Algoritmos, Inteligência Artificial e o Direito

(*) Paulo Sá Elias

O comitê de Ciência e Tecnologia do Parlamento da Inglaterra abriu inquérito[1] para examinar o uso crescente de algoritmos (e inteligência artificial) na tomada de decisões públicas e privadas, com o objetivo de avaliar como os algoritmos são formulados, os erros e possíveis correções – bem como o impacto que eles podem ter nos indivíduos e sua capacidade de entender ou desafiar decisões tomadas com base no uso da inteligência artificial. Após a leitura de todos os documentos (*que serviram como referência para a elaboração deste texto, bem como outros textos e pesquisas que tenho feito sobre o assunto*) resolvi escrever essas breves palavras sobre o tema.

Em primeiro lugar, é importante entendermos o que são os algoritmos (***Algorithms***) aplicados na informática e telemática, inteligência artificial (***Artificial Intelligence***), aprendizado de máquina (***Machine Learning***), aprendizado profundo (***Deep Learning***), redes neurais (***Neural Networks***), Internet das coisas (***Internet of Things***) e outros – que impressionam em razão dos recentes e impressionantes avanços e da importância cada vez maior que passaram (e passarão) a ter em nossas vidas.

Algoritmo (***algorithm***), em sentido amplo, é um conjunto de instruções, como uma receita de bolo, instruções para se jogar um jogo, etc. *É uma sequência de regras ou operações que, aplicada a um número de dados, permite solucionar classes semelhantes de problemas. Na informática e telemática, o conjunto de regras e procedimentos lógicos perfeitamente definidos que levam à solução de um problema em um número de etapas*. Em outras palavras mais claras: são as diretrizes seguidas por uma máquina. Na essência, os algoritmos são apenas uma forma de representar matematicamente um processo estruturado para a realização de uma tarefa. Mais ou menos como as regras e fluxos de trabalho, aquele passo-a-passo que encontramos nos processos de tomada de decisão em uma empresa, por exemplo.

Os sistemas algorítmicos estão presentes em todos os lugares, até mesmo nos sistemas de ABS (freios). São usados em computação há décadas, mas assumiram uma importância crescente em várias partes da economia e da sociedade na última década em virtude da disseminação de computadores. Vou falar mais sobre os algoritmos adiante.

A **inteligência artificial** (*Artificial Intelligence* – ou simplesmente AI), em definição bem resumida e simples, é a possibilidade das máquinas (*computadores,*

1 Veja os detalhes aqui: http://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/parliament-2017/algorithms-in-decision-making-17-19/

robôs e demais dispositivos e sistemas com a utilização de eletrônica, informática, telemática e avançadas tecnologias) executarem tarefas que são características da inteligência humana, tais como planejamento, compreensão de linguagens, reconhecimento de objetos e sons, aprendizado, raciocínio, solução de problemas, etc. Em outras palavras, é a teoria e desenvolvimento de sistemas de computadores capazes de executar tarefas normalmente exigindo inteligência humana, como a percepção visual, reconhecimento de voz, tomada de decisão e tradução entre idiomas, por exemplo.

O **aprendizado de máquina** (***Machine Learning***) é uma forma de conseguir a inteligência artificial. É um ramo da inteligência artificial que envolve a criação de algoritmos que podem aprender automaticamente a partir de dados. Ao invés de os desenvolvedores de *software* elaborarem enormes códigos e rotinas com instruções específicas para que a máquina possa realizar determinadas tarefas e conseguir resultados (e com isso limitar drasticamente o seu campo de atuação e resultados), no aprendizado de máquina *treina-se o algoritmo para que ele possa aprender por conta própria*, e até mesmo conseguir resultados que os desenvolvedores dos algoritmos nem mesmo poderiam imaginar. Neste treinamento, há o envolvimento de grandes quantidades de dados que precisam ser alimentadas para o algoritmo (*ou aos algoritmos envolvidos*), permitindo que ele (o algoritmo) se ajuste e melhore cada vez mais os seus resultados. Exemplo: o aprendizado de máquina foi utilizado para melhorar significativamente a visão por computadores (a capacidade de uma máquina reconhecer um objeto em uma imagem ou vídeo). Os seres humanos podem marcar imagens que têm um gato *versus* aquelas que não os possuem. O algoritmo tenta construir um modelo que pode marcar com precisão uma imagem como contendo um gato ou não, assim como um ser humano. Uma vez que o nível de precisão é alto o suficiente, a máquina agora "aprendeu" como é um gato, como ele se parece.

O **aprendizado profundo** (***Deep Learning***) é uma das várias abordagens para o aprendizado de máquinas. Outras abordagens incluem aprendizagem por meio de árvores de decisão (*decision tree learning*), programação de lógica indutiva (*inductive logic programming*), agrupamento (*clustering*), aprendizagem de reforço (*reinforcement learning*), redes bayesianas (*Bayesian networks*), entre outros. A aprendizagem profunda foi inspirada na estrutura e nas funções do cérebro humano, na interligação dos neurônios. As **redes neurais artificiais** (***Artificial Neural Networks – ANNs***) são algoritmos que imitam a estrutura biológica do cérebro humano. Nas ANNs, existem "neurônios" (entre aspas) que possuem várias camadas e conexões com outros "neurônios". Cada camada (*layer*) escolhe um recurso específico para aprender, como curvas e bordas no reconhecimento de uma imagem, por exemplo. A aprendizagem profunda tem o seu nome em razão dessas várias camadas. A profundidade é criada com a utilização de múltiplas camadas em oposição a uma única camada de aprendizado pelo algoritmo. Esses algoritmos de aprendizado profundo

formam as "redes neurais" e estas rapidamente podem ultrapassar a nossa capacidade de compreender todas as suas funções.

A inteligência artificial e a **Internet das coisas** (***Internet of things***)[2] estão intrinsecamente entrelaçadas. É como se fosse a relação entre cérebro e o corpo humano. Nossos corpos coletam as entradas sensoriais, como visão, som e toque. Nossos cérebros recebem esses dados e dão sentido a eles, por exemplo, transformando a luz em objetos reconhecíveis, transformando os sons em discursos compreensíveis e assim por diante. Nossos cérebros então tomam decisões, enviando sinais de volta para o corpo para comandar movimentos como pegar um objeto ou falar.

