



## **Nota técnica da Estratégia Latino-americana de Inteligência Artificial (ELA-IA) diante do Projeto de Lei nº 2338 de 2023, sobre o uso da Inteligência Artificial no Brasil**

Divulgado no início de maio de 2023, o Projeto de Lei nº 2338 de 2023 visa a regulação em torno do uso das tecnologias denominadas inteligências artificiais. O tema está cada vez mais em discussão nos últimos anos, tanto por conta do contínuo desenvolvimento de ferramentas de inteligência artificial quanto pelos impactos por vezes indesejados de suas aplicações. Considerar a regulação dessas aplicações pode no mínimo levar a uma melhor compreensão crítica sobre tais impactos.

Desse modo, apreciamos a continuidade desse debate e consideramos que o presente projeto de lei segue numa direção construtiva. Temos alguns pontos para questionar, com a intenção de contribuir para esse debate complexo, mas também propostas para endereçar satisfatoriamente esses pontos, viabilizando que o projeto nada perca em concretude.

**Art. 4º, item I, sobre a definição adotada para “sistema de inteligência artificial”.** Dentre os casos para aquilo que é produto do sistema, são amplamente conhecidos ao menos dois casos além dos três apontados no texto (“previsões, recomendações ou decisões”); podemos chamá-los, proporcionalmente, como os *materiais* e as *ações*.

Por *materiais* nos referimos às peças de mídia, como imagens, áudios, vídeos e afins, que podem ser também produzidos parcial ou totalmente por sistemas. A produção desses conteúdos por meios computacionais é uma realidade já há décadas, tomando diferentes formas ao longo do tempo, mas hoje há sistemas amplamente disponibilizados capazes de realizar isso. Há ainda uma gravidade adicional quando tais capacidades artificiais viabilizam a produção de peças de desinformação, como nas chamadas *deepfakes*, o que gera risco, pois o usuário é induzido a tomar a informação como precisa e verdadeira. A desconsideração desse caso na definição de “sistema de inteligência artificial” pode levar a uma brecha na regulação para tais aplicações.

Por *ações* nos referimos às ações mecânicas de sistemas sobre o ambiente, de modo autônomo ou integrado à participação humana, com base em lógicas, dados ou algoritmos disponíveis a tal sistema. Um exemplo desse caso são carros autônomos que operam através de inteligência artificial. Embora a ação de uma máquina possa ser tomada a partir de uma decisão



- o que tornaria esse caso redundante ao caso das “decisões”, já incluso no texto do projeto - isso nem sempre é uma necessidade, pois nem sempre há uma decisão sendo explicitamente tomada pela máquina no processo. Se um carro autônomo, por exemplo, tem capacidade de se mover por conta própria e toma a decisão de se mover, mas acidenta uma pessoa ao fazer isso, seria possível argumentar se a informação sobre a presença da pessoa e o risco do acidente eram conhecidos pelo sistema, e então argumentar - numa tentativa de proteger seus proprietários, fornecedores ou desenvolvedores - que o sistema não tenha decidido acidentiar a pessoa, mas apenas se mover (Somerville, 2018). Uma decisão que deixa de ser tomada explicitamente ainda pode inferir consequência explícita sobre o ambiente através da ação.

A distinção entre a tomada de decisões e de ações sobre o ambiente está presente há um tempo na literatura como algo que caracteriza diferentes formas de inteligência artificial (Russell & Norvig, 1995, e Nunes, 2023) e, se apenas decisões forem contempladas na definição aqui utilizada, mas sem conjunção com as ações, é possível que incidentes envolvendo máquinas autônomas estejam mais sujeitos a confusões. Embora máquinas como estas não estejam sendo amplamente empregadas hoje na sociedade brasileira, essa tratativa precaucionária contribuiria para tornar o projeto atemporal.

Cabe reforçar, em prol do princípio da praticidade, que esse questionamento em torno da definição utilizada não se trata de preciosismo epistemológico, mas de uma providência para que formas de inteligência artificial já conhecidas hoje sejam evidentemente contempladas pelo projeto.

