



Quatro polêmicas sobre o projeto que regula a inteligência artificial

29/02/2024

O Projeto de Lei nº 2338 de 2023, apresentado pelo senador Rodrigo Pacheco em maio do ano passado, tem como objetivo regulamentar o uso de tecnologias de inteligência artificial e tomou como referência o rascunho elaborado pela comissão de juristas formada em 2022. Quatro pontos do texto são polêmicos: definição para sistema de IA, sandbox regulatório, intervenção ou revisão humana e papéis dos sistema de IA e dos humanos.

Sistema de inteligência artificial

A definição tomou como referência tão somente as aplicações de natureza preditiva, ignorando os usos de natureza generativa como a criação de imagens, áudios, vídeos e textos.

Embora um pedido de criação de conteúdos de mídia através de linguagem natural possa ser enquadrado no trecho do inciso “atingir um dado conjunto de objetivos”, o final do mesmo trecho fixa uma redação que, ainda que não seja taxativa, é claramente limitadora ao intérprete da norma ao fixar “o objetivo de produzir previsões, recomendações ou decisões”.

Essa parte final cria com critério para a influência do conteúdo sobre o ambiente virtual ou real, o que fica evidenciado com as experiências com *deepfakes*, como a criação de imagens do papa ^[1], a criação de imagens da prisão de Donald Trump ^[2], as imagens do presidente francês Emmanouel Macron limpando as ruas de Paris ^[3] e de um pretenso ataque terrorista ao Pentágono. ^[4]

Outro trecho que detalha o conceito de “sistema computacional, com graus diferentes de autonomia” também precisa de maior acuidade.

Computadores possuem grau de autonomia para decidir em que momento a ventoinha será acelerada, de modo a reduzir a temperatura da placa-mãe de uma temperatura de 100º para 80º graus Celsius, que seria uma temperatura ideal de processamento.

Isso pode ser entendido como um grau de autonomia, mas seria caracterizada como uma automação e não como um sistema de inteligência artificial, já que não demandaria um grau de complexidade relevante para a tomada de uma ação.

Essa clareza ajudará a fixar qual ferramenta estará sujeita ao relatório de impacto algorítmico. A ISO/IEC 22989 poderia ser um parâmetro razoável à fixação de conceitos comuns ao mercado e ao legislador.

Sandbox regulatório

O Projeto não definiu o conceito, mas outras normas podem ser indicadas como referência para determinação desta estrutura de regulação, cabendo a adoção de termos já existentes de modo a evitar conflitos de conceito.

Ainda seria possível o uso de conceituações setoriais que melhor respondam às necessidades dos operadores, a exemplo do conceito trazido pela Portaria nº 75, de 29 de junho de 2020, da Comissão de Valores Mobiliários (CVM) sobre a composição e o funcionamento do Comitê de Sandbox (“CDS”) de que trata o artigo 2º, inciso III, da Instrução CVM n.º 626 ^[5], ou conforme normativa da Superintendência de Seguros Privados (Susep) ^[6].

Há outra problemática, vinculada à questão da competência de cada sandbox, especialmente na administração pública. Se por um lado é praticável e previsível a fixação de sandbox temático-setorial, por outro pode haver fixação de competência territorial para determinar se uma aplicação poderá ser contratada por um organismo público.

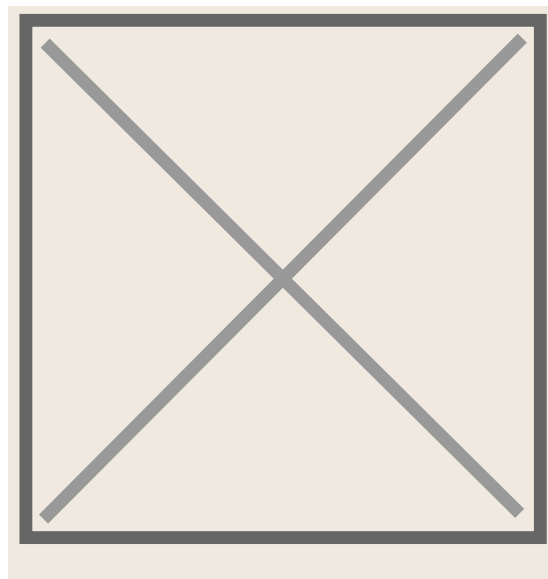
Neste caso um sandbox municipal poderia rejeitar uma aplicação que já foi aprovada por um sandbox regulatório estadual da unidade federativa em que está localizado?