Todos os sensores conectados que compõem a Internet das coisas (*Internet of things*) são como nossos corpos, eles fornecem os dados brutos do que está acontecendo no mundo. A inteligência artificial é como nosso cérebro, dando sentido a esses dados e decidindo quais ações executar. E os dispositivos conectados da Internet das coisas são novamente como nossos corpos, realizando ações físicas ou se comunicando com os outros.[3]

Os inúmeros dispositivos construídos atualmente, tais como aparelhos médicos, relógios inteligentes, veículos, eletrodomésticos, enfim, todos os itens construídos com componentes eletrônicos, *software*, sensores e que possuam a capacidade de coletar e transmitir dados à Internet, capazes de serem identificados de maneira única, formam o que é conhecido como a Internet das coisas (*Internet of things*). Cada vez mais as coisas estão conectadas à Internet, mesmo aquelas coisas que não possuíam este objetivo específico, como um fogão, uma geladeira, equipamentos de ginástica, lâmpadas, uma cama, prateleiras de supermercados e depósitos (que avisam automaticamente quando determinado produto está acabando), portas, etc. Tudo começa a ficar conectado. Na Wikipedia você encontra uma definição de fácil compreensão: "(...) *A Internet das Coisas, em poucas palavras, nada mais é que uma extensão da Internet atual, que proporciona aos objetos do dia-a-dia (quaisquer que sejam), mas com capacidade computacional e de comunicação, se conectarem à Internet. A conexão com a rede mundial de computadores viabilizará, primeiro, controlar remotamente os objetos e, segundo, permitir que os próprios objetos sejam acessados como provedores de serviços. Estas novas habilidades, dos objetos comuns, geram um grande número de oportunidades tanto no âmbito acadêmico quanto no industrial. Todavia, estas possibilidades apresentam riscos e acarretam amplos desafios técnicos e sociais*".

2 Também chamada de "*Internet of everything*" (Internet de tudo, de todas as coisas)

3 Ref. Calum McClelland / *AI/ML/DL – differences.*

O aprendizado de máquina (*Machine Learning*) e o aprendizado profundo (*Deep Learning*) trouxeram grandes avanços para a Inteligência Artificial nos últimos anos. Tanto o aprendizado de máquina como o aprendizado profundo exigem grande quantidade de dados para que possam funcionar adequadamente e estes dados estão sendo coletados pelos bilhões de sensores que continuam a entrar na Internet das coisas a cada dia.

O aprimoramento da Inteligência Artificial impulsiona a adoção da Internet das Coisas (e vice-versa), afinal, os algoritmos precisam de dados – e quanto maior o número de sensores e pontos de coleta de dados, melhor. É a Inteligência Artificial que torna a Internet das coisas útil, fazendo com que esses inúmeros dispositivos possam oferecer resultados úteis aos seus usuários e dados extremamente valiosos para quem os coleta.

O encolhimento dos *chips* de computadores e o aprimoramento das técnicas de fabricação, significa sensores cada vez mais poderosos e baratos. Há grandes avanços em relação à tecnologia das baterias e fontes de energia[4], fazendo com que esses sensores possam funcionar sem interrupção durante anos. As novas tecnologias de conexão sem fio (que estarão disponíveis em todos os lugares no futuro) e a inevitável redução dos valores de transmissão de grandes volumes de dados, bem como armazenamento e capacidade computacional para processá-los, abre portas para um futuro promissor. Mas há preocupações. Vou falar a respeito logo mais adiante, ainda neste texto.

O comitê de Ciência e Tecnologia do Parlamento da Inglaterra, a que fiz referência no início deste texto, convidou vários especialistas no assunto, bem como grandes empresas envolvidas com o tema, como é o caso da IBM, Microsoft, Google e outros. Na apresentação realizada pelos representantes dessas empresas, vários pontos importantes foram ressaltados.

O Google, destacou que os algoritmos, com o desenvolvimento da inteligência artificial *(Artificial Intelligence)* e o aprendizado de máquina *(Machine Learning)* revelam-se como ferramentas muito poderosas e que já começam a fornecer uma ajuda fundamental no avanço da ciência, melhorando até mesmo o acesso aos cuidados médicos, ajudam a resolver alguns dos desafios globais mais urgentes em meio ambiente, transporte e a condução de soluções mais inteligentes para os problemas quotidianos. Naquela empresa, por exemplo, os algoritmos são utilizados desde a sua criação, base fundamental de seu sistema de pesquisas. No *YouTube*, também produto da empresa, os algoritmos adicionam automaticamente legendas aos vídeos analisando o áudio e convertendo o discurso em texto. *Artificial Intelligence* (AI) e *Machine Learning* (ML) alimentam os algoritmos. No *Google Translator*, os algoritmos são

4 A propósito do tema da energia, recomendo ao leitor que conheça o projeto **Wendelstein 7-X** (http://www.ipp.mpg.de/w7x) – O futuro da energia pode estar aqui. *Stellarator technology (Fusion Power Plant*)

usados para aprender os padrões dos idiomas em texto e voz, traduzindo mais de 100 milhões de palavras por dia em 103 idiomas. A inteligência artificial é fundamental para conseguir avanços significativos como o reconhecimento de voz, atingindo níveis quase humanos de precisão. No Google Fotos, é possível procurar qualquer coisa nas imagens, pois o sistema utiliza algoritmos que categorizam objetos e conceitos em imagens.

A combinação de algoritmos cada vez mais avançados, dados e poder de computação mais baratos, como já disse, vai tornar ainda mais amplamente distribuídos os benefícios dos algoritmos, melhorando o dia-a-dia de nossas vidas em todas as áreas. *Machine Learning*, por exemplo, foi utilizado em diagnósticos de retinopatia diabética, uma das causas de maior crescimento de cegueira em todo o mundo.

Algoritmos robustos com uso de *machine learning* são dependentes em geral, de um conjunto de dados de alta qualidade. Muitas empresas que operam em setores regulados, geralmente em finanças, possuem algoritmos secretos para pontuação de crédito, aprovação de empréstimo, etc. Esses algoritmos fazem parte de sua vantagem competitiva. No entanto, muitos debatedores no parlamento inglês entendem que essas empresas precisam ser capazes de explicá-los aos reguladores e justificar as decisões e resultados, se necessário.