**Art. 4º, sobre a ausência de uma definição adotada para “sandbox regulatório”.** Foi prevista mais adiante no texto do projeto a criação de ambiente regulatório experimental (*sandbox regulatório*) como medida para fomentar a inovação da inteligência artificial. Mas o artigo não traz sua definição e o texto legal não prevê que a regulamentação específica deste ambiente seja realizada por órgãos e entidades públicas responsáveis pela regulação de setores específicos da atividade econômica. Propõe-se a inclusão da definição adaptada a partir das usadas pela CVM, BB e SUSEP com a inclusão do inciso IX no artigo 4º e a inclusão de texto no artigo 38:

*IX - Sandbox Regulatório - ambiente onde organizações são autorizadas por órgão ou entidade pública responsáveis por regulação setorial específica para testar, por período determinado, projeto de Sistemas de IA, observando um conjunto específico de disposições regulamentares que amparam a realização mediante o cumprimento de critérios e limites previamente estabelecidos e dá outras providências.*



*Art. 38. A autoridade competente ou órgãos e entidades públicas responsáveis pela regulação de setores específicos da atividade econômica definidos pelo poder executivo federal poderão autorizar o funcionamento de ambiente regulatório experimental para inovação em inteligência artificial (sandbox regulatório), por meio de regulamentação específica, para as entidades que o requererem e preencherem os requisitos especificados por esta Lei.*

**Art. 5º, item II, sobre o direito à explicação, e item III, sobre o direito à contestação.** São aqui explicitados os mesmos três casos de produtos de sistemas de inteligência artificial anteriormente mencionados, e cabem, portanto, as mesmas críticas anteriormente apresentadas. Ademais, é importante ressaltar a necessidade de incluir, no rol do Art. 5º, o direito à revisão das decisões automatizadas, baseadas em inteligência artificial, evitando riscos aos direitos fundamentais da pessoa humana, principalmente com caráter discriminatório.

Ressalta-se que muito embora a Lei Geral de Proteção de Dados, em seu Art. 20º, trate da revisão de decisões automatizadas, decisões automatizadas e Inteligência Artificial não são sinônimos, de forma que o direito à revisão também merece ser abordado no caso de decisões baseadas em inteligência artificial.

Para além dos prejuízos materiais, a exclusão gerada pelas decisões automatizadas baseadas em inteligência artificial configura violação dos direitos fundamentais do indivíduo, em especial à autonomia, personalidade e igualdade, ao passo que indivíduos em circunstâncias fáticas e jurídicas correlatas são tratados de forma divergente porque um sistema computacional considera que uma dessas pessoas não é digna dos mesmos direitos. No longo prazo, tal situação impossibilitaria que tais pessoas tivessem acesso às oportunidades (DONEDA, MENDES, SOUZA, ANDRADE, 2018).

**Art. 7º, item IV, sobre os papéis dos sistemas de inteligência artificial e dos humanos.** Além da menção restrita aos três casos anteriormente citados, como já aqui apontado, cabe pontuar que essa fronteira entre as responsabilidades dos dois conjuntos é de grande valor para a adequada aplicação dos direitos e responsabilidades entre fornecedores e operadores, além de contribuir para uma melhor compreensão, por parte das pessoas afetadas, sobre como estão sendo materialmente afetadas e podem legalmente proceder a respeito.



**Art. 7º, item VI, sobre as medidas adotadas.** O uso de métricas estatísticas para mensuração de resultados de algoritmos, embora recorrente, assume formatos muito diversificados, de acordo com o objetivo da aplicação do algoritmo e com a capacidade dos fornecedores e operadores de captar dados para a aplicação de métricas específicas. Mais ainda, os valores de tais métricas comumente flutuam ao longo do tempo, e a exposição de uma métrica “previamente à contratação ou utilização do sistema” tende a nada representar sobre seus valores atingidos após o início da referida contratação ou utilização.

Nesse sentido, a menção de métricas específicas no projeto não condiz com o que é necessário para compreender uma aplicação de um algoritmo, e tende a contribuir para percepções e expectativas incoerentes por parte das pessoas afetadas pelos sistemas. Contudo, caso existam métricas adotadas previamente e que sejam, por sua natureza, consistentes de modo atemporal, sua exposição às pessoas afetadas pode contribuir para uma maior segurança de sua parte. Isso poderia ocorrer de modo voluntário, porém, quando possível, recomendado pelo projeto.

