



UNIVERSIDADE DE BRASÍLIA

FACULDADE DE DIREITO

PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITO, REGULAÇÃO E
POLÍTICAS PÚBLICAS

**ESTRATÉGIAS DE GOVERNANÇA DO USO DA INTELIGÊNCIA
ARTIFICIAL NO PODER JUDICIÁRIO BRASILEIRO**

Amanda Lins Brito Faneco Amorim

Brasília
Março, 2024

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE DIREITO
PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITO, REGULAÇÃO E
POLÍTICAS PÚBLICAS

**ESTRATÉGIAS DE GOVERNANÇA DO USO DA INTELIGÊNCIA
ARTIFICIAL NO PODER JUDICIÁRIO BRASILEIRO**

Amanda Lins Brito Faneco Amorim

Dissertação apresentada ao Programa de Pós-graduação em Direito, Regulação e Políticas Públicas da Universidade de Brasília como requisito parcial para a obtenção do título de Mestre em Direito, Regulação e Políticas Públicas.

Orientador: Prof. Dr. Fabiano Hartmann Peixoto.

Brasília
Março, 2024

ESTRATÉGIAS DE GOVERNANÇA DO USO DA INTELIGÊNCIA ARTIFICIAL NO PODER JUDICIÁRIO BRASILEIRO

Amanda Lins Brito Faneco Amorim

Orientador: Prof. Dr. Fabiano Hartmann Peixoto

Dissertação apresentada ao Programa de Pós-graduação em Direito, Regulação e Políticas Públicas da Universidade de Brasília como requisito parcial para a obtenção do título de Mestre em Direito, Regulação e Políticas Públicas.

Brasília, 22 de abril de 2024.

Examinada por:

Avaliador

André Gualtieri de Oliveira

Universidade Mackenzie

Avaliador

Fernanda de Carvalho Lage

Universidade de Brasília - UnB

Orientador:

Fabiano Hartmann Peixoto

Universidade de Brasília - UnB

Le Lins Brito Faneco Amorim, Amanda
ESTRATÉGIAS DE GOVERNANÇA DO USO DA INTELIGÊNCIA
ARTIFICIAL NO PODER JUDICIÁRIO BRASILEIRO / Amanda Lins
Brito Faneco Amorim; orientador Fabiano Hartmann Peixoto. --
Brasília, 2024.
118 p.

Dissertação (Mestrado Profissional em Direito, Regulação
e Políticas Públicas) -- Universidade de Brasília, 2024.

1. Inteligência Artificial. 2. Judiciário. 3. Governança.
4. Regulação. 5. Comitê de ética. I. Hartmann Peixoto,
Fabiano, orient. II. Título.

RESUMO

A utilização de sistemas de Inteligência Artificial (IA) no Poder Judiciário brasileiro é uma realidade em crescente expansão. Considerando a necessidade de observância de parâmetros para o desenvolvimento e uso ético de IA, o Conselho Nacional de Justiça editou a Resolução nº 332/2020. A presente pesquisa tem por objetivo avaliar as estratégias de governança previstas no referido ato normativo, bem como outros mecanismos capazes de assegurar o uso ético da tecnologia e mitigar os riscos envolvidos, buscando subsidiar futuras iniciativas institucionais. Para tanto, no primeiro capítulo, são analisados os papéis desempenhados pela regulação, ética e governança e sua interseção no controle de riscos relacionados à utilização da tecnologia. No segundo capítulo, são examinados os riscos do uso IA no Poder Judiciário e a regulamentação pertinente, incluindo a análise de documentos internacionais como a Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais. No terceiro capítulo, é avaliada a criação de comitês de ética de IA como mecanismo para lidar com a complexidade do ciclo de desenvolvimento dos projetos de IA e as questões éticas inerentes à sua implementação no contexto sensível do Poder Judiciário. Ao final, é apresentada uma proposta de criação de comitês de ética de IA no âmbito do Poder Judiciário brasileiro e outros colegiados nacionais com detalhes sobre estrutura, composição, objetivos e atribuições.

Palavras-chave: Inteligência Artificial; Judiciário; Governança; Regulação; Comitê de Ética.

ABSTRACT

The use of Artificial Intelligence (AI) systems in the Brazilian Judiciary is an expanding reality. Recognizing the need to adhere to parameters for the ethical development and use of AI, the National Council of Justice issued Resolution No. 332/2020. This research aims to evaluate the governance strategies outlined in the aforementioned normative act, as well as other mechanisms capable of ensuring the ethical use of technology and mitigating the associated risks, aimed at supporting future institutional initiatives. To this end, the first chapter analyzes the roles played by regulation, ethics, and governance and their intersection in controlling risks related to technology use. In the second chapter, the risks of AI use in the Judiciary and the relevant regulations are examined, including an analysis of international documents such as the European Charter of Ethics on the Use of Artificial Intelligence in Judicial Systems. In the third chapter, the creation of AI ethics committees is evaluated as a mechanism to address the complexity of the development cycle of AI projects and the ethical issues inherent in their implementation in the sensitive context of the Judiciary. Finally, a proposal for the creation of AI ethics committees within the Brazilian Judiciary and other national boards is presented, detailing their structure, composition, objectives, and attributions.

Keywords: Artificial Intelligence; Judiciary; Governance; Regulation; Ethics Committees.

SUMÁRIO

Introdução	8
1. Inteligência artificial: regulação, ética e governança.....	15
1.1. Definindo inteligência artificial.....	15
1.2. O cenário regulatório nacional e global.....	19
1.3. Os princípios éticos da IA.....	25
1.4. O papel da governança.....	31
2. Governança do uso da Inteligência Artificial no Poder Judiciário Brasileiro	36
2.1. Riscos do uso da IA no Poder Judiciário	36
2.2. Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes.....	45
2.3. Outros documentos internacionais.....	52
2.4. Regulamentação da IA no Poder Judiciário brasileiro.....	55
2.5. Instrumentos previstos na Resolução n. 332/2020	59
2.5.1. A plataforma Sinapses (art. 10, III)	59
2.5.2. Painel de Projetos com Inteligência Artificial no Poder Judiciário (art. 11)	61
2.5.3. Comunicação de eventos adversos (art. 27)	65
2.5.4. Certificação de boas práticas (art. 8º, V)	70
2.5.5. Auditoria (art. 8º, V).....	73
3. Comitês de ética de IA e sua implementação no Poder Judiciário brasileiro	81
3.1. Comitê de ética de IA enquanto mecanismo de governança	81
3.2. Análise de experiências implementadas no Poder Judiciário brasileiro	87
3.3. Sugestão de criação de comitês de ética de IA no Poder Judiciário brasileiro	97
Conclusão	109
Referências	113

INTRODUÇÃO

O uso de sistemas de inteligência artificial (IA) no Poder Judiciário brasileiro é uma realidade em crescimento desde 2017. A partir de janeiro de 2021, o Conselho Nacional de Justiça (CNJ), em parceria com o Conselho da Justiça Federal (CJF) e o Programa das Nações Unidas para o Desenvolvimento (PNUD), por meio do Programa Justiça 4.0, passou a impulsionar a transformação digital e o uso de inteligência artificial buscando tornar a prestação de serviços de Justiça mais eficiente, eficaz e acessível à sociedade¹.

Em 2022, levantamento do CNJ identificou 111 projetos de IA desenvolvidos ou em desenvolvimento em 53 tribunais brasileiros, o que significou um aumento de 171% (cento e setenta e um por cento) em relação ao ano anterior, quando foram informados apenas 41 projetos espalhados em 32 tribunais².

As diferentes ferramentas de IA em utilização no Judiciário brasileiro são aplicadas nos mais variados momentos processuais e para desempenho de múltiplas funções, tais como: triagem da petição inicial; classificação de documentos e/ou processos; agrupamento de processos por similaridade; identificação de precedente aplicável; sugestão de minuta de decisão; e busca de bens em processo de execução.

As características essenciais das soluções de IA, capazes de processamento de grande volume de dados e reconhecimento de padrões, de fato, evidenciam que as soluções tecnológicas são capazes de contribuir muito para a celeridade e eficiência do Poder Judiciário.

Nesse contexto, contudo, ainda que as ferramentas existentes consistam em mecanismo de apoio à atividade decisória humana e não sua efetiva substituição, diversas experiências apontam para a necessidade de implementação de mecanismos de governança, parâmetros de uso ético e mitigação dos riscos envolvidos na utilização de sistemas de IA em uma área tão sensível como o Judiciário, com potencial impacto sobre direitos dos cidadãos e garantias processuais.

Assim, considerando a capacidade de sistemas de inteligência artificial para garantir agilidade e coerência do processo de tomada de decisão judicial e a ausência de regulação no país acerca do tema, o CNJ editou a Resolução nº 332/2020, que dispõe sobre ética,

¹ CONSELHO NACIONAL DE JUSTIÇA. Justiça 4.0. Brasília: **CNJ**, 2021. Disponível em: <<https://www.cnj.jus.br/wp-content/uploads/2021/08/cartilha-justica-4-0-20082021-web.pdf>>. Acesso em: 16 mar. 2024.

² CONSELHO NACIONAL DE JUSTIÇA; Programa das Nações Unidas para o Desenvolvimento. Relatório final gestão Ministro Luiz Fux: programa Justiça 4.0 [livro eletrônico]. Brasília: **CNJ**, 2022. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/09/af-pnud-relatorio-v3-web.pdf>. Acesso em: 10 maio 2023.

transparência e governança na produção e no uso de inteligência artificial no Poder Judiciário.

A referida norma optou por uma regulação com base em diretrizes gerais e princípios aplicáveis, nos moldes dos documentos internacionais, especialmente, na Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes da Comissão Europeia para Eficiência da Justiça (CEPEJ).

Apesar da notável iniciativa do CNJ e da importância do referido ato normativo, não há previsão de sanções ou mesmo preocupação com a aplicação das diretrizes postas (*enforcement*), comprometendo a efetividade da norma.

Passados mais de três anos da publicação da resolução, permanecem sem regulamentação as regras para realização de auditoria nos projetos de IA, assim como sequer há relatos sobre a realização de alguma auditoria nas diferentes ferramentas tecnológicas já em utilização nos tribunais brasileiros. Também não há informações sobre eventuais relatórios realizados para verificar o impacto da implementação do uso da tecnologia nas decisões judiciais, bem como nos direitos e garantias fundamentais, especialmente, no âmbito processual, tais como o contraditório e a ampla defesa.

Embora haja crescentes avanços no número de iniciativas com utilização de IA, pouco tem sido discutido, no âmbito do Poder Judiciário, sobre a criação de ferramentas de governança de IA com instrumentos eficientes de monitoramento contínuo de riscos.

É essencial que, paralelo ao avanço no desenvolvimento de novas soluções de IA, haja também avanço na regulamentação do tema, no aprimoramento de ferramentas de governança e na avaliação dos impactos decorrentes, não apenas no nível individual dos jurisdicionados, mas também no aspecto global do Estado de Direito.

A verdade é que a Resolução CNJ n. 332/2020, apesar de importante por estabelecer diretrizes principiológicas e indicar alguns instrumentos de governança, não apresenta mecanismo de *enforcement* efetivo, tampouco foi acompanhada de regulamentação posterior voltada para esse fim. Assim, nos termos em que editada, tem se mostrado insuficiente para garantir o monitoramento dos impactos e riscos envolvidos e a observância de direitos fundamentais.

Com efeito, a regulação do uso da IA, nos mais variados contextos, ainda é um assunto que gera grande debate tanto no cenário nacional quanto global. Nessa perspectiva, aguardar a evolução da discussão e da efetiva implementação da regulação no nível geral para lidar com questões éticas e riscos associados à disseminação da referida tecnologia no âmbito do sistema de justiça não parece ser prudente nem desejável. Pelo contrário, é crucial implementar, de

imediatamente, medidas de governança para identificar, compreender e mitigar possíveis impactos negativos na integridade do sistema de justiça.

A propósito, já existem diversos riscos documentados do uso da IA nas mais variadas áreas, especialmente envolvendo o direito a não discriminação, o direito à privacidade de dados e o direito à explicabilidade³.

Especificamente na área jurídica, um dos casos mais conhecidos de impactos negativos está relacionado ao uso do software COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), um sistema de avaliação de risco de reincidência utilizado para subsidiar a decisão de juízes em alguns estados dos Estados Unidos da América.

Um estudo realizado pela organização ProPublica, em que foram analisadas avaliações realizadas pelo COMPAS, no condado de Broward, na Flórida, concluiu pela existência de viés de discriminação racial, na medida em que “réus negros eram frequentemente classificados com maior chance de reincidir do que realmente eram” e “mesmo quando controlados crimes anteriores, risco de reincidência, idade e gênero, os réus negros tinham 45% mais probabilidade de receber notas de risco mais altas do que os réus brancos” (tradução livre)⁴.

Outro caso emblemático é o do *chatbot* Tay lançado pela Microsoft Corporation em março de 2016. Projetado para desenvolver interações online, em linguagem natural, com humanos, o sistema incorporava técnicas de aprendizado de máquina para melhorar suas habilidades de conversação com base nas interações com os usuários do Twitter. Nas primeiras 24 horas online, internautas interagiram com o sistema de IA com uma enorme quantidade de *tweets* (mensagens curtas) maliciosos, de modo que Tay aprendeu a reproduzir preconceitos, publicando mensagens discriminatórias, o que levou a empresa desenvolvedora a desativar o sistema⁵.

Recentemente, a rede americana de farmácia e drogarias *Rite Aid* foi proibida de usar tecnologia de reconhecimento facial pela *Federal Trade Commission* (FTC) em razão de ter implementado o seu uso sem adotar salvaguardas razoáveis para proteger a privacidade e a

³ KAUFMAN, Dora; JUNQUILHO, Tainá; REIS, Priscila. Externalidades negativas da inteligência artificial: conflitos entre limites da técnica e direitos humanos. **Revista de Direitos e Garantias Fundamentais**, v. 24, n. 3, p. 43-71, 2023. Disponível em: <https://doi.org/10.18759/rdgf.v24i3.2198>. Acesso em: 7 jan. 2024.

⁴ LARSON, Jeft et al. How we analyzed the COMPAS recidivism algorithm. [S. L.]: **ProPublica**, 23 may 2016. Disponível em: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Acesso em: 7 jan. 2024.

⁵ MICROSOFT CORPORATION. **Learning from Tay's introduction**. Disponível em: <https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/>. Acesso em: 7 jan. 2024.

segurança dos consumidores⁶.

Com a popularização da inteligência artificial generativa, novos problemas passaram a ser mapeados. Pouco tempo após a popularização do ChatGPT, ferramenta disponibilizada de maneira gratuita pela *OpenAI* em novembro de 2022, foram divulgadas notícias de diversos problemas decorrentes da sua utilização: estudantes acusados de plágio em trabalhos acadêmicos⁷; disseminação de informações falsas⁸; vazamento de dados e informações sigilosas⁹; difamação¹⁰; violação de direitos autorais¹¹; entres outros.

Especificamente na área jurídica, o caso dos advogados multados por apresentarem informações falsas, indicando precedentes judiciais inexistentes, em petição apresentada ao Tribunal de Nova York, chamou atenção em todo o mundo¹². O caso é emblemático para demonstrar a necessidade de os usuários da ferramenta, incluindo os profissionais da área jurídica, compreenderem as limitações dos sistemas de inteligência artificial antes de empregá-las no exercício da profissão.

No Brasil, um advogado também foi multado por litigância de má-fé em razão de ter protocolado uma petição com texto elaborado pelo ChatGPT, mesmo tendo ciência da inadequação do material, segundo o magistrado¹³. Na sequência, o advogado protocolou representação ao Conselho Nacional de Justiça buscando proibir magistrados de utilizarem a

⁶ VALINSKY, Jordan. **Rite Aid's 'reckless' use of facial recognition got it banned from using the technology in stores for five Years**. Disponível em: <https://www.ftc.gov/news-events/news/press-releases/2023/12/rite-aid-banned-using-ai-facial-recognition-after-ftc-says-retailer-deployed-technology-without>; <https://edition.cnn.com/2023/12/20/tech/rite-aid-ai-ftc-settlement/index.html>. Acesso em: 7 jan. 2024.

⁷ FABRO, Clara. **'Pode isso?' Quase 400 estudantes são investigados por usarem o ChatGPT**. 25 jul. 2023. Disponível em: <https://www.techtudo.com.br/noticias/2023/07/pode-isso-quase-400-estudantes-sao-investigados-por-usarem-o-chatgpt-ediviralizou.ghml>. Acesso em: 30 jul. 2023.

⁸ KURZWEIL, Ray. **A conversation on the singularity**. 1 jan. 2022. Disponível em: <https://www.thekurzweillibrary.com/dialogue-a-conversation-on-the-singularity>. Acesso em: 30 jul. 2023.

⁹ INFOMONEY. **CHATGPT confirma vazamento de dados de cartão de usuários; empresa afirma que bug foi corrigido**. 30 mar. 2023. Disponível em: <https://www.infomoney.com.br/negocios/chatgpt-confirma-vazamento-de-dados-de-cartao-de-usuarios-empresa-afirma-que-bug-foi-corrigido> e <https://www.forbes.com/sites/siladityaray/2023/05/02/samsung-bans-chatgpt-and-other-chatbots-for-employees-after-sensitive-code-leak/?sh=3a01a1bc6078>. Acesso em: 30 jul. 2023.

¹⁰ MELO, João Ozorio. **Especialistas discutem se é possível processar ChatGPT por difamação**. 17 abr. 2023. Disponível em: <https://www.conjur.com.br/2023-abr-17/especialistas-discutem-possivel-processar-chatgpt-difamacao>. Acesso em: 30 jul. 2023.

¹¹ ROSA, Giovanni Santa. **Escritores processam empresa do ChatGPT por infringir direitos autorais**. 10 jul. 2019. Disponível em: <https://tecnoblog.net/noticias/2023/07/10/escritores-processam-empresa-do-chatgpt-por-infringir-direitos-autoriais/>. Acesso em: 30 jul. 2023.

¹² STROM, Roy. **Fake ChatGPT Cases Cost Lawyers \$5,000 Plus Embarrassment**. 22 jun. 2023. Disponível em: <https://news.bloomberglaw.com/business-and-practice/fake-chatgpt-cases-costs-lawyers-5-000-plus-embarrassment>. Acesso em: 30 jul. 2023.

¹³ VITAL, Danilo. **TSE multa advogado por petição baseada em "fábula" escrita com ChatGPT**. 18 abr. 2023. Disponível em: <https://www.conjur.com.br/2023-abr-18/tse-multa-advogado-peticao-baseada-conversa-chatgpt>. Acesso em: 30 jul. 2023.

ferramenta de IA¹⁴.

No âmbito jurisdicional, um juiz colombiano ganhou destaque mundial ao afirmar que utilizou o ChatGPT para ajudar a decidir se o seguro de uma criança autista deveria cobrir todos os custos de seu tratamento médico¹⁵.

No território nacional, um caso em que um magistrado utilizou resultados formulados pela ferramenta de IA generativa, incluindo precedentes inexistentes do Superior Tribunal de Justiça (STJ), foi denunciado pelo advogado atuante junto à Corregedoria Regional da Justiça Federal da 1ª Região, que arquivou o caso após recomendação de que “não sejam utilizadas para a pesquisa de precedentes jurisprudenciais ferramentas de IA generativa abertas e não-homologadas pelos órgãos de controle do Poder Judiciário”¹⁶. O caso foi reaberto pelo Conselho Nacional de Justiça.

Os impactos adversos decorrentes da proliferação de sistemas de inteligência artificial nas mais diversas áreas foram, inclusive, apontados no Relatório Global de Riscos 2024 do Fórum Econômico Mundial como ponto de risco de significativa gravidade no longo prazo¹⁷. O documento aponta que juntamente com os benefícios de produtividade e avanços em campos diversos como saúde e educação, o avanço da IA traz consigo importantes riscos sociais, tais como: desinformação; redução significativa de empregos; uso criminoso e ciberataques; viés e discriminação; utilização em tomadas de decisão críticas por organizações e Estados; e integração da IA em armamentos e guerra¹⁸.

A multiplicação de casos de aplicação de ferramentas de IA com repercussão negativa levou a algumas empresas desenvolvedoras ou contratantes passassem a adotar cautelas, independentemente da existência de leis ou regulamentação sobre o tema, buscando a mitigação de riscos e de desenvolvimento de mecanismos de governança.

Ainda que a totalidade dos projetos em utilização ou em andamento no Poder Judiciário brasileiro esteja voltada a auxiliar a decisão humana e não a substituir por completo, exigindo, na grande maioria das vezes, a revisão da sugestão gerada pelo algoritmo, verifica-se que

¹⁴ VITAL, Danilo. **CNJ avalia se deve proibir juízes de usar ChatGPT para fundamentar decisões**. 26 abr. 2023. Disponível em: <https://www.conjur.com.br/2023-abr-26/cnj-avalia-proibir-juizes-usar-chatgpt-decisoes>. Acesso em: 30 jul. 2023.

¹⁵ TAYLOR, Luke. **Colombian judge says he used ChatGPT in ruling**. 3 feb. 2023. Disponível em: <https://www.theguardian.com/technology/2023/feb/03/colombia-judge-chatgpt-ruling>. Acesso em: 30 jul. 2023.

¹⁶ CONSULTOR JURIDICO. **CNJ vai investigar juiz que usou tese inventada pelo ChatGPT para escrever decisão**. 23 mar. 2023. Disponível em: <https://www.conjur.com.br/2023-nov-12/cnj-vai-investigar-juiz-que-usou-tese-inventada-pelo-chatgpt-para-escrever-decisao/>. Acesso em: 10 jan. 2024.

¹⁷ MCLENNAN, Marsh et al. **WORLD ECONOMIC FORUM**. 10 jan. 2024. Disponível em: <https://www.weforum.org/publications/global-risks-report-2024/>. Acesso em: 10 jan. 2024.

¹⁸ *Ibidem*.

permanece a necessidade de avaliação e monitoramento dos impactos decorrentes, a fim de assegurar o seu uso ético e a proteção aos direitos dos jurisdicionados.

De fato, o COMPAS, o caso mais comentado em todo mundo envolvendo riscos do uso da IA na atividade jurisdicional, não é mais do que mero auxílio à decisão humana, a qual permanece inteiramente nas mãos do magistrado.

Ademais, especificamente no âmbito da atividade judiciária no Brasil, que tem o Conselho Nacional de Justiça – CNJ como órgão regulador¹⁹, a dificuldade de aplicar sanções a tribunais que estiverem em desacordo com as regras estabelecidas, evidencia ainda mais a necessidade de reforço a um protagonismo do papel da ética e da governança na adoção das ferramentas de IA.

Assim, como empresas privadas têm adotado cautelas para o uso da IA, independentemente da existência de leis sobre o tema, parece razoável e recomendado que o Poder Judiciário também desenvolva e implemente mecanismos de controle ético do uso de suas ferramentas de IA criando uma verdadeira estrutura de governança apta a exercer fiscalização de resultados e *enforcement* das regras previstas na Resolução n. 332/2020.

Como reconhecido pelo próprio Conselho Nacional de Justiça, o grande desafio do momento “é a construção de uma adequada governança para todo o Poder Judiciário brasileiro, de modo a permitir sinergia no trabalho colaborativo para entrega de inovação, com o objetivo de melhorar a qualidade do serviço público judiciário”²⁰.

Nesse cenário, este trabalho se propõe a analisar e avaliar criticamente os instrumentos de governança estabelecidos na Resolução CNJ nº 332/2020, bem como apresentar sugestões de aperfeiçoamento com a finalidade de subsidiar futuras iniciativas institucionais.

Pretende-se, em específico, examinar se a criação de comitês de ética de IA no âmbito dos tribunais pode ser um caminho útil para garantir a observância de princípios éticos no desenvolvimento e implementação de ferramentas de IA no Poder Judiciário.

Para tanto, inicialmente serão analisados os papéis desempenhados pela regulação, ética e governança no controle de riscos relacionados à utilização de sistemas de IA, em especial no contexto do Judiciário. Será realizada uma breve análise do cenário regulatório nacional e global no que toca ao desenvolvimento e uso da IA, bem como dos princípios éticos correlatos.

Em seguida, serão examinados os riscos do uso da IA no Poder Judiciário e a

¹⁹ PAULINO, Ana Flávia Borges. **CNJ: o regulador da atividade judiciária**. Rio de Janeiro: Lumen Iuris, 2021.

²⁰ CONSELHO NACIONAL DE JUSTIÇA. *Inteligência artificial na Justiça*. Coord. José Antônio Dias Toffoli; Bráulio Gabriel Gusmão. Brasília: CNJ, 2019. Disponível em: <https://bibliotecadigital.cnj.jus.br/xmlui/handle/123456789/98>. Acesso em: 2 abr. 2022.

regulamentação pertinente, incluindo a análise de documentos internacionais, como a Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais, e o exame crítico dos instrumentos de governança previstos na Resolução CNJ nº 332/2020.

Por fim, será objeto de exame específico a estratégia de criação e implementação de comitês de ética de IA para garantir o controle e mitigação de riscos envolvidos na implementação do uso de sistemas de IA.

Destaca-se que o presente trabalho não pretende apresentar uma solução final sobre a governança de IA no Judiciário brasileiro, mas, sim, fomentar o debate acadêmico e institucional, oferecendo algumas perspectivas e deixando espaço para um diálogo construtivo sobre o tema. Ao final, espera-se contribuir para o avanço do entendimento acerca da matéria e para o aprimoramento das práticas relacionadas ao uso ético da inteligência artificial no contexto do Poder Judiciário.

1. Inteligência Artificial: Regulação, Ética e Governança

1.1. Definindo inteligência artificial

Para tratar do tema de governança de IA no Poder Judiciário, é necessário primeiro estabelecer uma compreensão clara do que é inteligência artificial.

Apesar da ausência de um consenso sobre o conceito técnico dessa expressão, em geral, associa-se à terminologia a sistemas de computador capazes de executar tarefas que normalmente exigiriam inteligência humana.

É importante ressaltar que a inteligência artificial não corresponde a uma única tecnologia, mas sim a diferentes técnicas que utilizam ferramentas computacionais com o objetivo de desenvolver máquinas que executem tarefas que normalmente exigem inteligência humana. Esse domínio de estudo vem se desenvolvendo em ritmo acelerado, com surgimento de novas técnicas e aplicações cada vez mais sofisticadas com o passar dos anos.

No âmbito regulatório, os documentos que tratam do tema geralmente definem sistemas de IA pelo que eles são capazes de fazer, sem atenção para as nuances das técnicas empregadas, oferecendo definições legais, em vez de científicas. Isto é, essas definições de sistema de IA propostas em textos regulatórios incluem diferentes categorias de técnicas, como aprendizado de máquina e abordagens baseadas em conhecimento, e áreas de aplicação, como visão computacional, processamento de linguagem natural, reconhecimento de fala, sistemas robóticos inteligentes, bem como a aplicação dessas ferramentas em diferentes domínios²¹.

Para fins do presente trabalho, mostra-se pertinente destacar algumas definições importantes no cenário nacional e internacional.

A Recomendação da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) sobre a Ética da Inteligência Artificial, aprovada em novembro de 2021, que expressamente destaca não ter “a ambição fornecer uma definição única de IA, uma vez que tal definição precisaria se modificar ao longo do tempo, em conformidade com os desenvolvimentos tecnológicos”, opta por abordar o tema conceituando “os sistemas de IA como sistemas que têm capacidade de processar dados e informações de uma forma que se assemelha ao comportamento inteligente e, normalmente, inclui aspectos de raciocínio,

²¹ OECD. Explanatory memorandum on the updated OECD definition of an AI System. OECD Artificial Intelligence Papers. **OECD Publishing, Paris**, 2024. Disponível em: <https://doi.org/10.1787/623da898-en>. Acesso em: 15 mar. 2024.

aprendizagem, percepção, previsão, planejamento ou controle”²².

Na mesma linha, é a definição utilizada pela Comissão Europeia de Inteligência Artificial, em 2018, segundo a qual “[o] conceito de inteligência artificial (IA) aplica-se a sistemas que apresentam um comportamento inteligente, analisando o seu ambiente e tomando medidas — com um determinado nível de autonomia — para atingir objetivos específicos”²³.

No Brasil, o Relatório Final apresentado pela Comissão de Juristas responsável por subsidiar a elaboração de substitutivo de Projeto de Lei sobre inteligência artificial (CJSUBIA), em dezembro de 2022, propôs a seguinte definição, incorporada no art. 4º, I, do Projeto de Lei n. 2.338/2023:

Sistema de inteligência artificial: sistema computacional, com graus diferentes de autonomia, desenhado para inferir como atingir um dado conjunto de objetivos, utilizando abordagens baseadas em aprendizagem de máquina e/ou lógica e representação do conhecimento, por meio de dados de entrada provenientes de máquinas ou humanos, com o objetivo de produzir previsões, recomendações ou decisões que possam influenciar o ambiente virtual ou real.²⁴

Em novembro de 2023, a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) atualizou a sua definição de sistema de inteligência artificial, realizando alguns ajustes para adequar ao avanço da tecnologia, em especial aos grandes modelos de linguagem, cujo emprego explodiu com a popularização do ChatGPT, para estabelecer que:

Um sistema de IA é um sistema baseado em máquina que, para objetivos explícitos ou implícitos, infere, a partir da entrada que recebe, como gerar resultados (*outputs*), tais como previsões, conteúdo, recomendações ou decisões, que podem influenciar ambientes físicos ou virtuais. Diferentes sistemas de IA variam em seus níveis de autonomia e adaptabilidade após a implantação (tradução livre)²⁵.

Em março de 2024, foi publicado memorando explicativo sobre a definição atualizada

²² UNESCO. **Recomendação sobre a Ética da Inteligência Artificial**. 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 23 nov. 2023.

²³ COMISSÃO EUROPEIA. **Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões: Inteligência artificial para a Europa**. Bruxelas: 2018. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A52018DC0237>. Acesso em: 5 ago. 2023.

²⁴ BRASIL. Congresso Nacional. Senado Federal. Projeto de Lei nº 21, de 2020: Relatório final da Comissão de Juristas responsável por subsidiar a elaboração de substitutivo sobre inteligência artificial. **Diário do Senado Federal**, ano 77, n. 204, suplemento n. B, p. 3-915, 9 dez. 2022.