Muitos dos algoritmos produzidos por meio de *Machine Learning*, notadamente aqueles baseados em "*deep learning*" ou *redes neurais (neural networks)*, não são totalmente compreendidos. Nenhum ser humano é capaz de dizer por quê determinado algoritmo desta natureza faz o que faz, nem pode prever totalmente o que o algoritmo poderá fazer em dados diferentes daqueles utilizados para o treinamento da máquina, ao longo do tempo.[5]

Até mesmos os maiores defensores desses sistemas, admitem essa fraqueza. Embora as redes neurais profundas (*DNN – Deep Neural Networks*) tenham demonstrado uma grande eficácia em uma ampla gama de tarefas, quando eles falham, muitas vezes falham espetacularmente, catastroficamente, produzindo resultados inexplicáveis e incoerentes que podem deixar o ser humano perplexo, sem conseguir entender a razão pela qual o sistema tomou tais decisões. A falta de transparência nos processos de decisão em redes neurais profundas ainda é um obstáculo significativo para a sua ampla adoção em determinados segmentos como saúde e segurança, onde a tolerância ao erro é muito baixa e a capacidade de interpretar, entender e tomar decisões confiáveis é um elemento crítico.[6]

5 Knight,W., *The Dark Secret at the Heart of AI*. https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/

6 Kumar,D., Wong,A. & Taylor,G.W., *Explaining the Unexplained: A CLass-Enhanced Attentive Response (CLEAR) Approach to Understanding Deep Neural Networks*. https://arxiv.org/abs/1704.04133

As empresas já contratam especialistas no tema. A Intel (conhecida por fabricar processadores de computadores), por exemplo, possui entre os seus cargos de maior importância na diretoria, a posição: "*Chief Algorithms Scientist*".

Algoritmos baseados em *Machine Learning* são utilizados para previsão do tempo, análise de compras realizadas pelos consumidores, ajudando a empresa decidir quais as mercadorias são as preferidas e onde elas devem estar localizadas nos supermercados, por exemplo. Até mesmo uma máquina separadora de tomates utiliza inteligência artificial para aprimorar a seleção. Na Internet, algoritmos determinam os resultados que os sistemas de busca retornam ao usuário, como é o caso do Google. Quais anúncios e conteúdos serão mostrados – e quando serão mostrados, para quem, onde, como, próximos de quais outros anúncios e conteúdos – e assim por diante.

Os algoritmos e a inteligência artificial também já estão sendo usados no Poder Judiciário. No caso *Zilly* (Angwin et al. 2016)[7] – uma das questões consideradas pelo *risk scoring algorithm* utilizado (algoritmo de pontuação de risco em matéria de execução penal) foi que um dos pais do acusado já havia sido preso. Sabemos que nunca um promotor de justiça ou um juiz vai aceitar a existência de um argumento como esse para pedir ou atribuir um período maior de prisão ao acusado pelo simples fato de que um dos seus pais teria sido preso anteriormente. Mas a máquina interpretou assim.

Por exemplo, ferramentas de mapeamento de crimes são cada vez mais utilizadas para examinar dados de crime e identificar *hotspots (locais com grande incidência de crimes)* para aumentar a eficiência na alocação de recursos policiais. No entanto, os dados de entrada (por exemplo, dados históricos da criminalidade) podem ser inclinados à parcialidade, provocada por policiais, cujas práticas podem não refletir a incidência real de crime, e em vez disso acabam influenciados pela segmentação dos grupos marginais.

Os professores devem estar atentos ao fato de que a Internet (em grande parte dirigida por processos algorítmicos) pode também ser injustamente prejudicial ou manifestar-se contra alguém ou alguma coisa em suas escolhas, especialmente em vídeos, textos e imagens selecionadas pelos algoritmos.[8] Os consultores escolares, bem como os de carreira, também devem estar atentos ao fato de que determinados serviços de Internet são suscetíveis de serem tendenciosos. Além do que, sabemos que essas

---

7 Angwin,J., Larson,J., Mattu,S. & Kirchner,L., *Machine Bias. There is software that is used across the county to predict future criminals. And it is biased against blacks*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

8 Kay,M., Matuszek,C. & Munson,S.A., *Unequal representation and gender stereotypes in image search results for occupations*. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, ACM, 2015, pp. 3819-3828.
https://github.com/mjskay/gender-in-image-search/blob/master/data/public/filtered_bls_occupations_with_search_gender_proportions.csv

empresas se tornaram grupos extremamente poderosos, com capacidade até mesmo de mudar regimes em determinados países, provocar protestos, influenciar a opinião pública, destruir a credibilidade e fechar várias portas a determinadas pessoas-alvo, etc.

Inteligência artificial, na aprendizagem de máquinas (*Machine Learning*) e em outras abordagens algorítmicas que fazem inferências baseadas em dados de sensores, tais como gravações de som, dados de sensores de telefones celulares, são cada vez mais utilizadas por órgãos governamentais e grupos privados, como já é notório. Algoritmos têm sido utilizados na tomada de decisão por muitas décadas, no entanto, no passado, operadores humanos eram envolvidos em todas as etapas. A discussão atual em relação aos riscos da inteligência artificial surge em parte por causa do crescente uso de dados e processos totalmente automatizados (com o elemento humano fora da equação) e também naqueles em que há utilização de algoritmos secretos, como os do Google.

Processos totalmente automatizados podem ser usados para substituir o julgamento individual de um operador de linha de frente e, portanto, podem remover uma fonte potencial de subjetividade e o viés tendencioso. Mesmo as pessoas bem-intencionadas têm preconceitos conscientes e inconscientes que afetam os seus julgamentos. Assim, acredita-se que a decisão algorítmica pode oferecer uma oportunidade para melhorar o aspecto de imparcialidade.

Mas atenção: algoritmos não são imparciais. Os próprios algoritmos podem conter os preconceitos presentes nos criadores do algoritmo ou nos dados que foram usados para treinar o algoritmo. O desempenho dos algoritmos depende muito dos dados utilizados para desenvolvê-los. Os preconceitos que estão presentes nos dados serão refletidos pelos algoritmos. Tais desvios, intencionais ou não, podem ser inerentes aos dados, como também oriundos do próprio desenvolvedor do algoritmo. Isso pode ter efeitos tão ruins como os vícios que eles pretendiam eliminar. Alguns denominam este fenômeno como "***Machine bias***", "***Algorithm bias****"* ou simplesmente, ***Bias***. É o viés tendencioso. A remoção de tal viés tendencioso em algoritmos não é trivial e é um campo de pesquisa em andamento. Os desvios são difíceis de serem descobertos se o algoritmo for muito complexo (como são os utilizados pelo Google), pior ainda se forem secretos. Se o algoritmo é simples e auditável, especialmente publicamente auditável, então haverá em tese (vou explicar adiante a razão do "em tese") maiores chances de que as decisões baseadas em tais algoritmos possam ser mais justas. Igualmente em relação aos dados utilizados para "treinar" o algoritmo. Se eles forem auditáveis (e anônimos quando apropriados) poderão ser identificados desvios desta natureza.

