princípios* pode apresentar complicações maiores para um agente inteligente. Na formulação clássica de Dworkin, os princípios são normas que se distinguem das regras por não serem aplicadas em um modo de *tudo ou nada* (ÁVILA, 2016, p. 56), em que uma colisão entre duas regras é resolvida pelo descarte de uma dessas. Em vez disso, os princípios contêm fundamentos para uma decisão, cujo conteúdo será determinado através da conjugação com os fundamentos fornecidos por outros princípios (ÁVILA, 2016, p. 57).

A aplicação de princípios exige, portanto, que o agente seja capaz de avaliar estes fundamentos e realizar operações de combinação entre eles. Se a valoração destes princípios necessariamente envolve a tomada de uma posição moral, como sustenta Dworkin (MARANHÃO, 2012, p. 57), então os sistemas inteligentes, por melhor que seja seu tratamento das regras jurídicas, não poderão, ao menos no futuro imediato, lidar de maneira adequada com princípios morais. Se, por outro lado, a conjugação de princípios pode ser produzida por operações lógicas formais — por exemplo, a partir de derivação abdutiva (MARANHÃO, 2012, p. 135–141) —, passa a ser possível construir

⁷ O impacto dos efeitos externos é mais visível na consideração *regras primárias*, isto é, aquelas que prescrevem a realização ou a não-realização de certos atos (NINO, 2010, p. 105). No caso das *regras secundárias*, que têm como objeto as regras primárias, o descumprimento não tem como efeito esperado uma sanção, mas ainda assim pode ser possível representar estes efeitos — por exemplo, a nulidade de uma norma emitida pelo agente automatizado — de uma forma que ele possa incorporar o efeito esperado ao decidir seu curso de ação.

representações lógico-formais de princípios jurídicos (PASSOS, 2016, p. 65–66), que por sua vez podem servir de bases para arquiteturas lógicas (MARANHÃO, 2017a) que produzam uma representação adequada do processo de interpretação jurídica. Caberia, pois, aos filósofos jurídicos tanto a identificação dos critérios segundo os quais a interpretação de princípios por agentes inteligentes seria ou não possível e, caso seja, participar da construção dos mecanismos formais que podem, então, ser implementados como parte de sistemas inteligentes.

IA e a aplicação de normas jurídicas

Sistemas inteligentes podem ser não apenas objeto de leis e regulações, mas também atuar diretamente na produção e aplicação de normas jurídicas. Como um exemplo, temos as ferramentas de *smartcontracts* empregadas para a gestão e execução automatizada de contratos; estes sistemas atualmente estão limitados a instruções procedimentais, mas o emprego de técnicas de inteligência artificial pode viabilizar que estes sistemas identifiquem de maneira mais ativa as posições negociais das partes em contratos cada vez mais complexos (MARANHÃO, 2017b), para a identificação das obrigações que competem a todos os envolvidos.

A difusão da automação de contratos no setor privado pode servir como um estímulo para que tarefas jurídicas desempenhadas pelo setor público sejam automatizadas. Num primeiro momento, surge a possibilidade, já mencionada na Seção 1, de automação de tarefas auxiliares, como a busca por jurisprudência e a classificação de decisões judiciais. Mas, com o refino das técnicas de inteligência artificial, é possível cogitar a automação de atividades mais diretamente relacionadas ao funcionamento do sistema jurídico.

Uma primeira possibilidade surge da substituição de humanos por sistemas inteligentes em papéis centrais no sistema judicial. Brennan-Marquez e Henderson (2018, p. 10–12, 19–24) mostram como muitos dos argumentos contrários à adoção de máquinas como jurados em casos criminais são circunstanciais. Em particular, as críticas baseadas na dificuldade de controlar os vieses envolvidos nas decisões automatizadas podem logo deixar de ser aplicáveis, conforme o avanço das técnicas de explicação do funcionamento de sistemas inteligentes permita a construção de máquinas que evitam muitos dos vieses presentes no julgamento humano.

Uma vez que as máquinas se mostrem capazes de tratar os réus sem incorrer nos vieses comuns entre jurados humanos⁸, é possível imaginar que isto sirva de argumento para a remoção de humanos do processo decisório. Para além dos problemas jurídicos que isso pode trazer⁹, Brennan-Marquez e Henderson (2018, p. 12–19) levantam uma outra forma de objeção: uma vez que, num Estado Democrático de Direito, a pessoa que aplica a norma não pode, em princípio, estar também fora do alcance dela, a adoção de jurado-robôs só seria legítima quando as máquinas pudessem também ser réus, uma vez que a possibilidade de inversão de papéis permitiria uma apreciação do aspecto interno do julgamento¹⁰.