²⁵ No original: “An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment”. STUART, Russell et al. Updates to the OECD’s definition of an AI system explained. **OECD AI Policy Observatory**. Disponível em: <https://oecd.ai/en/wonk/ai-system-definition-update>. Acesso em: 7 dez. 2024.

da OCDE que apresenta considerações sobre os pontos objeto de mudança, concluindo que:

A definição atualizada de IA é inclusiva e abrange sistemas que vão desde os mais simples até os mais complexos. IA representa um conjunto de tecnologias e técnicas aplicáveis a muitas situações diferentes. Técnicas específicas, como aprendizado de máquina, podem levantar considerações particulares para os formuladores de políticas, como viés, transparência e explicabilidade, e alguns contextos de uso (por exemplo, decisões sobre benefícios públicos) podem levantar preocupações mais significativas do que outros. Por esse motivo, quando aplicada na prática, critérios adicionais podem ser necessários para estreitar ou adaptar de outra forma a definição quando usada em um contexto específico (tradução livre)²⁶.

O texto da proposta de regulação europeia (*Artificial Intelligence Act - AI Act*) incorporou a definição revisada da OCDE da seguinte maneira:

Um sistema de IA é um sistema baseado em máquina projetado para operar com diferentes níveis de autonomia e que pode apresentar adaptabilidade após a implantação e que, para objetivos explícitos ou implícitos, infere, a partir da entrada (*input*) que recebe, como gerar resultados (*outputs*), como previsões, conteúdo, recomendações ou decisões que podem influenciar ambientes físicos ou virtuais (tradução livre)²⁷.

Do exame dos principais documentos no contexto global, nota-se que as definições de IA apresentadas em textos com propósito regulatório adotam um conceito mais amplo, sem a preocupação em apontar técnicas específicas da tecnologia empregada (tais como processamento de linguagem natural, aprendizado de máquina, árvore de decisão, redes adversariais etc.), focando no contexto em que utilizada a ferramenta, a fim de impedir que hipóteses não previstas fiquem fora do espoco da regulamentação²⁸.

Com efeito, destaca-se que a definição ampla utilizada na proposta de regulação europeia (*AI Act*), acompanhada no Projeto de Lei n. 2.338/2023, acaba incluindo sistemas que utilizam tecnologia muito mais simples, que estão em utilização há décadas, tendo como foco o impacto gerado pela sua aplicação.

No âmbito do Poder Judiciário brasileiro, considera-se modelo de Inteligência Artificial, nos termos do art. 3º, III, da Resolução n. 332/2020 do CNJ, o “conjunto de dados e algoritmos

²⁶ OECD. Explanatory memorandum on the updated OECD definition of an AI System. **OECD Publishing, Paris**, 2024. Disponível em: <https://doi.org/10.1787/623da898-en>. Acesso em: 15 mar. 2024.

²⁷ No original: “An AI system is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments”.

²⁸ Nessa linha: ZAVAGLIA COELHO, Alexandre; KLAFFKE, Guilherme Forma; MAITO, Deise Camargo; LATINI, Lucas Maldonado Diz; MARUCA, Giuliana; CHOW, Beatriz Graziano; FEFERBAUM, Marina. **Governança da Inteligência Artificial em Organizações**: Framework para Comitês de Ética em IA – versão 1.0. São Paulo: CEPI FGV Direito SP, 2023. p. 14-15.

computacionais, concebidos a partir de modelos matemáticos, cujo objetivo é oferecer resultados inteligentes, associados ou comparáveis a determinados aspectos do pensamento, do saber ou da atividade humana”²⁹.

Nessa linha de adotar uma definição mais ampla, o art. 2º, parágrafo único, da Portaria CNJ n. 271/2020 dispõe que:

São considerados como de inteligência artificial os projetos voltados a:

- I – Criar soluções para automação dos processos judiciais e administrativos e de rotinas de trabalho da atividade judiciária;
- II – Apresentar análise da massa de dados existentes no âmbito do Poder Judiciário; e
- III – Prover soluções de apoio à decisão dos magistrados ou à elaboração de minutas de atos judiciais em geral³⁰.

Assim, embora não exista consenso sobre o conceito de IA, para fins do presente trabalho será utilizada a perspectiva ampla adotada pelo CNJ na Resolução n. 332/2020 e na Portaria n. 271/2020, que não restringe a definição a sistemas que utilizam técnicas específicas de IA (por exemplo, aprendizado de máquina), mas adota delimitação que abrange qualquer solução tecnológica destinada à automação de atividade, à análise de dados ou ao apoio à tomada de decisões.

Para além de ser o caso da definição adotada na regulamentação específica do CNJ, para a perspectiva do presente trabalho, bem como para fins de debate especificamente sobre o tema do uso IA no domínio delimitado do uso pelo Poder Judiciário, faz sentido falar em IA não apenas no sentido técnico restrito, mas em sentido mais amplo, uma vez que, na presente discussão, o foco deve ser os impactos causados pela adoção da tecnologia, a garantia de respeito aos direitos fundamentais e à preservação de valores importantes do sistema de justiça.

Nessa perspectiva, ferramentas de automatização mais simples se aplicadas pelo Poder Judiciário em sua atividade fim são potencialmente capazes de impactar o próprio sistema de justiça e os direitos dos jurisdicionados, razão pela qual faz sentido a adoção dessa perspectiva ampla.

Ressalta-se que não se está afirmando que esse seja o melhor modelo a ser seguido nos documentos de regulação de IA em geral, nem mesmo em outras regulações específicas em

²⁹ CONSELHO NACIONAL DE JUSTIÇA. **Resolução n. 332, de 21 de agosto de 2020**. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. DJe/CNJ, nº 274, de 25/08/2020, p. 4-8. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 12 fev. 2024.

³⁰ CONSELHO NACIONAL DE JUSTIÇA. **Portaria n. 271, de 04 de dezembro de 2020**. Regulamenta o uso de Inteligência Artificial no âmbito do Poder Judiciário. DJe/CNJ nº 389/2020, de 9/12/2020, p. 2-4. E republicada no DJe/CNJ nº 393/2020, de 14/12/2020, p. 2-4. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em 12 fev. 2024.

outros setores da economia ou da Administração Pública. A posição ora adotada tem como foco, tão somente, o estudo específico do uso da IA pelo Poder Judiciário, buscando garantir o monitoramento de riscos e o respeito a direitos fundamentais.

1.2. O cenário regulatório nacional e global

A inteligência artificial está avançando incrivelmente rápido, oferecendo diversas oportunidades nas mais diferentes áreas do conhecimento, como saúde, educação, segurança pública, recrutamento de pessoal, entre outras, mas também envolve muitos riscos.

Nesse contexto, a regulação do desenvolvimento e uso de sistemas de IA está no centro do debate mundial sobre tecnologia e vem dominando a agenda de países e organizações internacionais importantes nos últimos anos.

Segundo Marcio Iorio Aranha, regulação consiste no conjunto de atividades voltadas a moldar o ambiente regulado “rumo à maior sintonia possível entre o modelo ideal de funcionamento do ordenado pelo Direito e o efetivo comportamento das atividades reguladas”³¹. Nessa linha, o autor define regulação pela combinação das funções normativas e administrativa, da seguinte forma:

A regulação, em síntese, é a presença de regras e atuação administrativa (*law and government*) de caráter conjuntural apoiadas sobre o pressuposto de diuturna reconfiguração das normas de conduta e dos atos administrativos pertinentes para a finalidade de redirecionamento constante do comportamento das atividades submetidas a escrutínio, tendo-se por norte orientador parâmetros regulatórios definidos a partir dos enunciados de atos normativos e administrativos de garantia dos direitos fundamentais.³²

De maneira mais simples, a regulação pode ser definida como o conjunto de medidas adotadas pelo Estado com o objetivo de influenciar o comportamento de indivíduos e organizações por meio de padrões e normas estabelecidos.

Ao lado da regulação estatal, fala-se em autorregulação referindo-se à “existência de regras comportamentais desenvolvidas com alguma participação do próprio destinatário ou exclusivamente desenvolvidas por ele”³³. No ponto, utiliza-se o termo autorregulação regulada para se referir a regulação que “pressupõe um diálogo entre as partes interessadas em um meio

³¹ ARANHA, Márcio Iorio. **Manual de Direito Regulatório: Fundamentos de Direito Regulatório**. 3ª ed. Londres: Laccademia, 2015. p. 269.

³² *Ibidem*. p. 268.

³³ *Ibidem*. p. 86.

termo entre as opções de autorregulação pura e de regulação burocrática central de comando e controle”³⁴.

Importante destacar que “há uma grande gama de opções regulatórias entre os extremos de controle burocrático estatal e de autorregulação pura” e que “os atores regulados podem também ser reconhecidos como reguladores, como também podem partilhar com o regulador estatal funções regulatórias várias”³⁵.

O debate regulatório envolvendo novas tecnologias envolve uma dificuldade adicional quanto aos detalhes do modelo a ser adotado: o pouco conhecimento sobre as consequências decorrentes do emprego da tecnologia e o risco de inibir a inovação e o desenvolvimento econômico.

Considerando essa peculiaridade, os países ao redor do mundo estão estudando e experimentando diferentes propostas de estrutura regulatória para mitigar os riscos do uso da inteligência artificial sem coibir a inovação.

Na União Europeia, após mais de cinco anos de intenso trabalho e debates, o *Artificial Intelligence Act (AI Act - Regulamento de Inteligência Artificial)*, cuja primeira versão foi publicada em abril de 2021, foi aprovado, com ampla maioria de votos, pelo Parlamento Europeu em 12 de março de 2024.

Desde o início, em 2018, os trabalhos realizados no âmbito do Parlamento Europeu para regulação da inteligência artificial buscaram promover a inovação, garantir a segurança e proteger os direitos fundamentais, tendo como foco assegurar que a IA seja centrada no ser humano e confiável.

A iniciativa europeia adota uma abordagem horizontal baseada em riscos considerando usos específicos de IA, de modo que estabelece regras diferentes em quatro níveis diferentes: risco mínimo, alto risco, risco inaceitável e risco de transparência específica. A regulamentação europeia também introduz regras específicas para modelos de IA de propósito geral.

No regulamento europeu, sistemas de inteligência artificial identificados como de alto risco devem atender a requisitos rigorosos, abrangendo medidas eficazes para mitigar riscos, garantir a qualidade dos conjuntos de dados utilizados, registrar atividades, apresentar documentação detalhada, fornecer informações claras aos usuários, manter o controle humano e alcançar elevados padrões de robustez, precisão e cibersegurança.

São exemplos de sistemas de IA considerados de alto risco: aqueles relacionados a

³⁴ *Ibidem*. p. 89.

³⁵ *Ibidem*. p. 91.

infraestruturas críticas nos setores de água, gás e eletricidade; dispositivos médicos; sistemas empregados para determinar o acesso a instituições educacionais ou utilizados no recrutamento de pessoal; e certos sistemas utilizados nas esferas da aplicação da lei, controle de fronteiras, administração da justiça e processos democráticos. Também são classificados como de alto risco os sistemas de identificação biométrica, categorização biométrica e reconhecimento de emoções.

Por sua vez, são considerados proibidos na proposta de regulação europeia os sistemas de IA que constituam uma ameaça clara aos direitos fundamentais das pessoas, tais como sistemas e aplicações de IA que manipulem o comportamento humano para iludir o livre arbítrio dos utilizadores, sistemas que permitam uma “classificação social” por parte dos governos ou das empresas e determinadas aplicações de policiamento preditivo. Serão também proibidas determinadas utilizações de sistemas biométricos, tais como sistemas de reconhecimento de emoções no local de trabalho, de categorização social ou de identificação biométrica remota em tempo real para efeitos da aplicação da lei em espaços acessíveis ao público (com exceções restritas).

Importante destacar que o *AI Act* (Regulamento de Inteligência Artificial) “faz parte de um pacote abrangente de medidas que visa resolver problemas decorrentes do desenvolvimento e da utilização da IA”³⁶.

O modelo de regulação europeu é amplamente reconhecido como uma referência mundial no debate sobre o tema. Embora tenha caráter vinculante apenas dentro da União Europeia, o “Efeito Bruxelas” (*Brussels Effect*) é frequentemente mencionado para descrever o impacto global da regulação estabelecida pelo bloco europeu, influenciando as práticas e padrões adotados por empresas e governos ao redor do mundo. Isso se deve ao fato de que as empresas multinacionais, visando manter o acesso ao vasto mercado europeu, optam por adotar voluntariamente essas regulamentações em suas operações globais³⁷.

Os Estados Unidos, por sua vez, têm adotado estratégia diferente, focada em estabelecer diretrizes setoriais (descentralizada) e contando com maior colaboração do setor privado (autorregulação).

A abordagem americana prioriza o fortalecimento de pesquisa e desenvolvimento da

³⁶ MELO, Ana Karolina Acris et al. **Regulação da Inteligência Artificial: benchmarking de países selecionados**. dez. 2022. Disponível em: <http://repositorio.enap.gov.br/handle/1/7419>. Acesso em: 10 mar. 2024.

³⁷ BRADFORD, Anu. *The Brussels Effect: How the European Union Rules the World*. Nova York: **Columbia Law School**, mar. 2020.

competitividade do país em IA buscando posicionar o país dentre os líderes mundiais na área³⁸. Há um enfoque no papel das agências reguladoras federais, que possuem relativa autonomia e conhecimento do domínio de implementação, adotando uma abordagem setorial para lidar com questões relacionadas à IA em suas áreas de competência.

Em outubro de 2022, o governo americano publicou o “*Blueprint for an AI Bill of Rights*”³⁹ da Casa Branca, que constitui uma declaração de princípios elaborada com o intuito de direcionar padrões para o design, desenvolvimento e implementação de sistemas de IA no país e foi assinado, em caráter voluntário, por diversas empresas privadas. O documento se baseia em cinco princípios: (I) sistemas seguros e eficazes; (II) proteção contra discriminação algorítmica; (III) privacidade de dados; (IV) aviso e explicação; (V) alternativas humanas, consideração e recurso.

Em 30 de outubro de 2023, o governo americano assinou uma ordem executiva para tratar dos riscos do uso da IA, o “*Executive Order on Safe, Secure and Trustworthy Artificial Intelligence*”, que agrupa um conjunto de diretrizes e estabelece como metas: a) desenvolver padrões para segurança e proteção de IA; b) proteger a privacidade dos americanos; c) promover a equidade e os direitos civis; d) defender consumidores, pacientes e estudantes; e) apoiar os trabalhadores; f) promover a inovação e a concorrência; g) avançar a liderança global americana; e h) garantir o uso responsável e eficaz da IA pelo governo⁴⁰.

Apesar do documento ser vinculante para os órgãos públicos, obrigando-os a estabelecer regras e normas, produzir relatórios e criar canais de financiamento, além de tornar obrigatória a observância do “*Blueprint for an AI Bill of Rights*” para o governo federal, não possui força de lei, de modo que o governo deverá contar com a legislação existente para concretizá-lo no que toca à iniciativa privada.

Nesse contexto, a “Lei de Produção de Defesa” de 1950, confere autoridade ao Presidente americano para obrigar as empresas privadas a apoiar esforços de segurança nacional, o que pode ser utilizado para exigir que provedores de computação em nuvem (como Amazon, Microsoft e Google) informem ao governo sobre seus clientes estrangeiros e que os

³⁸ MELO, Ana Karolina Acris et al. **Regulação da Inteligência Artificial: benchmarking de países selecionados**. dez. 2022. Disponível em: <http://repositorio.enap.gov.br/handle/1/7419>. Acesso em: 10 mar. 2024.

³⁹ THE WHITE HOUSE. **Blueprint for an AI Bill of Rights**. 24 jun. 2022. Disponível em: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>. Acesso em: 10 mar. 2024.

⁴⁰ THE WHITE HOUSE. **FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence**. 30 oct. 2023. Disponível em: <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>. Acesso em: 10 mar. 2024.

desenvolvedores de sistemas de IA compartilhem resultados de testes de segurança⁴¹.

Assim, a Ordem Executiva do presidente Biden navega por um caminho intermediário, “permitindo que o desenvolvimento da IA continue praticamente imperturbado, ao mesmo tempo em que implementa algumas regras modestas e sinaliza que o governo federal pretende ficar ‘de olho’ na indústria da IA nos próximos anos”⁴² (tradução livre).

No Reino Unido, a abordagem regulatória também é setorial, garantindo liberdade para regulamentação do tema em um contexto específico, assegurando atuação direcionada e que apoie a inovação e evite barreiras ao desenvolvimento⁴³. Há preocupação também de estabelecer princípios intersetoriais, que possam ser interpretados dentro de cada domínio, e incentivo à coordenação regulatória⁴⁴.

Desse modo, constata-se que os países utilizam abordagens regulatórias bastante diversas. Enquanto a União Europeia adota uma abordagem geral (horizontal) baseada em riscos, a estratégia americana tem sido apostar em uma regulação setorial (vertical) de monitoramento.

No Brasil, o debate sobre a regulação da IA ganhou contornos mais definidos com a proposição do Projeto de Lei nº 21/2020 e sua aprovação pela Câmara dos Deputados. No desenrolar do referido processo legislativo, no âmbito do Senado Federal, foi instituída uma Comissão de Juristas (CJSUBIA) responsável por subsidiar a elaboração de substitutivo para instruir a apreciação dos Projetos de Lei nºs 5.051/2019, 21/2020 e 872/ 2021, que têm como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil.

A referida comissão apresentou, em dezembro de 2022, um relatório de 908 páginas, contendo um resumo dos debates realizados e das contribuições recebidas, bem como minuta de substitutivo fortemente inspirado no modelo de regulação europeu. A proposta da CJSUBIA deu origem ao Projeto de Lei n. 2.338/2023.

Fortemente inspirado no projeto de regulamento de IA da União Europeia (AI Act), o

⁴¹ KAUFMAN, Dora. **Dez dias que abalaram o sistema de IA: declaração de Betchely, ordem executiva do governo Biden, órgão consultivo da ONU e artigo de acadêmicos**. 3 nov. 2023. Disponível em: <https://epocanegocios.globo.com/colunas/iagora/coluna/2023/11/dez-dias-que-abalaram-o-ecossistema-de-ia-declaracao-de-bletchely-ordem-executiva-do-governo-biden-orgao-consultivo-da-onu-e-artigo-de-academicos.ghtml>. Acesso em: 10 mar. 2024.

⁴² ROOSE, Kevin. **With Executive Order, White House Tries to Balance A.I.’s Potential and Peril**. 31 oct. 2023. Disponível em: <https://www.nytimes.com/2023/10/31/technology/executive-order-artificial-intelligence-regulation.html?smid=nytcore-ios-share&referringSource=articleShare>. Acesso em: 10 mar. 2024.

⁴³ MELO, Ana Karolina Acris et al. **Regulação da Inteligência Artificial: benchmarking de países selecionados**. dez. 2022. Disponível em: <http://repositorio.enap.gov.br/handle/1/7419>. Acesso em: 10 mar. 2024.

⁴⁴ *Ibidem*.

PL 2.338/2023, apresenta uma abordagem baseada em riscos e direitos, além de prever detalhadas medidas de governança. A redação proposta pela Comissão de Juristas aponta conceitos e princípios aplicáveis ao desenvolvimento e uso de sistemas de IA no Brasil, elenca direitos das pessoas afetadas e define regras para categorização dos riscos a serem identificados por meio de avaliação de impacto algorítmico.

Atualmente, está em andamento um trabalho da Comissão Temporária sobre IA (CTIA) no Senado criada para examinar os projetos concernentes ao relatório final aprovado pela Comissão de Juristas responsável por subsidiar a elaboração de substitutivo sobre Inteligência Artificial no Brasil, bem como eventuais novos projetos que disciplinem a matéria.

Em novembro de 2023, foi apresentado um substitutivo ao PL n. 2338/2023, composto por 22 artigos, que pretendem estabelecer princípios para o fomento, o desenvolvimento e o uso seguro, confiável e responsável da Inteligência Artificial (IA)⁴⁵.

Em fevereiro de 2024, foi apresentado o PL nº 210/2024, que dispõe sobre princípios para uso de IA no Brasil. Define sistemas de IA e afirma que o uso de sistemas de IA devem observar os princípios: (i) segurança e efetividade dos sistemas; proteção contra discriminação de algoritmo; (ii) garantia à privacidade dos dados e informações; (iii) direito à informação; (iv) direito à opção pelo tratamento humano; e (v) direito à contestação. Além disso, define que o indivíduo tem direito à proteção contra sistemas de IA que sejam inseguros ou ineficazes. Para demonstrar a segurança e eficácia, os sistemas de IA devem ser (i) desenvolvidos mediante consulta às comunidades, partes interessadas e especialistas para identificar riscos e potenciais impactos, e (ii) submetidos a testes de pré-implantação, identificação e mitigação de riscos.

Ainda em fevereiro de 2024, foi também apresentado o projeto de lei n. 266/2024, que dispõe sobre o uso de sistemas de IA para auxiliar a atuação de médicos, advogados e juízes. Em relação aos médicos, o PL define que sistemas de IA poderão ser utilizados para auxiliar a atuação do médico no campo da atenção à saúde, desde que preservada a autonomia do médico no desenvolvimento de suas ações profissionais. O texto do PL indica que o uso de sistemas de IA sem a supervisão de médico configuraria exercício ilegal da Medicina.

No que toca à advocacia, o PL n. 266/2024 propõe alterações no Estatuto da OAB a fim de incluir dispositivo que permite expressamente o uso de sistemas de IA para auxiliar a atividade de advocacia pública ou privada, devendo preservar a isenção técnica e a independência profissional do advogado, bem como enquadra como exercício ilegal da

⁴⁵ BRASIL. Senado Federal. **Projeto de Lei nº 2338, de 4 de maio de 2023**. Dispõe sobre o uso da Inteligência Artificial. Brasília: Senado Federal, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 17 jan. 2024.

advocacia o uso de referidos sistemas por pessoa não inscrita na OAB.

O referido projeto de lei propõe também acréscimo de dispositivo no Código de Processo Civil, de modo que passe a constar a previsão de possibilidade de uso de sistemas de IA para auxiliar a prática de atos processuais, bem como estabelece que “os pronunciamentos previstos no art. 203, quando elaborados com auxílio de sistemas de inteligência artificial, serão submetidos aos juízes para revisão e assinatura, sob pena de nulidade”.

Nos últimos cinco anos, foram apresentados, ao menos, 15 (quinze) projetos de lei abordando o uso de inteligência artificial no Brasil: PLs 3.592/2023; 2.338/2023; 5.691/2019; 5.051/2019; 21/2020; 872/2021; 266/2024; 145/2024; 210/2024; 146/2024; 262/2024; 390/2024; 303/2024; 349/2024; 370/2024⁴⁶.

1.3. Os princípios éticos da IA

Esse cenário de extenso debate em todo o mundo acerca do tema evidencia que regular o uso da IA não é tarefa simples.

Muito se fala, inclusive, que para estimular a inovação seria necessária uma legislação *ex post*, ou seja, que apenas seja elaborada após a efetiva implementação das ferramentas tecnológicas e quando estiver disponível mais conhecimento sobre o seu funcionamento e os riscos envolvidos diante de evidências e acontecimentos práticos.

De fato, costuma-se dizer que o Direito sempre vem a reboque das mudanças trazidas pelas tecnologias na sociedade. Conforme apontado por Caitlin Mulholland, “ao jurista não é concedida a capacidade de previsão antecipada de todas as consequências, benéficas ou maléficas, decorrentes do desenvolvimento tecnológico”⁴⁷.

Ocorre que deixar para regular apenas quando houver domínio abrangente e profundo sobre a tecnologia em si e evidências quanto seus impactos sociais, éticos, econômicos e ambientais pode tornar difícil a sua implementação e alteração do comportamento dos regulados, assim como o controle dos efeitos adversos.

De todo modo, no atual momento, o mapeamento de riscos identificados decorrentes da

⁴⁶ BELLI, Luca. Regulação da inteligência artificial para inglês ver?. **Jota**, 6 de março de 2024. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/ia-regulacao-democracia/regulacao-da-inteligencia-artificial-para-ingles-ver-06032024>. Acesso em: 10 mar. 2024.

⁴⁷ MULHOLLAND, Caitlin Sampaio. Mercado, Pessoa Humana e Tecnologias: a Internet das Coisas e a proteção do direito à privacidade. In: EHRHARDT JÚNIOR Marcos; CORTIANO JUNIOR Eroulths (coord.). Transformações no direito privado nos 30 anos da Constituição: estudos em homenagem a Luiz Edson Fachin. Belo Horizonte: **Fórum**, 2019, p. 103-116.

utilização de diferentes sistemas IA nas mais variadas áreas nos últimos anos tem demonstrado a necessidade de estabelecer mecanismos de controle para evitar prejuízos sociais e econômicos, entre outros.

Recentemente, o Relatório Global de Riscos do Fórum Econômico Mundial de 2024 apontou três riscos relacionados à tecnologia para o mundo nos próximos anos: a desinformação, a cibersegurança e os efeitos adversos da inteligência artificial.

Muito do que está em debate no cenário regulatório atual é decorrência de incidentes já mapeados de aplicações práticas de IA no contexto internacional, envolvendo principalmente vieses discriminatórios e violações de privacidade.

Nessa perspectiva, nos últimos anos, em paralelo ao debate sobre construção de uma regulação adequada para tratar do tema sem impedir a inovação, surgiram diversas iniciativas, da academia, de empresas privadas, do setor público, bem como de organizações não governamentais e órgãos internacionais, propondo princípios éticos para orientar a utilização de sistemas de IA de forma responsável.

A AlgorithmWatch, organização não governamental sem fins lucrativos, mapeou até abril de 2020, mais de 167 iniciativas, incluindo documentos do setor público, da iniciativa privada e de organizações não governamentais, de compilação de princípios ou diretrizes de ética de IA⁴⁸.

Levantamento realizado pelo Conselho da Europa aponta que, entre 2010 e 2016, identificou-se 89 documentos relacionados a iniciativas de IA, enquanto entre 2017 e novembro de 2021 foram identificados 417 documentos sobre a temática ao redor do mundo⁴⁹. Atualmente, a compilação aponta 605 iniciativas.

Dentre esses documentos, está o guia de Orientações Éticas para uma IA de Confiança, elaborado pelo Grupo de Peritos de Alto Nível sobre a Inteligência Artificial (GPAN IA), colegiado independente criado pela Comissão Europeia em 2018, que estabelece como ideia geral que os sistemas de IA sejam desenvolvidos e implementados de uma forma consentânea com os princípios éticos de respeito da autonomia humana, prevenção de danos, equidade e explicabilidade⁵⁰.

Nessa linha, foram enumerados sete requisitos que devem ser cumpridos para alcançar

⁴⁸ INVENTORY. **AI Ethics Guidelines Global Inventory**. apr. 2020. Disponível em: <https://inventory.algorithmwatch.org/>. Acesso em: 01 nov. 2023.

⁴⁹ COUNCIL OF EUROPE. **AI initiatives**. France, 2022. Disponível em: <https://www.coe.int/en/web/artificial-intelligence/national-initiatives>. Acesso em: 7 dez. 2023.

⁵⁰ EUROPEAN COMMISSION. **Ethics guidelines for trustworthy AI**. 8 apr. 2019. Disponível em: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai_. Acesso em: 01 nov. 2023.

uma IA de confiança, quais sejam: 1) ação e supervisão humanas; 2) solidez técnica e segurança; 3) privacidade e governação dos dados; 4) transparência; 5) diversidade, não discriminação e *fairness* (equidade); 6) bem-estar ambiental e societal; 7) *accountability* (responsabilização)⁵¹.

Em 2020, quando a publicação do documento de Considerações Setoriais sobre as Recomendações Políticas e de Investimento, o Grupo de Peritos de Alto Nível sobre a Inteligência Artificial (GPAN IA) apontou que, no que toca ao setor público de justiça e forças de segurança pública (*law-enforcement*), em geral as expectativas para o uso de sistemas de IA são muito altas e que o emprego das ferramentas em larga escala pode gerar riscos e oportunidades ainda não completamente compreensíveis. Concluíram, no ponto, que:

É necessários mais pesquisa, escrutínio e deliberação antes de formular orientações legais, éticas ou políticas. Portanto, seria importante iniciar um amplo debate de políticas na Europa (e além) sobre o desenvolvimento, uso e impacto de sistemas de tomada de decisões assistidos por IA e habilitados por IA na justiça e na aplicação da lei⁵² (tradução livre).

Outro importante documento no cenário mundial é a Recomendação sobre Inteligência Artificial do Conselho da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), de maio de 2019, que busca, por meio dos princípios éticos nela elencados, estabelecer uma base ética para o desenvolvimento e uso da inteligência artificial.

O Brasil, apesar de não ser membro da referida organização, aderiu ao documento, demonstrando um compromisso com os princípios estabelecidos, os quais, no entanto, não possuem caráter vinculante do ponto de vista legal.

A recomendação da OCDE estabelece os seguintes princípios para a gestão responsável da IA confiável: crescimento inclusivo, desenvolvimento sustentável e bem-estar; valores centrados no ser humano e *fairness* (equidade ou justiça); transparência e explicabilidade;

⁵¹ No documento português de Portugal as palavras em inglês foram traduzidas pelas palavras em parênteses.

⁵² No original: “*The expectations as to the capabilities of current AI systems to perform these tasks reliably are, however, generally too high. Although such systems are already used in specific law enforcement and justice contexts, the AI HLEG considers that their deployment at greater scale generates risks and opportunities that are not yet fully understood. More research, scrutiny and deliberation are needed prior to formulating legal, ethical or policy guidance. It would be, therefore, important to launch a wide-spread policy debate in Europe (and beyond) on the development, use and impact of AI-assisted and AI-enabled decision-making systems in justice and law enforcement*”. DERIVASHJ, Jola. **AI HLEG - Sectoral Considerations on Policy and Investment Recommendations for Trustworthy AI**. 23 July. 2020. Disponível em: <https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai?language=pt-pt>. Acesso em: 4 nov. 2023. BAHTTI, Sophia Adams. et al. **Artificial Intelligence (AI) and the Criminal Justice System**. 5-6 nov. 2020. Disponível em: <https://www.gov.si/assets/ministrstva/MP/Predsedovanje/2021-07-20-Konferenca-AI/ERA-AI-conference-5.11.-6.11.2020-compilation-Version-4.11.2020.pdf> Acesso em: 01 nov. 2023.

robustez e segurança; e *accountability* (prestação de contas e responsabilização)⁵³.

Em novembro de 2021, a Conferência Geral da UNESCO aprovou a Recomendação sobre a Ética da Inteligência Artificial⁵⁴, apontada como o primeiro instrumento global voltado para a construção de uma base ética de IA e controle de riscos. O texto do referido documento recebeu contribuição de 24 especialistas de todo o mundo e buscou “trazer um instrumento normativo aceito em âmbito global, com foco não apenas na articulação de valores e princípios, mas também em sua realização prática, por meio de recomendações políticas concretas”⁵⁵.