Como é possível notar, reitera-se: um dos maiores problemas em torno de novos algoritmos de aprendizagem de máquinas são os dados em que eles são treinados. Os

pontos fortes e fracos dos dados de entrada são, portanto, extremamente importantes, e devem ser considerados, bem como a lógica inerente ou formulação de usar um sistema automatizado ou analítico para resolver um determinado problema. Algoritmos não são, e podem nunca ser, neutros ou independentes da sociedade que os produziu. A sociedade deve estar ativamente envolvida na formação dos valores que estão em jogo no uso de algoritmos e da inteligência artificial. Os conjuntos de dados geralmente contêm traços de parcialidade. Se pudermos identificar essas falhas, e atingirmos uma definição matemática de justiça, quem sabe, conseguiremos estatisticamente, uma forma de mitigá-los. A maioria dos cientistas de dados não possuem essas habilidades, por isso, raramente esses cuidados são tomados. Afinal, como sabemos, existem diferenças enormes entre os fundamentos epistemológicos das *Geisteswissenchaften* (ciências do espírito) – onde está localizado o Direito em relação às *Naturwissenschaften* (ciências da natureza) – onde está localizada a Matemática, a Lógica, a Informática e a Telemática. As ciências do espírito fundamentam-se na realidade social e histórica e, por esta razão, são irredutíveis a modelos causalistas. Como ressalta Karl Jaspers: "Nas ciências humanas não podemos nos contentar com a constatação de algo que fisicamente existe, perceptível aos nossos sentidos, mensurável, avaliado através de experiências. Nas ciências humanas, temos de compreender a significação perseguida pelos seres que agem, pensam, prevêem e acreditam. Não podemos nos contentar com o conhecimento exterior das coisas, mas temos de apreender, no seu interior, o significado posto pelo homem." (JASPERS, Karl. *Introdução ao pensamento filosófico*. São Paulo: Cultrix, 1992. – Cf. DILTHEY, W . *Introduction à l'étude des sciences humaines*. Paris: Presses Universitaires de France, 1942.)

Estatísticas matemáticas bem conhecidas podem ser usadas para testar se as decisões de um determinado algoritmo tendem a variar com características protegidas. Não há um consenso claro sobre a melhor medida a fazer: medidas matemáticas padrão, como correlação ou informação mútua, podem ser relevantes. A "paridade estatística" é uma medida específica discutida em relação à equidade algorítmica. O ponto mais amplo é que os algoritmos de tomada de decisão podem ser auditados pela inspeção de suas entradas e saídas usando ferramentas matemáticas. Idealmente, tais testes seriam medidos em todas as características protegidas em conjunto, em vez de separadamente, porque pode haver preconceitos de nicho que dependem de critérios múltiplos.

Provedores de serviços de aplicações de Internet e também de acesso, cada vez mais reúnem grandes coleções de dados comportamentais que podem ser recolhidos a partir dos inúmeros sensores que estão acompanhando a Internet das coisas (*Internet of things – everything*), os telefones celulares, relógios inteligentes e demais produtos portáteis, tais como GPS, acelerômetros, etc. Esses dados geralmente produzem um perfil muito detalhado e pessoal, provavelmente maior do que a maioria dos cidadãos

pode imaginar e compreender. Existe, portanto, a capacidade de os algoritmos tomarem decisões com base em inferências como "determinada pessoa está frequentemente dirigindo na madrugada" entre outras coisas arrepiantes. *Freaking creepy*, para usar a excelente expressão inglesa.

Falando em expressões da língua inglesa, importante conceituar, ainda que brevemente, o que são:

Decision Trees (Árvores de decisão): Aprendizado de máquina (*machine learning*) e método de apoio a decisão que usa um modelo de árvore de sim-não para cada decisão. Exemplo: *A pessoa possui 18 anos ou mais? Sim ou não?*

Random Forests (Floresta aleatória): Método para a classificação e regressão que opera através da construção de um grande número de árvores de decisão.

Linear Regression (Regressão linear): Abordagem para modelar a relação entre variáveis como uma fórmula linear direta, por exemplo: cada cigarro diminui 11 minutos de sua expectativa de vida.

Logistic regression (Regressão logística): É um modelo de regressão relativamente simples amplamente utilizado para prever as variáveis categóricas (categorias).

Bayesian networks (Redes Bayesianas): Também chamadas de modelos gráficos probalísticos ou *belief networks*. São um tipo de modelo estatístico que representa variáveis aleatórias e suas dependências. Ao contrário de muitos outros métodos, as redes bayesianas modelam explicitamente nossas certezas e incertezas anteriores e as usam para inferir a certeza de seus resultados (*outputs*).

O êxito do desenvolvimento de novas tecnologias de dados, incluindo a utilização mais ampla de algoritmos, inteligência artificial, aprendizado de máquina (*Machine Learning*) e a *Internet das coisas (Internet of things),* como estamos percebendo, será essencial para o crescimento de qualquer país ou empresa nos próximos anos. O vasto volume de dados criado pela Internet e o crescimento gigantesco de dados coletados por inúmeros sensores nos mais variados itens e equipamentos mudará muito o mundo dos negócios e o dia a dia das pessoas nos próximos anos. E isto tem profundo impacto em relação a autodeterminação informativa, o direito constitucional da intimidade e a privacidade. Novas leis são necessárias, sim, em alguns aspectos específicos, como o projeto de lei de proteção de dados pessoais (Projeto de Lei 5.276/2016), capitaneado pelo excelente Danilo Doneda. Mas todo cuidado é pouco, tendo em vista o que ocorreu com o Marco Civil da Internet (Lei nº 12.965/2014) – que sofrendo influência do pernicioso *lobby* de empresas multinacionais, provedores de aplicações de Internet – criou uma aberração jurídica em relação a responsabilidade civil para protegê-los em detrimento de teoria de

responsabilidade civil adotada no Direito Brasileiro, o que exige, sem dúvida, urgente alteração para que se compatibilize com as normas de regência.[9]

Aliás é grave assistir jovens professores, que aparentemente desconhecendo a amplitude do direito brasileiro, defendem a criação de novas leis a todo instante. Mas por trás disso, existe o patrocínio de empresas interessadas, como vimos no recente escândalo envolvendo o Google, publicado pelo jornal *Wall Street Journal*.[10] Saiba mais detalhes também aqui: http://googletransparencyproject.org/

Em informática, técnicas preditivas (predição/antecipação), incluindo sistemas de aprendizagem baseados em modelos ou sistemas de aprendizagem ensinadas e que podem fazer predições baseadas em exemplos e dados ainda não analisados, foram desenvolvidas na década de 1960 e aprimoradas nas décadas posteriores.

