Inteligência artificial e abordagem via risquificação

Paola Cantarini

Doutora, (Direito e Filosofia, PUC-SP); Pós-Doutora (Direito, FD-USP) – Coordenadora. Pós doutora em Direito, Sociologia Jurídica, Filosofia, Arte e Pensamento Crítico; Doutora em Direito, Filosofia do Direito e em Filosofia; pesquisadora da Cátedra Oscar Sala, do Instituto Alan Turing, do Advanced Institute of IA, pesquisadora C4AI - Centro de Inteligência Artificial, Presidente e Pesquisadora no EthicAI - Grupo de Pesquisa em Inteligência Artificial. Membro Comissão de IA da OABMG. Pós-doutoranda em Ciências Sociais - PUCSP.

INTRODUÇÃO

Vivemos na sociedade da informação, sociedade de dados e sociedade 5.0 (Japão), atrelada aos conceitos de pós-humanismo e de transumanismo, falando-se ainda em "virada do não humano", um conceito macroscópico segundo Grusin, trazendo repercussões sociais de alta magnitude[1], com foco no descentramento do humano da biosfera, para se tornar verdadeira força geológica, a provocar a era do antropoceno.

Surgem ao mesmo tempo novos desafios e oportunidades com as novas tecnologias na interface com as humanidades, em especial com a utilização da chamada inteligência artificial (IA), sendo certo que as diretrizes éticas devem ir de mãos dadas com as questões legais, no âmbito da governança de algoritmos.

O Projeto de Lei 21/20 que cria o marco legal do desenvolvimento e uso da Inteligência Artificial (IA) é uma importante iniciativa de regulamentação da IA no Brasil, ao lado da Estratégia Brasileira de IA no Brasil, Instituída pela Portaria MCTI nº 4.617, de 6 de abril de 2021, apesar de algumas falhas e omissões, imprecisões técnicas, ausência de obrigações substantivas e processuais, ausência de parâmetros mínimos de procedimentalização e previsão de instrumentos de governança algorítmica, em especial se comparamos com as regulamentações internacionais. Há também uma falha a ser

[1] Grusin, Richard. Introduction. In: Grusin, Richard (org). *The nonhuman turn.* Minneapolis, mn: University of Minnesota Press, 2015. p. vii-xxxi, Grusin, 2015, Conferência realizada em 2012, "A Virada do Não Humano nos Estudos do Século XXI", Center for 21st Century Studies, Universidade de Wisconsin-Milwaukee.

destacada no tocante ao processo democrático de deliberação, já que houve um curto período de tempo para contribuições por parte da sociedade civil, ao contrário, por exemplo, do Marco Civil da Internet, Lei n° 12 965/2014, o qual contou com um período bem mais extenso de discussão democrática e inclusiva.

Um amplo período de debate envolvendo diversos grupos da sociedade civil é essencial e possui relação com o conceito de ética digital intercultural, trazendo ao diálogo os grupos vulneráveis e todos os setores da sociedade. Justamente ética digital intercultural e o estabelecimento de frameworks concretos para tradução de princípios éticos abstratos em práticas concretas são pontos a serem desenvolvidos, devendo contar com a contribuição de uma equipe interdisciplinar e *multistakeholder*, e, sobretudo, independente.

O Projeto de Lei 21/20 que cria o Marco Legal da IA no Brasil é uma importante iniciativa no sentido de regulamentação da IA, já que cada vez mais se fala no fim da era dos códigos de conduta (autorregulação), como bem aponta Luciano Floridi, no recente artigo "The end of an era: from self-regulation to hard law for the digital industry".[2]

Isto porque a autorregulação pelas empresas, não seria eficaz nem tampouco contribuiria para o aspecto da confiança, já que muitas vezes tal iniciativa colide com a busca de fins públicos e com a proteção de direitos fundamentais e humanos, voltando-se primordialmente para os valores de mercado, não sendo iniciativas pautadas na transparência e imparcialidade. Em muitos casos há aqui uma concepção proprietária dos diretos envolvidos, a busca da inovação e de valores econômicos acima de outros valores democráticos, envolvendo a elaboração de conteúdo unilateral e seletivo em termos de interesses, na linha de uma análise econômica do Direito, voltada para eficiência do mercado.

DESENVOLVIMENTO: IMPORTÂNCIA DO DEBATE DEMOCRÁTICO E INCLUSIVO E DE UMA ABORDAGEM VIA RISQUIFICAÇÃO

A maioria dos membros da diretoria das empresas entende que não estão preparados para lidar com questões de ética da IA, sendo necessário o estabelecimento de parcerias e alianças de forma colaborativa diante da complexidade das questões a serem

[2] Floridi, Luciano, *The end of an era: from self-regulation to hard law for the digital industry* (November 9, 2021). Available at SSRN: https://ssrn.com/abstract=3959766 or http://dx.doi.org/10.2139/ssrn.3959766

observadas, além do seu caráter de transversalidade, aproximando campos científicos os mais diversos, não jurídicos e disciplinas "transclássicas" (semiótica, cibernética, teoria de sistemas).

Recente pesquisa mostrou que 81% dos consumidores se sentem mais preocupados com a forma de tratamento de dados pessoais por parte das empresas, e 75% estão agora menos propensos a confiar às organizações seus dados pessoais[3].