Algumas soluções poderiam ser aplicadas, como a recepção, pelo sandbox municipal, de uma aprovação de um sandbox estadual; a fixação de critérios complementares, caso o sandbox estadual já tenha aprovado a ferramenta; a dispensa de avaliação caso qualquer outro sandbox já tenha aprovado a ferramenta.

Explicabilidade e contestação/intervenção ou revisão humana

Decisões automatizadas que evitem o exercício de um direito poderão ser apresentadas por um sistema de inteligência artificial, seja porque a decisão foi fundamentada, seja porque houve alguma anomalia, viés ou falha que levou à uma decisão do tipo.

A explicabilidade (artigo 21, IV do PL 2338), em especial nos casos de tratamento de dados pessoais, está fixada no artigo 20 da Lei Geral de Proteção de Dados brasileira ^[7], no artigo 22 da General Data Protection Regulation ^[8], além de normas vinculadas à Governança de TI e IA, como o item *Maintaining governance when introducing AI* da ISO/IEC 38507 (ainda não traduzida para o português).



Em ambos os artigos, é recomendável que duas características se mostrem presentes no caso de uso de mais de uma metodologia de explicabilidade, a *clareza* da explicação, através de diferentes metodologias e a *resiliência* da explicação. A segunda característica é exigida para que a descrição não seja deformada a cada nova apresentação ou a cada mudança de metodologia.

Papéis dos sistemas de IA e dos humanos

A notificação sobre o uso de IA em determinados contextos pode ser bem comparada com duas hipóteses:

- Filmagem de ambientes, hipótese na qual há a fixação de placa indicativa de que as imagens são coletadas, conforme lei local.
- Informe de uso de IA durante a interação do sujeito, como já ocorre com os informes de cookies em sites. No caso de cookies existe a coleta de consentimento do titular e o detalhamento sobre a necessidade indispensável de coleta dos cookies para o bom funcionamento do site.

Há uma crítica bastante contundente quanto à estafa do consentimento ou, ainda, um estado de conformidade que afasta o processo de tomada autônoma de decisão do titular. Durante a elaboração do texto do PL da IA pela comissão, este autor, juntamente com João Paulo Candia, encaminhou sugestões sobre as hipóteses de informe, os formatos de notificação e a não utilização de cookies para personalizar páginas jornalísticas.

Interações entre humanos e máquinas deverão ser claras, de modo que:

- *a) o ser humano tenha pleno discernimento de que está interagindo com outro humano ou com uma máquina, seja por meio escrito, audiofônico, pictórico, vídeo fotográfico, língua de sinais, braile ou outras formas de comunicação;*
- *b) seja destacado alerta que informe o uso de máquina dotada de capacidade linguística;*
- *c) seja criado protocolo para identificação de produções textuais e jornalísticas que evidencie que o criador é uma inteligência artificial;*
- *d) as fontes utilizadas pela inteligência artificial para produção textual sejam rastreáveis ou sejam previamente curadas;*
- *e) que os conteúdos de acesso público e midiático não sejam personalizados de acordo com o usuário, mantendo-se manchetes uniformes em todos os acessos.* [9]

O principal objetivo é garantir clareza quanto à forma de criação do conteúdo, como já ocorre na produção de conteúdos de mídia, cujo estilo artístico afeta a compreensão dos espectadores, como animações, que podem ter exageros irreais que contribuem para a narrativa e evitam o “vale da estranheza”. [10]

[1] Deep fake: Como lidar com desinformação gerada por inteligência artificial – 20/04/2023 – UOL Notícias

[2] Deepfake Trump arrest photos show disruptive power of AI – The Washington Post

[3] On a essayé de fabriquer un deepfake (et on est passé à autre chose) (lemonde.fr)

[4] Fake Trump arrest photos: How to spot an AI-generated image – BBC News

[5] Sandbox Regulatório (cvm.gov.br)

[6] Perguntas e Respostas sobre o Sandbox Regulatório – 1ª Edição — SUSEP – Superintendência de Seguros Privados (www.gov.br)

[7] L13709 (planalto.gov.br)

[8] Art. 22 GDPR – Automated individual decision-making, including profiling – General Data Protection Regulation (GDPR) (gdpr-info.eu)

[9] Relatório da Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre inteligência artificial no Brasil – Atividade Legislativa – Senado Federal, p. 807

[10] Momentos em que o conteúdo tenta simular a realidade, mas não cumpre com as exigências de qualidade visual que a tornariam verossímil.

Fonte: <https://conjur.jumps.com.br/2024-fev-29/quatro-polemicas-sobre-o-projeto-que-regula-a-inteligencia-artificial/>