Para quem ainda não entendeu, algoritmos são usados para determinar os resultados de uma série de *inputs* (entradas de dados) desde o início da era digital. Eles são fundamentalmente instruções sobre como combinar entradas para produzir uma saída específica. À medida que os sistemas se tornaram mais complexos, os algoritmos também se tornaram mais complexos. O que chamamos de algoritmo pode consistir em múltiplos sub-algoritmos conectados de múltiplas formas. Embora isso possa ser explicado em termos da estrutura do sistema, as entradas e saídas em cada etapa rapidamente se tornam mais complexas de serem explicadas. Desta forma, para garantir a transparência nos algoritmos, precisamos garantir que todas as etapas do algoritmo possam ser explicadas de maneira que um indivíduo não especializado possa entender.

Algoritmos podem ser criados "à mão", isto é: um humano projeta, testa e reconstrói o algoritmo até atingir o resultado desejado. Como já é possível perceber, algoritmos também podem ser criados por meio de *Machine Learning*. Para processos como esses, o diferencial fundamental é que o próprio sistema ajusta as ponderações de entrada do algoritmo. Para sistemas complexos, centenas, milhares ou milhões de ponderações são ajustadas em paralelo. Os sistemas resultantes podem ser explicados matematicamente, no entanto, as entradas para tais sistemas são captadas a partir dos dados em estado bruto para um ponto em que os números são praticamente insignificantes para qualquer observador externo.

Como destacamos há pouco em relação ao Poder Judiciário, os estados de New Jersey e Wisconsin, nos Estados Unidos, por exemplo, já utilizam algoritmos para tomar decisões na Justiça Criminal sobre questões incluindo execução penal, condenação, fiança, reincidência, etc. Empresas privadas como é o caso da Northpointe desenvolveram as ferramentas com a utilização de algoritmos para essas finalidades. Algoritmos também são usados para analisar casos ou transações incomuns e investigar

9 http://www.direitodainformatica.com.br/?p=1794

10 https://www.wsj.com/articles/paying-professors-inside-googles-academic-influence-campaign-1499785286

potenciais atividades fraudulentas. Na França, o governo utiliza um algoritmo para lidar com as atribuições dos professores nas escolas públicas. Os professores são designados às vagas disponíveis nas escolas por todo o país, com base em decisões tomadas por um algoritmo.

Empresas como a *Hire Predictive*, por exemplo, oferecem serviços a empresas de todo o mundo para fins de seleção de candidatos a empregos com utilização de algoritmos. Com base em dados de negócios existentes, são construídos modelos preditivos (modelos de predição/antecipação), dos quais são feitas decisões algorítmicas sobre quais candidatos devem ser contratados.

Há muito tempo, agências de referência de crédito e gestão de riscos, como a Experian (Serasa, no Brasil) utilizam algoritmos para decidir a pontuação de crédito das pessoas (*credit score*). Os credores também usam seus próprios algoritmos para decidir quais candidatos a empréstimos devem aceitar e quais as taxas de juros a serem definidas.

As seguradoras também utilizam algoritmos para tomar decisões sobre clientes potenciais. Os algoritmos analisam permanentemente os segurados para colocá-los em categorias de risco específicas e tomar decisões sobre a aceitação de pedidos e o nível de seguro a ser definido.

Classificadores baseados em redes neurais equivalem-se ou superam a precisão do nível humano em muitas tarefas comuns, no entanto, redes neurais são suscetíveis ao que se denomina como "*exemplos contraditórios*". Dados cuidadosamente adulterados podem provocar um comportamento ruim, levando a escolhas arbitrárias e equivocadas.[11]

Algoritmos de aprendizado de máquina (*Machine Learning*) também estão sendo desenvolvidos para aplicação em diversos setores da área de saúde (incluindo prevenção, diagnóstico, prognóstico, tratamento, gestão da demanda e alocação de recursos), muitas vezes na forma de ferramentas de apoio à decisão para ajudar e aperfeiçoar o trabalho dos profissionais humanos.

O que se sabe é que a maioria, se não todos os algoritmos usados hoje em dia pertencem ao campo da aprendizagem de máquina (*Machine Learning*), que nesta altura do texto já sabemos ser uma forma de inteligência artificial. Esses algoritmos de aprendizado de máquina são utilizados para aprender relações estatísticas dentro de dados, e podem ser aplicados a dados novos para criar "decisões" como a categorização ou previsão. Na área de saúde, como vimos, no diagnóstico médico, um algoritmo pode aprender a relação entre as características de diagnóstico de imagem e se as imagens pertencem a pacientes com câncer, e em seguida, são analisadas novas imagens de

11 https://arxiv.org/pdf/1707.07397.pdf

pacientes em geral, poderá ser usado para prever a probabilidade de câncer em novos pacientes. E isso é ótimo.

Na prática clínica, as ferramentas com o uso de algoritmos são desenvolvidas para aperfeiçoar os procedimentos de diagnóstico por meio da análise em alta velocidade de imagens médicas já recolhidos da prática clínica padrão. O aprendizado de máquina (*Machine Learning*) também está na vanguarda da medicina de precisão, ajudando a identificar subgrupos de pacientes e tratamentos-alvo apropriadamente. Em epidemiologia, está sendo aplicada a dados de saúde pública para detectar e rastrear surtos de doenças infecciosas. Também está sendo usado para aperfeiçoar o acompanhamento médico e para otimizar a gestão de demanda e alocação de recursos nos sistemas de saúde. Algoritmos nessas aplicações não são agentes autônomos, mas atuam como ferramentas de apoio à decisão com o objetivo de ajudar e aperfeiçoar o trabalho dos profissionais humanos da área de saúde.