Com o crescente avanço da utilização da IA nas diversas áreas de negócios, invadindo todos os aspectos de nossas vidas, com repercussões até mesmo na concepção de tempo, espaço, cultura e subjetividades, torna-se urgente o comprometimento com o requisito da confiança por parte das empresas que atuam com novas tecnologias, e neste sentido a adoção de boas práticas, práticas de *compliance* e de governança se tornam essenciais e um diferencial de mercado. A análise da ética e das regulamentações em tais áreas é tida como um diferencial competitivo das empresas, pois envolve a confiança e a transparência necessárias em qualquer relação jurídica.

Kai-Fu Lee aponta que as pessoas tendem a confiar em três principais fontes quando se trata de estudos acerca da IA: ficção científica, notícias na mídia e pessoas influentes. Tal observação revela a necessidade de mudança de mentalidade, de mudança de *mindset*, sendo essencial a aproximação das contribuições científicas, da área acadêmica, das empresas e das demais áreas[4]. No Brasil, da mesma forma que estamos ainda construindo uma nova cultura de proteção de dados, com respeito ao princípio da minimização, em especial, é urgente uma nova cultura e mentalidade também acerca da ética em IA, já que algumas pesquisas apontam para o baixo grau de preocupação neste setor do Brasil, ou seja, de apenas 15%, como também no resto da América Latina[5], em comparação com os EUA, por exemplo (63%), e Europa (47%). Da mesma forma, ainda é incipiente a regulamentação jurídica e propostas de certificações nesta seara no Brasil.

---

[3] Unpublished data from the 2018, *IBM Institute for Business Value Global Consumer Study.* IBM Institute for Business Value; Advancing AI ethics, beyond compliance From principles to practice, Brian Goehring, Francesca Rossi, Dave Zaharchuk

[4] Lee, Kai-Fu, *AI 2041: Ten Visions for Our Future*, Publisher Currency, 2021.

[5] Count is less than 20. Source: 2018 *IBM Institute for Business Value Global AI Ethics Study.* Q: Importance of AI ethics in your organization, N=1,247; Advancing AI ethics, beyond compliance From principles to practice, Brian Goehring, Francesca Rossi, Dave Zaharchuk.

Como alternativas à heterorregulação, devido à demora pelo Estado e falta de *expertise*, e à autorregulação, cada vez mais vem se destacando a autorregulação regulada, com a previsão de deveres procedimentais, havendo certo controle por meio do estabelecimento de requisitos e parâmetros mínimos de governança prefixados pelo Estado, a exemplo do "IA Act" da União Europeia, a depender do grau de risco a direitos fundamentais e liberdades fundamentais por parte da tecnologia, trazendo, pois uma clara abordagem via risquificação. Neste sentido, a prática do *compliance* relacionada à IA deverá se adequar à probabilidade e gravidade do risco (*risk-based approach*), por meio de um processo proativo, sistemático e contínuo de proteção aos direitos envolvidos, com destaque para a elaboração de relatório de impacto algorítmico e relatório de direitos humanos e fundamentais. Referidos documentos devem ser considerados obrigatórios, em especial no caso de aplicações de risco moderado e alto, bem como devem ser orientados para as características especiais do *Big Data* e para a proteção dos interesses públicos afetados pela sua utilização. Deverá haver uma avaliação permanente dos riscos e envolver um controle público com representantes da sociedade civil, permitindo a participação de setores vulneráveis.

Tais documentos amparam-se no princípio da precaução, o qual é considerado como um referencial capaz de mensurar e catalogar as salvaguardas necessárias para aplicações de IA de alto risco e de risco moderado. Trata-se de um *framework* importante para as aplicações de AI, ao lado das Avaliações de Impacto sobre Direitos Humanos/Direitos Fundamentais. Estes últimos documentos possuem previsão em algumas iniciativas, envolvendo uma mudança de foco: da ética na IA para um discurso com a moldura e o vocabulário dos Direitos Humanos/Direitos Fundamentais, bem como focando na proteção coletiva, e não apenas individual. Deverá haver um amplo escrutínio público, e envolver um processo de deliberação pública, com revisão por organizações ou consultores externos independentes com expertise em direitos fundamentais.

Neste aspecto cumpre ressaltar uma falha do PL ao prever apenas um artigo mencionando de forma genérica o Relatório de Impacto de IA e a adoção de padrões e de boas práticas (art. 13), sem maiores detalhamentos, bem como por prever que tal documento poderá ser solicitado pelo Estado, não tornando o mesmo obrigatório em todos os casos de aplicações de IA que envolvam um alto risco ou risco moderado, nem tornando obrigatória sua elaboração de forma preventiva, ou seja, antes do início da aplicação, quando do desenvolvimento do produto ou serviço. Este já é um problema

encontrado na LGPD – Lei Geral de Proteção de Dados ao trazer a possibilidade de diversas interpretações pela doutrina e jurisprudência no tocante à obrigatoriedade ou não do relatório de impacto de proteção de dados, bem como relativamente ao momento de sua elaboração e de quem seria a obrigação de sua elaboração, já que ao prever que tal documento poderá ser solicitado pela ANPD, poderá dar ensejo à uma interpretação literal e gramatical no sentido de não ser obrigatório, a não ser em caso de solicitação pela ANPD – Autoridade Nacional de Proteção de Dados. Tal intepretação, contudo, fere toda a lógica da sua realização qual seja, de englobar todo o ciclo do tratamento de dados, e do desenvolvimento da tecnologia, desde o início, bem como o princípio da prevenção de danos, não passando pelo crivo de uma análise funcional e sistemática da LGPD, contrariando também o entendimento do GDPR – Regulamento Geral de proteção e Dados da União Europeia, (RGPD) (UE) 2016/679, o qual foi o principal marco teórico orientador da LGPD, assim como o entendimento e diversas ANPD de diversos países, e orientações de órgãos consultivos como o WP29 - The Article 29 Working Party, e o EDPB - European Data Protection Board. Há pelo PL uma expressão muito vaga e genérica relacionada à elaboração do relatório de impacto, condicionando sua exigência "a justificação de sua necessidade", sem maiores comentários ou especificações, contribuindo para a insegurança jurídica e para uma proteção de nível fraco quanto aos direitos fundamentais envolvidos.