A automação judicial pode contribuir não só para a redução de vieses, mas também para a redução dos custos financeiros e humanos envolvidos no funcionamento de um sistema jurídico. Tal economia, por sua vez, pode servir como ponto de partida para novas formas de regulação das condutas humanas, como as “micro-diretivas” que CASEY; NIBLETT (2016) apontam como o futuro do direito. Estas diretivas indicariam, em tempo real, qual é a conduta esperada do indivíduo diante de uma situação concreta e seriam autoexecutáveis, uma vez que o mesmo sistema que gerou a norma hiper-individualizada teria as informações necessárias para julgar se houve violação do comando e se é pertinente a aplicação de alguma sanção jurídica.

Mesmo que este cenário esteja longe do estado da arte na inteligência artificial, os desenvolvimentos tecnológicos nesse sentido trazem questões relativas à *validade* e à *legitimidade* destes comandos individuais: até que ponto a criação pontual e instantânea das normas pode vincular os sujeitos de direito? Quais as bases para a criação de normas hiper-individualizadas? Não seria a autoexecução das normas incompatível com as garantias processuais típicas das principais tradições jurídicas, como o duplo grau de jurisdição? A possibilidade de automação em um sistema jurídico levanta, portanto, questões metajurídicas relevantes tanto no curto quanto no longo prazo.

Impactos da IA na filosofia jurídica

⁸ A respeito dos vieses de jurados humanos, ver, por exemplo, SUNDQUIST (2018).

⁹ Por exemplo, tanto o artigo 22 da GDPR — regulação de proteção de dados da União Europeia —, quanto o artigo 20 da Lei Geral de Proteção de Dados Pessoais hoje restringem as possibilidades de tomada de decisão automatizada, bem como estabelecem a possibilidade de intervenção humana neste tipo de processo decisório.

¹⁰ Para que isso ocorra, é necessária a possibilidade de sancionamento discutida na Subseção 3.1, mas ela pode não ser suficiente. Por questões de espaço, não tratarei deste tema aqui.

A adoção de sistemas inteligentes, como discutido nas seções anteriores, promete trazer mudanças tanto nos objetos das relações sociais tuteladas pelo direito quanto nas ferramentas utilizadas para o funcionamento e o estudo dos sistemas jurídicos. Estas mudanças, por sua vez, levantam novas questões para o filósofo interessado em questões jurídicas, ou ao menos novos ângulos para exame, ao mesmo tempo em que traz ferramentas potencialmente úteis para a prática filosófica. Como, então, o filósofo jurídico pode lidar com estas mudanças?

Uma primeira reação possível é a indiferença, já que os problemas centrais do campo continuam a existir, ainda que com novas roupagens ou, em alguns casos, novas dimensões. Mesmo as vantagens que a adoção de métodos apoiados em IA podem trazer não são imbatíveis, como mostra o exemplo da filosofia da ciência, em que abordagens “tradicionais” prosperam mesmo depois de décadas de trabalhos em linhas computacionais (THAGARD, 1993). Portanto, um filósofo jurídico que decida não se ocupar de questões específicas da inteligência artificial continuará a dispor de amplas possibilidades de investigação.

Já os filósofos que pretendam abordar questões trazidas pela inteligência artificial ou que sofram modificações diante da popularização de sistemas inteligentes precisarão, de alguma forma, se familiarizar com os aspectos técnicos e sociais destas tecnologias. Em uma perspectiva reducionista, esta familiarização pode ser realizada com o objetivo de traduzir os aspectos da IA que são salientes em um dado problema para a linguagem da tradição de filosofia jurídica em que o filósofo opera; feita essa operação de tradução, a discussão da IA se daria nos termos já estabelecidos. No outro extremo, temos a possibilidade de que um pesquisador enfatize as formas como a IA pode tornar inadequados os conceitos jurídicos e metajurídicos presentes em uma discussão, para, a partir disso, propor alterações que englobem as diferenças introduzidas pela inteligência artificial.