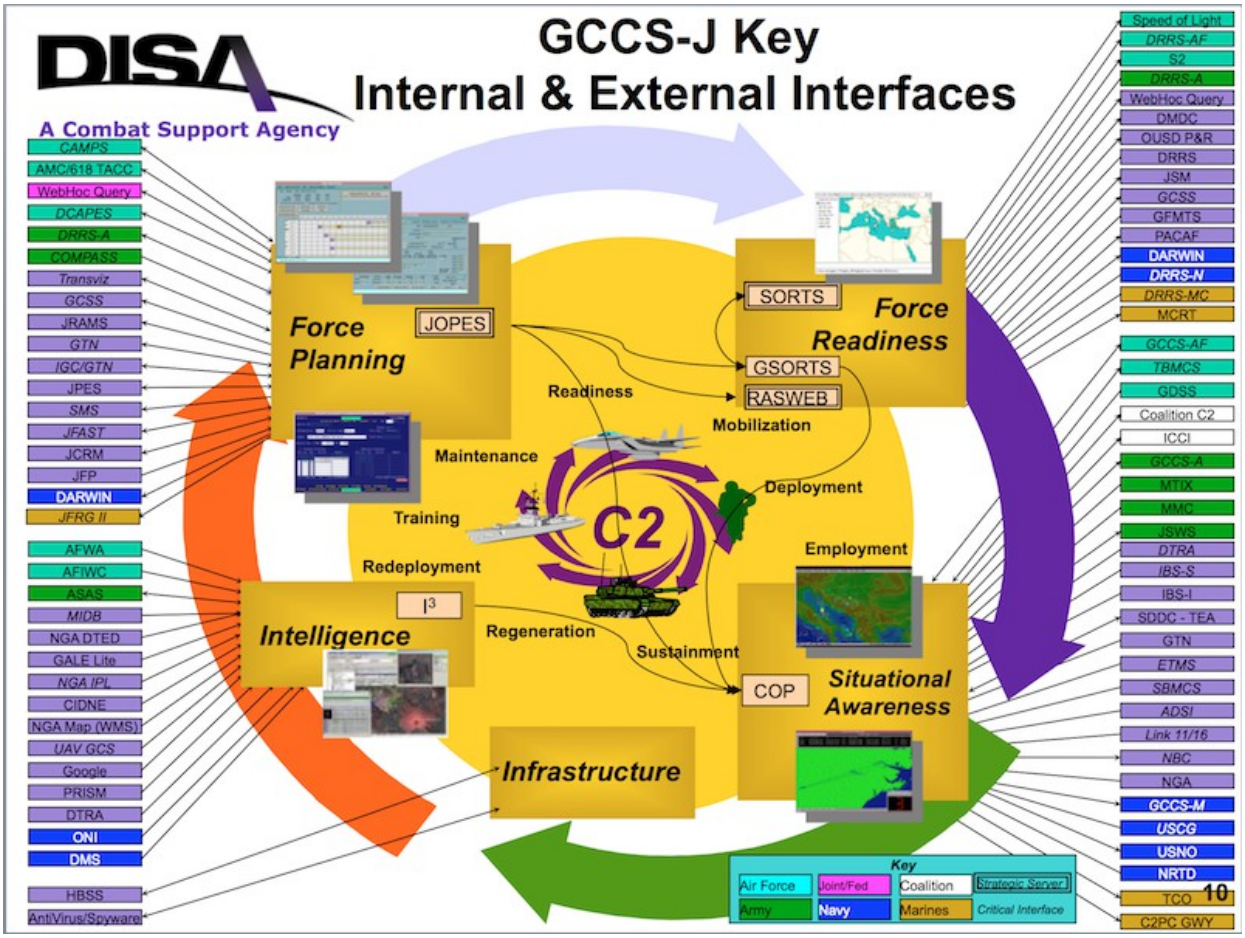
No inquérito em curso na Inglaterra, destacou-se que o uso de algoritmos no setor público poder levar a policiamento discriminatório e monitoramento indiscriminado, bem como a ações de agências de inteligência (inclusive aquelas que estão disfarçadas como empresas convencionais, mas são órgãos de inteligência **#G11000**), influência comportamental e invasões em larga escala de privacidade. O uso de algoritmos para tomar decisões no setor privado também pode levar à discriminação em áreas como recrutamento e emprego, acesso a serviços e finanças, preços diferenciados, etc. A tendência crescente de utilização de algoritmos na tomada de decisões pode pressionar indivíduos e serviços no sacrifício da privacidade e deteriorar ainda mais os direitos relativos à intimidade, privacidade e autodeterminação informativa. Todos devem se lembrar do resultado inadequado dos algoritmos utilizados pelo Google, quando classificou os negros como gorilas. (*Artificial Intelligence's White Guy Problem* – Kate Crawford, The New York Times, 25th June 2016).

Uma das questões de maior preocupação é a utilização dos algoritmos para policiamento e coleta de dados para inteligência. Há um potencial considerável de abusos secretamente praticados por governos e empresas privadas.

É notório que os algoritmos estão sendo utilizados extensivamente pelas agências de segurança e inteligência para analisar o público e a comunicação dos cidadãos. A coleta o processamento das comunicações e dados de milhões de pessoas deveria ocorrer para monitorar conhecidos alvos e suspeitos – e também para descobrir novos – mas com certas restrições e controle judicial, no entanto, a técnica de encontrar uma agulha no palheiro, invertendo a relação entre a tradicional vigilância de suspeitos e suas relações para o monitoramento indiscriminado é extremamente preocupante. Quem assistiu o excelente filme "O inimigo do Estado" muitos anos antes de Edward

Snowden, viu a ficção tornar-se realidade com o episódio envolvendo a NSA e o programa PRISM.

A propósito, naquele período do escândalo envolvendo a NSA – *National Security Agency* e o analista Edward Snowden, enquanto o mundo voltava os olhos para o programa PRISM da agência de inteligência norte-americana, uma apresentação em *Powerpoint* que teria sido supostamente vazada pela inteligência francesa – e cuja autenticidade já foi confirmada por vários especialistas, mostrou uma informação muito preocupante e que rapidamente foi apagada da Internet. São poucos os lugares em que encontramos este material *online* hoje em dia. O trecho que mais chamou atenção é esse:



Observe que na apresentação da DISA (*Defense Information Systems Agency*) – órgão integrante do sistema de inteligência militar norte-americano, o programa **PRISM** aparece como a interface e a empresa **Google**, como a responsável pela coleta das informações no item "*Intelligence*". Há um livro muito esperado, escrito pelo excelente jornalista Yasha Levine chamado "*Surveillance Valley – The Secret Military History of the Internet*"[12] que promete trazer revelações importantes. Sugiro ao leitor que também assista o vídeo *Secret History of Silicon Valley* oferecido pelo *Computer*

12 https://surveillancevalley.com/

History Museum, apresentado pelo professor Steve Blank.[13] E também o texto elaborado por Julian Assange do Wikileaks – "*Google Is Not What It Seems*".[14]

A análise do *big data* e o processamento algorítmico deveriam ser realizados em dados relevantes de grupos de suspeitos, como mencionei, no entanto, nossos telefones registram onde vamos, com quem falamos, mídias sociais armazenam até como nos sentimos, nossos cartões bancários e de crédito registram uma grande quantidade de informações da nossa atividade, medidores inteligentes até gravam quando estamos em casa e quanto de energia utilizamos. Snowden garantiu (e comprovou pelos documentos que vazou) que os aparelhos de celular inteligentes podem gravar o ambiente, mesmo desligados. Grande parte desses dados estão disponíveis para o Estado que procura incansavelmente "pegadas digitais abrangentes" em vez de partir de uma avaliação cuidadosa das evidências relevantes.