Não há também no PL qualquer previsão acerca de níveis de potencial dano das aplicações de IA ao contrário de diversas regulamentações da EU neste sentido, o que contribuiria para a abordagem baseada em risco.

Diante da insuficiência da heterorregulação, diante da possibilidade de lavagem ética por outro lado, se tem apostado na autorregulação regulada, sujeitando as boas práticas e códigos de conduta, certificações a diversas precauções materiais e processuais, por lei e tratados internacionais, trazendo incentivos para uma melhor concepção tecnológica, a abertura do acesso e possibilidades de certificações. É essencial o envolvimento de representantes da sociedade civil a fim de democratizar a discussão, inclusive com poderes para controlar o cumprimento dos compromissos voluntários por parte das empresas.

Na linha da proteção pela técnica, com a utilização de remédios tecnológicos teríamos a utilização de ferramentas tecnológicas de governança na própria construção dos sistemas de decisão automatizada, de forma a dar efetividade, por exemplo, ao direito

de revisão de decisões automatizadas, através de uma abordagem preventiva (*privacy by design* (PbD). Tal perspectiva traria a obrigação de respeito aos direitos fundamentais como um objetivo central do processo de construção de software, devendo ser observada durante todo o ciclo de vida do sistema, como um requisito para a viabilidade de tal projeto, a exemplo do disposto no artigo 25 da GDPR.

A recente proposta de 04.2021 da EU segue a ótica de uma regulamentação via risquificação, traçando uma análise de risco e separando em diversos patamares e níveis de risco as aplicações de IA, de alto risco, moderado-limitado, baixo risco e risco inaceitável, envolvendo aplicações que jamais deveriam ser desenvolvidas. Fundamenta-se no incremento de regulamentações "ex ante" com base no princípio da precaução, como códigos de conduta, certificações, auditorias independentes, elaboração de documentos como DPIA – Relatório de Impacto de proteção de dados e LIA – Avaliação do Legítimo interesse, e na área da IA, relatório de impacto algoritmo, ou relatórios de direitos fundamentais e humanos, seguindo-se a ótica do direito regulatório e do direito ambiental, já presente na área de proteção de dados, como se observa da GPDR, da LGPD. Referida abordagem está também presente nas propostas europeias de regulamentação da IA, com destaque para o "White paper On Artificial Intelligence - A European approach to excellence and trust", de 19.02.2020, com foco em um prisma mais complexo do direito regulatório, envolvendo técnicas de prevenção e mitigação de riscos a direitos e liberdades fundamentais, bem como, apontando-se para uma preocupação com a proteção não apenas individual, mas também coletiva e social.

Segundo o "White paper on IA", é considerado um alto risco, quando a utilização envolver riscos significativos, em especial, com relação à proteção da segurança, dos direitos dos consumidores e dos direitos fundamentais. Uma aplicação de IA deverá ser considerada de alto risco se preencher os dois critérios cumulativos: a aplicação de IA é utilizada num setor em que, dadas as características das atividades tipicamente realizadas, se pode esperar que ocorram riscos significativos. A avaliação do nível de risco de uma determinada utilização poderá basear-se no impacto nas partes afetadas.

São consideradas como de alto risco as aplicações de IA para os processos de recrutamento, situações que afetem os direitos dos trabalhadores, bem como para efeitos de identificação biométrica à distância e de outras tecnologias de vigilância intrusivas. Como previsão de regulação "ex ante", referido documento exige que os sistemas devam

ser tecnicamente robustos e exatos, sendo desenvolvidos de forma responsável, mediante uma análise prévia de riscos. No caso de aplicações de baixo risco há a previsão de um regime de rotulagem voluntária, optando por vincular-se aos requisitos legais, ou a requisitos semelhantes, especialmente criados nesta área, condição para o recebimento de um selo de qualidade para as aplicações de IA, confirmando que determinada empresa estaria em conformidade com determinados padrões objetivos.

Por sua vez, segundo a nova proposta regulamentadora do AI Act de 2021, são considerados riscos inaceitáveis no caso de uma possível ameaça clara aos cidadãos europeus como aplicações que possam manipular comportamentos, opiniões e emoções humanas, principalmente tendo em vista setores vulneráveis da população, como ocorre com brinquedos com assistentes de voz, por manipular mais facilmente crianças e adolescentes. Também o recurso a sistemas de pontuação social por parte de governos, a exemplo da China é visto como um risco inaceitável. Em tais casos há a proibição expressa da utilização da aplicação de IA. Também são consideradas como de alto risco, a utilização de algoritmos na área de seleção e recrutamento, avaliações de solvabilidade, distribuição de benefícios da segurança social ou pedidos de asilo e vistos, ou ajudar os juízes a tomar decisões.