Na prática, muitas posições cairão em algum ponto entre estes dois extremos, ao identificar tanto continuidades quanto rupturas nas estruturas conceituais existentes. Para ambos os objetivos, contudo, o filósofo jurídico precisará se familiarizar com a literatura da disciplina de inteligência artificial. A leitura de manuais da área, como o livro de RUSSELL; NORVIG (2010) a respeito de inteligência artificial ou o guia de GILBERT; TROITZSCH (2005) sobre simulações em contextos de ciências sociais, pode ser útil tanto para o pesquisador que deseje tomar um contato direto com a

programação e construção de sistemas computacionais quanto para aquele que deseja apenas o contato com as noções técnicas fundamentais para o entendimento dos em questão.

O impacto da inteligência artificial, no entanto, não pode ser entendido apenas por uma perspectiva interna aos sistemas inteligentes, uma vez que o interesse da filosofia jurídica muitas vezes está na forma como eles impactam relações sociais tuteladas pelo direito. Para esta perspectiva interna, pode ser importante travar contato com as literaturas de estudos de ciência, tecnologia e sociedade, bem como da filosofia da tecnologia. Em particular, o livro editado por HILDEBRANDT; ROUVROY (2011) traz interfaces entre a filosofia da tecnologia e a filosofia jurídica que podem ser relevantes para o filósofo interessado nos efeitos da inteligência artificial enquanto tecnologia inserida em um contexto social.

Nos casos em que a IA é utilizada como ferramenta de apoio à filosofia jurídica, há uma preocupação adicional: como expor os artefatos computacionais, e seus produtos, à crítica? Para que estes sejam submetidos ao mesmo grau de escrutínio empregado em relação aos argumentos filosóficos, é importante que os códigos e dados eventualmente usados estejam disponíveis para o exame dos interessados, o que pode ser feito, por exemplo, através da disponibilização de uma interface ou mesmo do compartilhamento dos materiais por um *link* no próprio material publicado. Com isso, passa a ser possível avaliar as hipóteses, as técnicas e as premissas envolvidas no uso da IA, bem como testar a robustez do sistema (MULDOON, 2007) a partir da variação de configurações do sistema.

Conclusão

A difusão de técnicas de inteligência artificial nas sociedades contemporâneas terá impacto no funcionamento dos sistemas jurídicos destas sociedades, o que por sua vez trará novos pontos de consideração para a filosofia jurídica. Mesmo que estas novas questões muitas vezes possam ser expressas em termos já habituais a uma ou outra tradição jusfilosófica, o seu tratamento adequado exigirá que o filósofo jurídico tome contato com noções oriundas de outras áreas do conhecimento, em particular da disciplina de inteligência artificial.

Não segue disso, no entanto, que a tarefa da filosofia jurídica será reduzida à discussão da inteligência artificial ou que o filósofo do direito deverá internalizar o conhecimento destes outros campos. Se a inteligência artificial trará uma mudança de paradigma para (setores da) filosofia jurídica, esta se dará menos pela mudança de métodos — ainda que os aportes oriundos de sistemas inteligentes possam ser úteis — e mais pelo estabelecimento de um diálogo interdisciplinar, em que o filósofo jurídico travará contato com as contribuições de outros campos e as situará nas discussões jusfilosóficas relevantes.

Agradecimentos

O autor gostaria de agradecer a Renata Vaz Shimbo, João Gabriel Arato Ferreira, Victor Luís Nascimento Barroso, Guilherme Paulino Passos, Andrea Faggion e Daniela Rigotto pelas questões e comentários feitos.

Referências

ALMADA, M.; ATTUX, R. Ethical design of social simulations. In: XXXVIII Congresso Da Sociedade Brasileira De Computação, Natal, 2018. **3º Workshop sobre aspectos sociais, humanos e econômicos de Software**. Porto Alegre: SBC, 2018.

ÁVILA, H. Teoria dos princípios: da definição à aplicação dos princípios jurídicos. 17ª ed. São Paulo: Malheiros, 2016.

BRENNAN-MARQUEZ, K.; HENDERSON, S. E. Artificial Intelligence and Role-Reversible Judgment. 2018.

BROWNLEE, K.; STEMPOWSKA, Z. Thought Experiments. In: BLAU, A. (Ed.). **Methods in Analytical Political Theory**. Cambridge: Cambridge University Press, 2017.

BRYSON, J. J.; DIAMANTIS, M. E.; GRANT, T. D. Of, for, and by the people: the legal lacuna of synthetic persons. **Artificial Intelligence and Law**, v. 25, n. 3, p. 273–291, 2017.

CASEY, A. J.; NIBLETT, A. Self-driving laws. **University of Toronto Law Journal**, v. 66, n. 4, p. 429–442, 2016.