O referido documento, de fato, vai além da enunciação de princípios e aponta ações concretas a serem desenvolvidas pelos Estados-membros para implementar os valores e preceitos da Recomendação, a fim de que lidem de forma responsável com os impactos conhecidos e desconhecidos das tecnologias de IA sobre seres humanos, sociedades, meio ambiente e ecossistemas.

Com base na recomendação da UNESCO, em julho de 2022, o Comitê de Alto Nível sobre Programas (*High-Level Committee on Programmes* - HLCP), responsável pela coerência, cooperação e coordenação nas dimensões programáticas de questões estratégicas enfrentadas pelo sistema das Nações Unidas, aprovou documento apontando os princípios para o uso ético da IA no âmbito das entidades integrantes do sistema das Nações Unidas.

Os princípios elencados pelo HLCP são: não causar danos; propósito definido, necessidade e proporcionalidade; segurança; equidade (*fairness*) e não discriminação; sustentabilidade; direito à privacidade, proteção de dados e governança de dados; autonomia humana e supervisão; transparência e explicabilidade; responsabilidade e prestação de contas (*accountability*); e inclusão e participação⁵⁶.

Mais recentemente, em 30 de outubro de 2023, os dirigentes do G7 firmaram acordo sobre princípios orientadores internacionais e um código de conduta voluntário para Organizações Desenvolvendo Sistemas de IA Avançada⁵⁷. Os referidos documentos foram

⁵³ OECD. **Recommendation of the Council on Artificial Intelligence**, OECD/LEGAL/0449. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#relatedInstruments>. Acesso em: 01 nov. 2023.

⁵⁴ UNESCO. **Recomendação sobre a Ética da Inteligência Artificial**. 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 23 nov. 2023.

⁵⁵ *Ibidem*.

⁵⁶ HLCP. **Principles for the Ethical Use of Artificial Intelligence in the United Nations System**. 2022. Disponível em: https://unsceb.org/sites/default/files/2022-09/Principles%20for%20the%20Ethical%20Use%20of%20AI%20in%20the%20UN%20System_1.pdf. Acesso em: 01 de nov. 2023.

⁵⁷ UNIVERSITY OF TORONTO. **G7 Leaders' Statement on the Hiroshima AI Process**. 1 dec. 2023. Disponível em: <http://www.g7.utoronto.ca/ict/2023-statement-2.html#:~:text=We%20are%20dedicated%20to%20promoting,trustworthy%20AI%20and%20risk%20managem ent>. Acesso em: 01 de nov. 2023.

elaborados para serem atualizados e apresentam uma abordagem mais prática com orientações de ação para os desenvolvedores de modelo de IA.

Em 1 de novembro de 2023, o Reino Unido, os EUA, a União Europeia, o Brasil, a China, o Japão, e outros 22 países que participavam da Cúpula de Segurança de IA (*AI Safety Summit*) assinaram a Declaração de Bletchely (*Bletchley Declaration on AI Risks*), cujo objetivo é promover a cooperação internacional na área da IA visando o crescimento econômico inclusivo, o desenvolvimento sustentável e a inovação, proteger os direitos humanos e fortalecer a confiança pública nos sistemas de IA para realizar plenamente seu potencial⁵⁸.

Os Estados signatários concordam no documento que os modelos de IA apresentam potencial de sério risco de dano, inclusive catastrófico, e apontam que a cooperação terá como foco: identificar os riscos de segurança da IA que são de preocupação compartilhada e construir uma compreensão científica e baseada em evidências desses riscos no contexto de uma abordagem global para compreender o impacto da IA na sociedade; bem como construir políticas de gestão de riscos nos países signatários para garantir a segurança diante desses riscos, o que inclui maior transparência por parte dos atores privados, a definição de métricas de avaliação apropriadas, ferramentas para testes de segurança e o desenvolvimento do setor público e pesquisa científica.

Uma análise detalhada dos principais documentos que estabelecem princípios éticos para a inteligência artificial permite verificar que muito do seu conteúdo se repete com diferentes nomenclaturas e pequenas alterações conceituais, cuidando basicamente dos seguintes temas: privacidade; *accountability* (prestação de contas e responsabilização); segurança; transparência e explicabilidade; *fairness* e não discriminação; controle humano; responsabilidade profissional e promoção de valores humanos⁵⁹.

Com o objetivo de condensar os princípios éticos de IA, em 2019, Luciano Floridi et al. resumiu-os em cinco, dos quais quatro são princípios da bioética reinterpretados dentro do contexto da IA, quais sejam: beneficência, não maleficência, autonomia, justiça, e explicabilidade, que por sua vez engloba as ideias de inteligibilidade e *accountability*⁶⁰.

⁵⁸ POLICY PAPER. **The Bletchley Declaration by Countries Attending the AI Safety Summit**. 1 nov. 2023. Disponível em: <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>. Acesso em: 4 nov. 2023.

⁵⁹ FJELD, Jessica et al. **Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI**. Berkman Klein Center Research Publication, n. 2020-1, 2020.

⁶⁰ FLORIDI, Luciano; COWLS, Josh. A Unified framework of five principles for AI In Society. **Harvard Data Science Review**, Issue 1.1, Summer 2019. Disponível em: <https://hdsr.mitpress.mit.edu/pub/10jsh9d1/release/8>. Acesso em: 4 nov. 2023.

Assim, é possível afirmar que há algum tempo já se tem um certo consenso em relação aos princípios éticos da inteligência artificial.

Os avanços para concretizar esses princípios na prática, contudo, ainda ocorrem lentamente, de modo que nos últimos anos essa abordagem principiológica vem sendo criticada ante sua insuficiência, assim como tem sido empregados esforços para traduzi-los em recomendações de ações específicas e concretas.

De fato, a eficiência do estabelecimento de princípios éticos para conformação de padrões de comportamento no âmbito do desenvolvimento de sistemas de IA não tem se mostrado tão eficiente como ocorreu no campo da bioética⁶¹.

Pode-se apontar, como fator de distinção, o fato de que os profissionais e entidades atuantes das áreas da saúde já são fortemente regulamentados por leis e conselhos de profissionais capazes de aplicar sanções concretas em caso de inobservância das diretrizes éticas aplicáveis, o que não ocorre no caso de profissionais da ciência da computação. Ademais, o consenso sobre padrões éticos na área da saúde se estende por muitos séculos (o juramento de Hipócrates remonta ao século V a.C.) enquanto a construção de uma base ética sobre o uso de aplicações de inteligência artificial ainda está em evolução, sendo um desafio mundial contemporâneo.

Em 2021, Floridi já apontava que, para além dos princípios e diretrizes, é necessária a criação de regras e sanções para garantir o uso ético das ferramentas de inteligência artificial⁶².

Para assegurar um ambiente em que haja mitigação de riscos e garantia de direitos fundamentais, em especial em uma área sensível como o sistema de justiça, faz-se necessário encontrar uma posição intermediária entre a lentidão da regulação estatal por meio da edição de normas legislativas vinculantes e as diretrizes éticas destituídas de força cogente e de mecanismos de controle.

Nesse contexto, encontra-se a governança como “uma camada intermediária entre a ética e a regulação, como forma de reduzir os potenciais prejuízos gerados a indivíduos em cenários de uso dessas tecnologias na contemporaneidade”⁶³.

⁶¹ CORRÊA, Nicholas Kluge; DE OLIVEIRA, Nythamar Fernandes; MASSMANN, Diogo Fernando. Sobre a eficiência da ética como ferramenta de governança da inteligência artificial. **Veritas (Porto Alegre)**, v. 67, n. 1, p. e42584-e42584, 2022.; MITTELSTADT, Brent. Principles alone cannot guarantee ethical AI. *Nature machine intelligence*, v. 1, n. 11, p. 501-507, 2019.

⁶² FLORIDI, Luciano. The end of an era: From self-regulation to hard law for the digital industry. **Philosophy & Technology**, v. 34, p. 619-622, 2021.

⁶³ DA SILVA, Alexandre Pacheco; FEFERBAUM, Marina. Crônica de um problema anunciado. In **Ética, Governança e Inteligência Artificial**. Almedina Brasil, 2023. pp. 19-36. p. 23

1.4. O papel da governança

Com efeito, nesse cenário de debate sobre princípios e regulação do uso da IA nos últimos anos, o termo governança tem sido utilizado com frequência para se referir a práticas e processos não cogentes destinados a direcionar a utilização de IA de maneira ética e responsável.

O conceito de governança, oriundo do âmbito corporativo, pode ser compreendido como “um sistema formado por princípios, regras, estruturas e processos pelo qual as organizações são dirigidas e monitoradas, com vistas à geração de valor sustentável para a organização, para seus sócios e para a sociedade em geral”⁶⁴.

A governança das organizações está, assim, relacionada à forma de direção, supervisão e controle, bem como ao cumprimento de valores previamente definidos, com o objetivo de entrega de resultados às partes interessadas.

Uma boa governança, portanto, depende da definição e da implementação de mecanismos e práticas que garantam aos interessados que o funcionamento da organização será alinhado a valores e princípios previamente estabelecidos de modo a balancear adequadamente os interesses envolvidos.

Desse modo, ainda que a estrutura de governança possa ter um papel significativo para assegurar conformidade com normas legais, vai muito além do controle da observância de regras cogentes, pois tem o papel de estabelecer a direção a ser seguida e de implementar um sistema de gerenciamento de riscos.

No âmbito das organizações públicas, governança pode ser definida como:

A aplicação de práticas de liderança, de estratégia e de controle, que permitem aos mandatários de uma organização pública e às partes nela interessadas avaliar sua situação e demandas, direcionar a sua atuação e monitorar o seu funcionamento, de modo a aumentar as chances de entrega de bons resultados aos cidadãos, em termos de serviços e de políticas públicas⁶⁵.

No que toca às instituições privadas, os mecanismos de governança asseguram a existência de formas de controle por parte dos acionistas sobre as decisões e o desempenho das empresas, bem como são essenciais para garantir a reputação das companhias.

⁶⁴ IBGC. Instituto Brasileiro de Governança Corporativa. Disponível em: <https://www.ibgc.org.br/conhecimento/governancacorporativa#:~:text=Governan%C3%A7a%20corporativa%20%C3%A9%20um%20sistema,para%20a%20sociedade%20em%20geral>. Acesso em: 19 jan. 2024.

⁶⁵ BRASIL. Tribunal de Contas da União. Referencial básico de governança aplicável a organizações públicas e outros entes jurisdicionados ao TCU. 3 ed. **Brasília: TCU, Secretaria de Controle Externo da Administração do Estado – Secex Administração**, 2020. p. 15.

No contexto público, a implementação de uma estrutura de governança proporciona diversos benefícios, dentre os quais se destacam: assegurar que a organização seja percebida como responsável perante os cidadãos; promover transparência ao manter a sociedade informada sobre decisões tomadas e riscos envolvidos; compilar e empregar informações de qualidade e mecanismos robustos de apoio à tomada de decisão; estabelecer um diálogo com a sociedade e prestar contas a ela; garantir a presença de um sistema eficaz de gestão de riscos; e disponibilizar aos cidadãos dados e informações de qualidade⁶⁶.

A implementação de estrutura de governança possibilita, assim, que as organizações públicas atendam ao interesse público, reportando-se aos cidadãos sobre estratégias, políticas internas e monitoramento de resultados.

Quando se trata do desenvolvimento e uso de sistemas de IA, é essencial que as organizações obtenham uma compreensão aprofundada das implicações éticas associadas à utilização dessa tecnologia em suas atividades para que possam direcionar a implementação e gerenciar os riscos envolvidos.

No caso de governança de IA, dado o caráter essencialmente multidisciplinar da tecnologia, bem como o fato de a regulação ainda estar em estágios iniciais, os princípios éticos são frequentemente incorporados às políticas de governança.

Assim, enquanto a ética de IA contempla a reflexão sobre “padrões amplamente aceitos de certo e errado para orientar a conduta moral no desenvolvimento e uso de tecnologias de IA”⁶⁷, a regulação estabelece o que é legal e ilegal, por meio de normas cogentes, constituindo um padrão ético mínimo obrigatório. A governança, por sua vez, concentra-se em direcionar e monitorar a gestão da organização para a realização de propósitos específicos.

No âmbito da regulação da IA, a ética contribui para a formulação de diretrizes que orientam a conduta apropriada em diversas situações. A governança, então, é concebida como um meio para efetivar os princípios éticos dentro das organizações, mediante a implementação de políticas, práticas e processos direcionados ao monitoramento, avaliação e alcance dos resultados desejados, considerando objetivos, interesses e riscos envolvidos.

Nessa perspectiva, os conceitos de ética, regulação e governança estão interrelacionados, conforme bem apontado por Luciano Floridi, em 2018, no artigo “*Soft Ethics*

⁶⁶ *Ibidem*. pp. 29-30.

⁶⁷ Tradução livre de LESLIE, David. Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. **The Alan Turing Institute**, 2019. Disponível em: <https://doi.org/10.5281/zenodo.3240529>. Acesso em: 19 mar. 2024.

and the Governance of the Digital”⁶⁸.

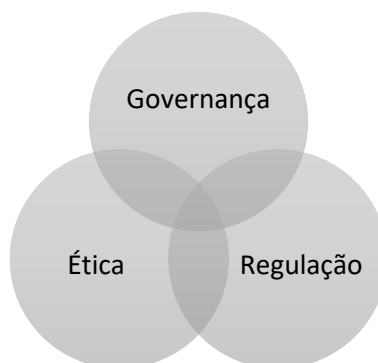


Figura 1: A relação entre ética digital, regulação digital e governança digital

Fonte: Elaborado pela autora a partir da imagem de Luciano Floridi em *Soft Ethics and the Governance of the Digital*. *Philosophy & Technology*. v. 31, n. 1, p. 1–8, 2018. p. 3.

Sobre o tema, indicando a influência da ética IA na regulação, o professor Fabiano Hartmann aponta que:

Uma IA ética impacta e deve se relacionar indissociavelmente aos planos normativo e social, influenciando-os, isto é, deve ser fundamento axiológico para o processo de normatização (que inclusive deve ser aberto o suficiente para receber as demandas concretas da visão ética ao longo do processo contínuo de inovação); bem como deve ser socialmente relevante em termos de benefícios, proteção e segurança social⁶⁹.

Desse modo, em especial em um cenário de construção e consolidação de normas regulatórias, bem como de evolução acelerada da técnica, a governança de IA é importante para construir uma cultura de uso da tecnologia de forma responsável, permitindo ou facilitando uma posterior conformidade com a futura legislação e assegurando fortalecimento da confiança das pessoas nas instituições.

Diante das diversas iniciativas buscando regular o uso de sistemas de IA e das recomendações governamentais e internacionais, desenvolveu-se uma profusão de documentos que apontam práticas de governança com o objetivo de assegurar que o desenvolvimento e utilização dessas ferramentas observem os princípios éticos estabelecidos, bem como que seja assegurada *accountability* (prestação de contas e responsabilidade).

Frequentemente são apontados como instrumentos de governança de IA: elaboração de carta de princípios; adoção de um código de conduta; treinamento de pessoal; existência de comitê ou comissão de ética ou cargo específico com responsabilidade de supervisão;

⁶⁸ FLORIDI, Luciano. *Soft Ethics and the Governance of the Digital*. *Philosophy & Technology*. v. 31, n. 1, p. 1–8, 2018. Disponível em: <https://doi.org/10.1007/s13347-018-0303-9>. Acesso em: 12 ago. 2023.

⁶⁹ PEIXOTO, Fabiano Hartmann. Inteligência artificial e direito: convergência ética e estratégica. *Alteridade Editora*, 2020. p. 153.

auditorias; avaliação de impacto; e certificação.

Em um contexto em que a discussão sobre uso ético e responsável de IA está em alta, não raramente instituições adotam medidas superficiais com o objetivo de parecer genuinamente preocupadas com os riscos éticos sem implementar mudanças substanciais para abordar essas questões, focando apenas na aparência enganosa de comprometimento ético.

Essa prática, denominada de “*ethics bluwashing*”, “*ethics washing*” ou apenas “*bluwashing*”, assemelha-se à conduta de instituições que buscam passar uma imagem de preocupação com o meio ambiente para melhorar sua reputação, mas que não passa de uma fachada (*greenwashing*)⁷⁰.

De fato, a sinalização da existência de ferramentas de governança nem sempre acarretará um efetivo gerenciamento de riscos. Nessa perspectiva, pesquisa realizada com base nas informações públicas divulgadas por 254 empresas chegou à conclusão de que o volume de atividades éticas de IA relatadas é baixo e que os sinais típicos de governança, incluindo a existência de princípios éticos de IA, não estão associados à efetiva implementação dessas práticas⁷¹.

Segundo os dados coletados, cerca de um terço (27% - vinte e sete por cento) de todas as empresas que exibem pelo menos um sinal de governança não apresentou nenhuma atividade de implementação⁷². No caso de empresas que possuem um a dois tipos de sinais de governança, metade não apresentou atividades de implementação, enquanto os restantes 30% (trinta por cento) apresentaram exatamente um tipo de atividade de implementação⁷³.

Os dados apresentados evidenciam que não basta a adoção formal de medidas de governança, como a publicação de princípios éticos ou estabelecimento de comitês de revisão destituídos de poder ou avaliação crítica, sem efetiva modificação na prática da organização, bem como que a referida prática é mais comum do que se imagina.

Floridi aponta que *ethics bluwashing* é especialmente atraente no contexto da IA devido a inúmeros desafios éticos, altos custos associados a fazer a coisa certa, incerteza sobre normas aplicáveis ao caso e confusão generalizada⁷⁴. O filósofo italiano defende que “a melhor

⁷⁰ FLORIDI, Luciano. Translating principles into practices of digital ethics: Five risks of being unethical. *Philosophy & Technology*, v. 32, n. 2, p. 185-193, 2019.

⁷¹ DOTAN, Ravi et al. Evaluating AI Governance. *Insights from Public Disclosures*. 2023. Disponível em: <https://www.ravitdotan.com/evaluating-ai-governance>. Acesso em: 19 mar. 2024.

⁷² *Ibidem*.

⁷³ *Ibidem*.

⁷⁴ No original: “*These days, ethics bluwashing is especially tempting in the context of AI, where the ethical issues are many, the costs of doing the right thing may be high, and normative uncertainty or sometimes confusion are widespread*”. FLORIDI, Luciano. Translating principles into practices of digital ethics: Five risks of being unethical. *Philosophy & Technology*, v. 32, n. 2, p. 185-193, 2019. P. 3-4.

estratégia contra o *bluewashing* é a mesma já adotada para o caso do *greenwashing*: transparência e educação”⁷⁵.

Nesse contexto, recomenda-se que as instituições preocupadas em evitar a aparência de envolvimento em más práticas de *bluewashing* devem priorizar a transparência pública, responsável e baseada em evidências sobre práticas éticas e que envidem esforços educacionais para seus executivos e membros de conselhos consultivos⁷⁶.

⁷⁵ *Ibidem.* p. 4.

⁷⁶ *Ibidem.* p. 4.

2. Governança do Uso da Inteligência Artificial no Poder Judiciário Brasileiro

2.1. Riscos do uso da IA no Poder Judiciário

A inteligência artificial é um campo de estudo que se desenvolveu exponencialmente nos últimos anos. A popularização do uso da referida tecnologia nos mais diferentes setores ocorre por se mostrar uma poderosa aliada no aumento da eficiência por meio de análise de grande volume de dados com rapidez, identificação de padrões complexos, automatização de tarefas repetitivas, geração de textos, entre outras possibilidades.

Um das vantagens do uso da IA é exatamente maximizar a produtividade e minimizar os custos envolvidos, aumentando a capacidade de processar rapidamente mais informações a fim de atender demandas crescentes.

Os grandes avanços nas ferramentas de processamento de linguagem natural (*Natural Language Processing – NLP*), subcampo da IA que se dedica à compreensão da linguagem humana pela máquina, mostrou-se muito útil para a área jurídica, que se operacionaliza por meio da linguagem, contribuindo para a crescente demanda por ferramentas tecnológicas.

O uso de inteligência artificial no sistema de justiça cria oportunidades significativas para lidar com deficiências e falhas conhecidas na administração da Justiça, como a vagariedade e a falta de isonomia entre as decisões judiciais.

Com efeito, a falta de uniformidade de julgamento é apontada como um problema nos sistemas de justiça em todo o mundo⁷⁷.

Segundo Daniel Kahneman, psicólogo e economista vencedor do Prêmio Nobel de Ciências Econômicas de 2002, “as decisões dos seres humanos não são confiáveis porque são fortemente influenciadas por fatores irrelevantes, tais como o humor atual, o tempo desde sua última refeição e o clima. Chamamos a chance de variabilidade de julgamentos de *ruído*”⁷⁸.

De fato, conforme relatam Dierle Nunes, Natanel Lud e Flávio Quinaud Pedron, a crença na racionalidade decisória humana “foi extremamente enfraquecida pelos avanços da psicologia comportamental e por estudos empíricos que demonstram o impacto dos atalhos cognitivos (heurísticas) que nos fazem decidir com deturpações e sem levar em consideração toda a

⁷⁷ O ponto foi trabalhado anteriormente em AMORIM, Amanda Faneco; PEIXOTO, Fabiano Hartmann. Ruído, vieses e algoritmos: benefícios e riscos do uso da inteligência artificial nos sistemas de Justiça. In: Direito, governança e novas tecnologias II [Recurso eletrônico on-line] organização CONPEDI. Coordenadores: Irineu Francisco Barreto Junior; Jonathan Cardoso Régis; José Renato Gaziero Cella. Florianópolis: CONPEDI, 2022. pp. 157-175.

⁷⁸ KAHNEMAN, Daniel; ROSENFELD; Andrew. M.; GANDHI, Linnea; BLASER, Tom. Noise: how to overcome the high, hidden cost of inconsistent decision making. **Harvard Business Review**, oct. 2016. Disponível em: <<https://hbr.org/2016/10/noise>>. Acesso em: 25 fev. 2024.

informação relevante”⁷⁹.

Essa variabilidade não desejada e assistemática em julgamentos humanos que deveriam, em termos ideais, ser idênticos é chamada de ruído, fenômeno que gera injustiça, desconfiança e custos econômicos⁸⁰.

No ponto, o uso de algoritmos no processo decisório pode ao mesmo tempo aumentar a celeridade no julgamento e a uniformidade de decisões, eliminando a variabilidade não desejada e reduzindo vieses cognitivos humanos.

Por outro lado, embora tenha capacidade de melhorar a eficiência, a utilização de IA pode comprometer elementos essenciais do sistema de justiça como a análise de contextos e a interpretação das normas, bem como outros valores, tais como independência, imparcialidade e efetivo contraditório, que ultrapassam a mera uniformidade de decisões para casos semelhantes ou idênticos. Ademais, ainda que o uso de algoritmos contribua para a elaboração de melhores decisões, casos empíricos apontam, sem margem de dúvidas, que os sistemas utilizados estão sujeitos a erros dos mais variados tipos.

Conforme ressalta Ruth Chang, não há como determinar que dois casos são idênticos ou semelhantes e merecem igual julgamento sem antes estabelecer que as diferenças entre eles são irrelevantes para a questão em debate e esse é um julgamento que depende de uma avaliação jurídica prévia⁸¹.

Com efeito, um dos problemas éticos do uso da IA no Judiciário é tratar casos iguais de modo diferente ou tratar casos diferentes de forma semelhante, violando a ideia de equidade, que é ponto central do sistema de justiça. Identificar se as diferenças entre os casos comparados justificam tratamento jurídico diferente é uma análise que demanda compreensão e interpretação de contextos, causalidades, contrafactuais e constrições, que dependem do julgamento humano qualitativo⁸².

De fato, segundo explica Adriana Reis de Albuquerque:

⁷⁹ NUNES, Dierle; LUD, Nathanael; PEDRON, Flávio. Desconfiando da imparcialidade dos sujeitos processuais. **Salvador: JusPodivm**, 2018. p. 10.

⁸⁰ KAHNEMAN, Daniel; SIBONY, Olivier; SUNSTEIN, Cass R. **Ruído: uma falha no julgamento humano**. Rio de Janeiro: Objetiva, 2021.

⁸¹ CHANG, Ruth. “Algorithms Eliminate Noise (and That Is Very Good)”. 25 mar. 2021. **Part of the Colloquium on AI Ethics series presented by the Institute of Ethics in AI**. Apresentação: John Tasioulas. Palestrante: Cass Sustein. Comentários: Ruth Chang e Nigel Shadbolt. Disponível em: <<https://podcasts.ox.ac.uk/algorithms-eliminate-noise-and-very-good>>. Acesso em: 1 aug. 2022.

⁸² FRAZÃO, Ana. Discriminação algorítmica: resgatando os aspectos positivos dos julgamentos humanos. A importância das narrativas, das contextualizações e das molduras cognitivas. **JOTA**, 2021. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica-resgatando-os-aspectos-positivos-dos-julgamentos-humanos-01092021>. Acesso em: 17 jan. 2024.

[c]ompreender imagens e textos, especialmente a partir do contexto – normas, precedentes e tudo o que Direito envolve – requer a capacidade de diferenciar, quando diante de uma determinada situação, o que é nela essencial, relevante, onde está sua essência, de modo a afastar o que pode ser ignorado⁸³.

Nessa perspectiva, nas palavras da autora, “a insensibilidade dos sistemas computacionais ao elemento contexto impacta o processo interpretativo e, portanto, decisório, também porque a linguagem – veículo de construção do conhecimento humano – é, em muitas ocasiões, dotada de ambiguidade e vagueza”⁸⁴.

Ademais, especificamente no que toca ao Direito, a aplicação de entendimento de decisões anteriores para resolver um determinado caso depende não apenas do exame das semelhanças fáticas entre as situações, mas da identificação do princípio aplicável capaz de aproximar os episódios comparados⁸⁵.

Nesse contexto, verificar se determinado precedente se aplica a um caso concreto sob julgamento não é tarefa tão simples, pois exige não apenas uma análise factual, mas a análise da presença “[d]os mesmos critérios axiológicos e teleológicos que moldaram a própria construção da *ratio essendi* a ser replicada”⁸⁶.

Muitos autores utilizam diferentes maneiras para apontar no mesmo sentido de que “capturar alguma essência da realidade através de um modelo mental a fim de criar um curso de ação eficaz” é algum natural para os seres humanos, mas que os sistemas computacionais não conseguem reproduzir⁸⁷.

Sobre o tema, Ana Frazão destaca:

A partir do momento em que desprezamos as habilidades do raciocínio humano e passamos a atribuir a julgamentos algorítmicos a palavra única ou final, passamos a correr diversos riscos, que vão desde a possibilidade da prevalência e da cristalização dos valores ou dos frames dos seus programadores – sem qualquer discussão ou escrutínio – até o empobrecimento da própria compreensão e da interpretação dos dados, os quais serão vistos a partir de correlações que não estarão submetidas às análises de contextos, sentidos, causalidades, contrafactuais, constrições e tantas outras que dependem do julgamento humano qualitativo, pois dificilmente podem ser traduzidas em números e métricas⁸⁸.

⁸³ DE ALBUQUERQUE, Adriana Reis. **Poder Artificial de Tributar: Limites e Requisitos à Utilização (Adequada) da Inteligência Artificial pela Administração Tributária**. Editora Thoth, 2022. p. 144.

⁸⁴ ALBUQUERQUE, Adriana Reis. **Poder Artificial de Tributar**. Op. Cit. p. 142.

⁸⁵ *Ibidem*. pp. 144-148.

⁸⁶ *Ibidem*. p. 150.

⁸⁷ CUKIER, Kenneth; MAYER-SCHÖNBERGER, Viktor; DE-VERICOURT, Francis. **Framers: A vantagem humana em uma era de tecnologia e instabilidade**. 2023. Ebook.

⁸⁸ FRAZÃO, Ana. Discriminação algorítmica: resgatando os aspectos positivos dos julgamentos humanos. A importância das narrativas, das contextualizações e das molduras cognitivas. **JOTA**, 2021. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica-resgatando-os-aspectos-positivos-dos-julgamentos-humanos-01092021>. Acesso em: 17 jan. 2024.

Assim, embora sistemas de IA tenham potencial de contribuir com o processo de tomada de decisão em razão de sua maior capacidade de processamento de grande volume de dados, não se pode ignorar as limitações da tecnologia, que acarretam, inevitavelmente, o risco de erros de julgamento.

Outrossim, a interpretação das normas realizada pelos juízes, quando dos julgamentos de casos concretos é também o que proporciona a evolução do pensamento jurídico para acompanhar as mudanças sociais, econômicas e/ou legislativas.

De fato, algoritmos podem sugerir inúmeras decisões rapidamente, mas não são capazes de levantar questões nunca formuladas, aplicar a imaginação para criar novas soluções e ponderar as consequências, pois trabalham sempre com a reprodução de dados do passado.

Sistemas de IA também não interagem com o meio ou com outros julgadores, não promovem o estudo de situações inéditas, não são sensíveis aos anseios sociais. Assim, o preço pago por eliminar o ruído pode ser um processo decisório mecânico, limitado e falho.

A propósito dos riscos do uso da IA no sistema de justiça, Dierle Nunes e Ana Luiza Marques apontam que:

A implementação de sistemas de IA, sem a devida reflexão sobre seus impactos e seu modo de funcionamento, apresenta grandes riscos para o direito, do ponto de vista sistêmico, por tornar indiscutíveis – ou, ao menos, restringir drasticamente o âmbito de discussão – decisões enviesadas, com fundamento em argumentos puramente quantitativos, em especial o aumento da rapidez proporcionado pelo uso de tais tecnologias⁸⁹.

Sobre o tema, Fabiano Hartmann alerta que o uso da IA no Direito “vem com o risco de aumentar decisões erradas, tendenciosas, casuísticas ou preconceituosas, que contariam com mais uma ferramenta retórica de justificativa pela associação ao rigor técnico-científico de um modelo tecnológico de processamento de dados”⁹⁰.

O professor da Universidade de Brasília (UnB), pesquisador do Programa das Nações Unidas para o Desenvolvimento – PNUD e coordenador acadêmico de diversos projetos de IA desenvolvidos para o Poder Judiciário brasileiro, tais como Victor (STF) e Mandamus (TJRR), indica entre os riscos que “devem sempre estar sendo monitorados, registrados e metodologicamente prevenidos”: redução de direitos fundamentais; detrimento do devido processo legal; quebra da paridade de armas; violação da privacidade; redução-combate à

⁸⁹ NUNES, Dierle; MARQUES, Ana Luiza Pinto Coelho. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. In: **Revista de Processo**. 2018. p. 421-447.