Há expressa proibição do uso de tecnologias de vigilância, com exceção da utilização por órgãos governamentais para a prática de investigação de crimes graves, como no caso da utilização do reconhecimento facial no combate ao terrorismo. Em se tratando de risco elevado, como aplicações de IA nas áreas de transportes, educação, produtos de segurança, recrutamento e acesso a emprego, serviços essenciais ou públicos, reforço da lei ou em situações de migrações, concessão de asilo ou controlo de fronteiras é possível sua utilização, mas desde que sujeita à observância de regras e obrigações rígidas. Os sistemas de reconhecimento facial e identificação biométrica são considerados como de risco elevado, proibindo-se como regra sua utilização em espaço público, salvo exceções "rigidamente definidas e reguladas", sob a condição de prévia autorização judicial, limitando o tempo, a localização e os dados utilizados. A proposta também pretende combater práticas discriminatórias, os denominados vieses, ou “bias”, propondo que os conjuntos de dados não "incorporem quaisquer preconceitos intencionais ou não intencionais" que possam levar à discriminação.

Como aplicações de IA com risco mínimo e limitado, temos, por exemplo, a utilização de “chatbots”, bem como o uso gratuito de jogos ou filtros de spam com inteligência artificial.

Seguindo-se a abordagem da risquificação, há a definição de obrigações vinculando-se a uma avaliação adequada do risco e mitigação dos sistemas, um maior controle na qualidade dos conjuntos de dados que alimentam os sistemas, registro das atividades de forma a garantir a rastreabilidade dos resultados, documentação detalhada, sendo obrigatória a supervisão humana.

A recente proposta de regulamentação da IA pela Comissão Europeia (AI Act de 04.2021) reflete, pois, a análise entre diversas possibilidades regulatórias do setor, e de articulação com a já existente legislação setorial europeia, com foco na “GDPR – Regulamento Geral de Proteção de Dados da EU”, no “Digital Services Act”, no “Digital Markets Act”, no “White paper on IA” e no “Regulamento relativo à responsabilidade civil pelo uso da IA”, em preparação. A Comissão Europeia entende, em suma, que a nova regulamentação é imprescindível para se possibilitar a inovação tecnológica e os progressos científicos, garantindo a necessária vantagem competitiva e liderança tecnológica da UE, em um contexto de forte concorrência mundial, mas sem deixar a preservação de direitos fundamentais e humanos, e de valores básicos consagrados pela EU, colocando a tecnologia a serviço dos cidadãos europeus.

A proposta de regulamentação da IA via IA ACT de 2021 segue a estratégia europeia para a IA apresentada em 04/2018 denominada “Inteligência artificial para a Europa” (COM/2018/237), com foco nos valores europeus como forma de enfrentamento dos novos desafios da IA. Diante da ausência de regulamentações pela maior parte dos países, inclusive por parte dos países da UE, a não ser em casos pontuais para certos setores de aplicação da IA, como algumas regulamentações esparsas, diante de tal vácuo normativo, visando evitar fragmentações e antinomias diante de iniciativas pontuais isoladas, a Comissão Europeia, braço executivo da União Europeia, apontou para a necessidade de uma regulamentação geral para todos os países da UE, de forma a promover o desenvolvimento da IA e ao mesmo tempo enfrentar os riscos potencialmente elevados para a segurança e os direitos fundamentais e humanos. Um dos principais objetivos a ser destacado, é a possibilidade da UE tornar-se economicamente competitiva em tal setor, disputando o mercado atualmente dominado pelos EUA e China principalmente, concorrendo de igual para igual com tais países, tendo apresentado, neste

sentido, em 21.04.2021, uma proposta de regulamentação do uso da inteligência artificial na União Europeia, além da previsão de investimentos e financiamento no setor de aproximadamente 20 bilhões de euros por ano na próxima década para o desenvolvimento de tecnologias que utilizam Inteligência Artificial.

Trata-se do denominado "efeito Bruxelas" das regulamentações da EU na área digital. O impacto da nova regulamentação da União Europeia vai muito além das fronteiras dos países membros, o que já ocorre com a GDPR, modelo que inspirou demais iniciativas legislativas em outros países, a exemplo da Lei Geral de Proteção de Dados Pessoais no Brasil (LGPD), Lei 13.709/18, bem como o Digital Services Act (DAS), o Digital Markets Act (DMA) e a regulamentação do discurso de ódio e fake News, com o estabelecimento do Código de Conduta voluntário da Comissão sobre o Combate à Ilegalidade Hate Speech Online. O efeito Bruxelas ("Brussels Effect") seria uma manifestação da europeização do ambiente regulador global, no sentido de eficácia extraterritorial do direito europeu e a influência e impacto mundial de sua regulamentação[6]. Trata-se do poder de influência mercadológica global dos regulamentos da EU, tendo sido majorado significativamente nas últimas duas décadas, particularmente na área da economia digital.[7]

O PL 20-21 por sua vez traz princípios, direitos, deveres, mas quase nenhuma previsão acerca de instrumentos de governança algorítmica. Contudo, poderá ser considerada uma abordagem relacionada à risquificação, na linha da regulamentação da União Europeia, com destaque para o Regulamento de IA (AI ACT), abandonando a anterior ideia de atribuição de personalidade eletrônica, tal como era prevista na Resolução do parlamento europeu com orientações de Direito Civil sobre Robótica, ao prever que de um lado há a preocupação na proteção dos Direitos Humanos e Direitos Fundamentais, e princípios, valores democráticos, fundamentos específicos do PL (art. 4), ao lado da não discriminação, pluralidade, livre iniciativa e proteção de dados, e de outro de forma a não obstar a inovação. Ambas as propostas tentam trazer um equilíbrio entre proteção aos direitos fundamentais de um lado, enfrentamento dos riscos potenciais, e de outro lado, a promoção do desenvolvimento da IA, tornando o país economicamente competitivo em tal setor, não trazendo obstáculos à inovação. Portanto, o PL na linha do

[6] Disponível em https://valor.globo.com/opiniao/coluna/efeito-bruxelas-atinge-big-techs.ghtml. Acesso em 05/04/2021.
[7] Bradford, Anu. *The Brussels Effect.* Oxford University Press, 2020.