Essa abordagem preocupante emprega algoritmos para analisar dados passivamente, tratando todos os cidadãos como suspeitos até que seja comprovado o contrário. Os algoritmos podem tratar como alvos pessoas inocentes, criminalizando-as por associação equivocada e impactando negativamente essas pessoas com base em pouco mais de um palpite apoiado por fatos não-criminais. Dados indiscriminadamente coletados em massa nas mãos do Estado para processamento algorítmico, da forma como estão sendo construídos, jamais poderiam ser compatíveis com a legislação de direitos humanos. São práticas que não são saudáveis em uma democracia. Independente da alegada precisão de tais processamentos algorítmicos, o cenário é muito preocupante.

Conjuntos de dados de teste para esses algoritmos, ou os dados recolhidos da sociedade, podem ser reflexo de padrões de discriminação e as desigualdades já existentes. Na medida em que a sociedade contém desigualdades, exclusão ou outros vestígios de discriminação, o mesmo poderá acontecerá com os dados. De fato, há grande preocupação que tais desigualdades sociais possam ser perpetuadas por meio de processos algorítmicos, gerando efeitos jurídicos negativos e significativos sobre indivíduos no contexto da aplicação da lei.

Não temos conhecimento de quaisquer processos formais para a supervisão independe dos algoritmos utilizados por empresas como o Google e a comunidade de inteligência. No inquérito em curso na Inglaterra, vários especialistas acreditam fortemente que deveria haver uma supervisão independente dos mecanismos e operação de algoritmos em qualquer contexto que possa resultar em efeitos jurídicos e outros significativos na vida das pessoas.

---

13 https://www.youtube.com/watch?v=ZTC_RxWN_xo

14 https://wikileaks.org/google-is-not-what-it-seems/

A disponibilidade de novas fontes de dados, incluindo dados de redes sociais, biometria e *software* de reconhecimento facial, cria oportunidades para interferências relativas ao direito à privacidade em nível individual e social.

Existem processos algorítmicos sendo utilizados para influenciar comportamentos de pessoas. Processamento algorítmico não deve ser a única base para uma decisão que produza efeitos jurídicos ou possa impactar os direitos de qualquer indivíduo.

Os criadores de algoritmos devem sempre manter a capacidade de fornecer transparência em relação a todo o processo algorítmico envolvido e explicações para as decisões e resultados atingidos. As decisões algorítmicas que envolvem os direitos e as liberdades dos indivíduos devem ser sempre desafiáveis. É preciso cautela contra qualquer tomada de decisão algorítmica que envolva a coleta de dados não consentida, ou mesmo que consentida, não compreendida, bem como compartilhamento ou análise de dados pessoais. A importância dos direitos constitucionais de intimidade, da privacidade e autodeterminação informativa como um direito individual e uma característica de uma sociedade democrática deve ser cuidadosamente mantida ao longo dos desenvolvimentos tecnológicos. Não há como não ficar preocupado com a aprendizagem por máquinas, pela qual um computador aprende e extrai algoritmos para a realização de determinadas tarefas, apresentando resultados, a partir da transferência de dados fornecidos sem que o indivíduo possa compreender a qualidade e extensão destes.

No inquérito inglês, a Microsoft concorda que o poder computacional dos algoritmos pode aumentar a criatividade humana, permitindo o aumento da produtividade e que pessoas poderão tomar decisões mais precisas, rapidamente. No entanto, mesmo que formalmente, reconhecem a necessidade de maior transparência.

As ferramentas de inteligência artificial, como vimos, são conduzidas por uma combinação complexa de algoritmos. Alguns defendem que colocar estes tipos de códigos e algoritmos complicados no domínio público para que todos possam inspecioná-los (auditá-los) provavelmente pouca utilidade poderá oferecer na aferição de responsabilidades pelo mal-uso da tecnologia. Vulnerabilidades conhecidas em determinados algoritmos, por exemplo, de código aberto, disponibilizados publicamente e amplamente utilizados e auditados, levaram anos para que pudessem ser identificadas. Além disso, conhecer o funcionamento de um fragmento de código algorítmico não oferece necessariamente condições para que seja possível realmente compreender todas as suas funções, a menos que as entradas dos algoritmos (seu processo de coleta de dados) também sejam auditáveis. Em outras palavras, a informação detalhada sobre o algoritmo, por si só, seria de pouco valor na compreensão de como o algoritmo realmente funciona, já que muitos deles conseguem acesso a um

conjunto enorme de dados que são de desconhecimento das pessoas, pois protegidos por cláusulas de confidencialidade, segredo industrial e outros mecanismos do gênero.

Muitos sistemas de aprendizado de máquina são verdadeiras “caixas pretas”, cujos métodos são difíceis de interpretar. Embora esses sistemas possam produzir resultados estatisticamente confiáveis, o usuário final não será necessariamente capaz de explicar como esses resultados foram gerados ou quais características específicas de um caso têm sido importantes para chegar a uma decisão final, causando desafios de interpretabilidade e transparência. Os algoritmos de aprendizado de máquina são apenas programas informáticos, e o alcance e a extensão do seu uso são extremamente amplos e extremamente diversos. Seria estranho, pesado e intrusivo sugerir governança para todos os usos de programação de computadores, e o mesmo argumento geral se aplica a todos os usos de aprendizagem de máquina.

Em muitos ou mais contextos, a aprendizagem por máquinas é geralmente incontroversa e não precisa de um novo quadro de governança. Como é o caso de uma empresa que usa o aprendizado de máquina para melhorar seu uso de energia ou instalações de armazenamento, o que não parece exigir mudanças significativas na governança. A aprendizagem de máquinas deve ser, obviamente, sujeita à lei, e também envolve o uso de dados pessoais coletados de forma apropriada e transparente.

Muitas das questões em torno dos algoritmos de aprendizagem de máquinas são muito específicas e complexas, por isso alguns defendem que seria inútil criar uma estrutura geral de governança ou um corpo de governança para todas as aplicações de aprendizagem em máquina. No entanto, as questões relacionadas com a segurança e os testes adequados em aplicações de transporte provavelmente serão melhor manipuladas pelos órgãos existentes nesse setor. Questões sobre validação de aplicações médicas de aprendizagem de máquinas por órgãos reguladores médicos existentes. Aqueles relacionados à aprendizagem de máquina em finanças pessoais pelos reguladores financeiros, e assim por diante.