⁹⁰ PEIXOTO, Fabiano Hartmann. *Op. cit.* p. 160.

discriminação e enfraquecimento da democracia⁹¹.

Destaca-se que, ainda que a sugestão do algoritmo tenha de ser objeto de revisão humana, inserindo-se em um cenário de apoio à tomada de decisão humana e não de efetiva substituição, como ocorre na quase totalidade dos casos de utilização de IA no Judiciário brasileiro, não é possível desconsiderar o enviesamento causado pela indicação do sistema.

Na medida em que os servidores e magistrados estão sujeitos a metas de produtividade, há nítida tendência de que seja reproduzido o entendimento sugerido pelo algoritmo. Esse fenômeno também pode ser explicado pelo viés de automação, que consiste na tendência humana de acreditar que resultados indicados por ferramentas tecnológicas são corretos e confiáveis.

A doutrina aponta que, mesmo nos casos em que há supervisão humana de sistemas automatizados, a confiança na tecnologia diminui o senso de responsabilidade e *accountability* do usuário-supervisor⁹².

A implementação de sistemas de IA no Judiciário levanta outras questões como a influência exercida por atores externos na independência do juiz, inclusive na separação dos poderes. O controle sobre os dados de treinamento, a infraestrutura usada para armazenar e processar os dados e o próprio *design* dos algoritmos podem suscitar dúvidas sobre a imparcialidade e autonomia do magistrado, assim como a confiança pública no sistema⁹³. No caso de aquisição de sistemas desenvolvidos pela iniciativa privada, essas preocupações com eventuais influências indiretas sobre o Judiciário são ainda maiores.

Assim, não é razoável falar em desenvolvimento e implementação de sistemas de IA no Judiciário sem que haja o adequado dimensionamento e monitoramento dos riscos envolvidos em uma perspectiva ampla para a sociedade em geral. A propósito, leciona Fabiano Hartmann:

Não é possível se falar adequadamente em um projeto de IA sem um projeto ético (que coloque centralmente as discussões sobre ética e valores envolvidos), sem uma reflexão das melhores formas de comunicação e esclarecimento sobre os objetivos e impactos do projeto, sem um sistema de controle (tanto para aferição de benefícios como para identificação e gestão de danos)⁹⁴.

⁹¹ *Ibidem*. p. 160-161.

⁹² GREEN, Ben; CHEN, Yiling. **Disparate Interactions**: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments, in Conference on Fairness, Accountability, and Transparency (FAT* '19). Atlanta, GA, USA; 2019. Disponível em: <https://scholar.harvard.edu/bgreen/publications/disparate-interactions-algorithm-loop-analysis-fairness-risk-assessments>. Acesso em: 13 de maio de 2023.

⁹³ GENTILE, Giulia. AI in the courtroom and judicial independence: An EU perspective. **EUIdeas**, 22 aug. 2022. Disponível em: <https://euideas.eui.eu/2022/08/22/ai-in-the-courtroom-and-judicial-independence-an-eu-perspective/> Acesso em: 15 de maio de 2023.

⁹⁴ PEIXOTO, Fabiano Hartmann. *Op. Cit.* p. 84-85.

No que toca especificamente às ferramentas de IA generativa, que alcançaram grande popularidade em 2023, embora apresentem capacidade promissora e revolucionária na geração de textos bem elaborados e convincentes, estudos apontam que as “alucinações”, termo técnico utilizado para descrever resultados produzidos sem justificativa nos dados de treinamento, são generalizadas e perturbadoras.

Large Language Model (LLM) ou, em português, modelo de linguagem de grande escala, como o famoso GPT-3, da *OpenIA*, destacam-se pela sua capacidade de gerar textos longos e coesos, semelhantes aos escritos por um ser humano. Tais modelos funcionam a partir da detecção de padrões e relacionamentos entre pontos nos dados de treinamento, de modo a prever a próxima palavra diante de uma sequência de palavras anteriores por meio de probabilidade e estatística. Assim, um LLM, como o ChatGPT, gera texto por meio de manipulação de formas linguísticas de acordo com probabilidades sem compromisso com a realidade.

Em 2021, um artigo científico liderado por doutoras cientistas da computação chamou atenção e popularizou o termo “papagaio estocástico” nos campos da computação, engenharia e estudos sociais da tecnologia⁹⁵. Na referida publicação, os modelos de linguagem foram definidos como sistemas que “aleatoriamente costuram sequências de formas linguísticas observadas em seus vastos dados de treinamento, de acordo com informação probabilística sobre como combinam, mas sem nenhuma referência a significado” (tradução livre)⁹⁶.

Em 2023, Konstantine Arkoudas, PhD em Ciência da Computação e com mais de dez anos de experiência em IA, especializado em processamento de linguagem natural, publicou artigo intitulado “GPT-4 Não Consegue Raciocinar” (tradução livre), no qual critica métodos atuais empregados para avaliar a destreza de raciocínio dos modelos de linguagem (LLMs) e conclui que, “não obstante as instâncias intermitentes de brilhantismo analítico, o estado atual do GPT-4 o torna fundamentalmente incapaz de se envolver em um raciocínio genuíno” (tradução livre)⁹⁷.

⁹⁵ BENDER, Emily M. et al. On the dangers of stochastic parrots: Can language models be too big?. In: **Proceedings of the 2021 ACM conference on fairness, accountability, and transparency**. 2021. p. 610-623

⁹⁶ No original: “Contrary to how it may seem when we observe its output, an LM is a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning: a stochastic parrot”. BENDER, Emily M. et al. On the dangers of stochastic parrots: Can language models be too big?. In: **Proceedings of the 2021 ACM conference on fairness, accountability, and transparency**. 2021. p. 610-623

⁹⁷ ARKOUDAS, Konstantine. *GPT-4 Can't Reason*. **Preprints**. 2 aug. 2023. Disponível em: <https://doi.org/10.20944/preprints202308.0148.v2>. Acesso em: 13 ago. 2023.

Na citada publicação, o autor realiza uma extensa análise qualitativa do desempenho do GPT-4 em desafios de raciocínio diversos e, com base nos resultados dessa análise, apresenta, entre outras conclusões, a de que:

[o] uso de IA generativa no desenvolvimento de software (ou na ciência e engenharia em geral) para qualquer coisa além de tarefas tediosas (como uma espécie de preenchimento automático turbinado para perguntas de codificação ricas em conhecimento) está repleto de riscos graves (tradução livre)⁹⁸.

No caso do ChatGPT, a própria empresa desenvolvedora adverte os usuários de que a ferramenta não produz informações confiáveis e que não deve ser utilizada sem a conferência da veracidade do resultado. Dependendo do comando (*prompt*) utilizado, a ferramenta pode fazer alguns alertas diretos sobre a necessidade de verificar as informações fornecidas ou de buscar ajuda especializada acerca do tema.

Novas ferramentas que utilizam a tecnologia do GPT-4, mas treinadas especificamente para desempenho de atividades jurídicas já estão despontando e muitas outras ainda irão surgir. Em tais casos, o alerta de necessidade de conferência dos resultados do modelo de IA muitas vezes é inexistente.

Mesmo nesse caso em que o modelo de IA é treinado e refinado especificamente para o uso no meio jurídico, não é possível eliminar completamente as “alucinações” e, por essa razão, todos os usuários do sistema, em um escritório de advocacia que utiliza sistema de IA para apoio das atividades jurídicas, são orientados a seguir algumas regras no uso da ferramenta, a mais importante delas: checar e validar todos os resultados⁹⁹. Para além das alucinações, “parte da prática do direito envolve compreender as circunstâncias particulares do seu cliente, então o resultado raramente será ideal” (tradução livre)¹⁰⁰.

Estudo recente de pesquisadores da Universidade de Stanford e de Yale concluiu que as alucinações jurídicas são generalizadas nos modelos de linguagem de última geração, alcançando taxas que variam de 69% a 88% em resposta a consultas jurídicas específicas¹⁰¹.

⁹⁸ *Ibidem*.

⁹⁹ SOTEL-WALKER, Chris. **Generative AI is coming for Lawyers**. 12 feb. 2023. Disponível em: <https://www.wired.co.uk/article/generative-ai-is-coming-for-the-lawyers>. Acesso em: 30 jul. 2023.

¹⁰⁰ Tradução livre. Disponível em: <https://www.wired.co.uk/article/generative-ai-is-coming-for-the-lawyers>. Acesso em: 30 jul. 2023.

¹⁰¹ No original: “We find that legal hallucinations are alarmingly prevalent, occurring between 69% of the time with ChatGPT 3.5 and 88% with Llama 2, when these models are asked specific, verifiable questions about random federal court cases”. DAHL, Matthew et al. Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models. Disponível em: **arXiv preprint arXiv:2401.01301**, 2024. Acesso em: 30 jul. 2023.

Os resultados da pesquisa apontam que, em tarefas mais complexas que exigem uma compreensão diferenciada de questões jurídicas ou interpretação de textos, os resultados dos modelos de linguagem não se saem melhor do que adivinhação aleatória (50%), indicando que os grandes modelos de linguagem (LLMs) não são capazes de realizar o tipo de raciocínio jurídico que os advogados e magistrados empregam quando avaliam a relação entre o caso examinado e os precedentes¹⁰².

Além das alucinações e da aleatoriedade nos resultados, não há clareza em como realmente funcionam os LLMs, como o seu desempenho varia ao longo do tempo ou como são atualizados, existindo necessidade de monitoramento contínuo dos modelos.

No caso do GPT-4, a problemática é ainda maior, dada a opacidade do modelo, pois a *OpenAI* “não revelou a fonte dos materiais de treinamento, o código-fonte, os pesos da rede neural ou mesmo um artigo descrevendo sua arquitetura” (tradução livre)¹⁰³.

Um estudo realizado por pesquisadores da Universidade de Stanford e Universidade da Califórnia, Berkeley, que comparou a performance do ChatGPT em março e junho de 2023, apontou existência de mudanças significativas no desempenho do GPT-3.5 e GPT-4.

No artigo intitulado “Como o comportamento do ChatGPT está mudando ao longo do tempo?” (tradução livre), aponta-se que a precisão na identificação de números primos e compostos diminuiu no GPT-4 de março de 2023 (84%) para junho de 2023 (51%), enquanto a performance do GPT-3.5 foi muito melhor em junho do que em março na mesma tarefa¹⁰⁴. O estudo assinala também que o GPT-4 se tornou menos disposto a responder a perguntas sensíveis e pesquisas de opinião em junho do que em março e que tanto o GPT-4 quanto o GPT-3.5 tiveram mais erros de formatação na geração de código em junho do que em março.

Ainda que os grandes modelos de linguagem tenham um grande potencial de aplicação na prática jurídica, as limitações da tecnologia evidenciam que o emprego dessas ferramentas

¹⁰² No original: “The results for the high complexity tasks (Table 4) confirm this general pattern of poor performance, but must be interpreted slightly differently. First, the Doctrinal agreement task is another binary classification task, so the LLMs’ hallucination rates on this task—near 0.5—represent little improvement over random guessing. This suggests that LLMs are not yet able to perform the kind of legal reasoning that attorneys perform when they assess the precedential relationship between cases—a core purpose of legal research”. DAHL, Matthew et al. Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models. *arXiv preprint arXiv:2401.01301*, 2024. Acesso em: 12 ago. 2023.

¹⁰³ EDWARDS, Benj. **Study claims ChatGPT is losing capability, but some experts aren’t convinced. Either way, experts think OpenAI should be less opaque about its AI model architecture.** 7 aug, 2023. Disponível em: <https://arstechnica.com/information-technology/2023/07/is-chatgpt-getting-worse-over-time-study-claims-yes-but-others-arent-sure/>. Acesso em: 12 ago. 2023.

¹⁰⁴ CHEN, Lingjiao; ZAHARIA, Matei; ZOU, James. **How is ChatGPT's behavior changing over time?**. 31 oc. 2023. Disponível em: <https://arxiv.org/abs/2307.09009>. Acesso em: 5 ago. 2023.

exige supervisão, atenção e prévia compreensão humana, não apenas sobre o assunto em questão, mas também sobre as capacidades e limites da IA.

Mesmo que o controle de decisões judiciais continue a existir de maneira endoprocessual, por meio da possibilidade recorrer dos pronunciamentos decisórios, o avanço do uso de ferramentas de IA nos tribunais aumenta a preocupação com as consequências decorrentes, inclusive no tocante ao direito ao contraditório, incluída a possibilidade de influir no convencimento do juiz.

Com efeito, mesmo sistemas mais simples, como de classificação de processos e identificação de causas repetitivas, podem acarretar decisões injustas, em especial no contexto em que a tecnologia é empregada para garantir eficiência e alcançar metas de produtividade focadas em números.

Assim como o uso da IA possibilita a produção de tarefas em grande escala, as implicações éticas decorrentes também adquirem uma magnitude proporcional, assumindo dimensões consideráveis. No atual cenário de transformação na prática jurídica, os conflitos éticos estão por toda parte.

No âmbito dos sistemas de justiça, eventuais deficiências e falhas apresentam perigo não apenas na esfera individual dos jurisdicionados, mas no próprio sistema e na ideia de Estado de Direito e de devido processo legal.

A propósito, destacando a necessidade de avaliar os riscos que a implementação da tecnologia apresenta do ponto de vista sistêmico, Dierle Nunes e Ana Luiza Marques destacam:

(...) no sistema processual brasileiro, tendo em vista a perspectiva do devido processo constitucional, a utilização de mecanismos ocultos para as partes do processo na tomada de decisões viola garantias processuais constitucionais, como o contraditório e a ampla defesa, as quais somente podem ser efetivadas mediante o conhecimento dos pressupostos utilizados no julgamento.

(...)

Verifica-se, portanto, que, com o rápido desenvolvimento das tecnologias, as ferramentas de inteligência artificial terão suas funções ampliadas. Contudo, a implementação de sistemas de IA, sem a devida reflexão sobre seus impactos e seu modo de funcionamento, apresenta grandes riscos para o direito, do ponto de vista sistêmico, por tornar indiscutíveis – ou, ao menos, restringir drasticamente o âmbito de discussão – decisões enviesadas, com fundamento em argumentos puramente quantitativos, em especial o aumento da rapidez proporcionado pelo uso de tais tecnologias.

(...)

Principalmente, ao se constatar a virada tecnológica no direito, torna-se imperativa a rediscussão do próprio conteúdo dos seus princípios fundamentais, a começar pelo devido processo constitucional, para que estes possam agir contrafaticamente (de modo corretivo) de forma a controlar os poderes decisórios, agora tecnológicos, com redução de sua opacidade e gerando sua necessária transparência (*accountability*)¹⁰⁵.

¹⁰⁵ NUNES, Dierle; MARQUES, Ana Luiza Pinto Coelho. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. In: **Revista de Processo**. 2018. p. 421-447

Destaca-se que muitos dos problemas éticos surgem exatamente de consequências não intencionais do emprego da tecnologia. À medida que as ferramentas de IA se tornam mais comuns no Judiciário brasileiro, despontam várias questões: quais os riscos envolvidos? Quais decisões e quanto dessas decisões delegar à IA? Quem está se beneficiando do uso da IA?

Assim, é necessário conciliar abordagens de custo e benefício, avaliação de impacto, gestão de riscos e proteção a direitos e garantias fundamentais. Para tanto, faz-se necessário que os tribunais que utilizam ferramentas tecnológicas implementem estruturas de governança que direcionem o caminho a ser seguido, orientem aplicações e garantam direitos dos jurisdicionados.

Nessa linha, o presente capítulo se dedica a analisar os mecanismos de governança de IA empregados nos sistemas de justiça no cenário global e brasileiro com o objetivo de avaliar o modelo proposto pela Resolução CNJ n. 332/2020, sua implementação prática e identificar pontos aperfeiçoamento com a finalidade de subsidiar futuras iniciativas institucionais.

2.2. Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes

Apesar da existência de mais de uma centena de documentos elaborados ao redor do mundo para tratar de diretrizes éticas para o desenvolvimento e uso de IA nos diferentes setores, poucos são aqueles destinados a analisar especificamente a aplicação da referida tecnologia nos sistemas de justiça.

O primeiro e mais importante é a Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes, adotada pela Comissão Europeia para a Eficácia da Justiça (*European Commission for the Efficiency of Justice - CEPEJ*) na sua 31ª reunião plenária, em 3 dezembro de 2018¹⁰⁶.

Em razão da sua grande importância para o estudo do tema, assim como por ter sido inspiração declarada para a edição da Resolução CNJ n. 332/2020, mostra-se oportuno realizar um breve estudo do referido documento antes de adentrar no exame da regulamentação da matéria no âmbito do Conselho Nacional de Justiça.

A referida Carta de Ética do CEPEJ é um documento supranacional não vinculante,

¹⁰⁶ LUPO, Giampiero. The Ethics of Artificial Intelligence: An Analysis of Ethical Frameworks Disciplining AI in Justice and Other Contexts of Application. *Oñati Socio-Legal Series*, vol. 12, n. 3, Junho 2022, pp. 614-53. 30 mar. 2022. Disponível em: <https://opo.iisj.net/index.php/osls/article/view/1366>. Acesso em: 12 ago. 2023.

destinada às partes interessadas públicas e privadas responsáveis pela concepção e implementação de ferramentas de inteligência artificial que envolva o tratamento de decisões e dados judiciais, bem como “aos decisores públicos responsáveis pelo quadro legislativo ou regulamentar, desenvolvimento, auditoria ou utilização dessas ferramentas e serviços”¹⁰⁷.

Apesar de considerar que a aplicação da IA no domínio da justiça pode ser uma oportunidade para melhorar a eficiência e qualidade da Justiça, com potencial para melhorar a coerência das decisões judiciais, a Carta salienta a necessidade de que a tecnologia seja desenvolvida e empregada de forma responsável e com respeito aos direitos fundamentais e aos princípios nela enunciados.

Cabe destacar que, no momento da sua elaboração, em 2018, não se tinha conhecimento de ferramentas de IA de aplicação significativa em utilização pelos sistemas de justiça dos Estados-Membros do Conselho da Europa¹⁰⁸. Em verdade, o estudo técnico que segue como apêndice do documento aponta a existência de alguns softwares de justiça preditiva, principalmente direcionados à iniciativa privada, destinados a calcular as probabilidades de sucesso (ou de fracasso) de um caso perante um tribunal.

Aponta-se, de maneira expressa, que “em 2018, a utilização de algoritmos de inteligência artificial nos sistemas judiciais europeus continua a ser essencialmente uma iniciativa comercial do setor privado destinada às companhias de seguros, departamentos jurídicos, advogados e particulares” (tradução livre)¹⁰⁹.

A Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes, em si, é extremamente concisa e se limita a identificar cinco macro princípios para serem observados no desenvolvimento e implementação de modelos de IA nos sistemas

¹⁰⁷ CEPEJ. Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente. Estrasburgo, 3 dez. de 2018. Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 12 ago. 2023.

¹⁰⁸ “In 2018, the use of artificial intelligence algorithms in European judicial systems remains primarily a private-sector commercial initiative aimed at insurance companies, legal departments, lawyers and individuals”; “The question of the use of AI in judicial systems was dealt with in a specific online survey, launched in April 2018 for representatives of the CEPEJ member States and civil society. The response level was relatively low and did not allow clear trends to be identified. Some private operators did not seem very receptive to this survey and the members of the CEPEJ, who belong for the most part to ministries of justice or higher councils of justice, were able to quote only the tools currently used by the public sphere”. RONSIN, Xavier. Et al. Appendix I In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data. In CEPEJ. European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, Estrasburgo, 3 dec. 2018, p. 42. Disponível em <https://rm.coe.int/ethicalcharter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 12 ago. 2023.

¹⁰⁹ XAVIER, Rosi. In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data. in: CEPEJ. European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, Estrasburgo, 3 dec. 2018, p. 42. Disponível em <https://rm.coe.int/ethicalcharter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 12 ago. 2023.

de justiça, quais sejam:

1. **Princípio de respeito aos direitos fundamentais:** assegurar que o *design* e a implementação de ferramentas e serviços de inteligência artificial sejam compatíveis com os direitos fundamentais.
2. **Princípio de não-discriminação:** prevenir especificamente o desenvolvimento ou a intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos.
3. **Princípio de qualidade e segurança:** em relação ao processamento de decisões e dados judiciais, utilizar fontes certificadas e dados intangíveis com modelos elaborados de forma multidisciplinar, em ambiente tecnológico seguro.
4. **Princípio da transparência, imparcialidade e equidade (*fairness*):** tornar os métodos de tratamento de dados acessíveis e compreensíveis, autorizar auditorias externas.
5. **Princípio "sobre o controle do usuário":** impedir uma abordagem prescritiva e garantir que os usuários sejam atores informados e estejam no controle das escolhas feitas¹¹⁰.

O citado documento, ao tratar do princípio do respeito dos direitos fundamentais, indica a necessidade de assegurar o respeito à proteção dos dados pessoais e às garantias processuais, como a paridade de armas e o respeito ao contraditório, assim como a independência e a imparcialidade dos juízes, bem como aos princípios essenciais ao Estado de Direito.

Ao abordar o princípio da qualidade e da segurança, a Carta cita a multidisciplinariedade como uma condição importante desde a concepção dos modelos de IA. Afirma-se que a área técnica responsável pela elaboração do modelo deve ser capaz de “aproveitar amplamente a experiência dos profissionais dos sistemas de justiça (juízes, promotores, advogados, etc.) e pesquisadores/professores das áreas de direito e ciências sociais (por exemplo, economistas, sociólogos e filósofos)”¹¹¹. Aponta-se que “[a]s salvaguardas éticas existentes devem ser constantemente compartilhadas por essas equipes de projeto e aprimoradas por meio de *feedback*”¹¹². Demonstra-se, ainda, preocupação com a integridade e intangibilidade do sistema, assim como a rastreabilidade do processo decisório.

No ponto em que trata do princípio da transparência, imparcialidade e *fairness*, indica-se como primeira opção a completa transparência, incluindo o código-fonte do modelo, mas se admite também que os sistemas sejam explicados em linguagem clara e familiar, comunicando como os resultados são produzidos, a natureza do serviço oferecido, as ferramentas desenvolvidas, sua performance e os riscos de erro. É interessante também a previsão de que

¹¹⁰ CEPEJ. Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente. Estrasburgo, 3 dez. de 2018. Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 12 ago. 2023.

¹¹¹ *Ibidem*.

¹¹² *Ibidem*.

“autoridades ou peritos independentes podem ser encarregados de certificar e auditar métodos de processamento ou de prestar aconselhamento prévio”, assim como “autoridades públicas poderiam conceder certificação, que seria regularmente revista”¹¹³.

Com relação ao princípio “sobre o controle do usuário”, vale destacar a previsão de que a autonomia dos utilizadores da ferramenta deve ser aumentada e não diminuída, devendo ser assegurado “programas de alfabetização informática para os usuários e debates envolvendo profissionais do sistema de justiça”¹¹⁴. Assegura-se, ainda, que o usuário externo deve ser claramente informado sobre o uso de IA antes ou durante o processo judicial, bem como deve ser garantido o direito de se opor à referida utilização, “de modo que seu caso possa ser ouvido diretamente por um tribunal, de acordo com o significado do artigo 6 da Convenção Europeia de Direitos do Homem”¹¹⁵.

Apesar de não possuir caráter obrigatório, a Carta Europeia de Ética sobre o uso da IA nos sistemas de justiça e seus ambientes apresenta, em um de seus quatro apêndices, duas listas de verificação (*checklist*) de conformidade (para autoavaliação da integração dos princípios nela enunciados) e estabelece, logo em sua introdução, que “autoridades independentes mencionadas na Carta poderiam ser responsáveis por avaliar periodicamente o nível de adesão de todos os intervenientes aos princípios da Carta e por propor melhorias para a adaptar à evolução das tecnologias e das utilizações dessas tecnologias”¹¹⁶. Assim, fica evidenciada a preocupação, desde a concepção do documento, com a avaliação e monitoramento contínuo da implementação dos princípios enunciados.

Como apêndices à Carta, constam ainda outros três documentos, além das listas de verificação, entre eles um aprofundado estudo desenvolvido por especialistas sobre a utilização de IA nos sistemas judiciais, que analisa o funcionamento de ferramentas de aprendizado de máquina aplicado no contexto de decisões judiciais, suas limitações e os riscos envolvidos, assim como uma lista de indicação de casos de usos classificados em encorajados; possíveis, mas que exigem precauções metodológicas ou que demandam estudos científicos adicionais; e usos a serem considerados com as mais extremas reservas.

O referido estudo (Apêndice I da Carta) desenvolve reflexões sobre o uso da tecnologia e a necessidade de garantia de observâncias de direitos fundamentais e proteção de dados, tratando em especial de ferramentas de justiça preditiva e aponta, com certo destaque, para a

¹¹³ *Ibidem.*

¹¹⁴ *Ibidem.*

¹¹⁵ *Ibidem.*

¹¹⁶ *Ibidem.*

experiência frustrada realizada, por iniciativa do Ministério da Justiça da França, nos tribunais de recurso de Douai e Rennes, em 2017, que tinha por objetivo reduzir a excessiva variabilidade das decisões judiciais em nome da isonomia e segurança jurídica, mas que concluiu pela ausência de valor agregado do *software* desenvolvido por considerar que produzia resultados inadequados ou aberrantes¹¹⁷.

Ressalta-se, ainda, no referido estudo técnico, a necessidade de debater, testar e avaliar a aplicação de ferramentas de IA antes da implementação de políticas públicas, assim como revisar, de maneira contínua, o impacto no trabalho dos profissionais de justiça. Indica-se também a criação de um *framework* ético para promover os princípios da transparência, equidade (*fairness*) e neutralidade desde a concepção dos modelos e sua implementação.

Embora a Carta aponte a multidisciplinariedade das equipes como uma condição prévia importante, o documento não tratou do envolvimento das partes interessadas (usuários externos) seja durante o desenvolvimento dos modelos seja no processo de fiscalização dos resultados, o que poderia contribuir tanto para melhorar as soluções, por meio de críticas e sugestões, quanto para ampliar as perspectivas de aceitação da ferramenta tecnológica e confiança na sua aplicação¹¹⁸.

O documento examinado também não abordou a questão da responsabilização dos envolvidos, limitando-se a encorajar a realização de auditorias externas realizadas por profissionais independentes.

Ressalta-se também a ausência, na Carta, de um órgão consultivo (conselho, comitê ou entidade semelhante) destinada a avaliar os riscos da introdução de ferramentas de IA no sistema de justiça.

Esse ponto foi corrigido, em 2022, com a criação do Conselho Consultivo de IA do CEPEJ (*Artificial Intelligence Advisory Body - AIAB*), responsável por “acompanhar o aparecimento efetivo de aplicações de IA nos sistemas de justiça, debater as problemáticas atuais e propor novas estratégias relativas à utilização da IA no sistema judicial, respeitando os

¹¹⁷ XAVIER, Rosi. In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data. in: European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, Estrasburgo, CEPEJ – Commission Européenne pour l'Efficacité de la Justice, 3 dec. 2018, p. 42. Disponível em: <https://rm.coe.int/ethicalcharter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 12 ago. 2023.

¹¹⁸ LUPO, Giampiero. The Ethics of Artificial Intelligence: An Analysis of Ethical Frameworks Disciplining AI in Justice and Other Contexts of Application. **Oñati Socio-Legal Series**, vol. 12, n. 3, june 2022, pp. 614-53. Disponível em: <https://opo.iisj.net/index.php/osls/article/view/1366>. Acesso em: 12 ago. 2023.

direitos fundamentais” (tradução livre)¹¹⁹.

O referido Conselho Consultivo, composto por cinco especialistas, selecionados entre integrantes da academia, da indústria, do sistema de justiça e do terceiro setor, para mandatos de dois anos, atua como uma entidade fundamental para garantir monitoramento da implementação da "Carta Ética Europeia sobre o Uso de Inteligência Artificial em Sistemas Judiciais e seu ambiente”.

Importante destacar que, dois anos após a publicação da Carta Ética, em 8 dezembro de 2020, a CEPEJ publicou um estudo de viabilidade para a possível introdução de um mecanismo de certificação de soluções de IA de acordo com os princípios propostos na referida carta¹²⁰. Desde então o CEPEJ vem trabalhando em elaborar instrumento de avaliação da conformidade de ferramentas de IA com a sua Carta Ética.

Em dezembro de 2021, a CEPEJ publicou um “Roteiro revisado para assegurar um seguimento adequado (*follow-up*) da Carta Ética do CEPEJ sobre a utilização da inteligência artificial nos sistemas judiciais e nos seus ambientes” (tradução livre). Nesse documento, além do trabalho em criar um instrumento de avaliação de conformidade da Carta Ética e do papel o Conselho Consultivo de IA do CEPEJ (AIAB), apontou-se a necessidade de realização de atividades de formação sobre o conteúdo da Carta tendo como público alvo desenvolvedores de ferramentas de IA e potenciais usuários (juízes, procuradores, advogados), assim como “pessoas responsáveis pela certificação por outras organizações, permitindo-lhes compreender e considerar melhor os aspectos relacionados aos direitos fundamentais” (tradução livre)¹²¹.

Por fim, destaca-se que, apesar de toda essa estrutura em desenvolvimento para garantir observância dos princípios estabelecidos na Carta Ética, dados coletados pelo CEPEJ, no relatório de avaliação de 2022 (dados 2020), indicam que não há, no âmbito dos 44 países da União Europeia analisados e 3 Estados observadores (Israel, Cazaquistão e Marrocos), ferramentas de IA empregadas efetivamente na atividade jurisdicional além de projetos em experimentação ou automação de atividades simples¹²².

¹¹⁹ CEPEJ. **Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment**. 9 dec. 2021. Disponível em: <https://rm.coe.int/cepej-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 15 nov. 2023.

¹²⁰ CEPEJ. **Possible introduction of a mechanism for certifying artificial intelligence tools and services in the sphere of justice and the judiciary**: Feasibility Study. 8 dec. 2020. Disponível em: <https://rm.coe.int/feasability-study-en-cepej-2020-15/1680a0adf4>. Acesso em: 12 ago. 2023.

¹²¹ CEPEJ. **Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment**. 9 dez. 2021. Disponível em: <https://rm.coe.int/cepej-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 15 nov. 2023.