CASTRO JÚNIOR, M. A. de. **Personalidade Jurídica do Robô e sua efetividade no Direito**. 2009. Tese (Doutorado em Direito) – Universidade Federal da Bahia, Salvador.

EDMUNDSON, W. A. Coercion. In: MARMOR, A. (Ed.). **The Routledge Companion to Philosophy of Law**. New York: Routledge, 2012. p. 451–466.

ELSEN BROICH, C.; GILBERT, N. **Modelling Norms**. NY: Springer Publishing, 2013.

FELIN, T. *et al.* The law and big data. **Cornell Journal of Law and Public Policy**, 2017.

GENDLER, T. S. **Thought experiment: On the powers and limits of imaginary cases**. Tradução. New York: Garland Publishing, 2000.

GILBERT, N.; TROITZSCH, K. G. **Simulation for the Social Scientist**. Maidenhead: Open University Press, 2005.

HILDEBRANDT, M.; ROUVROY, A. **Law, Human Agency and Autonomic Computing: The Philosophy of Law Meets the Philosophy of Technology**. London: Routledge, 2011.

HOLMES, O. W. The Path of the Law. **Harvard Law Review**, v. 110, n. 5, mar. 1997.

HUMPHREYS, P. **Extending ourselves: Computational science, empiricism, and scientific method**. Oxford: Oxford University Press, 2004.

LOPES, M. F.; ROBERTO, E. Quando um carro autônomo atropela alguém, quem responde? **El País Brasil**, abr. 2018.

MARANHÃO, J. **Positivismo jurídico lógico-inclusivo**. São Paulo: Marcial Pons, 2012.

_____. A logical architecture for dynamic legal interpretation. In: **Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law**, 2017a.

_____. A pesquisa em inteligência artificial e Direito no Brasil. **Conjur**, dez. 2017b.

MARMOR, A. **The Routledge Companion to Philosophy of Law**. NY: Routledge, 2012.

MONTEIRO BRITO JUNIOR, V. A filosofia analítica do Direito é etnograficamente limitada? **Revista Direito GV**, v. 14, n. 1, p. 49–78, 2018.

MULDOON, R. Robust simulations. **Philosophy of Science**, v. 74, n. 5, p. 873–883, 2007.

MULDOON, R. **Social contract theory for a diverse world: Beyond tolerance**. NY: Routledge, 2016.

MULDOON, R.; BORGIDA, M.; CUFFARO, M. The conditions of tolerance. **Politics, Philosophy & Economics**, v. 11, n. 3, p. 322–344, 2012.

NEVEJANS, N. **European civil law rules in robotics**. Brussels: European Parliament, 2016.

NINO, C. S. **Introdução à análise do direito**. São Paulo: Editora WMF Martins Fontes, 2010.

OLIVEIRA, R. B. De; WASSERMANN, R. Utilização de Ontologia para Busca em Base de Dados de Acórdãos do STF. In: **X Seminar on Ontology Research in Brazil**, Brasília. Brasília: CEUR-WS, 2017.

PASSOS, G. P. **Modelos Lógicos em Direito**. 2016. Monografia (Bacharelado em Direito) — Universidade Estadual do Rio de Janeiro, Rio de Janeiro, 2016.

PRIEL, D. NotAll Law is an Artifact: Jurisprudence Meets the Common Law. In: BURAZIN, L.; HIMMA, K. E.; ROVERSI, C. (Ed.). **Law As an Artifact**. Oxford: Oxford University Press, 2018.

RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. [S. l.]: Prentice Hall, 2010.

SAAM, N. J. What is a Computer Simulation? A Review of a Passionate Debate. **Journal for General Philosophy of Science**, v. 48, n. 2, p. 293–309, 2017.

SORENSEN, R. A. **Thought Experiments**. New York: Oxford University Press, 1992.

STRANG, L. J. How Big Data Can Increase Originalism's Methodological Rigor: Using Corpus Linguistics to Reveal Original Language Conventions. **UCDL Rev.**, v. 50, p. 1181, 2016.

SUNDQUIST, C. The Science of Juror Racial Bias. **Denver University Law Review**, 2018.

THAGARD, P. **Computational Philosophy of Science**. Cambridge: MIT Press, 1993.

TRAZZI, M.; YAMPOLSKIY, R. V. Building Safer AGI by introducing Artificial Stupidity. **arXiv**, 2018.

TURING, A. M. Computing machinery and intelligence. **Mind**, v. 59, n. 236, p. 433, 1950.

WEISBERG, M. **Simulation and similarity: Using models to understand the world**. Oxford: Oxford University Press, 2012