AI Act da EU segue a perspectiva de “human rights by design”, “beneficial AI”, “AI for good” e “HumanCentered AI”, ou seja, visa-se trazer um balanceamento entre o desenvolvimento tecnológico, de modo a não obstar a inovação, de um lado, e a proteção dos valores democráticos, direitos humanos e fundamentais, de outro lado. Parte-se da abordagem “centrada no ser humano", trazendo o eixo valorativo da pessoa humana e da dignidade humana, respeitando o estímulo e desenvolvimento da inovação e tecnologia, mas desde que haja, por outro lado, o respeito aos direitos fundamentais e aos direitos humanos.

O PL 20-21, contudo, da mesma forma que a Estratégia Brasileira de IA vem sendo bastante criticado por não ser talvez a melhor proposição em especial no que tange a falta de *enforcement* ao não prever sanções, não especificando questões de governança de algoritmos, não trazendo critérios para a autorregulação regulada, não sendo muito preciso quanto ao sistema de responsabilidade civil adotado, não contribuindo para a necessária segurança jurídica, trazendo um desequilíbrio em tais relações jurídicas já assimétricas, no caso de prevalecer o entendimento acerca da responsabilidade civil subjetiva como regra, senão vejamos. Neste caso, teríamos um ônus demasiado ao cidadão comum sem conhecimento técnico específico, tornando a produção de uma prova diabólica, já que excessivamente difícil ou impossível de ser produzida. Assim como ocorre com o CDC onde há a presunção da vulnerabilidade do consumidor, diversos autores vêm apontando para tal vulnerabilidade também na seara da proteção de dados, e ainda em maior escala quando se fala de *big data*, falando-se em uma hipervulnerabilidade, sendo que na ótica consumerista teríamos a responsabilidade objetiva (artigos 12 e 14), assim como no direito ambiental.

O PL traria algumas definições simplistas e não unânimes, tais como o conceito de IA, trazendo certa falta de técnica legislativa, não sendo preciso ou detalhista, da mesma forma ao prever os autores envolvidos em tais relações jurídicas. Não traz padrões mínimos de governança vinculantes, sendo muito fluido, trazendo possibilidade de interpretações heterogêneas, o que poderia desprestigiar o investimento no país, bem como o próprio desenvolvimento de negócios e da tecnologia, não contribuindo para a aplicação de instrumentos na linha da autorregulação regulada.

Entre os críticos ao PL destaca-se a carta aberta endereçada ao Senado assinada por diversos juristas, com destaque para Ana Frazão, Anderson Schreiber, Bruno Bioni,

Bruno Miragem, Caitlin Sampaio Mulholland, Danilo Doneda, Gustavo Tepedino, Ingo Wolfgang Sarlet, Laura Schertel Mendes, Maria Celina Bodin de Moraes, Milena Donato Oliva, Rafael Zanatta. Em sentido contrário, destaca-se a importante contribuição do jurista Juliano Maranhão[8] apontando para a falha em se atribuir a responsabilidade objetiva como regra no PL, por inviabilizar a inovação e investimentos no país, bem como apontando para a falta de detalhamento de regras de condutas específicas para reafirmarem ou tornarem efetivos os princípios trazidos, pois caso contrário o PL seria contraproducente, tornando impossível a aplicação e implementação dos princípios, trazendo insegurança jurídica e inviabilizando o desenvolvimento da tecnologia no país. Destaca a importância de se diferenciar aplicações de alto risco, de médio e de baixo risco, diferenciando a regulamentação conforme o grau de risco da aplicação da IA.

A carta dos juristas traz uma importante crítica ao artigo 6º, inciso VI, do PL por trazer no seu entender o estabelecimento da responsabilidade civil subjetiva como regra, inviabilizando a produção probatória por parte das vítimas de danos causados por Inteligências Artificiais, já que em muitos casos se tornaria impossível a produção probatória, e, por consequência, comprometendo a garantia dos direitos fundamentais[9].

Onde atualmente consta:

> Artigo 6º: VI – responsabilidade: normas sobre responsabilidade dos agentes que atuam na cadeia de desenvolvimento e operação de sistemas de inteligência artificial devem, salvo disposição legal em contrário, se pautar na responsabilidade subjetiva, levar em consideração a efetiva participação desses agentes, os danos específicos que se deseja evitar ou remediar, e como esses agentes podem demonstrar adequação às normas aplicáveis por meio de esforços razoáveis compatíveis com padrões internacionais e melhores práticas de mercado.

Na carta há a sugestão de alteração do artigo para a seguinte redação proposta:

> Artigo 6º: VI – responsabilidade: normas sobre responsabilidade dos agentes que atuam na cadeia de desenvolvimento e operação de sistemas de inteligência artificial devem, salvo disposição legal em contrário, levar em consideração a tipologia da inteligência artificial, o risco gerado e seu grau de autonomia em relação ao ser humano, além da natureza dos agentes envolvidos, a fim de se determinar, em concreto, o regime de responsabilidade civil aplicável.

---

[8] *Marco Legal da Inteligência Artificial* – Conversações, 18.11.21, Cátedra Oscar Sala, artigo; https://politica.estadao.com.br/blogs/gestao-politica-e-sociedade/o-debate-sobre-o-marco-legal-da-inteligencia-artificial-no-brasil/.