¹²² No original: “It is, of course, essential to look at emerging ICT tools that could be used in the European judicial systems, such as artificial intelligence (AI) and blockchain. Can these technologies be useful? Where and how could they be used? Which are the risks they pose in the justice domain? To answer these questions, attention

Com efeito, do documento oficial “Intercâmbio de experiências entre a União Europeia e o Brasil sobre e-Justice: relatório final”, de 2022, observa-se que a experiência com ferramentas de IA no sistema de justiça no contexto europeu está muito atrás dos mais de 111 projetos de andamento no Brasil, sendo mais da metade deles em utilização¹²³.

No referido documento, foram apontados como ferramentas de IA, em desenvolvimento ou em utilização, no contexto europeu: 1) IZZY: *chatbot* sobre direitos dos consumidores relacionados às compras on-line ou off-line e está disponível apenas em francês; 2) Anonimização de decisões judiciais (Áustria): projeto com o objetivo de anonimizar (dados pessoais diretos e indiretos) as decisões judiciais e assim promover a transparência e a publicidade; 3) Uso de Inteligência Artificial no registro (Áustria): prevê um reconhecimento automático do tipo concreto de processo, alocando os referidos arquivos ao órgão competente do Judiciário; 4) Uso de Inteligência Artificial durante o processamento do arquivo (Áustria): “identificação automática dos tipos de documentos e uma proposta automatizada para o título do documento”; “Futuramente, as áreas de aplicação incluirão a geração automática de trechos de conteúdo pertinentes a partir de arquivos e modelos de fluxo de trabalho”; “Os futuros campos de aplicação podem incluir a identificação e apresentação de conexões (ocultas)”; 5) *Chatbot* Justicia (Áustria): implementado em uma plataforma online que oferece vários serviços judiciais aos cidadãos. O *chatbot* responde a questões legais e também guia o usuário através de diferentes tipos de procedimento; 6) SALME – Assistente de reconhecimento de voz (Estônia): o sistema realiza transcrições de gravações em tempo real ou off line; 7) Guardtime (Estônia): busca garantir que os modelos IA sejam imparciais, permitir o controle e a auditoria sobre os conjuntos de treinamento de IA, verificar as informações de entrada e outras configurações, resistência a ataques, precisão dos modelos, qualidade e integridade dos dados, controle de acesso, transparência, responsabilidade, conformidade, auditoria de processos e compartilhamento de dados/modelos ; 8) Bürokratt (Estônia): plataforma baseada em IA para a interação entre o Estado e seus cidadãos e segundo apontado “[n]este momento, os serviços

should be paid to policies and practical experiences and distinguish these technologies’ “trendy” effects from their real potential. However, data in this sector do not show an actual deployment of such tools that goes beyond initial experimentation or automation of simple activities (e.g. the use of AI for speech to text conversion)”. CEPEJ. European judicial systems - CEPEJ Evaluation Report – 2022. Evaluation cycle (2020 data). p. 112. Disponível em: <https://rm.coe.int/cepej-report-2020-22-e-web/1680a86279>. Acesso em 12 ago. 2023.

¹²³ CONSELHO NACIONAL DE JUSTIÇA. Delegação da União Europeia no Brasil. Intercâmbio de experiências entre a União Europeia e o Brasil sobre e-Justice: relatório final. Brasília: CNJ, 2022. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/09/seminario-e-justice-v6.pdf>. Acesso em: 24 jan. 2023.

relacionados ao Judiciário não parecem estar incluídos”¹²⁴.

2.3. Outros documentos internacionais

Posteriormente à edição da Resolução CNJ n. 332/2020, em 23 de novembro de 2021, foi aprovada a Recomendação da UNESCO sobre o uso da IA, primeiro documento a nível global sobre o tema adotada por 193 países.

Trata-se de documento extenso que aponta, além de valores e princípios, diversas áreas de ação incluindo, expressamente, o uso da sistemas de IA no Judiciário. No ponto, a recomendação dispõe que nesse caso, assim como outras áreas sensíveis aos direitos humanos, os Estados-membros “devem fornecer mecanismos para monitorar os impactos sociais e econômicos de tais sistemas por autoridades supervisoras adequadas, incluindo autoridades independentes com a função de proteção de dados, supervisão setorial e órgãos públicos responsáveis pela supervisão”¹²⁵.

Ainda especificamente quanto ao uso de IA no sistema de justiça, o documento estabelece a necessidade de estabelecer salvaguardas suficientes “para garantir, entre outros, a proteção dos direitos humanos fundamentais, do Estado de direito, da independência judicial, bem como o princípio da supervisão humana, assim como para garantir que o desenvolvimento e o uso de sistemas de IA no próprio Judiciário sejam confiáveis, orientados ao interesse público e centrado no ser humano”¹²⁶.

Mais recentemente, dois anos depois, em novembro de 2023, a UNESCO publicou documento específico sobre o uso de ferramentas de IA nos sistemas de justiça buscando ampliar o entendimento dos atores envolvidos (juízes, promotores, procuradores, advogados públicos, servidores públicos, acadêmicos de direito) sobre as implicações do emprego da tecnologia na sociedade e os benefícios e os riscos do uso da IA em seu trabalho¹²⁷.

A iniciativa tem caráter nitidamente educativo, redigida em formato de curso destinado à capacitação e desenvolvimento de uma compreensão sobre o papel apropriado dos atores no uso e implementação de ferramentas de IA, bem como oferece uma visão sobre os direitos

¹²⁴ CONSELHO NACIONAL DE JUSTIÇA. Delegação da União Europeia no Brasil. Intercâmbio de experiências entre a União Europeia e o Brasil sobre e-Justice: relatório final. Brasília: CNJ, 2022. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/09/seminario-e-justice-v6.pdf>. Acesso em: 24 jan. 2023.

¹²⁵ UNESCO. **Recomendação sobre a Ética da Inteligência Artificial**. 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 23 nov. 2023.

¹²⁶ *Ibidem*.

¹²⁷ UNESCO. **Global Toolkit on AI and the Rule of Law for the Judiciary**. 2023. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000387331>. Acesso em: 23 nov. 2023.

humanos envolvidos e princípios e orientações que fundamentam o uso responsável da tecnologia.

Em dezembro de 2023, o Judiciário do Reino Unido publicou documento oficial de orientação aos seus membros, servidores e integrantes de equipes de apoio sobre o uso de inteligência artificial¹²⁸. Ele apresenta os principais riscos e problemas associados ao uso da IA generativa, em especial *chatbots* disponíveis ao público – ChatGPT, Bard e similares, e algumas sugestões para minimizá-los.

O documento desenvolve basicamente as seguintes diretrizes: 1) entenda as capacidades e limitações das ferramentas de IA; 2) mantenha a confidencialidade e a privacidade (trate todas as ferramentas de IA como capazes de tornar público tudo o que for inserido nelas); 3) garanta *accountability* e acurácia (qualquer informação obtida deve ser verificada); 4) esteja atento aos vieses; 5) mantenha a segurança (utilize contas e dispositivos profissionais); 6) assuma a responsabilidade (titulares de cargos judiciais são pessoalmente responsáveis pelo material produzido em seu nome); 7) esteja ciente de que os usuários do tribunal podem ter usado ferramentas de IA (atente-se que ferramentas de IA estão sendo usadas para produzir material falso, incluindo texto, imagens e vídeos).

A referida orientação é apontada como sujeita a futuras atualizações à medida que a tecnologia evolua, assim como apenas o primeiro passo de trabalhos futuros para auxiliar o Judiciário em suas interações com a IA.

Em janeiro de 2024, o Tribunal de Justiça da União Europeia (*Court of Justice of the European Union - CJEU*) publicou sua Estratégia de Inteligência Artificial, documento que aponta suas metas e objetivos, estabelece princípios, aponta os riscos envolvidos e possíveis estratégias de mitigação, bem como estabelece uma proposta de governança¹²⁹.

No ponto, apesar de apresentar uma nomenclatura diferente os princípios elencados pelo Tribunal de Justiça da União Europeia não diferem muito daqueles apontados na Carta Ética da CEPEJ, cabendo destacar a previsão específica de que as soluções de IA devem ser auditáveis, da necessidade de mecanismos de *accountability*, bem como de melhoria contínua. Confira-se:

1. **Equidade (*fairness*), imparcialidade e não-discriminação:** Tanto os dados quanto os algoritmos criados ou adotados devem evitar viés e ser orientados pelo princípio da imparcialidade, para que todas as partes recebam tratamento igual durante

¹²⁸ JUDICIARY OF UK. **Artificial Intelligence (AI) – Judicial Guidance**. 12 dec. 2023. Disponível em: https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance/#related_content. Acesso em: 24 jan. 2024.

¹²⁹ COURT OF JUSTICE OF THE EUROPEAN UNION. **Artificial Intelligence Strategy**. Disponível em: https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf. Acesso em: 24 jan. 2024.

o processo judicial ou administrativo. Uma solução de IA produzida ou utilizada não deve discriminar qualquer indivíduo ou grupo com base em fatores como raça, gênero ou status socioeconômico.

2. **Transparência:** A lógica por trás dos algoritmos de IA deve ser clara e compreensível, tanto para aqueles criados internamente quanto para os adquiridos.

3. **Rastreabilidade:** As soluções de IA precisam ser auditáveis e explicativas. Deve haver mecanismos claros de responsabilidade para garantir que os sistemas de IA operem de acordo com padrões legais e éticos

4. **Privacidade e Proteção de Dados:** As soluções de IA criadas ou adotadas devem respeitar os direitos de privacidade e proteção de dados dos indivíduos. Elas precisam garantir que os dados pessoais sejam tratados de maneira segura e ética.

5. **Supervisão Humana:** Qualquer ferramenta de IA deve estar sob supervisão humana rigorosa, de forma contínua. O pensamento crítico humano é essencial e deve ser continuamente desenvolvido e incentivado.

6. **Melhoria Contínua:** Os padrões legais e éticos aplicáveis estão em constante evolução, e, portanto, as soluções de IA devem ser continuamente atualizadas para refletir essa evolução¹³⁰.

Destaca-se que o grande avanço do referido documento foi apontar, de fato, uma estratégia não só para implementar soluções de IA, mas para criar uma estrutura de governança para apoiar que o desenvolvimento, a aquisição e o uso das ferramentas respeitem integralmente a visão e os princípios definidos.

A propósito, o modelo de governança proposto no referido documento está estruturado em quatro órgãos que possuem diferentes funções: a) conselho de gerenciamento de IA (*A.I. Management Board*), que será criado com a missão de garantir que o desenvolvimento ou a aquisição de ferramentas de IA respeite os princípios declarados, especialmente ética e direitos humanos. Esse conselho deverá emitir uma carta de ética e direitos fundamentais que será usada como base de avaliação para qualquer decisão sobre aquisição ou desenvolvimento de qualquer ferramenta de IA, bem como pode estabelecer áreas de negócios ou ferramentas que envolvem risco muito alto e devem ser vedadas; b) comitê diretivo de informática (*Informatics Steering Committee*), já existente, deverá alinhar as ambições com os recursos disponíveis; c) Rede IA+ (*AI+ Network*), criada em fevereiro de 2020 e composta por representantes de cada departamento, dos dois registros e das câmaras dos Presidentes do Tribunal e do Tribunal Geral, terá como objetivo identificar as áreas em que as ferramentas trarão benefícios, bem como testar e avaliar os protótipos ou pilotos projetados; d) Conselhos de Arquitetura e Governança de Dados - o primeiro é um conselho técnico formado por especialistas da ciência da computação que toma decisões sobre os componentes técnicos da arquitetura geral¹³¹. O conselho de governança de dados será responsável por garantir a correção e a qualidade dos dados que serão

¹³⁰ *Ibidem*. p. 16.

¹³¹ COURT OF JUSTICE OF THE EUROPEAN UNION. **Artificial Intelligence Strategy**. 2024. Disponível em: https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf. Acesso em: 24 jan. 2024.

utilizados nos processos de aprendizado de máquina, com um foco especial na representatividade do conjunto de dados de treinamento e na prevenção de vies.

Além dos órgãos acima, o documento também menciona a existência do Laboratório de Inovação, que consiste em uma espécie de um fórum no qual ideias são discutidas e testadas em colaboração com a equipe de TI, bem como que garante a disseminação de informação entre os órgãos componentes do modelo de governança no momento apropriado.

2.4. Regulamentação da IA no Poder Judiciário brasileiro

No Brasil, a Portaria CNJ nº 197, de 22/11/2019, instituiu grupo de trabalho, formado por seis magistrados, um consultor legislativo do Senado Federal, um professor da Universidade de Brasília e uma servidora do CNJ, destinado à elaboração de estudos e propostas voltadas à ética na produção e uso da inteligência artificial no Poder Judiciário.

Como resultado foi apresentada a minuta do que virou a Resolução CNJ nº 332 de 21/8/2020, nitidamente inspirada na Carta Ética do CEPEJ, havendo, inclusive, expressa referência em seu texto ao referido documento.

É preciso destacar a importância da iniciativa de regulamentar o tema, por ato normativo de caráter vinculante, buscando garantir a observância de princípios éticos no desenvolvimento de ferramentas de IA já em amplo desenvolvimento no Judiciário nacional, quando ausente regulamentação da matéria no âmbito legislativo no país.

Enquanto resolução do CNJ, a referida regulamentação possui caráter vinculante para todo o Poder Judiciário brasileiro, exceto o Supremo Tribunal Federal, constituindo verdadeira regulação setorial, na medida em que o CNJ vem se consolidado como verdadeira agência reguladora do poder judiciário¹³².

Apesar de nitidamente inspirada na Carta de Ética do CEPEJ, a Resolução n. 332/2020 optou por implementar um modelo comunitário ou colaborativo de desenvolvimento de projetos de IA no Judiciário, buscando evitar o desperdício de recursos, e estabeleceu alguns instrumentos de governança, prevendo ferramentas de controle, sem paralelo no modelo de inspiração, que serão detalhadamente examinados no próximo tópico do presente estudo.

Em uma visão geral da Resolução n. 332/2020, observa-se que, apesar de asseverar a necessidade de respeito aos direitos fundamentais, não há preocupação específica com as garantias processuais do contraditório, ampla defesa, independência e imparcialidade dos

¹³² PAULINO, Ana Flávia Borges. **CNJ: o regulador da atividade judiciária**. Rio de Janeiro: Lumen Iuris, 2021.

juízes, pontos tratados em detalhes no Apêndice I da Carta Ética do CEPEJ, assim como com o treinamento de magistrados e servidores acerca das capacidades e limitações da IA, bem como dos dilemas éticos envolvidos.

De fato, quanto à própria definição do conteúdo ético, embora a Resolução n. 332/02020 adote como guia os princípios elencados na Carta Ética da CEPEJ não houve desenvolvimento, nem mesmo por regulamentação ou diretrizes posteriores, quanto aos contornos precisos para avaliação do que seria adequado.

Mesmos os cursos disponibilizados pelo CNJ para servidores e magistrados sobre o tema focam apenas o aspecto técnico, como conceitos básicos de IA e de *machine learning*¹³³, sem desenvolver os dilemas éticos, o que deveria ser uma preocupação na capacitação dos atores do sistema de justiça, como realizado no documento da UNESCO (*Global toolkit on AI and the Rule of Law for the Judiciary*, 2023).

Verifica-se também que, enquanto a formação de equipes multidisciplinares é uma preocupação presente da Carta Europeia, ao tratar da composição das equipes, o art. 20 da Resolução n. 332/2020 focou mais em assegurar diversidade “em seu mais amplo espectro, incluindo gênero, raça, etnia, cor, orientação sexual, pessoas com deficiência, geração e demais características individuais” do que no “caráter interdisciplinar”, esse citado no § 4º apenas para indicar que a formação das equipes “deverá considerar seu caráter interdisciplinar, incluindo profissionais de Tecnologia da Informação e de outras áreas cujo conhecimento científico possa contribuir para pesquisa, desenvolvimento ou implantação do sistema inteligente”, sem apontar quais áreas do conhecimento seriam essas, deixando de reconhecer, assim, a importância de aproveitar a experiência de profissionais/pesquisadores das áreas de ciências humanas.

Chama atenção também que, apesar de citar alguns mecanismos de governança como a realização de auditoria e certificação de boas práticas, não houve preocupação em estabelecer a necessidade de avaliação de impacto das diferentes ferramentas de IA no sistema de justiça, tampouco de monitoramento contínuo dos riscos envolvidos durante a fase de operacionalização ou do controle acerca da efetiva implementação de suas regras e/ou avaliação da conformidade com as diretrizes estabelecidas.

Poucos meses depois da publicação da Resolução CNJ n. 332/2020 foi publicada a Portaria n. 271, de 4 de dezembro de 2020, buscando “uniformizar o processo de criação, armazenamento e disponibilização de modelos de inteligência artificial no âmbito do Poder

¹³³ CONSELHO NACIONAL DE JUSTIÇA. Curso online introduz conceitos de Inteligência Artificial para o Judiciário. Brasília: CNJ, 22 jun. 2023. Disponível em: <https://www.cnj.jus.br/curso-online-introduz-conceitos-de-inteligencia-artificial-para-o-judiciario/>. Acesso em: 24 jan.2024.

Judiciário” e “otimizar recursos humanos e financeiros aplicados em projetos de tecnologia da informação”, tal como consta em seus considerando¹³⁴.

A referida portaria tratou de implementar o Sinapses enquanto plataforma oficial do Poder Judiciário, buscando incentivar a colaboração, a transparência, o aprimoramento e a divulgação dos projetos.

Na sequência, em janeiro de 2021, com a implementação do Programa Justiça 4.0, o CNJ passou a adotar nítida política de impulsionamento da transformação digital e do uso de IA para tornar a prestação de serviços de justiça mais eficiente, eficaz e acessível à sociedade¹³⁵.

Nesse contexto, foi sendo criado um ambiente de incentivo à inovação e transformação digital com a publicação da Lei n. 14.129/2021 (Lei do Governo Digital), de 29/3/2021, e Resolução CNJ nº 395 de 07/06/2021, que instituiu a Política de Gestão da Inovação no âmbito do Poder Judiciário.

Assim, é inegável que a edição da Resolução nº 332/2020 está inserida em um contexto de estímulo à inovação, e foi seguida por implementação de projetos destinados a promover e concretizar aplicação de tecnologia em serviços do Poder Judiciário.

Ocorre que, desde a publicação do referido ato normativo, pouco tem sido feito para verificar a implementação na prática das diretrizes estabelecidas na Resolução nº 332/2020 ou para monitoramento dos resultados.

Enquanto no âmbito da CEPEJ a Carta Ética foi apenas o primeiro passo no desenvolvimento progressivo de uma estrutura (ainda em construção) voltada a garantir observância aos direitos humanos e princípios éticos nas aplicações de IA nos sistemas de justiça, no âmbito do CNJ o grupo de trabalho responsável por elaborar a resolução foi dissolvido oficialmente em 2022 e não foi criado nenhum conselho, comitê ou órgão responsável por propor estratégias de implementação de suas normas ou fiscalizá-las.

Ressalta-se que, apesar de inspirada na Carta Ética da CEPEJ, o CNJ não implementou ciclos de avaliação com relatórios públicos. Também não se tem notícia da aplicação prática de instrumentos de governança previstos no ato normativo, em especial as auditorias. Em verdade,

¹³⁴ CONSELHO NACIONAL DE JUSTIÇA. **Portaria n. 271, de 04 de dezembro de 2020.** Regulamenta o uso de Inteligência Artificial no âmbito do Poder Judiciário. DJe/CNJ nº 389/2020, de 9/12/2020, p. 2-4. E republicada no DJe/CNJ nº 393/2020, de 14/12/2020, p. 2-4. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em 12 fev. 2024.

¹³⁵ CONSELHO NACIONAL DE JUSTIÇA. O Programa Justiça 4.0 – Inovação e Efetividade na Realização da Justiça para Todos é desenvolvido em parceria entre o CNJ, o Programa das Nações Unidas para o Desenvolvimento (PNUD) e o Conselho da Justiça Federal (CJF), com apoio do Tribunal Superior Eleitoral (TSE), Superior Tribunal de Justiça (STJ) e Conselho Superior da Justiça do Trabalho (CSJT) Brasília: CNJ, 14 dez. 2023. Disponível em <https://www.cnj.jus.br/tecnologia-da-informacao-e-comunicacao/justica-4-0/> Acesso em: 01 abr. 2022.

até mesmo a transparência, princípio sobre o qual é possível afirmar existir um certo consenso quanto à sua aplicação no campo do uso ético da IA, está longe de ser observada na prática, ainda que nos parâmetros estabelecidos no ato normativo, como será visto adiante.

Assim, verifica-se que, na prática, não obstante a importância da Resolução n. 332/2020, o referido documento assemelha-se mais a carta de intenções do que ato normativo cogente.

Nesse sentido, Dierle Nunes aponta que, apesar de ser um passo importante, a referida resolução ainda é insuficiente, pois seu conteúdo ainda precisa alcançar a prática na sua implementação desde o design dos modelos de IA até o controle de seus resultados¹³⁶.

Em um contexto de debate regulatório efervescente no Brasil e no mundo, bem como de avanço de técnicas da tecnologia, em especial diante da disseminação de ferramentas de inteligência artificial generativa e da pretensão, oficialmente divulgada pelo Supremo Tribunal Federal, de contratação de soluções tecnológicas para gerar resumos de processos¹³⁷, foi publicada, em 30 de novembro de 2023, pelo CNJ, a Portaria Presidência nº 338, que institui Grupo de Trabalho sobre inteligência artificial no Poder Judiciário.

O objetivo do referido grupo, formado por magistrados, advogados e membros da academia e do Ministério Público, consta, de maneira expressa, no art. 2º da portaria, nos seguintes termos:

Art. 2º O Grupo de Trabalho tem por objetivo realizar estudos e apresentar proposta de regulamentação do uso de sistemas de inteligência artificial generativa baseada em grandes modelos de linguagem no Poder Judiciário que disponha sobre:

- I – modelo de governança para gestão do processo de desenvolvimento, sustentação e uso de soluções de inteligência artificial, orientado pela transparência de auditabilidade;
- II – colaboração e compartilhamento de informações acerca do uso das soluções de inteligência artificial;
- III – auditoria de modelos e soluções de inteligência artificial sob as perspectivas da segurança da informação, proteção de dados, performance, robustez, confiabilidade, vieses, correlação entre entradas e saídas, conformidade legal e ética, dentre outros;
- IV – mapeamento e gerenciamento de riscos;
- V – práticas e casos de uso permitido, regulado e proibido; e
- VI – revisão da Resolução CNJ nº 332/2020.

Verifica-se que há um movimento oficial de revisão da Resolução nº 332/2020, buscando aprimorar a regulamentação atual, para tratar especificamente de sistemas de IA generativa baseada em grandes modelos de linguagem, de modo especial abordando

¹³⁶ BAHIA, Alexandre; NUNES, Dierle; PEDRON, Flávio. **Teoria Geral do Processo**. Salvador: Juspodivm, 2021. p. 144.

¹³⁷ SUPREMO TRIBUNAL FEDERAL. STF faz chamamento público para projetos de inteligência artificial que automatizem resumos de processos. Brasília: STF, 07 nov. 2023. Disponível em: <https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=518467&ori=1> Acesso em: 24 jan. 2023.

mecanismos de governança como auditorias, mapeamento e gerenciamento de riscos.

Nessa perspectiva, o presente trabalho pode contribuir na medida em que busca indicar pontos de melhoria na atual regulamentação da matéria, em especial na estrutura de governança.

2.5. Instrumentos previstos na Resolução n. 332/2020

O tema governança foi previsto expressamente na Resolução n. 332/2020 no capítulo V, “Da governança e da qualidade”.

No referido capítulo há expressa referência à: a) necessidade de comunicação ao CNJ da pesquisa, o desenvolvimento, a implantação ou o uso da Inteligência Artificial, bem como os respectivos objetivos e os resultados que se pretende alcançar (art. 10, I); b) promoção de esforços para atuação em modelo comunitário, com vedação a desenvolvimento paralelo quando a iniciativa possuir objetivos e resultados alcançados idênticos a modelo de Inteligência Artificial já existente ou com projeto em andamento (art. 10, II,); c) depósito do modelo de IA no Sinapses (art. 10, III).

No mencionado capítulo há também referência à divulgação, pelo CNJ, da relação dos modelos de IA desenvolvidos ou utilizados pelos órgãos do Poder Judiciário (art. 11) e necessidade de que os modelos desenvolvidos apresentem “interface de programação de aplicativos (API) que permitam sua utilização por outros sistemas” (art. 12).

Há, ainda, outros mecanismos de governança previstos na lei de maneira esparsa como a necessidade de comunicação ao CNJ de “todos os registros de eventos adversos no uso da Inteligência Artificial” (art. 27) e “apresentação dos mecanismos de auditoria e certificação de boas práticas” (art. 8º, V).

Passa-se à análise pormenorizadas dos referidos instrumentos.

2.5.1. A plataforma Sinapses (art. 10, III)

Ao tratar especificamente de governança, o art. 10 da Resolução n. 332/2020 estabelece expressamente o dever de depositar os modelos de IA criados no Sinapses. O Sinapses, por sua vez, é definido no art. 3º, III, da referida resolução como sendo “solução computacional, mantida pelo Conselho Nacional de Justiça, com o objetivo de armazenar, testar, treinar, distribuir e auditar modelos de Inteligência Artificial”.

O Sinapses é, portanto, a plataforma nacional para gerenciamento e treinamento de

modelos de IA, disponibilizada e administrada conjuntamente pelo CNJ e pelo Tribunal de Justiça do Estado de Rondônia, por meio da qual se busca concretizar o modelo comunitário ou colaborativo para desenvolvimento de soluções de IA, que almeja evitar o desperdício de recursos financeiros e humanos e promover a integração do Judiciário.

De fato, trata-se de uma “plataforma para desenvolvimento e disponibilização em larga escala de modelos de Inteligência Artificial, também comumente conhecido como ‘Fábrica de Modelos de IA’”¹³⁸. O termo técnico utilizado na área da Tecnologia da Informação (TI) é *pipeline* de desenvolvimento, que corresponde ao conjunto de tarefas sequenciais necessárias para o desenvolvimento e entrega de modelos de IA.

Segundo os seus desenvolvedores, o Sinapses já nasceu com a ideia de automatizar os passos de criação de um modelo de IA para agilizar o desenvolvimento de soluções. A plataforma pode ser utilizada para treinamento do modelo, o que permite gerenciar dados e realizar certa rastreabilidade, possibilitando a realização de auditoria, bem como funciona como repositório de modelos de IA¹³⁹.

A utilização do Sinapses, portanto, agiliza o trabalho da equipe responsável por desenvolver soluções tecnológicas, permite o controle do fluxo de desenvolvimento, garante maior segurança, bem como possibilita que, além de construir e compartilhar seus próprios modelos, sejam adotados modelos à disposição desenvolvidos por outros tribunais.

A adoção do Sinapses como plataforma nacional pelo CNJ buscou evitar que diferentes tribunais envidassem esforços para desenvolver ferramentas de IA visando o mesmo objetivo, proporcionando escalabilidade e economicidade. Uma vez que o modelo é liberado para produção fica disponível na prateleira da plataforma para que os tribunais interessados possam desenvolver integração do modelo aos seus sistemas processuais. Por essa razão, fala-se que funciona como uma espécie de mercado para micro serviços de IA.

A plataforma foi objeto de regulamentação na Portaria n. 271, de 4 de dezembro de 2020, que estabelece que a equipe de administração do Sinapses poderá “criar, alterar e extinguir indicadores para os modelos”, “com a finalidade de ranquear os modelos quanto à sua aplicação e benefícios” (art. 6º), bem como que “[o] CNJ poderá prestar auxílio técnico aos tribunais” na tarefa de integração dos modelos aos seus sistemas (art. 7º)¹⁴⁰.

¹³⁸ TRIBUNAL DE JUSTIÇA DE RONDÔNIA. Sinapses: termo de cooperação técnica com o CNJ completa 1 ano. 19 dez. 2019. Disponível em: <https://www.tjro.jus.br/noticias/item/11953-sinapses-termo-de-cooperacao-tecnica-com-o-cnj-completa-1-ano>. Acesso em: 3 jun. 2022.

¹³⁹ Entrevista com Alcides Fernando Faria Campos, Pablo Filetti Moreira e Mikael Barbosa de Araújo parte do módulo 2 do curso do CNJ “Introdução à Inteligência Artificial para o Poder Judiciário - 2023/2”.

¹⁴⁰ CONSELHO NACIONAL DE JUSTIÇA. **Portaria n. 271, de 04 de dezembro de 2020**. Regulamenta o uso de Inteligência Artificial no âmbito do Poder Judiciário. DJe/CNJ nº 389/2020, de 9/12/2020, p. 2-4. E republicada

A norma prevê, ainda, que o órgão responsável pelo projeto ou modelo de IA deve disponibilizar na plataforma a documentação pertinente, “indicando claramente as necessidades que ensejaram sua criação e os objetivos que se pretende alcançar”, bem como que cabe ao CNJ “apresentar, periodicamente, levantamento dos projetos de inteligência artificial que estão sendo desenvolvidos e o número de processos judiciais beneficiados”.

Em junho de 2023, indicava-se que 150 modelos de IA ativos, produzidos por 29 tribunais e conselhos, estavam depositados na plataforma Sinapses, sem que fossem divulgadas quaisquer informações sobre as soluções disponíveis¹⁴¹.

O Sinapses é instrumento essencialmente interno ao CNJ, de modo que ao adotar a referida plataforma como meio para acompanhamento e avaliação das ferramentas de IA foca apenas na governança interna da instituição, acarretando nítida falta transparência com o público externo.

Assim, para fortalecer a confiança da sociedade na utilização de IA no Judiciário brasileiro, sugere-se uma abordagem mais proativa em relação à transparência.

No ponto, o CNJ poderia divulgar a lista de todos os modelos hospedados na plataforma disponíveis para utilização, destacando os objetivos alcançados, os indicadores de sucesso e os benefícios observados, assim como publicar sumários executivos ou versões resumidas das avaliações realizadas por meio do Sinapses. Essa prática proporcionaria uma visão acessível e compreensível para o público em geral, reforçando a prestação de contas à sociedade (*accountability*).

Desse modo, recomenda-se que seja desenvolvido um ponto focal acessível ao público para informações estruturadas e confiáveis sobre sistemas de IA hospedados no Sinapses.

2.5.2. Painel de Projetos com Inteligência Artificial no Poder Judiciário (art. 11)

Para garantir transparência quanto à utilização das ferramentas de IA, a Resolução 332/2020, em seu artigo 11, estabeleceu o dever de publicação, pelo CNJ, em área própria de seu sítio na rede mundial de computadores, de relação dos modelos de IA desenvolvidos ou utilizados pelos órgãos no Poder Judiciário.

no DJe/CNJ nº 393/2020, de 14/12/2020, p. 2-4. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em 12 fev. 2024.