**Art. 8º, Art. 9º, Art. 10º, Art. 11º, Art. 12º.** Ocorre nos textos principais desses artigos mais uma vez a menção apenas aos três casos de sistemas de inteligência artificial presentes no texto. Isso se repete no texto do parágrafo segundo do Art. 9º.

**Art. 10º, sobre a intervenção ou revisão humana.** Ainda que esteja prevista a revisão, não houve a previsão da revisão humana de eventuais decisões baseadas em sistemas de inteligência artificial (IA) como um direito. Este direito, assim como o direito à explicação são tratados no artigo 20 da LGPD, visando mitigar os riscos dos vieses em decisões automatizadas que se utilizam de dados pessoais e pessoais sensíveis. As IA não se utilizam apenas desta categoria de dados, por isto se justifica a previsão do direito nesta matéria. Propõe-se a inclusão do inciso VII no art 5º:

*VII - direito à revisão humana de decisões baseadas em sistema de inteligência artificial que afetem de alguma forma, direitos e garantias fundamentais previstas na constituição*

**Art 13º, sobre a avaliação preliminar de risco.** Tendo em vista que tecnologias são desenvolvidas de modo generalista que viabilize certa amplitude de suas aplicações, o que muito se distingue de uma aplicação de tecnologia existente (Headrick, 2009), que tem objetivos e contextos mais específicos e pode para isso necessitar de adequações à tecnologia existente -



pequenas o bastante para não caracterizar uma nova tecnologia a partir da anterior -, os riscos relacionados a uma tecnologia tendem a muito variar de acordo com cada aplicação em questão (Nunes, 2023), tanto em sua presença quanto em seus impactos possíveis, portanto a avaliação preliminar de riscos que não leve em consideração a aplicação sendo realizada se torna incompleta e não garante o adequado seguimento dos fundamentos do Art. 2º.

Um exemplo simples para isso está no entendimento diferentemente aprofundado sobre o funcionamento de um sistema de inteligência artificial entre o fornecedor e o operador. Uma tecnologia fornecida para auxiliar na contratação de pessoas por uma empresa, por exemplo, sobre a qual se afirma que “selecione as melhores pessoas candidatas para um processo seletivo”, mas que tecnicamente selecione dentre as pessoas candidatas aquelas que estatisticamente tendem a permanecer por mais tempo na organização, pode acabar selecionando apenas pessoas dentro de um recorte demográfico que, inclusive por questões de discriminação, permaneça em média por mais tempo naquela organização; enquanto para o operador pode ser compreendido apenas que “as melhores pessoas candidatas estejam sendo selecionadas”, o que o sistema de fato estaria fazendo poderia contribuir para um agravamento dos aspectos discriminatórios presentes na cultura daquela organização (Nunes, 2023). O risco de agravamento da discriminação a partir de aplicações dessa tecnologia varia, finalmente, de uma organização para outra, sendo recomendável que aconteça uma avaliação ao menos parcial a cada uma de suas aplicações. Em alguns casos esse risco pode chegar a sequer existir; em outros, pode ter um impacto possível alto demais, caracterizando-o como excessivo.

Ao mesmo tempo em que essa avaliação por aplicação e não por tecnologia pulverizaria o processo e aumentaria o trabalho necessário para sua implementação, traria resultados mais coerentes e materiais, podendo inclusive incentivar uma maior criação de empregos associados a essa avaliação.

***Art. 14º, item II, sobre implementação e uso de sistema de inteligência artificial de risco excessivo, que explorem vulnerabilidades de grupos específicos, em razão da idade :*** É importante ressaltar que, na atualidade, as crianças não têm contato apenas com sistemas de Inteligência Artificial implementados em produtos e serviços digitais desenvolvidos para o público infantil.

Ademais, os dados pessoais de crianças são considerados dados especiais, pois exigem um tratamento diferenciado em termos de cuidados. (PECK, 2018).

Entendemos que, dessa forma, é necessário que qualquer sistema de IA com que crianças possam interagir – direta ou indiretamente -, ou que possam de alguma maneira impactá-las, respeite os direitos infantis com absoluta prioridade. Entre as ações que devem ser tomadas,



mas não limitando-se a elas, estão: proteger a privacidade dos menores, promover sua saúde física e mental e evitar riscos potenciais.