[9] https://www.conjur.com.br/2021-out-27/especialistas-questionam-artigo-pl-marco-legal-ia.

Há bastante divergência na doutrina nacional e internacional acerca de qual sistema de responsabilidade civil adotar, alguns apostando na melhor proteção dos direitos fundamentais e humanos envolvidos, via responsabilidade objetiva, na linha do direito ambiental e consumerista, responsabilidade objetiva pelo fato da coisa – no caso de máquinas ou robôs que tomem decisões – ou também pelo risco. Os agentes econômicos que auferem lucros altíssimos ainda mais no sistema do *big data* com o superávit comportamental[10], ou seja, amparados na enorme quantidade de dados pessoais obtidos de forma gratuita, deveriam também ser responsáveis pela criação e supervisão da tecnologia, de forma a minorar os riscos assumidos, com imposição do dever de diligência e cuidado dos gestores.

A previsão do regime de responsabilidade civil para aplicações em IA envolve a questão de quem responderá por danos causados por um robô, o fabricante, o programador, ou a empresa que celebra contrato diretamente com o consumidor? Quais os riscos da previsão de uma personalidade jurídica eletrônica (*epersonality*)?

As regulações da EU mais recentes (White paper, AI Act) abandonam a ideia de imposição de uma personalidade eletrônica no caso de aplicações de IA, pois ao invés de facilitar a indenização das vítimas estava contribuindo para uma maior dificuldade, sendo objeto de críticas por trazer a naturalização da ideia de inteligência, envolvendo as questões da falácia androide e retórica antropomórfica. Já a anterior Resolução do Parlamento europeu de 20.10.2020 trazendo recomendações à Comissão Europeia sobre o regime de responsabilidade civil para aplicações de IA, traz em seu considerando 7 a afirmação de não ser mais necessário conferir personalidade jurídica aos sistemas de IA.

Com o estabelecimento da personalidade eletrônica poderia ocorrer a excessiva valorização da autonomia da Ia, podendo dar ensejo a alguma excludente de responsabilidade – caso fortuito ou força maior, considerando a IA a única responsável pelo dano causado. Haveria certa dificuldade de se extrair o estabelecimento da relação causal entre os danos e a atividade humana, em razão do grau de autonomia do robô. Trata-se de um juízo de merecimento de tutela a ser realizado casuisticamente em um futuro ainda desconhecido, a depender de um grau de autonomia da IA hoje inexistente.

[10] Zuboff, Shoshana, *Capitalismo de vigilância. A luta por um futuro humano na nova fronteira de poder*. Editora Intrínseca, 2021.

Anteriormente, a Resolução com recomendações sobre regras de Direito Civil e Robótica, de 16.02.2017 (2015/2103/INL) adotada pelo Parlamento Europeu, trazia uma recomendação para a criação de uma espécie de personalidade jurídica para robôs, isto é, de um estatuto jurídico específico, em longo prazo, de modo que pelo menos os robôs autônomos mais sofisticados pudessem ser detentores do estatuto de pessoas eletrônicas, responsáveis por danos. Seria o caso, por exemplo, de robôs que produzem decisões autônomas, ou em que interagem por qualquer outro modo com terceiros de forma independente.

A resolução mencionava ainda duas importantes iniciativas relacionadas ao desenvolvimento de robôs inteligentes:

(i) adoção de um registro obrigatório dos robôs;

(ii) criação de um seguro para fazer frente às hipóteses de danos causados pelos robôs.

A Resolução de 2017 trazia ainda uma lista enumerando quais aplicações de IA seriam consideradas como objeto da personalidade eletrônica, envolvendo a discussão se tal lista é exemplificativa ou taxativa por parte da doutrina; a lista previa as seguintes aplicações de IA: veículos autônomos, drones inteligentes, robôs assistentes de idosos ou enfermos e robôs médicos, algoritmos de processamento e análise de dados que possam causar práticas discriminatórias. Há a menção de que os robôs podem ser dotados de capacidades adaptativas e de aprendizagem que integram certo grau de imprevisibilidade no seu comportamento, uma vez que aprendem de forma autônoma, com a sua experiência própria vaiável e interagem com o seu ambiente de um modo único e imprevisível. Houve crítica por parte da doutrina acerca da exigência da construção de um vultoso patrimônio mínimo como condição para a operação de determinadas aplicações de IA, entendendo-se que poderia criar um entrave excessivamente oneroso ao desenvolvimento tecnológico, bem como monopólio de mercado, gerando uma vantagem competitiva para as *big five*, na mesma linha da discussão envolvendo a temática da proteção de dados e o princípio da minimização de dados ou da necessidade.

Destaca-se a previsão no parágrafo 57 da Resolução uma possível alternativa relativa à responsabilidade pelos danos causados por robôs, qual seja, o estabelecimento de um regime de seguros obrigatórios como já ocorre com os carros.

Ainda de acordo com o documento, o futuro instrumento legislativo deverá basear-se numa avaliação aprofundada da Comissão que determine se a abordagem a

aplicar deve ser a da responsabilidade objetiva ou a da gestão de riscos. Deverá ser criado um regime de seguros obrigatório, que poderá ter basear-se na obrigação do produtor de subscrever um seguro para os robôs autónomos que produz. O regime de seguros deverá ser complementado por um fundo a fim de garantir que os danos possam ser indenizados caso não exista qualquer cobertura de seguro.