¹⁴¹ UNITED NATIONS DEVELOPMENT PROGRAMME. Plataforma Sinapses reúne 150 modelos de inteligência artificial. 13 July. 2023. Disponível em: <https://www.undp.org/pt/brazil/news/plataforma-sinapses-reune-150-modelos-de-inteligencia-artificial>. Acesso em: 17 jun. 2023.

Em sentido semelhante, mas indo um pouco além, em 2021, no âmbito europeu, o “Roteiro revisado para assegurar um acompanhamento adequado (*follow-up*) da Carta Ética do CEPEJ sobre a utilização da inteligência artificial nos sistemas judiciais e no seu ambiente” propôs a criação de uma espécie de banco de dados acessível ao público, regularmente atualizado, em que devem ser indicadas todas as aplicações de IA existentes no domínio da justiça e no qual podem ser acrescentadas informações adicionais sobre a experiência do usuário, avaliações e certificações¹⁴².

A fim de concretizar esse dever de publicação, o Conselho Nacional de Justiça passou a disponibilizar, a partir de 2021, anualmente, um painel de *Business Intelligence* com o resultado de pesquisa enviado aos órgãos do Poder Judiciário denominado “Painel de Projetos com Inteligência Artificial no Poder Judiciário”.

No ponto, o referido painel, embora cumpra com o papel de consolidar e divulgar ao público os dados obtidos pelo Conselho Nacional de Justiça acerca dos projetos de IA em utilização ou em desenvolvimento está longe de cumprir o ideal de transparência.

Com efeito, os dados divulgados consistem em mero compilado das informações brutas obtidas a partir do formulário da pesquisa realizada pelo Conselho Nacional de Justiça (CNJ) junto aos tribunais “sem a verificação técnica de sua confiabilidade e avaliação qualitativa por terceiro tecnicamente habilitado e neutro do ponto de vista do interesse no objeto avaliado”¹⁴³.

Em estudo dedicado tema, o magistrado Emiliano Zapata Leitão destaca que:

O exame do conteúdo “Painel de Projetos com Inteligência Artificial no Poder Judiciário” em sua atualização de maio de 2022 demonstra que ele é a simples translação para o formato de painel de *Business Intelligence* das respostas à pesquisa que lhe deu origem. Essas respostas são aquelas apresentadas pelos Tribunais e Conselhos do Poder Judiciário brasileiro ali indicados às perguntas formuladas pelo Conselho Nacional de Justiça (CNJ) em formulário de pesquisa a eles apresentado no primeiro semestre de 2022.

(...)

a) Há impossibilidade de adequada interpretação das respostas da pesquisa realizada pelo Conselho Nacional de Justiça (CNJ) apresentadas como dados no painel por não ter sido demonstrado que, na formulação das perguntas que lhes deram origem, houve a adequada explicitação dos conceitos técnicos utilizados em sua formulação;

b) Há ausência, na forma de apresentação dessas respostas no painel analisado, de elementos descritivos ou exemplificativos mais detalhados necessários à aferição da congruência de compreensão, pelos respondentes da pesquisa, do conteúdo técnico mais preciso das perguntas formuladas e, portanto, também, das respostas

¹⁴² CEPEJ. **Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment**. 9 dec. 2021. Disponível em: <https://rm.coe.int/cepej-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 15 nov. 2023.

¹⁴³ LEITÃO, Emiliano Zapata. “Painel de Projetos com Inteligência Artificial no Poder Judiciário” do Conselho Nacional de Justiça (CNJ): avaliação qualitativa da transparência governamental. 2022. 129 f. Dissertação (Mestrado em Direito). **Escola Nacional de Formação e Aperfeiçoamento de Magistrados**: Brasília, 2022. p. 99.

apresentadas;

c) Não há possibilidade de compreensão adequada da validade ou não das inferências que podem ser extraídas dos dados apresentados em função do caráter lacônico das respostas divulgadas como dados nesse painel;

d) Os dados apresentados em referido painel são interpretações apresentadas pelos próprios interessados aos questionamentos formulados na pesquisa do Conselho Nacional de Justiça (CNJ) que deu origem aos dados divulgados no painel analisado; não têm, assim, esses dados característica de dados brutos ou primários ou não tratados para fins de atendimento das necessidades do público mais especializado na aferição adequada da qualidade e validade dessas informações e das inferências delas extraíveis;

e) Há obstáculo à extração de inferências técnicas válidas a partir dos dados divulgados no painel em face da não apresentação de documentos primários descritivos das questões técnicas objeto dos questionamentos da pesquisa realizada pelo Conselho Nacional de Justiça (CNJ), na forma como avaliadas e implementadas originalmente em cada projeto de uso de inteligência artificial ali indicados¹⁴⁴.

Conforme apontado em documento oficial, produzido pelo Conselho Nacional de Justiça em conjunto com a Delegação da União Europeia no Brasil, afirma-se que “[n]ão é possível dizer que estas [ferramentas listadas no painel do CNJ] são todas as iniciativas em andamento, pois a atualização depende das informações fornecidas pelos próprios Tribunais e da classificação adequada do que seja ou não uma ferramenta de AI”¹⁴⁵.

De fato, com facilidade, é possível encontrar notícias de sistemas de IA em utilização ou em treinamento em órgãos do Poder Judiciário brasileiro que não constam do painel disponibilizado ao público pelo CNJ.

Nesse sentido, a título ilustrativo, é possível apontar as seguintes ferramentas de IA, todas elas apontadas em notícias divulgadas em sites oficiais dos tribunais ou do CNJ até 2021 e que não constam no “Painel de Projetos com Inteligência Artificial no Poder Judiciário” em sua atualização de maio de 2022: a) Poti, Clara e Jerimum do Tribunal de Justiça do Estado do Rio Grande do Norte (TJRN)¹⁴⁶; b) Radar do Tribunal de Justiça do Estado de Minas Gerais (TJMG)¹⁴⁷; c) LEIA Precedentes, desenvolvida pela empresa Softplan e utilizada em diversos

¹⁴⁴ *Ibidem*, p. 94-112.

¹⁴⁵ CONSELHO NACIONAL DE JUSTIÇA. Delegação da União Europeia no Brasil. Intercâmbio de experiências entre a União Europeia e o Brasil sobre e-Justice: relatório final. Brasília: CNJ, out. 2022. p. 81. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/09/seminario-e-justice-v6.pdf>. Acesso em: 11 jan. 2024.

¹⁴⁶ CONSELHO NACIONAL DE JUSTIÇA. Judiciário ganha agilidade com uso de inteligência artificial. Brasília: CNJ, 3 abr. 2019. Disponível em: <https://www.cnj.jus.br/judiciario-ganha-agilidade-com-uso-de-inteligencia-artificial/>. Acesso em: 11 jan. 2024.

¹⁴⁷ TRIBUNAL DE JUSTICA DO ESTADO DE MINAS GERAIS. TJMG é destaque em projeto do STF e STJ. Programa de inteligência artificial Radar foi apresentado durante o projeto Sextas Inteligentes. 2 out. 2021. Disponível em: <https://www.tjmg.jus.br/portal-tjmg/noticias/tjmg-e-destaque-em-projeto-do-stf-e-stj.htm>. Acesso em: 11 de janeiro de 2024; Judiciário ganha agilidade com uso de inteligência artificial. Disponível em: <https://www.cnj.jus.br/judiciario-ganha-agilidade-com-uso-de-inteligencia-artificial/>. Acesso em: 11 jan. 2024.

tribunais, tais como o Tribunal de Justiça do Estado do Acre¹⁴⁸, Tribunal de Justiça do Mato Grosso do Sul¹⁴⁹, Tribunal de Justiça do Estado do Amazonas¹⁵⁰.

No ponto, cabe destacar que, conforme apontado por Emiliano Zapata Leitão, os dados apresentados no Painel de Projetos do CNJ não podem ser considerados como dados primários ou brutos, mas apenas interpretações dos próprios interessados do formulário de pesquisa e sequer é disponibilizado ao público o questionário utilizado para obtenção dos dados¹⁵¹.

Ressalta-se, ainda, que o conceito de transparência é relacional, de modo que o conteúdo informativo deve levar em consideração o receptor dessas informações e o uso respectivo, isto é, se público externo ou interno, com ou sem conhecimento técnico na área de IA, diretamente afetado ou apenas exercício de controle democrático. Tal ponto é desconsiderado na política de transparência implementada pelo CNJ através do “Painel de Projetos com Inteligência Artificial no Poder Judiciário”, que se limita a apresentar uma única versão com apresentação, sem o devido tratamento, de informações colhidas de respostas ao questionário apresentado.

Nessa linha, o documento “*Algorithmic accountability for the public sector*”, elaborado, em 2021, pelo Instituto Ada Lovelace em parceria com o AI Now Institute e Open Government Partnership, aponta como lição para órgãos públicos que mecanismos que se concentram na transparência “devem ser detalhados e adequados ao público para fundamentar a *accountability*” (tradução livre)¹⁵².

A propósito, Maranhão et al. propõem uma matriz de transparência para os modelos de IA existentes no Poder Judiciário brasileiro, apontando questões relevantes a serem respondidas a depender do destinatário da informação (usuários, cidadãos, magistrados e servidores da atividade-fim, servidores de TI, comitê decisório sobre tecnologia e CNJ) em relação aos

¹⁴⁸ PODER JUDICIÁRIO DO ESTADO DO ACRE. TJAC utiliza Inteligência Artificial para identificar processos vinculados a precedentes. 2 dez. 2019. Disponível em: <https://www.tjac.jus.br/2019/12/tjac-utiliza-inteligencia-artificial-para-identificar-processos-vinculados-a-precedentes/>. Acesso em: 11 jan. 2024.

¹⁴⁹ Projeto LEIA Precedentes. Disponível em: <https://sti.tjms.jus.br/confluence/display/GPS/Projeto+LEIA+Precedentes>. Acesso em: 11 jan. 2024.

¹⁵⁰ TJAM adota ferramenta de inteligência artificial capaz de identificar processos que podem ser vinculados a Temas de Precedentes dos Tribunais Superiores. Disponível em: <https://www.tjam.jus.br/index.php/menu/sala-de-imprensa/2190-tjam-adota-ferramenta-de-inteligencia-artificial-capaz-de-identificar-processos-que-podem-ser-vinculados-a-temas-de-precedentes-dos-tribunais-superiores>. Acesso em: 11 jan. 2024.

¹⁵¹ LEITÃO, Emiliano Zapata. *Op. cit.* No original: “Em realidade, do ponto de vista acadêmico, por exemplo, esses dados não são dados primários, mas interpretações apresentadas pelos próprios interessados aos questionamentos formulados na pesquisa do Conselho Nacional de Justiça (CNJ) que deu origem aos dados divulgados no painel analisado. Não têm, assim, a característica de dados brutos ou primários ou não tratados. Esta melhor atenderia aos interesses do público mais especializado referido na avaliação adequada da validade das informações apresentadas nesse painel e das inferências válidas passíveis de serem delas extraídas”.

¹⁵² ADA LOVELACE INSTITUTE; AI NOW INSTITUTE; OPEN GOVERNMENT PARTNERSHIP. [2021]. *Algorithmic Accountability for the Public Sector*. p. 4. Disponível em: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>. Acesso em: 13 ago. 2023.

aspectos “por que”, “o que”, “quando” e “como”¹⁵³.

De fato, há informações que, apesar de importantes para o monitoramento de riscos, podem eventualmente não ser adequadas para divulgação pública, mas é importante que haja transparência sobre a própria política de transparência da organização¹⁵⁴.

Partindo da proposta apresentada por Maranhão et al., indica-se como informações essenciais para o público externo (usuários e cidadãos em geral): objetivo da ferramenta e seu efeito na qualidade dos serviços judiciais; tipo de aplicação e função do sistema de IA e seus possíveis resultados específicos na condução e tramitação do processo.

Vale salientar que a transparência, além de ser um dos princípios orientadores da Carta Ética da CEPEJ, é um dever reconhecido tanto na Resolução n. 332/2020, quanto na Portaria n. 271/2020, consistindo, ainda, em elemento fundamental para que exista um controle social quanto a observância dos demais princípios éticos, assim como dos impactos da implementação da tecnologia.

Destaca-se, ainda, a necessidade atualização das informações de modo permanente. No atual cenário, o CNJ apenas tem divulgado a consolidação dos dados obtidos junto aos tribunais uma vez ao ano, ficando as informações desatualizadas até nova publicação. A atualização contínua do painel é medida viável, extremamente simples, que muito contribui para a transparência.

De fato, notícias mais recentes apontam que o próximo levantamento realizado pelo CNJ terá o objetivo de produzir um relatório e uma base de dados passíveis de atualização contínua¹⁵⁵.

Outrossim, para além da consolidação das informações pelo CNJ, os Tribunais também necessitam observar o dever de transparência, publicando informações em seu próprio site oficial, em local de fácil visualização e acesso, sobre as ferramentas utilizadas, seus objetivos e resultados esperados.

2.5.3. Comunicação de eventos adversos (art. 27)

¹⁵³ DE ALBUQUERQUE MARANHÃO, Juliano Souza; JUNQUILHO, Tainá Aguiar; TASSO, Fernando Antonio. Transparência sobre o emprego de Inteligência Artificial no Judiciário: um modelo de governança. **Suprema-Revista de Estudos Constitucionais**, v. 3, n. 2, p. 145-187, 2023.

¹⁵⁴ *Ibidem*.

¹⁵⁵ CONSELHO NACIONAL DE JUSTIÇA. Com a plataforma Sinapse, Judiciário assume protagonismo no desenvolvimento de soluções de IA. Brasília: **CNJ**, 26 jun. 2023. Disponível em: <https://www.cnj.jus.br/com-a-plataforma-sinapses-judiciario-assume-protagonismo-no-desenvolvimento-de-solucoes-de-ia/>. Acesso em: 11 jan. 2024.

Segundo disposto no art. 27 da Resolução n. 332/2020, “os órgãos do Poder Judiciário informarão ao Conselho Nacional de Justiça todos os registros de eventos adversos no uso da Inteligência Artificial”.

Trata-se de previsão similar, em alguma medida, à comunicação de incidente de segurança, prevista no art. 48 da Lei n. 13.709/2018¹⁵⁶ (Lei Geral de Proteção de Dados – LGPD), e ao instrumento de comunicação de incidentes graves do art. 31 do PL nº 2338/2023¹⁵⁷.

No ponto, a proposta do PL elaborada pela comissão de juristas segue exatamente a linha da proposta do *AI Act* da União Europeia, apontando a necessidade de comunicação à autoridade competente de incidentes graves, considerados esses “quando houver risco à vida e integridade física de pessoas, a interrupção de funcionamento de operações críticas de infraestrutura, graves danos à propriedade ou ao meio ambiente, bem como graves violações aos direitos fundamentais” (art. 31).

Nota-se que, no ponto, a redação dada pela Resolução CNJ n. 332/2020 foi bastante ampla ao se referir à comunicação de “todos os registros de eventos adversos”, o que certamente vai além de incidentes de segurança, podendo incluir, entre outros casos de consequências negativas, discriminação algorítmica e verificação de erros.

Em verdade, o referido dispositivo da resolução carece de maior regulamentação, explicitando, prazo para comunicação, canal apropriado e, inclusive, o próprio conceito de “evento adverso”.

¹⁵⁶ “Art. 48. O controlador deverá comunicar à autoridade nacional e ao titular a ocorrência de incidente de segurança que possa acarretar risco ou dano relevante aos titulares. § 1º A comunicação será feita em prazo razoável, conforme definido pela autoridade nacional, e deverá mencionar, no mínimo: I - a descrição da natureza dos dados pessoais afetados; II - as informações sobre os titulares envolvidos; III - a indicação das medidas técnicas e de segurança utilizadas para a proteção dos dados, observados os segredos comercial e industrial; IV - os riscos relacionados ao incidente; V - os motivos da demora, no caso de a comunicação não ter sido imediata; e VI - as medidas que foram ou que serão adotadas para reverter ou mitigar os efeitos do prejuízo. § 2º A autoridade nacional verificará a gravidade do incidente e poderá, caso necessário para a salvaguarda dos direitos dos titulares, determinar ao controlador a adoção de providências, tais como: I - ampla divulgação do fato em meios de comunicação; e II - medidas para reverter ou mitigar os efeitos do incidente.”. BRASIL. **Lei n. 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União de 15/08/2018, pág. nº 59. Disponível em: https://www.planalto.gov.br/CCIVIL_03/_Ato2015-2018/2018/Lei/L13709.htm. Acesso em: 17 jan. 2024.

¹⁵⁷ “Art. 31. Os agentes de inteligência artificial comunicarão à autoridade competente a ocorrência de graves incidentes de segurança, incluindo quando houver risco à vida e integridade física de pessoas, a interrupção de funcionamento de operações críticas de infraestrutura, graves danos à propriedade ou ao meio ambiente, bem como graves violações aos direitos fundamentais, nos termos do regulamento. § 1º A comunicação será feita em prazo razoável, conforme definido pela autoridade competente. § 2º A autoridade competente verificará a gravidade do incidente e poderá, caso necessário, determinar ao agente a adoção de providências e medidas para reverter ou mitigar os efeitos do incidente”. BRASIL. Senado Federal. **Projeto de Lei nº 2338, de 4 de maio de 2023**. Dispõe sobre o uso da Inteligência Artificial. Brasília: Senado Federal, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 17 jan. 2024.

Destaca-se que a medida de compilar casos em que verificadas falhas ou consequências negativas do uso de ferramentas de IA, apesar de aparentemente simples, tem o potencial, se adotada estrutura apta para organizar as informações coletadas, de gerar aprendizado, dentro da perspectiva de trabalho colaborativo incentivado pelo CNJ, permitindo que se extraia informações úteis a partir dos casos relatados para o aprimoramento dos modelos em desenvolvimento e utilização.

Em perspectiva semelhante, em 2022, documento da OCDE tratou do tema e indicou que desenvolver uma estrutura comum para tratar de incidentes de IA, especialmente aqueles negativos ou prejudiciais, assim como controvérsias, permite construir uma base de evidências sobre riscos que se materializaram em incidentes ou quase incidentes, de modo que possam ser observados, documentados e estudados, assim como permite compartilhar práticas para prevenir e mitigar prejuízos¹⁵⁸.

No contexto tratado no âmbito da OCDE, incidente é uma terminologia utilizada para descrever problemas ou falhas nos sistemas de IA tendo como foco a ocorrência de prejuízo a pessoas, propriedade ou ao meio ambiente, bem como violação a direitos humanos e valores fundamentais.

Em novembro de 2023, foi lançada uma versão beta de banco de monitoramento de incidentes de IA da OCDE (no original, *OECD AI Incidents Monitor - AIM*), que se propõe a documentar, em tempo real, incidentes para auxiliar formuladores de políticas, profissionais e todas as partes interessadas a obter *insights* sobre os riscos que concretizam as ameaças da IA¹⁵⁹. Aponta-se que, com o tempo, a iniciativa irá ajudar a identificar padrões e estabelecer uma compreensão coletiva dos incidentes de IA, esses entendidos enquanto consequências negativas, permitindo prevenir acontecimentos semelhantes no futuro.

Na prática, essa base de dados da OCDE realiza a compilação de artigos de notícias que relatam incidentes e “espera-se que as informações de incidentes provenientes de artigos de notícias sejam complementadas por decisões judiciais e decisões de autoridades supervisoras públicas, sempre que existirem” (tradução livre)¹⁶⁰.

¹⁵⁸ OECD. **OECD Framework for the classification of AI systems**. 22 feb. 2022. Disponível em: <https://doi.org/10.1787/cb6d9eca-en>. Acesso em: 11 nov. 2023.

¹⁵⁹ OECD. **OECD AI Incidents Monitor (AIM)**. Disponível em: https://oecd.ai/en/incidents?search_terms=%5B%5D&and_condition=false&from_date=2014-01-01&to_date=2023-11-14&properties_config=%7B%22principles%22:%5B%5D,%22industries%22:%5B%5D,%22harm_types%22:%5B%5D,%22harm_levels%22:%5B%5D,%22harmed_entities%22:%5B%5D%7D&only_threats=false&order_by=date&num_results=20. Acesso em: 11 nov. 2023.

¹⁶⁰ OECD. **OECD AI Incidents Monitor (AIM)**. Disponível em: <https://oecd.ai/en/incidents-methodology>. Acesso em: 16 fev. 2024.

Iniciativa semelhante já havia sido implementada, em 2021, no “*AI Incident Database*”, idealizado por Sean McGregor com a missão de “identificar, definir e catalogar incidentes de Inteligência Artificial” (tradução livre)¹⁶¹.

A comunicação obrigatória de incidentes foi também apontada, no Relatório Global de Riscos do Fórum Econômico Mundial, como uma das soluções propostas para lidar com a preparação para riscos e a redução dos resultados adversos das tecnologias de IA¹⁶².

Também o *framework* de governança para IA generativa elaborado por órgãos do governo de Singapura (*AI Verify Foundation* e *Infocomm Media Development Agency* - IMDA) aponta a comunicação de incidentes como prática capaz de proporcionar melhoria contínua dos sistemas de IA por meio de *insights*, descoberta de vulnerabilidades, problemas funcionais e correções de falhas¹⁶³.

No âmbito do Judiciário brasileiro, não obstante a existência de alguns incidentes já noticiados pela mídia nacional¹⁶⁴, não há registro de implementação de ferramenta semelhante, que busque compilar os eventos adversos, seja com base em notícias jornalísticas, seja com base nos comunicados oficiais dos tribunais ao Conselho Nacional de Justiça.

Na perspectiva do trabalho que vem sendo desenvolvido no âmbito da OCDE sobre monitoramento de incidentes de IA, é importante ressaltar que a mera comunicação de eventos adversos por parte dos tribunais ao CNJ é insuficiente para que sejam obtidos os benefícios potenciais da criação de um banco de dados para extrair informações úteis a partir dos casos relatados.

Com efeito, para que sejam extraídos ainda mais benefícios da compilação de informações parece ser essencial também que seja criado um canal para que usuários externos (advogados, membros das funções essenciais à Justiça e jurisdicionados em geral) possam comunicar a ocorrência de incidentes para tratamento e análise por parte do Conselho Nacional de Justiça.

Nesse mesmo sentido, Eunice Maria Batista Prado apresentou a proposta de criação do

¹⁶¹ MCGREGOR, Sean. Preventing repeated real world AI failures by cataloging incidents: The AI incident database. 18 apr. 2021. Disponível em: <https://ojs.aaai.org/index.php/AAAI/article/view/17817>. Acesso em: 18 jul. 2022.

¹⁶² WORLD ECONOMIC FORUM. **Global Risks Report 2024**. 19th Edition. Disponível em: <https://www.weforum.org/publications/global-risks-report-2024/>. Acesso em: 18 jul. 2022.

¹⁶³ AI VERIFY FOUNDATION. Proposed Model AI Governance Framework for Generative AI: Fostering a Trusted Ecosystem. 2024. Disponível em: <https://aiverifyfoundation.sg/news-and-resources/#proposed-model-governance-framework-for-generative-ai>. Acesso em: 18 jul. 2022.

¹⁶⁴ LIMA, Daniela. “Juiz usa inteligência artificial para fazer decisão e cita jurisprudência falsa; CNJ investiga caso”. 13 nov. 2023. Disponível em: <https://g1.globo.com/politica/blog/daniela-lima/post/2023/11/13/juiz-usa-inteligencia-artificial-para-fazer-decisao-e-cita-jurisprudencia-falsa-cnj-investiga-caso.ghml>. Acesso em: 18 jul. 2022.

Banco de Incidentes de IA do Judiciário (BIIAJud), como parte integrante de uma política pública judiciária de gerenciamento dos riscos da IA. No ponto, a magistrada destaca que:

Na essência, a possibilidade de ter um canal oficial, à disposição de usuários leigos em tecnologia, para reportar ao CNJ experiências negativas relacionadas ao uso da IA pelo Poder Judiciário, vai muito além de ser apenas uma medida de gerenciamento de riscos.

Significa estabelecer um mecanismo para que a sociedade brasileira seja ouvida e participe da transformação digital do Poder Judiciário do seu país, ajudando a legitimar as mudanças tão profundas e disruptivas que estão sendo feitas no âmbito da chamada Justiça 4.0.

Significa concretizar o comando normativo expresso no art. 7º, V, da Resolução nº 395/2021, pelo qual cabe ao LIODS/CNJ “abrir espaço para a participação cidadã na concepção de projetos inovadores no Poder Judiciário ou que contribuam para a efetividade da Agenda 2030”¹⁶⁵.

De fato, a criação de um banco de dados de monitoramento de incidentes de IA, de comunicação obrigatória para usuários internos (servidores e magistrados) e facultativa para usuários externos, contribui não só para o gerenciamento de riscos, representando fonte valiosa de informações, que devem ser devidamente tratadas e analisadas no âmbito do CNJ, mas também para a construção de confiança nas aplicações de IA no âmbito do Poder Judiciário.

Acrescente-se que a busca regular por *feedback* das partes interessadas (que podem ser direta ou indiretamente afetadas pelos sistemas de IA) é frequentemente apontada como importante prática para assegurar a participação no desenvolvimento de ferramentas confiáveis¹⁶⁶.

Ademais, colher informações por parte dos usuários externos permite o melhor monitoramento dos resultados e gerenciamento de riscos, contribuindo para o alcance dos objetivos estabelecidos e consolidação da governança da instituição.

Outrossim, especificamente no caso de instituições públicas, “deve-se lembrar que o *feedback* da sociedade civil permite melhorar procedimentos em direções não necessariamente previstas e rever democraticamente os processos que afetam a vida das pessoas” e que “transparência não é satisfeita apenas com a publicação de informações, mas também com mecanismos participativos, nos quais as pessoas podem expor suas preocupações diretamente à autoridade”¹⁶⁷.

¹⁶⁵ PRADO, Eunice Maria Batista. Banco de incidentes de inteligência artificial do judiciário. 2022. 262 f. Dissertação (Mestrado em Direito). **Escola Nacional de Formação e Aperfeiçoamento de Magistrados**: Brasília, 2022. p. 221.

¹⁶⁶ EUROPEAN COMMISSION. **The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment**. 16 july. 2020, <https://data.europa.eu/doi/10.2759/002360>. Acesso em: 11 jan. 2024.

¹⁶⁷ ARÁNGUIZ VILLAGRÁN, Matías. **Auditoria algorítmica para sistemas de tomada de decisão ou suporte à decisão**. Apr. 2022. Disponível em: <http://dx.doi.org/10.18235/0004154>. p. 13. Acesso em: 11 jan. 2024.

2.5.4. Certificação de boas práticas (art. 8º, V)

Um instrumento de governança de IA frequente apontado em documentos internacionais é a certificação para sistemas ou selo de aprovação, de modo a atestar que os modelos testados por terceiro independente estão de acordo com regras ou padrões estabelecidos.

Nesse sentido, a Recomendação da UNESCO estabelece que os Estados-membros são encorajados “a considerar formas brandas de governança, tais como um mecanismo de certificação para sistemas de IA e o reconhecimento mútuo de sua certificação”, instrumento esse que “pode incluir diferentes níveis de auditoria de sistemas, dados e adesão às diretrizes éticas e aos requisitos procedimentais considerando os aspectos éticos”¹⁶⁸.

O processo de certificação por instituições especializadas, sejam elas públicas ou privadas, de fato, é desenhado para gerar confiança em produtos, serviços e sistemas que passaram por uma avaliação que atesta o cumprimento de determinados padrões pré-estabelecidos em normas e regulamentos técnicos.

Nessa linha, o PL n. 2338/2023, inspirado no processo de certificação previsto no *AI ACT*, ao abordar a criação da autoridade competente pela implementação e fiscalização da lei indicou algumas de suas competências, dentre elas expedir normas para regulamentação, inclusive sobre “procedimento para certificação do desenvolvimento e utilização de sistemas de alto risco” (art. 32, parágrafo único, inciso VI)¹⁶⁹.

No Brasil, sistema similar existe no âmbito do INMETRO, que acredita (credencia) um organismo de certificação (entidade privada), reconhecendo-o competente para avaliar um produto, processo, serviço, sistemas de gestão ou ainda um profissional, com base em regras preestabelecidas.

No âmbito global, recentemente, em dezembro de 2023, a *International Organization for Standardization* (ISO), organização internacional independente e não governamental, que reúne especialistas para compartilhar conhecimentos e desenvolver padrões internacionais para produtos, serviços e sistemas, com o objetivo de garantir qualidade, segurança e eficiência, publicou a ISO/IEC 42001, que apresenta normas para implementação de um Sistema de Gestão

¹⁶⁸ UNESCO. **Recomendação sobre a Ética da Inteligência Artificial**. 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 23 nov. 2023.

¹⁶⁹ BRASIL. Senado Federal. **Projeto de Lei nº 2338, de 4 de maio de 2023**. Dispõe sobre o uso da Inteligência Artificial. Brasília: Senado Federal, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 17 jan. 2024.

de Inteligência Artificial (SIGA).

A certificação ISO, fornecida por um terceiro imparcial acreditado (credenciado), atesta que a entidade avaliada está seguindo as normas estabelecidas por entidade de reputação internacional promovendo confiança.

No caso da ISO/IEC 42001, a norma define requisitos para estabelecer, implementar, manter e melhorar continuamente um sistema de gerenciamento de IA com o objetivo de ajudar a organização a desenvolver, fornecer ou usar sistemas de IA de forma responsável na busca de seus objetivos e atender aos requisitos aplicáveis.

No âmbito do Judiciário brasileiro, o art. 8º, V, da Resolução CNJ n. 332/2020, aponta entre os elementos do conceito de transparência “apresentação dos mecanismos de auditoria e certificação de boas práticas”.

A previsão normativa é fortemente inspirada na Carta Ética da CEPEJ, que prevê que autoridades independentes ou especialistas poderiam ser encarregados de certificar e auditar os modelos de IA ou fornecer orientação prévia, bem como autoridades públicas poderiam conceder certificação, a ser revisada regularmente.

Assim, razoável concluir que, segundo a Resolução n. 332/2020, a certificação de boas práticas contribui para que seja atendido o critério de transparência, exigido de todas as ferramentas de IA utilizadas no Poder Judiciário brasileiro.

A certificação de boas práticas, por sua vez, exigiria que fosse elaborado um conjunto de regras de condutas e padrões para servir de base para o processo de certificação a ser realizado por terceiros a fim de atestar que o modelo de IA atende às diretrizes estabelecidas.