**Art. 20º, item V, sobre as medidas de governança para sistemas de inteligência artificial de alto risco.** Entendemos que as informações acerca dos sistemas de inteligência artificial, quando requisitados pelo interessado, precisam ser, além de adequadas, concisas, completas, corretas, claras, acessíveis e compreensíveis para os utilizadores, de modo que as informações estejam em linguagem de fácil compreensão.

**Art. 24º, sobre a metodologia da avaliação de impacto.** Entendendo que os riscos e impactos associados a uma aplicação de inteligência artificial podem ser muito diversificados em intensidades e naturezas, assim como considerando que uma especificação muito detalhada dos riscos e impactos a serem endereçados pela lei acabaria prejudicando a atemporalidade de seu conteúdo, a ausência total de referências sobre que tipos de riscos e impactos precisam ser considerados pode levar a avaliações incompletas. A adoção de algum esquema mínimo poderia contribuir para um melhor endereçamento de riscos e impactos.

Para a compreensão do impacto de uma aplicação de tecnologia sobre o ambiente psicossocial, há um esquema para mapear e simular esse impacto através de três eixos (Ribeiro, 2018, e Nunes, 2023): um subjetivo, que contempla os aspectos psicológicos e sociais que não são mensuráveis, mas participam da experiência humana e são estudados pelas ciências; um comportamental, que contempla os comportamentos dos seres vivos; e um ambiental, que contempla o lugar físico e simbólico onde acontecem as relações humanas. Esse esquema, além de servir para representar e analisar a ética de uma população, pode ser também utilizado para identificar as alterações nessa ética causadas por uma aplicação de tecnologia. A recomendação no projeto de um esquema similar viabilizaria que todos os três tipos de riscos e impactos fossem minimamente considerados como possibilidades na construção das avaliações.

**Art. 27º, sobre a inversão do ônus da prova.** Em que pese a inversão do ônus da prova estar disposta apenas no § 2º do Art. 27, entendemos que o ônus deve ser invertido, não só quando não se tratar de sistema de inteligência artificial de alto risco, mas sempre que o sistema de inteligência artificial for utilizado, tendo em vista a vulnerabilidade da vítima.

**Art. 28º, item I, sobre a ausência de responsabilidade:** Não é razoável limitar a responsabilidade dos agentes de inteligência artificial às ações de: “colocar em circulação, empregar ou tirar proveito econômico”. Pois seria possível que houvesse dano, mesmo que o



sistema de inteligência artificial não estivesse em circulação (em testes, por exemplo) e ainda que não houvesse tido proveito econômico.

## Referências

DONEDA, Danilo Cesar Maganhoto; MENDES, Laura Schertel; SOUZA, Carlos Affonso Pereira de; ANDRADE, Norberto Nuno Martin Becerra Gomes de. **Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal**. Pensar - Revista de Ciências Jurídicas, 2018.

HEADRICK, D. **Technology: A World History**. Oxford: Oxford University Press, 2009.

NUNES, L. J. **Artifício: Um estudo introdutório dos aspectos psicossociais da artificialização das inteligências**. São Paulo: Dialética, 2023.

PACHECO, Rodrigo. **Projeto de Lei nº 2338, de 2023**. Disponível em <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233> . Último acesso em 24/05/2023.

PECK PINHEIRO, Patricia. **Proteção de dados pessoais: comentários à Lei n. 13.709/2018 (LGPD)** - São Paulo: Saraiva Educação, 2018.

RIBEIRO, S. M. P. **Lições preliminares para o estudo do *ethos* humano contemporâneo**. São Paulo: 2018.

RUSSELL, S. & NORVIG, P. **Artificial Intelligence - A Modern Approach**. New Jersey: Prentice-Hall, 1995.

SOMERVILLE, H. **Motorista de carro autônomo do Uber estava assistindo "The Voice" antes de acidente fatal, diz polícia**, no jornal **Reuters**, em 22/06/2018. Disponível em <https://www.reuters.com/article/tech-uber-acidente-autonomo-idBRKBN1J12BB-OBRIN>. Último acesso em 24/05/2023.