É fundamental o diálogo e a colaboração via contribuições da sociedade civil ao PL, a fim de democratizar a discussão, e com isso encontrarmos uma maior proporcionalidade entre inovação e responsabilidade, por meio de audiências públicas, democratizando o debate, e dando maior legitimidade à discussão, dando voz a todos, em especial de grupos vulneráveis. A regulamentação não traz a necessária criação de um *framework* ético em IA que leve em consideração e seja sensível às diferenças culturais presente na sociedade brasileira, na linha do conceito de ética digital intercultural.

Como bem é apontado no livro "The rise of big data policing: surveillance, race, and the future of law enforcement"[11] a governança de algoritmos deveria se pautar em algumas questões essenciais, de modo a se evitar o determinismo tecnológico, tais como: é possível identificar os riscos que tecnologia escolhida está tentando endereçar? é possível defender os inputs do sistema (acurácia dos dados e idoneidade da metodologia)? É possível defender os outputs do sistema e como eles impactarão as políticas em prática e as relações comunitárias? É possível testar a tecnologia, oferecendo *accountability* e alguma medida de transparência? A política de uso da tecnologia respeita a autonomia das pessoas que elas irão impactar?

Destaca-se, como a proposta de Wolfgang Hoffmann-Riem, apontando para a importância dos direitos fundamentais no sentido de essencial para a proteção da autonomia, e no sentido de se compatibilizar, a proteção aos direitos fundamentais, princípios, responsabilização, e de outro lado, não impedir a inovação, a denominada "responsabilidade pela inovação", ou "innovation forcing"[12]. Trata-se da definição normativa de objetivos ou padrões que ainda não podem ser cumpridos sob o padrão de desenvolvimento atual, mas que são plausíveis de serem cumpridos no futuro. Caso não haja tal implementação dentro de determinado período o desenvolvimento e uso da aplicação de IA em questão devem ser abandonados. É o que destaca também Laura Mendes em sua apresentação ao livro: "o

[11] Ferguson, Andrew Guthrie. *The rise of big data policing: surveillance, race, and the future of law enforcement.* Nova Iorque: New York University Press, 2017.
[12] Hoffmann-Riem, Wolfgang. *Teoria Geral do Direito Digital*, Forense, ed. kindle, pp. 13-14; p. 150 e ss.

professor Hoffmann-Riem nos ensina que a preocupação com a preservação e atualização dos direitos fundamentais deve ser constante, enxergando o Direito como um instrumento de limitação de poderes e de regulação da inovação, de acordo com os objetivos e os valores firmados no ordenamento jurídico, especialmente, os princípios constitucionais". [13]

O PL traz como princípios em seu artigo 6º para o uso responsável de inteligência artificial no Brasil, destacando-se, pois como uma lei principiológica, na esteira da LGPD, a finalidade, com o fim de aumentar as capacidades humanas, e com isso reduzir as desigualdades sociais e promover o desenvolvimento sustentável; a centralidade no ser humano: respeito à dignidade humana, à privacidade e à proteção de dados pessoais e aos direitos trabalhistas; a não discriminação: impossibilidade de uso dos sistemas para fins discriminatórios, ilícitos ou abusivos; transparência e explicabilidade: garantia de transparência sobre o uso e funcionamento dos sistemas de inteligência artificial e de divulgação responsável do conhecimento de inteligência artificial, observados os segredos comercial e industrial, e de conscientização das partes interessadas sobre suas interações com os sistemas, inclusive no local de trabalho; segurança: utilização de medidas técnicas e administrativas, compatíveis com os padrões internacionais, aptas a permitir a funcionalidade e o gerenciamento de riscos dos sistemas de inteligência artificial e a garantir a rastreabilidade dos processos e decisões tomadas durante o ciclo de vida do sistema; responsabilização e prestação de contas: demonstração, pelos agentes de inteligência artificial, do cumprimento das normas de inteligência artificial e da adoção de medidas eficazes para o bom funcionamento dos sistemas, observadas suas funções.

Apesar da proposta via risquificação do PL, na linha da regulamentação europeia, trazendo a importante regulamentação de diversas aplicações de IA conforme o grau de risco potencial, no entender de Marck Coeckelbergh[14], apesar de apontar para a importância das questões fundamentalmente éticas, tais como, o que nós como sociedade entendemos ser importante, e que as novas tecnologias, em especial a IA, poderia contribuir para nos ajudar na compreensão de questões filosóficas essenciais, tal como "o que é o ser humano", "como queremos viver", "quais valores são importantes", tal como também aponta Kai-Fu Lee, fazendo uma importante ressalva quanto à previsão estática nos sistemas de regulamentação da IA que se fundamentam na análise de risco, a exemplo da União europeia (White paper on AI, regulação de 04-2021, da Recomendação do

[13] Mendes, Laura, *Ibidem*, p. 04 e ss.
[14] Coeckelbergh, Marck, 2º. Congresso de Ia da PUCSP-TIDD, palestra proferida em 17.11.21.

Conselho da Europa de 2010 e do IA ACT de 2021), ao tratar, especificamente de estabelecimentos fixos de diversos patamares e risco, alto, baixo, moderado quanto a aplicações de IA. No seu entender seria mais adequada uma abordagem mais flexível, pois a depender do contexto e do caso concreto uma aplicação, antes classificada como de baixo risco poderá se tornar de alto risco, e vice-versa.