Nessa linha, desde 2020, a CEPEJ vem trabalhando para desenvolver um mecanismo de certificação de soluções de IA a fim de atestar a observância dos princípios propostos na sua Carta Ética¹⁷⁰.

No âmbito do CNJ, contudo, mais uma vez não há regulamentação acerca do tema. Não há sequer indicativo de trabalhos sendo desenvolvidos para elaborar orientações de desenvolvimento ou implementação de modelos de IA além daquelas já previstas na Resolução n. 332/2020 e Portaria n. 271/2020.

Em verdade, a utilização do termo “boa prática” no âmbito do Judiciário brasileiro possui conotação bem específica, definida expressamente na Portaria CNJ n. 140/2019 como “experiência, atividade, ação, caso de sucesso, projeto ou programa, cujos resultados sejam

¹⁷⁰ CEPEJ. **Possible introduction of a mechanism for certifying artificial intelligence tools and services in the sphere of justice and the judiciary**: Feasibility Study. 8 dec. 2020. Disponível em: <https://rm.coe.int/feasability-study-en-cepej-2020-15/1680a0adf4>. Acesso em: 22 jan. 2024.

notórios pela eficiência, eficácia e/ou efetividade e contribuam para o aprimoramento e/ou desenvolvimento de determinada tarefa, atividade ou procedimento no Poder Judiciário” (art. 4º, I).

Em 2019, foi instituído o Portal CNJ de Boas Práticas do Poder Judiciário com o intuito de disseminar iniciativas e projetos inovadores, possíveis de serem replicados, em busca da melhoria da gestão e da prestação jurisdicional.

Conforme previsão normativa, as boas práticas apenas serão publicadas no Portal após avaliação realizada pela equipe técnica do Departamento de Gestão Estratégica (DGE) e posterior aprovação pelo Plenário do CNJ. Na fase de avaliação, é verificado se foram atendidos os seguintes critérios: eficiência; qualidade; criatividade; exportabilidade; satisfação do usuário; alcance social; e desburocratização (art. 12).

Cabe destacar que, nos termos da Portaria n. 140/2019, a inclusão da prática no Portal de Boas Práticas “não confere atestado de regularidade ou certificação do CNJ sobre a gestão ou a conduta do(s) respectivo(s) dirigente(s) ou de seus servidores”, consistindo apenas em “um reconhecimento meramente de natureza técnica, a fim de disseminar práticas que visem o aprimoramento do Poder Judiciário, não se confundindo com a atuação correicional do CNJ” (art. 21).

No que toca a ferramentas de IA, foi incluído no Portal o SIGMA, sistema de IA, do Tribunal Regional Federal da 3ª Região (TRF3), que auxilia a produção de minutas de processos ordenando os modelos de atos judiciais por relevância, considerando informações extraídas de peças processuais selecionadas. O Projeto SIGMA – Ranqueamento de modelos de atos judiciais a partir de peças processuais também foi a iniciativa premiada em 2021 pelo Prêmio Innovare na categoria Tecnologia.

Para estimular esse tipo de prática inovadora, além do conhecido Prêmio Innovare, que reconhece e divulga ações transformadoras no sistema de Justiça, foi criado o Prêmio Justiça e Inovação com o objetivo de incentivar a pesquisa acadêmica e a inovação tecnológica capazes de contribuir para o aperfeiçoamento da prestação jurisdicional e a solução de desafios práticos enfrentados pelo Judiciário brasileiro.

Ressalta-se, contudo, a importância de adotar medidas concretas para implementar a certificação de boas práticas, nos moldes em que prevista na Resolução n. 332/2020, inspirada na Carta Ética da CEPEJ. Se devidamente aplicada essa ferramenta de governança garante uma avaliação criteriosa das práticas envolvendo sistemas de IA promovendo a transparência e a confiança necessárias para o uso ético e responsável.

2.5.5. Auditoria (art. 8º, V)

A previsão de implementação de auditorias nos sistemas de IA em utilização no Poder Judiciário está expressamente incluída no art. 8º, V, da Resolução n. 332/CNJ, enquanto elemento do conceito de transparência. Também há expressa referência à realização de auditoria nos arts. 12 e 13 da Portaria n. 271/2020, bem como indicação, na Portaria n. 338/2023, de que a proposta de regulamentação do uso de IA generativa no Judiciário disponha sobre auditoria de modelos e soluções de IA.

A propósito da auditabilidade de sistemas de IA, Hartmann aponta que “[q]ualquer envolvimento de um sistema autônomo na tomada de decisões judiciais deve fornecer uma explicação satisfatória auditável por uma autoridade humana competente, de modo a verificar se os compromissos de justiça estão sendo observados”¹⁷¹.

Não obstante a implementação da plataforma Sinapses como mecanismo divulgado como sendo apto a viabilizar a auditoria de modelos de IA, não se tem notícia de realização de auditoria interna ou externa nas ferramentas em utilização ou em desenvolvimento no Poder Judiciário brasileiro.

As auditorias e avaliações de risco ou de impacto (*assessments*) são instrumentos clássicos de governança, especificamente de gerenciamento de riscos, que frequentemente são incorporados na regulação estatal, mas que não se confundem.

É possível definir auditoria como “um processo sistemático, independente e documentado, com o qual se busca coletar e avaliar evidências para determinar o grau em que determinados critérios previamente determinados são atendidos”¹⁷².

Assim, resulta em uma avaliação objetiva que verifica se estão sendo observados procedimentos, padrões ou critérios previamente estabelecidos, permitindo o melhor gerenciamento de riscos e se mostrando essencial para a efetiva responsabilização pelos danos.

Existem auditorias de diversas naturezas como contábil/financeira, jurídica, operacional, de conformidade, entre outras. Com a multiplicação de sistemas de IA em diversas áreas, passou-se a falar cada vez mais em auditoria algorítmica com o objetivo de avaliar os riscos envolvidos na utilização dessas ferramentas tecnológicas nos contextos em que aplicadas.

Uma auditoria algorítmica, portanto, pretende avaliar um sistema automatizado de

¹⁷¹ PEIXOTO, Fabiano Hartmann. *Op. Cit.* p. 155.

¹⁷² ARÁNGUIZ VILLAGRÁN, Matías. **Auditoria algorítmica para sistemas de tomada de decisão ou suporte à decisão**. Apr. 2022. p. 5. Disponível em: <http://dx.doi.org/10.18235/0004154>. Acesso em: 11 jan. 2024.

tomada de decisão ou suporte à decisão e seu processo de desenvolvimento, “incluindo a formulação e os dados usados para treinar o sistema”, “em termos de precisão, justiça algorítmica, vieses, discriminação, privacidade e segurança, entre outros”¹⁷³.

Apesar da definição acima apresentada, a verdade é que as opiniões sobre o modo de realização de auditorias de sistemas automatizados, seu escopo e momento adequado para sua realização, assim como periodicidade, são tão diversas que não existe uma visão unânime, estando o tema em processo de discussão e evolução ao redor do mundo.

Os pontos objeto de exame em uma auditoria de sistemas automatizados dependem das finalidades da parte interessada, do contexto em que ele é utilizado, da importância da função desempenhada, entre outros. Um relatório de auditoria pode ser exigido por órgãos reguladores para avaliar o atendimento aos padrões legais, pelo desenvolvedor para melhorar os resultados e mitigar os riscos ou por terceiros para tomar decisões informadas sobre como se envolver com empresas específicas¹⁷⁴.

A realização de auditorias permite seja mapeado não apenas o processo de desenvolvimentos dos serviços e sistemas como também os mecanismos de controle/ estruturas de governança e como a implementação da tecnologia impacta os usuários, contribuindo para a tomada de decisão, antecipação de problemas e identificação de questões a serem resolvidas e/ou riscos a serem mitigados.

No âmbito do Poder Público, a realização de auditoria externa permite, ainda, o controle democrático, fornecendo informações essenciais para o cumprimento do ideal de transparência e publicidade. Por essas razões, a auditoria de modelos de IA é frequentemente apontada como uma das ferramentas de operacionalizar *accountability* (responsabilização e prestação de contas).

As auditorias, em regra, são realizadas após a implementação do sistema, de forma independente, e podem ser internas ou externas. A auditoria interna é aquela realizada no âmbito interno da organização, enquanto a externa é realizada por terceiros e, em geral, tem como destinatário os usuários afetados, órgãos reguladores ou a sociedade em geral. Em qualquer caso, é requisito indispensável que o responsável pela auditoria não tenha estado envolvido no desenvolvimento do sistema.

No momento, as práticas de auditoria de sistemas de IA estão em estágio embrionário,

¹⁷³ *Ibidem*, p. 6.

¹⁷⁴ BROWN, Shea; DAVIDOVIC, Jovana; HASAN, Ali. 2021. The algorithm audit: Scoring the algorithms that score us. **Big Data & Society**, 8(1). Disponível em: <https://doi.org/10.1177/2053951720983865>. Acesso em: 11 jan. 2024.

apesar da crescente implementação, ainda há carência de definições de escopo, de métricas abrangentes e robustas e de modelos que permitam uma padronização¹⁷⁵.

Recentemente, foi publicada a ISO/IEC 42001, a norma internacional que estabelece requisitos para avaliar sistemas de gerenciamento de IA com o objetivo de ajudar organizações a desenvolver, fornecer ou usar sistemas de IA de forma responsável.

Ao lado dos processos de auditorias, as avaliações de riscos ou de impacto (*risks or impact assessments*) também são muito populares como instrumento de governança e gerenciamento de riscos de sistemas de IA. Em geral, diferem-se das auditorias por envolver o exame de questões mais abrangentes como o impacto social da utilização da ferramenta.

Como apontam Gualtieri e López, as avaliações (*assessments*), “no Brasil, costumam ser traduzidas pela legislação por relatórios, como o Relatório de Impacto Ambiental e o Relatório de Impacto à Proteção de Dados Pessoais”, e “se, por um lado, implicam maior subjetividade e menor independência, por outro, permitem um maior diálogo e parceria, em um *loop de feedback* que permite melhoria contínua”¹⁷⁶.

Avaliações de impacto algorítmico (AIAs) envolvem o estudo do uso de um sistema algorítmico em um contexto específico buscando “entender, categorizar e responder melhor aos possíveis danos ou riscos apresentados pelo uso desses sistemas” (tradução livre)¹⁷⁷.

Nos termos da ISO/IEC 42001, avaliação do impacto do sistema de IA é o “processo formal e documentado pelo qual os impactos sobre indivíduos, grupos de indivíduos, ou ambos, e sociedades são identificados, avaliados e abordados por uma organização que desenvolve, fornece ou usa produtos ou serviços que utilizam inteligência artificial” (tradução livre)¹⁷⁸.

Nesse cenário, avaliações de impacto são instrumentos adequados para identificar, além do grau de acurácia do sistema, potenciais prejudicados por erros ou falhas que venham a ocorrer, investigar o contexto e formas de uso em que a ferramenta será empregada, o grau de

¹⁷⁵ LAM, Khoa et al. **A Framework for Assurance Audits of Algorithmic Systems**. 26 jan. 2024. Disponível em: <https://arxiv.org/abs/2401.14908>. Acesso em: 20 mar. 2024.

¹⁷⁶ GUALTIERI, André; LÓPEZ, Nuria. Decisões automatizadas: auditoria sobre aspectos discriminatórios do art. 20 da LGPD. In: PALHARES, Felipe; FRANCOSKI, Denise de Souza Luiz. Coord. Temas Atuais de Direito Digital. São Paulo: **Revista dos Tribunais**, 2023.

¹⁷⁷ No original: “*Algorithmic impact assessments (AIAs) are an emergent policy mechanism being utilised by public agencies, involving studying the potential use of an algorithmic system in context, and seeking to better understand, categorise and respond to the potential harms or risks posed by the use of these systems*”. ADA LOVELACE INSTITUTE; AI NOW INSTITUTE; OPEN GOVERNMENT PARTNERSHIP. [2021]. **Algorithmic Accountability for the Public Sector**. Disponível em: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>. Acesso em: 13 ago. 2023

¹⁷⁸ No original: “*AI system impact assessment: formal, documented process by which the impacts on individuals, groups of individuals, or both, and societies are identified, evaluated and addressed by an organization developing, providing or using products or services utilizing artificial intelligence*”. *Ibidem*.

supervisão humana, entre outros pontos importantes, configurando, portanto, elemento essencial nas organizações que se propõem a utilização de IA de maneira ética e responsável.

Trata-se de ferramenta frequentemente apontada como mecanismo de governança destinado a mapear os riscos ou danos gerados por um sistema e recebeu papel de destaque nas projeções regulatórias americana e europeia, assim como no PL 2338/2023, que dedicou uma seção inteira para o seu tratamento (arts. 22 a 26).

A propósito do tema, a Recomendação da UNESCO sobre a Ética da Inteligência Artificial estabelece que “[o]s Estados-membros devem criar marcos para a realização de avaliações de impacto, como avaliação de impacto ético, para identificar e avaliar os benefícios, as preocupações e os riscos dos sistemas de IA, bem como medidas adequadas de prevenção, mitigação e monitoramento de riscos, entre outros mecanismos de garantia”¹⁷⁹.

O referido documento ainda orienta que a avaliação de impacto ético para identificar e avaliar os benefícios, as preocupações e os riscos dos sistemas de IA “também deve estabelecer mecanismos de supervisão apropriados, incluindo auditabilidade, rastreabilidade e explicabilidade, que permitam a avaliação de algoritmos, dados e processos de planejamento, bem como incluir a revisão externa dos sistemas de IA” e que “[e]ssas avaliações também devem ser multidisciplinares, multiparceiros, multiculturais, pluralistas e inclusivas”¹⁸⁰.

No que toca às organizações públicas, a Recomendação sobre a Ética da IA determina, ainda, que “[a]s autoridades públicas devem ser obrigadas a monitorar os sistemas de IA implementados e/ou utilizados por essas autoridades, introduzindo mecanismos e ferramentas adequados”¹⁸¹.

A propósito do tema, o governo do Canadá desenvolveu um questionário online, o *Algorithm Impact Assessment* (AIA), a fim de auxiliar a avaliação e a mitigação de riscos associados ao desenvolvimento e implantação de sistemas de decisões automatizadas¹⁸².

No mesmo sentido, o governo americano também disponibilizou uma versão ainda em teste e desenvolvimento de questionário elaborado com o mesmo objetivo e que recebeu idêntica denominação¹⁸³.

¹⁷⁹ UNESCO. **Recomendação sobre a Ética da Inteligência Artificial**. 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 23 nov. 2023.

¹⁸⁰ *Ibidem*.

¹⁸¹ *Ibidem*.

¹⁸² GOVERNMENT OF CANADA. *Algorithmic Impact Assessment tool*. 25 apr. 2023. Disponível em: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html>. Acesso em: 17 jan. 2024.

¹⁸³ ALGORITHMIC IMPACT ASSESSMENT. Disponível em: <https://www.cio.gov/aia-eia-js/#/>. Acesso em: 17 jan. 2024.

A avaliação ou relatório de impacto pode ser considerada instrumento consideravelmente difundido entre os documentos que estabelecem diretrizes para o funcionamento de sistemas de IA seguros e responsáveis.

Em pesquisa realizada com 106 (cento e seis) documentos de princípios éticos de IA, observou-se que a avaliação de impacto é medida recomendada em 69,4% (sessenta e nove vírgula quatro por cento) das publicações, indicando que “a ferramenta de avaliação do impacto é essencial para reduzir os riscos associados a uma utilização incorreta da IA ou a um mau funcionamento, o que também é válido em contextos extremamente sensíveis como o Judiciário” (tradução livre)¹⁸⁴.

A proposta de regulação do PL 2338/2023 determina a obrigação de realizar avaliação de impacto algorítmico sempre que o sistema for considerado como de alto risco¹⁸⁵.

O projeto de lei estabelece ainda que a AIA deve ser sempre realizada por profissional ou equipe com independência funcional, podendo haver casos, conforme posterior regulamentação da autoridade competente, em que há necessidade de o responsável pela avaliação ser externo ao fornecedor.

Houve a preocupação em apontar, precisamente, elementos mínimos que devem constar da avaliação de impacto, quais sejam:

- a) **riscos conhecidos e previsíveis associados** ao sistema de inteligência artificial à época em que foi desenvolvido, **bem como os riscos que podem razoavelmente dele se esperar;**
 - b) benefícios associados ao sistema de inteligência artificial;
 - c) **probabilidade de consequências adversas**, incluindo o número de pessoas potencialmente impactadas;
 - d) **gravidade das consequências adversas, incluindo o esforço necessário para mitigá-las;**
 - e) lógica de funcionamento do sistema de inteligência artificial;
 - f) processo e resultado de testes e avaliações e medidas de mitigação realizadas para verificação de possíveis impactos a direitos, com especial destaque para **potenciais impactos discriminatórios;**
 - g) **treinamento e ações de conscientização dos riscos associados ao sistema de inteligência artificial;**
 - h) medidas de mitigação e indicação e justificação do risco residual do sistema de inteligência artificial, acompanhado de testes de controle de qualidade frequentes;
 - i) **medidas de transparência ao público**, especialmente aos potenciais usuários do sistema, a respeito dos riscos residuais, principalmente quando envolver alto grau de nocividade ou periculosidade à saúde ou segurança dos usuários, nos termos dos artigos 9º e 10 da Lei nº 8.078, de 11 de setembro de 1990 (Código de Defesa do Consumidor);
- § 2º Em atenção ao princípio da precaução, quando da utilização de sistemas de

¹⁸⁴ LUPO, Giampiero. *Op. cit.* p. 628.

¹⁸⁵ BRASIL. Senado Federal. **Projeto de Lei nº 2338, de 4 de maio de 2023**. Dispõe sobre o uso da Inteligência Artificial. Brasília: Senado Federal, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 17 jan. 2024.

inteligência artificial que possam gerar impactos irreversíveis ou de difícil reversão, **a avaliação de impacto algorítmico levará em consideração também as evidências incipientes, incompletas ou especulativas.**

§ 3º A autoridade competente poderá estabelecer outros critérios e elementos para a elaboração de avaliação de impacto, incluindo a participação dos diferentes segmentos sociais afetados, conforme risco e porte econômico da organização¹⁸⁶.

(sem destaques no original)

Outro ponto interessante que cabe destacar na proposta de regulamentação do PL 2338/2023 é a previsão de que a AIA “consistirá em processo iterativo contínuo, executado ao longo de todo o ciclo de vida dos sistemas de inteligência artificial de alto risco, requeridas atualizações periódicas” (art. 25), bem como que “[a] atualização da avaliação de impacto algorítmico contará também com participação pública, a partir de procedimento de consulta a partes interessadas, ainda que de maneira simplificada” (art. 25, §2º).

Enquanto as auditorias, em regra, são realizadas apenas após a implementação do sistema, as avaliações de impacto podem e devem ser aplicadas antes, a fim de que as consequências envolvidas possam ser previstas e, notadamente, mitigados os riscos ao máximo, mas também durante todo o ciclo de vida, com atualização de forma periódica, de modo que sejam monitoradas as consequências advindas da utilização de sistemas de IA.

Apresentadas as diferenças entre as auditorias e as avaliações de impacto, verifica-se que essas últimas são mais abrangentes, buscando prever consequências, mitigar riscos e abordar implicações sociais, entre outras.

Nada obstante, as regulamentações do uso da IA no Judiciário brasileiro apenas tratam de mencionar a realização de auditorias, sem sequer discorrer sobre escopo, órgão responsável, periodicidade, critérios a serem observados e nada apontam sobre a necessidade de realização de avaliações de impacto.

O art. 8º da Resolução CNJ n. 332/2020 ao tratar da transparência, elenca como essencial a “documentação dos riscos identificados e indicação dos instrumentos de segurança da informação e controle para seu enfrentamento”, sem sequer apontar a necessidade de divulgação de tais informações.

Nota-se a limitação da previsão normativa que apenas aponta a necessidade de indicação de instrumento de “segurança da informação”, como se o único risco envolvido na implementação de ferramentas de IA envolvesse a proteção de dados ou cibersegurança, ignorando em absoluto os impactos sociais de utilização da referida tecnologia no sistema de

¹⁸⁶ BRASIL. Senado Federal. **Projeto de Lei nº 2338, de 4 de maio de 2023**. Dispõe sobre o uso da Inteligência Artificial. Brasília: Senado Federal, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 17 jan. 2024.

justiça.

Destaca-se, ainda, que mesmo que o ato normativo não tenha pretendido dar publicidade à referida documentação, também não se tem notícia da produção de tal documentação e apresentação ao CNJ. Conforme apontado em tópico anterior, mais uma vez falta transparência quanto à política de transparência adotada no âmbito do Poder Judiciário.

De fato, a realização de auditoria nos modelos, nos termos em que apontado na Resolução n. 332/2020, demanda maior regulamentação por parte do CNJ, esclarecendo, ao menos, pontos essenciais como: responsável pela sua realização; periodicidade; pontos a serem avaliados, bem como as métricas de avaliação.

Parece adequado também haja previsão de dever de realização de AIA no âmbito dos projetos em desenvolvimento do Judiciário a fim de que sejam mapeados os potenciais efeitos adversos do uso da ferramenta, inclusive no que toca à proteção e ao gozo de direitos fundamentais e garantias processuais, assim como dos valores essenciais ao sistema de justiça.

No que toca ao uso da IA generativa, tratando especificamente do uso de grandes modelos de linguagem (LLM), como o GPT-3 e similares, o Instituto Ada Lovelace aponta a necessidade de realização de auditoria por terceiros independentes em todos os modelos utilizados pelo Poder Público, seja em caso de soluções desenvolvidas internamente seja em caso de utilização de ferramentas desenvolvidas pela iniciativa privada¹⁸⁷.

De fato, auditorias e avaliações de impacto são ferramentas importantes para implementar na prática o princípio da responsabilização e prestação de contas, permitindo uma melhor gestão de riscos pelas organizações assim como o controle democrático pela sociedade. Conforme foi apontado, contudo, o tema no que toca à sistemas de IA ainda está em desenvolvimento, havendo muita discussão e pouco consenso acerca da matéria. Nada obstante, em uma área sensível como o Poder Judiciário, deve-se seguir uma postura de cautela de modo a implementar medidas que viabilizem a avaliação das ferramentas de IA e riscos associados, ainda que eventualmente possam vir a ser melhoradas ao longo do processo de implementação.

Ademais, na hipótese, para fins de gestão de riscos no âmbito do sistema de justiça brasileiro, dado o contexto atual, parece mais adequada, antes da regulamentação e implementação do processo de auditorias, o efetivo implemento de avaliações de impacto focadas na identificação de riscos conhecidos e previsíveis, probabilidade de consequências adversas e verificação de possíveis impactos a direitos.

¹⁸⁷ DAVIES, Matt; JONES, Elliot. **Foundation models in the public sector**. Ada Lovelace Institute. 12 oct. 2023. Disponível em: <https://www.adalovelaceinstitute.org/evidence-review/foundation-models-public-sector/>. Acesso em: 17 nov. 2023.

3. Comitês de ética de IA e sua implementação no Poder Judiciário brasileiro

3.1. Comitê de ética de IA enquanto mecanismo de governança

Conforme analisado no capítulo 1, governança corresponde a um caminho para concretização dos princípios éticos dentro das organizações por meio de políticas, práticas e processos para direcionar, monitorar e avaliar os resultados e riscos envolvidos.

Nesse contexto, para implementar as diretrizes éticas na prática faz-se necessária uma estrutura formada por mecanismos e processos devidamente estabelecidos.

Dentre as ações de governança de IA, a nível estratégico, está a criação de um colegiado voltado para apoiar a implementação de políticas e diretrizes por meio da identificação e mitigação dos riscos envolvidos no uso e desenvolvimento de ferramentas de IA, como um conselho ou comitê de ética¹⁸⁸.

Um modelo de supervisão baseado em comitês já tem sido utilizado em diversas áreas, em especial nas pesquisas envolvendo seres humanos. No Brasil, a Resolução do Conselho Nacional de Saúde nº 196/96 determinou que toda pesquisa envolvendo seres humanos deve ser submetida à apreciação de um Comitê de Ética em Pesquisa, colegiado de natureza interdisciplinar e independente, cujas atividades inclui identificar conflitos de valores, refletir criticamente sobre os dilemas éticos e contribuir para que a pesquisa seja realizada dentro de padrões éticos, tendo como fundamento básico a proteção da dignidade do ser humano.

A ideia da criação de comitês de ética ou órgão semelhante está em reunir pessoas com variadas habilidades para analisar, avaliar e responder problemas complexos. Esse tipo de estrutura, dentro de um contexto de governança no âmbito interno das organizações, possibilita que sejam protegidos os valores chave da instituição, permite identificar problemas, desenvolver diretrizes e processos a serem observados, bem como ajustar os projetos em andamento considerando os seus potenciais impactos, não apenas internamente, mas na

¹⁸⁸ “Em um nível mais elevado, a função do comitê é simples: identificar de forma sistemática e abrangente e ajudar a mitigar os riscos éticos dos produtos de IA que são desenvolvidos internamente ou adquiridos de terceiros” (tradução livre). No original: “*At a high level the function of the committee is simple: to systematically and comprehensively identify and help mitigate the ethical risks of AI products that are developed in-house or purchased from their party vendors*”. BLACKMAN, Reid. **Why You Need an AI Ethics Committee**, Harvard Business Review, July-Aug. 2022. Disponível em: <https://hbr.org/2022/07/why-you-need-an-ai-ethics-committee>. Acesso em: 17 nov. 2023. Também tratando do tema, entre outros: MORLEY, Jessica, et al. Ethics as a service: A pragmatic operationalization of AI ethics. **Minds and Machines**, v. 31, n. 2, p. 239-256, 2021. Disponível em: <https://doi.org/10.1007/s11023-021-09563-w>. Acesso em: 17 nov. 2023.; TIELL, Steven. **Create an ethics committee to keep your AI initiative in check**. 15 nov. 2019. Disponível em: <https://hbr.org/2019/11/create-an-ethics-committee-to-keep-your-ai-initiative-in-check>. Acesso em: 17 nov. 2023.

sociedade como um todo¹⁸⁹.

Esse comitê, conselho ou comissão também pode promover transparência e *accountability*, assim como exercer importante papel de comunicação entre as partes interessadas, em especial atores externos, possibilitando a concretização do princípio da participação, um dos princípios da governança para o setor público, que estimula a promoção de diálogo com a sociedade, de modo que diferentes perspectivas sejam consideradas na tomada de decisão¹⁹⁰.

Trata-se de colegiado que pode desempenhar relevante papel em um programa de conformidade, contribuindo para que as atividades da organização observem as leis e demais normas aplicáveis, fornecendo orientações para a cúpula da instituição, assim como promovendo uma cultura organizacional que valorize a comportamentos éticos e responsáveis.

Nesse contexto, incentivar que as organizações criem conselhos de revisão ou comitês de ética está entre as ações estratégicas do eixo 2 (Governança de IA) da Estratégia Brasileira de Inteligência Artificial (EBIA)¹⁹¹.

O documento elaborado pelo Grupo de Especialistas de Alto Nível em Inteligência Artificial do Conselho da Europa, “*Assessment List for Trustworthy AI*” – ALTAI, que consolida questionamentos concretos como forma de autoavaliação das diretrizes éticas estabelecidas, aponta, ainda que de maneira indireta, a instalação de um conselho de revisão ética independente (ou mecanismo semelhante) para discutir responsabilidade e questões éticas como ferramenta de gerenciamento de risco e *accountability*¹⁹².

O Fórum Econômico Mundial, na publicação *Empowering AI Leadership* (2022), também indica a criação de conselhos ou comitês, externos e internos, como componente

¹⁸⁹ “As organizações que optam por estabelecer um comitê de ética não estão simplesmente lidando com o viés humano, viés de dados e equidade (*fairness*); o escopo de suas preocupações é geralmente muito mais amplo e frequentemente inclui o desejo de aumentar a maturidade organizacional em relação a como seus produtos e serviços afetam as partes interessadas, incluindo as sociedades civis” (tradução livre). No original: “*Organizations that choose to establish an ethics committee aren’t simply addressing human bias, statistical bias, and fairness; their scope of concerns is typically much wider, and often includes a desire to increase organizational maturity concerning how their products and services impact stakeholders, including civil societies*”. TIELL, Steven. **Create an ethics committee to keep your AI initiative in check**. 15 nov. 2019. Disponível em: <https://hbr.org/2019/11/create-an-ethics-committee-to-keep-your-ai-initiative-in-check>. Acesso em: 17 nov. 2023.

¹⁹⁰ TCU. **Referencial básico de governança aplicável a organizações públicas e outros entes jurisdicionados ao TCU**. 3 ed. Brasília, 2020. p. 46. Disponível em: <https://portal.tcu.gov.br/imprensa/noticias/tcu-publica-a-3-edicao-do-referencial-basico-de-governanca-organizacional.htm> Acesso em: 17 nov. 2023.

¹⁹¹ BRASIL. Ministério da Ciência, Tecnologia e Inovações. **Estratégia Brasileira de Inteligência Artificial**. 6 abr. 2021. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligencia-artificial-estrategia-repositorio>. Acesso em: 17 nov. 2023.

¹⁹² No original: “*Did you consider establishing an AI ethics review board or a similar mechanism to discuss the overall accountability and ethics practices, including potential unclear grey areas?*”. ALA-PIETILÄ, Pekka et al. **The assessment list for trustworthy artificial intelligence (ALTAI)**. European Commission, 2020. 2020. Doi:10.2759/002360. Acesso em: 17 nov. 2023.

essencial para operacionalizar a governança de IA dentro de uma organização, tendo como papel, inclusive, desenvolver ferramentas, práticas e treinamento para tratar de questões éticas¹⁹³.

Também a proposta de modelo de governança de IA elaborada por instituições governamentais de Singapura aponta a sugestão de criação de um conselho de revisão ética para tratar de considerações éticas sobre o uso da tecnologia nas organizações¹⁹⁴.

O documento elaborado pelo Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE), renomada organização profissional americana dedicada ao avanço da tecnologia em benefício da humanidade, *Ethically Aligned Design* (EAD), também recomenda a criação de conselho de revisão, formado por especialistas imparciais com diversidade de conhecimentos e experiências, como um recurso importante para avaliar os riscos de projetos e seu progresso, bem como para assegurar *accountability*, e sugere que, em caso de órgãos governamentais cujo uso de IA pode impactar o público geral, a composição do conselho deve incluir a participação de diversos grupos de cidadãos a fim de garantir legitimidade¹⁹⁵.