Deve-se reconhecer que a explicação de algoritmos, incluindo novos algoritmos de aprendizagem profunda, em virtude de sua amplitude e profundidade, deve ser elaborada por pesquisadores e desenvolvedores, não só por pressões regulatórias. Há quem defenda, como vimos, que os padrões de transparência algorítmica não podem ser legislativamente definidos, já que as especificidades da tecnologia, algoritmos e sua aplicação variam muito.

Ainda na área jurídica, sistemas de raciocínio jurídico automatizado (com a aplicação de algoritmos e inteligência artificial) já foram considerados no passado. Mas verificou-se que é impossível a sua aplicação autônoma no Direito. A formação da convicção de um magistrado não está restrita somente a dados objetivos. A aplicação das leis é muito mais do que a aplicação de um conjunto de regras e jurisprudência. A independência de cada magistrado na formação de sua convicção, em cada caso

concreto, jamais pode ser ameaçada ou influenciada equivocadamente por máquinas. Os juízes não aplicam a lei como robôs. A hermenêutica e a interpretação exercem grande influência no Direito. Os algoritmos ainda não conseguem imitar o raciocínio jurídico. É claro que poderão auxiliar muito os juízes, advogados, promotores e demais profissionais da Justiça. Mas jamais substituir o elemento humano na equação. Os algoritmos são ferramentas e não devem ser tratados como substitutos completos para o julgamento humano.

Como vimos, os algoritmos são treinados e trabalham com um conjunto gigantesco de dados que os alimentam. Reitera-se: se esses dados são tendenciosos ou direcionados propositalmente, os algoritmos serão impactados e os sistemas tornar-se-ão involuntariamente parciais devido ao conjunto de regras que eles seguem. Até mesmo os desenvolvedores dos algoritmos muitas vezes não conseguem compreender a lógica que os algoritmos por eles mesmos desenvolvidos começaram a seguir para tomar certas decisões e produzir determinados resultados, já que nem todos os processos, inclusive as fontes e formas de coleta de informação são transparentes para os desenvolvedores.

Não sou misoneísta, os algoritmos mudaram a forma como o mundo funciona. Sem eles não teríamos a Internet, computadores e vários avanços tecnológicos que beneficiam a sociedade. E que beneficiarão ainda mais. Mas não podemos ser ingênuos. A velocidade com que os algoritmos processam uma grande quantidade de dados pode dar uma falsa impressão de exatidão para os menos atentos. Precisamos estar cientes de que os algoritmos estão longe de serem infalíveis e muito mais longe ainda de estarem livres da influência humana, de governos, interesses políticos, militares, econômicos e vários outros que você já consegue imaginar.

A regulação, como sugerida por alguns, parece ser utópica. Impossível regulamentar um segmento que sofre modificações impressionantes a cada dia. O que sabemos, é que os algoritmos, embora "vendidos" por essas grandes empresas como sendo imunes ao viés político, tendencioso, aos interesses econômicos, políticos, militares, estratégicos, de inteligência e tudo mais – na realidade não são. É difícil provar a afirmação dessas empresas (que subestimam a inteligência de alguns) sem uma supervisão significativa dos algoritmos complexos, muitas vezes secretos, incluindo o conjunto de dados, a forma e as fontes de coleta que consideraram para chegar a determinados resultados. É necessária uma supervisão constante dos processos algorítmicos mais sensíveis, como os que impactam a formação de opinião, por exemplo.

Uma das conclusões parciais do inquérito inglês é que os algoritmos podem replicar aspectos prejudiciais e discriminatórios da tomada de decisão humana e potencialmente criar novos tipos de danos. Por conseguinte, são necessários mais

trabalhos para identificar e prever novos tipos de danos e desenvolver métodos que detectem possam detectá-los de forma eficaz na medida em que ocorrerem. Sem o desenvolvimento dessas áreas, indivíduos e grupos afetados têm pouca perspectiva de retificar os erros quando eles ocorrerem.

A auditoria algorítmica na inteligência artificial é uma dessas áreas que deve exigir muita atenção dos especialistas em Direito, ética, cientistas de dados, formuladores de políticas e órgãos reguladores.

Abordagens atuais para auditar a funcionalidade e impacto dos algoritmos de decisões em áreas sensíveis incluem soluções complexas e avançadas, como a criação de funções de alerta incorporados nesses sistemas e algoritmos. A auditoria pode criar um registro processual para demonstrar viés ideológico, político, estratégico, preconceitos contra indivíduos (e grupos particulares) e ajudará os controladores de dados a atender os requisitos de responsabilidade, detectando quando as decisões produzem efeitos prejudiciais, explicando como eles ocorreram e em que condições eles podem ocorrer de novo. No entanto, a auditoria não é viável sem um forte apoio regulamentar ou cooperação de prestadores de serviços. Talvez sejam, no futuro, obrigados por decisões judiciais (com apoio em legislação específica) a colaborar. Para que a tomada de decisão algorítmica seja significativa, responsável, transparente e justa, os controladores de dados devem cooperar em auditorias. Da mesma forma, devem ser realizados mais trabalhos para entender como ocorre a governança nesses processos. No inquérito inglês, recomenda-se a criação de um grupo de trabalho para determinar se há necessidade da criação de um órgão específico de supervisão e quais os requisitos e os termos de referência a serem adotados por esses órgãos.

Mesmo que seja compreensível que os controladores de dados tenham um interesse legítimo em não divulgar segredos comerciais, esse interesse não deve superar o desejo de entender como os algoritmos tomam decisões sobre os seres humanos. Para equilibrar esses interesses concorrentes, governos devem explorar a necessidade da criação de um *Watchdog (órgão fiscalizador)* em Inteligência Artificial, ou outro organismo regulador de confiança e totalmente independente. Qualquer um desses órgãos precisaria ser equipado com os conhecimentos adequados (de direito, ética e informática), recursos e autoridade de auditoria (para fazer inspeções) - tudo para garantir que a tomada de decisão algorítmica seja justa, imparcial e transparente. Bem como os resultados apresentados aos cidadãos.

Há um grande potencial para que os algoritmos e a inteligência artificial possam ser usados para o bem de toda a sociedade. Na verdade, há uma oportunidade para tornar o mundo mais justo e menos tendencioso com a utilização dos algoritmos e da inteligência artificial. As leis não devem travar e dificultar a inovação, mas, reitera-se, não podemos ser ingênuos e deslumbrados com as novas tecnologias, deixando de perceber o que está por trás de tudo isso.

(*) **Paulo Sá Elias**, 46 anos, é advogado e professor universitário. Especialista em Direito da Informática e Tecnologia da Informação. Mestre em Direito pela UNESP. Mantém o site: www.direitodainformatica.com.br e o perfil da rede Twitter @paulosaelias