Na abordagem via risquificação ocorre a reformatação jurídica a partir da ampliação da tutela coletiva, a disseminação de instrumentos regulatórios *ex ante* e o uso intensivo de metodologias de gestão de risco e calibragem entre riscos, inovações, aproximando-se de um processo de "negociação coletiva", trazendo semelhanças com o direito ambiental, como se percebe pela ideia de poluição de dados, ou seja, quando há um vazamento de dados não se está afrontando apenas direitos individuais do titular dos dados, mas muitas vezes valores democráticos, e o próprio Estado Democrático de Direito. Tal abordagem possui aproximação com a característica essencial de todo e qualquer direito fundamental, qual seja, sua múltipla dimensionalidade, ou seja, há um aspecto individual mas também coletivo e social dos direitos fundamentais. Exemplo de tal abordagem já estava presente no GDPR ao afirmar uma "identificação dos riscos relacionados com o tratamento", sua "avaliação em termos de origem, natureza, probabilidade e gravidade", bem como a "identificação das melhores práticas para atenuação dos riscos", os quais poderão ser obtidas por códigos de conduta aprovados, certificações aprovadas, e orientações profissionais fornecidas pelo encarregado pela proteção de dados pessoais.

A partir do reconhecimento da tríplice dimensão ou multidimensionalidade de todo Direito Fundamental, há o reconhecimento dos seus aspectos individual, coletivo e social, já que relacionados à cidadania e à igualdade material dos tutelados. Trata-se do reconhecimento de que um vazamento de dados ocorre como se fosse um sistema de poluição de dados, afetando não apenas o titular, mas causando danos coletivos e sociais, devendo haver uma conjugação das formas de responsabilização *ex post* e *ex ante*. Tal perspectiva também envolve a consideração que o impacto dos sistemas de IA não se limita a aspectos individuais, devendo envolver uma perspectiva coletiva e social, isto é, os problemas relativos ao capitalismo de vigilância e, pois, ao *big data* são coletivos. Nesta linha de abordagem, a utilização de sistemas de IA é reconhecida com um potencial de desempenhar um importante papel na realização dos Objetivos de Desenvolvimento Sustentável e na preservação do processo democrático e dos direitos sociais. As

tecnologias digitais como a IA são um fator fundamental para a realização dos objetivos do Pacto Ecológico, demandando uma abordagem via risquificação e via teoria dos DF, essenciais para se postular e implementar uma proposta de *human-centered IA*, vinculando-se à proposta de regulação *precaucionária*, reconhecendo-se a importância de contestação coletiva, no sentido de dar voz a todos os grupos vulneráveis da sociedade, em especial.

Omri Ben-Shahar[15] aponta para o arranjo regulatório preventivo e coletivo na área de proteção de dados, vinculado à ideia de "poluição de dados", na linha de um "direito ambiental da proteção de dados pessoais", associando a formas de responsabilização *ex post* de danos coletivos, inspirado pela legislação ambiental, com enfoque em medidas de mitigação e em recalibragem das regras de responsabilização civil para gerar novos incentivos aos agentes poluentes. Quando há um vazamento de dados ou outro incidente nesta área os danos não são apenas individuais, mas coletivos, já que todo o ecossistema de dados é afetado. Segundo tal análise há o reconhecimento de uma dimensão pública e centrada na poluição dos dados, uma dimensão social e coletiva, portanto, não apenas individual. A proposta tem por enfoque medidas de mitigação e em recalibragem das regras de responsabilização civil para gerar novos incentivos aos agentes poluentes, propondo uma análise sobre as medidas de mitigação que afetem o coletivo, dispensando a comprovação da ocorrência de danos concretos e individuais, colocando em destaque o princípio da prevenção. Destaca-se, por outro lado, importante contribuição relacionada com tais temáticas acerca da discussão sobre o dano moral coletivo reconhecido pelo Superior Tribunal de Justiça, como expõem Felipe Teixeira Neto e José Luiz Faleiros Junior[16].

Para Rafael Zanatta, trata-se de modelo potencialmente promissor no Brasil, em especial se for superado o caráter *voluntário* dos "relatórios de impacto à proteção de dados", como seria ainda a perspectiva principal da LGPD, não trazendo a lógica predominante da risquificação, já que traz poucas previsões acerca de regulação *ex ante*, aproximando-se mais do modelo teórico da autodeterminação informacional. Em suas palavras: "apesar da rica experiência brasileira no campo ambiental, ainda não foi feita a

[15] Ben-Shahar, Omri. Data Pollution, *Journal of Legal Analysis*, Volume 11, 2019, p. 133.
[16] Teixeira Neto, Felipe; Faleiros Junior, José Luiz. Dano moral coletivo e vazamentos massivos de dados pessoais: uma perspectiva luso-brasileira, *Revista de Direito da Responsabilidade*, ano 3, 2021, p. 265-287.

conexão entre os dois mundos, adaptando-se os instrumentos de análise de impacto e o farto uso de ações civis públicas”[17].

Diversos autores vêm apontando para uma mudança de paradigma ou ponto de virada na moldura teórica quanto à proteção de dados e IA, por meio da adoção de uma arquitetura de gerenciamento dos riscos, precaucionaria de danos quando da utilização da IA, com destaque para Serge Gutwirth & Yves Poullet, Claudia Quelle, Alessandro Spina e Nadezhda Purtova, contudo, segundo Rafael Zanatta não se trataria propriamente de uma mudança de paradigma, envolvendo uma ruptura normativa de abordagem, mas de uma fricção, uma nova abordagem da proteção de dados pessoais centrada na regulação do risco, com a intensificação da regulação *ex ante*.

[17] Zanatta, Rafael, REDE 2017, I Encontro da Rede de Pesquisa em Governança da Internet, *Proteção de dados pessoais como regulação de risco: uma nova moldura teórica?*.