Nesse sentido de estabelecer um órgão específico para garantir a adoção de um modelo de governança apropriado, a estratégia de IA do Tribunal de Justiça da União Europeia (TJUE) prevê a criação de um conselho de gestão de IA (“*AI management board*”), que tem a missão de garantir que o desenvolvimento ou aquisição de ferramentas de IA respeitem os princípios éticos estabelecidos¹⁹⁶.

O referido conselho, que ainda não foi formalmente criado, terá, dentre suas atribuições, o papel de elaborar uma carta de ética, que será usada como base de avaliação na tomada de decisão sobre o desenvolvimento e aquisição de qualquer ferramenta de IA, bem como deverá definir as áreas em que a tecnologia não deve ser utilizada em razão do grau de risco, as áreas de atenção especial e as áreas de baixo risco¹⁹⁷.

Estudo detalhado da FGV sobre o tema compilou os benefícios de instalar um ambiente colegiado específico, como um conselho ou comitê, para abordar a preocupação com questões éticas dentro das organizações, entre eles aponta-se: a) mitigação de danos por meio de revisão

¹⁹³ WORLD ECONOMIC FORUM. Empowering AI Leadership: AI C-Suite Toolkit. 2022. Disponível em: <https://www.weforum.org/publications/empowering-ai-leadership-ai-c-suite-toolkit/>. Acesso em: 17 nov. 2023.

¹⁹⁴ AI VERIFY FOUNDATION; Infocomm Media Development Authority of Singapore (IMDA). Model Artificial Intelligence Governance Framework. Second Edition. 21 jan. 2020. Disponível em: <http://go.gov.sg/ai-gov-mf-2>. Acesso em: 17 nov. 2023.

¹⁹⁵ IEE. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. 14 march. 2022. Disponível em: http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html. Acesso em: 17 nov. 2023.

¹⁹⁶ COURT OF JUSTICE OF THE EUROPEAN UNION. **Artificial Intelligence Strategy**. 2024. Disponível em: https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf. Acesso em: 24 jan. 2024.

¹⁹⁷ *Ibidem*.

das etapas de validação dos desenvolvimento de modelos de IA, bem como verificação de normas de privacidade e proteção de dados; b) intermediar a comunicação entre os diferentes departamentos da organização e/ou outras partes interessadas, encarregando-se do *feedback* sobre os projetos de IA; c) apoiar o desenvolvimento de uma política interna com diretrizes éticas e processos a seres observados, bem como outros mecanismos de monitoramento e avaliação dos projetos de IA, inclusive em caso de aquisições de sistemas por terceiros¹⁹⁸.

Em geral, esse conselho ou comitê, que recebe diferentes nomenclaturas, é instalado dentro da estrutura da organização, podendo também ser externo e independente, formado com uma abordagem multidisciplinar, e é responsável por apoiar o desenvolvimento de uma cultura de uso da IA de maneira ética e responsável.

É muito importante assegurar diversidade na composição do colegiado para abordar a variedade de questões que a tecnologia pode apresentar em diferentes domínios, incorporando indivíduos com formação técnica e não técnica, bem como assegurando a inclusão de perspectivas de indivíduos que serão impactados.

A efetividade do comitê de ética, no entanto, é bastante variada, dependendo de como foi projetado, da finalidade para o qual foi criado e os objetivos almejados, da sua composição, estrutura, atribuições, poderes e papéis desempenhados.

Nesse contexto, apesar das múltiplas funções que um comitê de ética de IA pode desenvolver dentro da organização, a decisão de instalação de tal organismo deve ser acompanhada da definição precisa de suas responsabilidades (objetivos, direitos e deveres) e processos envolvidos, bem como efetiva implementação prática do seu papel sob pena de caracterizar prática de *ethics washing*, isto é, uma percepção enganosa de que existe preocupação com questões éticas sem que se reflita em ações práticas.

Com efeito, para que comitês de ética possam desempenhar seu papel de examinar todo o espectro de danos, riscos e impactos que podem surgir da implementação de sistemas de IA é necessário que estejam bem estruturados.

É possível que no caminho para a concretização de seus objetivos os referidos colegiados enfrentem alguns desafios, entre eles: falta de recursos, expertise e treinamento necessários para abordar adequadamente os impactos; dificuldade na aplicação de princípios éticos nas pesquisas envolvendo IA; e falta de transparência em relação a seus processos e/ou

¹⁹⁸ COELHO, Alexandre Zavaglia et al. **Governança da inteligência artificial em organizações: framework para comitês de ética em IA: versão 1.0**. CEPI FGV Direito SP, 2023. Disponível em: <https://repositorio.fgv.br/items/6c919903-ff09-4749-9d7e-0d3c75bc58a2>. Acesso em: 17 nov. 2023.

decisões¹⁹⁹.

Experiências anteriores demonstram o potencial da implementação de comitês de éticas nas organizações, mas também os desafios envolvidos.

Um caso interessante de sucesso seguido de fracasso é do Conselho de Ética de IA e Tecnologia de Policiamento (*AI and Policing Technology Ethics Board*) da Axon Enterprise, uma empresa sediada nos Estados Unidos, que desenvolve tecnologia e produtos de armas para militares, policiais e civis, incluindo câmeras corporais vendidas para órgãos de segurança pública.

O referido Conselho de Ética foi criado, em 2018, como um órgão consultivo externo, com alto grau de independência, composto por onze membros externos, especialistas em diversas áreas, incluindo inteligência artificial, ciência da computação, privacidade, aplicação da lei, liberdades civis e políticas públicas, com o objetivo de aconselhar a empresa sobre o desenvolvimento de produtos e serviços considerando o seu impacto na sociedade²⁰⁰.

Em 2019, o Conselho emitiu seu primeiro relatório, apontando considerações sobre o seu papel e suas atividades, apresentando um *framework* para avaliação de produtos da Axon e oferecendo recomendações sobre as tecnologias de reconhecimento facial, que conluam no sentido dessas ferramentas não serem suficientemente confiáveis ou eticamente justificáveis para serem comercializadas²⁰¹. Com base nas recomendações do Conselho, a empresa decidiu não oferecer serviços de reconhecimento facial nas câmeras corporais por ela comercializadas²⁰².

Três anos depois, em 2022, o Conselho de Ética de Tecnologias de IA e Policiamento da Axon foi efetivamente descontinuado após a renúncia de nove de seus membros em decorrência do anúncio público da empresa de que começaria a desenvolver drones equipados com Tasers (uma arma de choque elétrico, que interfere nos sinais nervosos, causando uma perda temporária de controle muscular) para posicioná-los em diversas escolas e locais públicos

¹⁹⁹ PETERMANN, M. et al. **Looking before we leap: Expanding ethical review processes for AI and data science research**. 2022. 13 dec. 2022. Disponível em: <https://www.adalovelaceinstitute.org/report/looking-before-we-leap/>. Acesso em: 8 jan. 2024.

²⁰⁰ AXON. **First Report of the Axon AI & Policing Technology Ethics Board**. June. 2019. Disponível em: https://static1.squarespace.com/static/58a33e881b631bc60d4f8b31/t/5d13d7e1990c4f00014c0aeb/1561581540954/Axon_Ethics_Board_First_Report.pdf. Acesso em: 8 jan. 2024.

²⁰¹ *Ibidem*.

²⁰² Em pronunciamento público, o CEO da empresa apontou que seguiria a recomendação do Conselho de Ética no momento e agradeceu aos seus membros pelo trabalho desempenhado, bem como apontou que a equipe de IA continuaria avaliando o estado da tecnologia, mantendo o conselho informado. SMITH, Rick. *The future of face matching at axon and AI ethics board report*. Axon, June, v. 27, 2019. Disponível em: <https://www.axon.com/company/news/ai-ethics-board-report>. Acesso em: 8 jan. 2024.

(alvos potenciais de tiroteios em massa), a serem ativados em resposta à vigilância persistente alimentada por IA com parte de uma estratégia de longo prazo contra tiroteios em massa²⁰³.

O referido anúncio, aparentemente motivado pelo trágico tiroteio em uma escola em Uvalde, no Texas, ocorreu menos de três semanas após o Conselho de Ética expressamente recomendar que a Axon não avançasse com um programa piloto restrito destinado a avaliar a ideia de drones equipados com Tasers. Os ex-membros do colegiado que renunciaram apresentaram uma nota pública apontando que o anúncio da Axon indica que, após vários anos de trabalho, a empresa falhou em abraçar os valores que tentaram transmitir e que perderam a confiança na sua capacidade de ser um parceiro responsável²⁰⁴.

Após a renúncia dos membros do Conselho, a empresa divulgou nota anunciando que “em função do feedback recebido, estamos pausando o trabalho neste projeto e direcionando nossos esforços para engajar mais profundamente com as partes interessadas, a fim de explorar completamente o melhor caminho a seguir” (tradução livre), bem como lamentando que “alguns membros do painel consultivo de ética da Axon tenham optado por se retirar do envolvimento direto nessas questões antes de ouvirmos ou termos a oportunidade de responder às suas perguntas técnicas” (tradução livre)²⁰⁵.

Em janeiro de 2023, os nove ex-membros do Conselho de Ética em IA da Axon, em colaboração com o *The Policing Project* da Faculdade de Direito da Universidade de Nova York, divulgaram um relatório detalhando suas preocupações com o desenvolvimento da tecnologia de drones equipados com Tasers pela empresa²⁰⁶.

Em novembro de 2022, pouco tempo após a compra do Twitter, Elon Musk anunciou a

²⁰³ OLUMHENSE, Ese. Axon's Ethics Board Resigned Over Taser-Armed Drones. Then the Company Bought a Military Drone Maker: The CEO's vision for Taser-equipped drones includes a fictitious scenario in which the technology averts a shooting at a day care center. *Wired*. 8 sep. 2023. Disponível em: <https://www.wired.com/story/axon-taser-drones-ethics-board/>. Acesso em: 8 jan. 2024.

²⁰⁴ POLICING PROJECT. Statement of resigning axon ai ethics board members. *Policing project*. 6 jun. 2022. Disponível em: <https://www.policingproject.org/statement-of-resigning-axon-ai-ethics-board-members>; Acesso em: 8 jan. 2024.

Policing project, former ethics board members, release report on axon enterprise's proposal for taser-equipped drones; express concerns over potential misuse. *Policing project*. 24 jan. 2023. Disponível em: <https://www.policingproject.org/report-on-axon-enterprises-proposal>. Acesso em: 8 jan. 2024.

A Report on Axon's Proposal for Taser-Equipped Drones. Jan. 2023. Disponível em: <https://static1.squarespace.com/static/58a33e881b631bc60d4f8b31/t/63c9e9c86aa1c92a6d0d276a/1674176981131/ION+REPORT.pdf>. Acesso em: 8 jan. 2024.

²⁰⁵ SMITH, Rick. Axon committed to listening and learning so that we can fulfill our mission to protect life, together. 5 jun. 2022. Disponível em: <https://www.axon.com/news/technology/axon-committed-to-listening-and-learning>. Acesso em: 8 jan. 2024.

²⁰⁶ POLICING PROJECT. A Report on Axon's Proposal for Taser-Equipped Drones. *Policing project*. jan. 2023. Disponível em: <https://static1.squarespace.com/static/58a33e881b631bc60d4f8b31/t/63c9e9c86aa1c92a6d0d276a/1674176981131/ION+REPORT.pdf>. Acesso em: 8 jan. 2024.

demissão de milhares de funcionários, dentre eles pesquisadores que estavam trabalhando para tornar os algoritmos da empresa mais transparentes e justos²⁰⁷. Essa se tornou uma conduta popular entre as grandes empresas de tecnologia como Microsoft, Meta, Google e Amazon, que também diminuíram suas equipes encarregadas de assistir as companhias sobre o uso de IA de forma ética e responsável²⁰⁸.

O caso do conselho de ética da Axon evidencia como essas estruturas são relativamente frágeis, mas possuem um potencial considerável para influenciar positivamente o desenvolvimento de projetos de IA destacando questões éticas emergentes, propondo políticas e práticas responsáveis e, especialmente, garantindo que as decisões sejam informadas por uma variedade de perspectivas. Embora enfrentem desafios, seu verdadeiro impacto pode ser percebido quando há um compromisso genuíno das partes envolvidas em promover o desenvolvimento de tecnologias de IA de modo ético e responsável.

3.2. Análise de experiências implementadas no Poder Judiciário brasileiro

No campo do Judiciário brasileiro, estrutura semelhante a um comitê de revisão foi implementada no Tribunal Regional Federal da 3ª Região (TRF 3ª Região) para o propósito específico de realizar análise de conformidade das soluções de IA.

O Grupo de Validação Ético-Jurídica de Soluções de Inteligência Artificial (GVEJ) do Laboratório de Inteligência Artificial Aplicada do TRF 3ª Região (LIAA-3R) foi criado com a finalidade de analisar as soluções elaboradas no âmbito do laboratório a partir de critérios ético-jurídicos a fim de atender exigência prevista no item 4-c do Anexo à Portaria CNJ nº 25/2019, então vigente²⁰⁹.

Assim, com o objetivo de garantir conformidade com as regras em vigor e assegurar a auditabilidade das soluções de IA, foram destacados alguns integrantes do LIAA-3R para, além de realizar a atividade de avaliação das ferramentas, definir as diretrizes e dar as orientações

²⁰⁷ KNIGHT, Will. Elon Musk Has Fired Twitter's 'EthicalAI' Team. 4 nov. 2022. Disponível em: <https://www.wired.com/story/twitter-ethical-ai-team/>. Acesso em: 12 fev. 2024.

²⁰⁸ CRIDDLE, Cristina; MURGIA, Madhumita. Inteligência artificial: Microsoft, Amazon e Google demitem equipes de ética. Tradução de Sabino Ahumada. 3 abr. 2023. Disponível em: <https://valor.globo.com/brasil/noticia/2023/04/03/inteligencia-artificial-microsoft-amazon-e-google-demitem-equipes-de-etica.ghtml>. Acesso em: 12 fev. 2024.

²⁰⁹ LABORATÓRIO DE INTELIGÊNCIA ARTIFICIAL APLICADA DA 3ª REGIÃO (LIAA-3R). **Diretrizes de auditabilidade e conformidade no desenvolvimento e testes de soluções de IA no âmbito do LIAA-3R**. 2. ed., rev. e atual. São Paulo: LIAA-3R, 2022. Disponível em: https://www.trf3.jus.br/documentos/adeq/Inova/LIAA-3R/MANUAL/DIRETRIZES_DE_AUDITABILIDADE_2_ED_FICHA_ATUALIZADA.pdf. Acesso em: 12 fev. 2024.

necessárias para que a equipe de desenvolvimento observasse os preceitos que seriam depois analisados na etapa da validação ético-jurídica. Com esse propósito foi publicado o “Manual de Diretrizes de Auditabilidade e Conformidade no Desenvolvimento e Testes de Soluções de IA no âmbito do LIAA-3R”²¹⁰.

Segundo o referido documento, o GVEJ tem o dever de “zelar para que o escopo e a finalidade dos projetos fiquem devidamente documentados, a fim de que os eventuais usuários das soluções de IA possam compreender com clareza os casos de uso recomendados e os casos de uso não recomendados ou não abrangidos pelo projeto”²¹¹.

O referido grupo, contudo, não possui atribuição de “conjecturar os riscos ou os problemas que poderiam resultar de eventuais desvios no uso das soluções de IA em relação à finalidade para as quais foram concebidas”²¹².

Além do manual de diretrizes, o GVEJ também elaborou o parecer de validação dos sistemas SIGMA e da SINARA, realizando a verificação da compilação da documentação pertinente e a análise de conformidade, incluindo exame quanto ao respeito aos direitos fundamentais; não discriminação; publicidade e transparência; governança, qualidade e segurança; controle do usuário; pesquisa, desenvolvimento e implantação e prestação de contas e responsabilização.

Destaca-se que embora pretenda realizar uma validação denominada “ético-jurídica”, o escopo do GVEJ é tão somente verificar a conformidade com as regras em vigor, sem que sejam examinadas questões éticas e possíveis impactos sociais envolvidos, especialmente após a efetiva implementação prática das ferramentas examinadas.

Com efeito, o Manual de diretrizes de auditabilidade elaborado pelo GVEJ aponta, expressamente, que alguns dos requisitos de conformidade estão fora do escopo desenvolvido pelo grupo tendo em vista que o LIAA-3R “não desenvolve produtos finais, mas apenas protótipos”, bem como não exerce atividade de controle²¹³.

Consiste, assim, em um grupo de trabalho criado de modo informal dentro da estrutura do Laboratório de Inovação, composto exclusivamente por servidores vinculados ao Tribunal, que se voluntariaram para participar, com função de produzir documentação na fase inicial do ciclo de vida do sistema de IA.

Não obstante o caráter louvável da iniciativa, sem outra semelhante no âmbito nacional,

²¹⁰ *Ibidem.*

²¹¹ *Ibidem.*

²¹² *Ibidem.*

²¹³ *Ibidem.*

diante da análise de outras experiências em diferentes organizações, bem como em razão do caráter extremamente sensível do uso de IA no Judiciário, conclui-se que a ampliação dos objetivos do grupo, de modo a englobar todo o ciclo de vida das ferramentas desenvolvidas ou adquiridas, incluindo o monitoramento e tratamento de *feedback* dos usuários internos e externos, poderia contribuir para a melhoria da governança de IA no tribunal.

De fato, a doutrina é firme em defender que os impactos éticos do uso de ferramentas de IA devem ser examinados continuamente, ao menos em três fases distintas: validação (fase inicial de planejamento), verificação (fase de desenvolvimento) e avaliação (fase de operação)²¹⁴.

Quanto à sua estrutura, o fato de o grupo estar vinculado ao Laboratório de Inovação não se mostra uma boa alternativa. Isso porque o ambiente de incentivo à inovação, intencionalmente mais flexível, apesar de essencial para que se desenvolvam novas ideias, não parece compatível com a análise de resultado, antecipação de riscos e avaliação de potenciais impactos, em especial sob uma perspectiva ética, que demanda a formulação de perguntas, identificação de conflitos de valores e reflexão crítica sobre dilemas.

Não parece adequado também que os envolvidos, ainda que indiretamente, no desenvolvimento de sistemas de IA participem da atividade avaliativa dos potenciais riscos decorrentes da sua implementação.

Um órgão colegiado vinculado diretamente à Presidência do Tribunal ou suas secretarias, composto por uma equipe multidisciplinar, integrada por representantes de diferentes áreas da organização, poderia garantir maior imparcialidade e legitimidade na avaliação realizada, além de proporcionar um ambiente mais apropriado para o debate e reflexão.

No âmbito das Cortes Superiores, a Portaria STJ/GDG n. 365, de 26 de maio de 2021, instituiu o Comitê Gestor de Ética de Inteligência Artificial do Superior Tribunal de Justiça com o objetivo de assegurar um ambiente de discussão dos algoritmos em uso no tribunal na área de IA, seus modelos, premissas, limitações e possíveis consequências.

O normativo estabelece como atribuições do comitê: avaliar e homologar os algoritmos e modelos de IA de forma identificar tendências discriminatórias em seu funcionamento; identificar riscos na utilização das ferramentas e indicar os instrumentos de segurança e controle para seu enfrentamento; avaliar, antes do seu desenvolvimento, a viabilidade de modelos de IA

²¹⁴ MORLEY, Jessica et al. Ethics as a service: a pragmatic operationalisation of AI ethics. **Minds and Machines**, v. 31, n. 2, p. 239-256, 2021.

quanto ao custo-benefício de sua implantação.

Há também expressa previsão de que a entrada em produção de novos modelos de IA estar condicionada à sua homologação pelo comitê, bem como que os modelos já em uso devem receber atenção prioritária para fins de avaliação em relação os novos que podem vir a ser desenvolvidos ou implementados.

A multidisciplinariedade da composição do comitê foi, em tese, assegurada, garantido participação de integrantes de diferentes setores do Tribunal, de modo que é composto por um servidor de cada uma das seguintes unidades: I) Assessoria de Inteligência Artificial; II) Secretaria de Tecnologia da Informação e Comunicação; III) Secretaria Judiciária; IV) Núcleo de Admissibilidade e Recursos Repetitivos; V) Núcleo de Gerenciamento de Precedentes; VI) Secretaria de Gestão Estratégica; VII) Secretaria de Jurisprudência.

A iniciativa merece destaque por instituir, de maneira formal, colegiado específico, com diferentes visões da atividade desenvolvida na instituição, para o debate acerca dos impactos éticos do uso de IA.

Mais uma vez, contudo, não houve preocupação em o monitoramento de riscos de modo contínuo e coleta de *feedback*, focando apenas na avaliação preventiva nas fases de desenvolvimento e implementação.

Destaca-se que, ainda que o nome do grupo instituído seja “Comitê Gestor de Ética”, não foi atribuída nenhuma função de apoio à melhoria da estrutura de governança de IA do Tribunal, ou seja, não foi prevista a possibilidade elaboração de orientações de uso ou desenvolvimento de treinamento de pessoal e medidas de conscientização.

Cumpram também ressaltar que, não obstante o referido comitê tenha sido criado em maio de 2021, não há notícia pública de atividades realizadas, sequer reuniões para discussão quanto aos sistemas em uso ou em desenvolvimento, o que se mostra problemático ante os numerosos alertas de pesquisadores de que, frequentemente, as organizações criam instrumentos de governança que não chegam a ser implementados na prática, caracterizando prática de *ethics washing*.

No caso do Supremo Tribunal Federal, não obstante a implementação do sistema Victor, desenvolvido em parceria com a Universidade de Brasília, ainda em 2017, e desenvolvimento de duas novas ferramentas de IA, RAFA 2030 e VitórIA, pela equipe interna do tribunal, não há estrutura orgânica específica para governança e gestão de riscos das soluções tecnológicas implementadas, mas apenas o órgão encarregado de operacionalizar e executar os projetos de IA.

Saindo do cenário específico de IA, no âmbito nacional, interessante mencionar a estratégia de governança desenvolvida pelo CNJ relativa à gestão e ao uso de tecnologia da informação e comunicação.

A Resolução CNJ n. 370, de 28 de janeiro de 2021, determinou que todos os órgãos do Poder Judiciário deverão constituir ou manter Comitê de Gestão de Tecnologia da Informação e Comunicação (TIC), cujas atribuições são:

- I – envolver a alta administração nas decisões estratégicas que incidem sobre os serviços de TIC;
- II – aprovar planos táticos e operacionais junto a alta administração, disseminando a importância da área de TIC nos tribunais;
- III – monitorar a execução orçamentária e financeira de TIC;
- IV – planejar, priorizar e monitorar as contratações de TIC;
- V – **acompanhar o andamento das iniciativas estratégicas bem como seus desdobramentos;**
- VI – **apoiar na estruturação de escritório de projetos que favoreça o emprego das melhores práticas de gestão de projetos preconizadas pelos padrões nacionais e internacionais;**
- VII – definir a carteira de projetos e a gestão de portfólio de serviços de TIC;
- VIII – estabelecer plano de ação para iniciativas de curta duração ou escopo simplificado;
- IX – **promover recomendações e a adoção de boas práticas;**
- X – **propor modelos e padrões referentes à Governança de Tecnologia da Informação e Comunicação** por meio de campanhas institucionais;
- XI – promover a participação coletiva na elaboração de propostas e admissão de projetos;
- XII – **analisar, organizar e estruturar o atendimento das demandas de TIC**²¹⁵.
(grifos acrescidos)

Além dos comitês existentes em cada órgão, há o Comitê Nacional de Gestão de Tecnologia da Informação e Comunicação do Poder Judiciário, responsável por acompanhar o cumprimento da estratégia nacional e propor medidas necessárias à melhoria do desempenho. Cabe ao CNJ realizar anualmente avaliações e diagnósticos para aferir o nível de cumprimento das diretrizes estratégicas, bem como divulgar o resultado dos diagnósticos com objetivo de promover a transparência, a integração e o compartilhamento de informações entre os órgãos e as áreas de TIC do Poder Judiciário.

Iniciativa semelhante de instituir um órgão encarregado da gestão nacional de uma política específica no âmbito do Poder Judiciário foi realizada pela Resolução CNJ n. 395, de 7 de junho de 2021, que instituiu a Rede de Inovação do Poder Judiciário Brasileiro (RenovaJud), integrada pelos seguintes órgãos: I) Comitê Gestor Nacional da Inovação do

²¹⁵ CONSELHO NACIONAL DE JUSTIÇA. **Resolução n. 370, de 28 de janeiro de 2021**. Estabelece a Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário (ENTIC-JUD). DJe/CNJ nº 158/2021, de 22 de junho de 2021, p. 2-11 (republicação). Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3706>. Acesso em: 12 fev. 2024.

Poder Judiciário; II) Laboratório de Inovação e dos Objetivos de Desenvolvimento Sustentável (LIODS/CNJ); III) Laboratórios de Inovação instituídos no âmbito dos órgãos do Poder Judiciário; e IV) Conselho Consultivo Nacional da Inovação do Poder Judiciário.

Ao Comitê Gestor Nacional da Inovação foram atribuídas as seguintes competências:

- I – fomentar o desenvolvimento, pelos tribunais, de projetos inovadores que utilizem ferramentas de interação, cocriação, empatia e troca de conhecimento;
- II – **disseminar a cultura da inovação**, incentivando pesquisas, estudos e ações de capacitação na temática;
- III – **manter contato** com o LIODS e demais Laboratórios de Inovação do Poder Judiciário, **auxiliando-os em questões de amplo interesse ou relativas a políticas judiciárias nacionais**;
- IV – **estabelecer comunicação** com a Rede de Governança Colaborativa do Poder Judiciário, regulamentada pela Portaria CNJ nº 59/2019;
- V – **identificar problemas ou necessidades passíveis de solução** por meio das metodologias de inovação, encaminhando-os ao Laboratório de Inovação do Poder Judiciário para tratamento;
- VI – **estabelecer interlocução com agentes externos ao Poder Judiciário**, nos termos do art. 2º, IV, da Lei no 10.973/2004, visando à promoção da inovação; e
- VII – **propor ao CNJ normas relacionadas à gestão da inovação no Poder Judiciário**²¹⁶.

(grifos acrescidos)

Quanto à sua composição, integram o referido comitê 20 (vinte) indivíduos, todos integrantes de diferentes órgãos do Poder Judiciário, assegurando múltiplas visões, incluindo membros e servidores da Justiça Federal, Estadual, Eleitoral, Militar e do Trabalho, bem como do Conselho Nacional de Justiça.

Por sua vez, o Conselho Consultivo Nacional da Inovação, vinculado diretamente à Presidência do CNJ, tem por objetivo subsidiar a atuação do Comitê Gestor Nacional da Inovação (art. 13 da Resolução CNJ n. 395/2021).

O normativo determina que o referido colegiado será composto por 12 (doze) membros, representantes de órgãos e entidades dos setores público e privado, da academia e do terceiro setor, com renomada experiência e atuação na área da inovação, indicados pela Presidência do CNJ para o respectivo mandato.

A Portaria CNJ nº 328, de 16 de dezembro de 2021, que designou os membros do Conselho Consultivo Nacional da Inovação, contudo, apenas nomeou 7 (sete) integrantes, sendo três deles da academia, dois da iniciativa privada, um da sociedade civil e um

²¹⁶ CONSELHO NACIONAL DE JUSTIÇA. **Resolução n. 395, de 07 de junho de 2021**. Institui a Política de Gestão da Inovação no âmbito do Poder Judiciário. DJe/CNJ nº 148/2021, de 9 de junho de 2021, p. 3-7. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 12 fev. 2024.

magistrado²¹⁷. Um novo membro, professor e, à época, servidor do Poder Judiciário, foi incluído Portaria nº 332, de 20/12/21²¹⁸.

Conforme previsão normativa, a atuação tanto no Comitê Gestor quanto no Conselho Consultivo é realizada em caráter honorífico e não remunerado.

Vale ressaltar a iniciativa de criação de uma estrutura formal composta por membros externos ao Poder Judiciário para contribuir para a reflexão sobre problemas enfrentados pelo sistema de justiça, que interessam toda a sociedade.

No ponto, a abertura do Judiciário à participação democrática, além de ampliar as perspectivas sobre os desafios envolvidos, possibilitando a adoção de correções e melhorias, favorece a *accountability* por meio da participação de terceiros externos à instituição.

Não obstante a existência de um fórum para debate, formalmente estabelecido na estrutura do CNJ, verifica-se que a prática demonstra a dificuldade de implementar a política de gestão participativa e democrática.

No caso, nem o ato que instituiu o referido conselho, nem o que designou parte de seus membros (deixando vagas quatro das doze cadeiras previstas), estabeleceu os objetivos almejados, as responsabilidades dos envolvidos e tampouco o procedimento para desempenho de suas atribuições, isto é, se mediante reuniões periódicas, mediante provocação ou ação de ofício.

A Resolução n. 395/2021 apenas aponta que o colegiado (Conselho Consultivo Nacional da Inovação) tem como propósito subsidiar a atuação do Comitê Gestor Nacional da Inovação. Ocorre, contudo, que fornecer subsídios ao comitê gestor é forma pela qual o conselho atua e não precisamente a meta a ser perseguida. Melhor seria se o ato apontasse precisamente a razão pela qual faz-se necessário de subsídio de uma visão externa da instituição, como por exemplo: contribuir para a política de inovação do Poder Judiciário. Importante também que fossem indicados os direitos e deveres dos membros do Comitê Gestor e sua forma de atuação.

De fato, “um componente essencial da criação de um comitê de ética eficaz é estabelecer procedimentos claros para o uso do comitê, bem como processos eficientes pelos quais o comitê realiza o seu trabalho” (tradução livre)²¹⁹. Nesse contexto, é necessário definir como o comitê

²¹⁷ CONSELHO NACIONAL DE JUSTIÇA. **Portaria n. 328, de 16 de dezembro de 2021**. Designa os integrantes do Conselho Consultivo Nacional da Inovação do Poder Judiciário. DJe/CNJ nº 320/2021, de 17/12/2021, p. 6-7.

²¹⁸ CONSELHO NACIONAL DE JUSTIÇA. **Portaria n. 332, de 20 de dezembro de 2021** Altera a Portaria no 328/2021, que designa os integrantes do Conselho Consultivo Nacional da Inovação do Poder Judiciário. DJe/CNJ nº 322/2021, de 22/12/2021, p. 2.

²¹⁹ SANDLER, Ronald; BASL, John; TIELL, Steven. Building data and AI ethics committees. **Accenture & Northeastern University**, 2019. Disponível em: https://ethics.harvard.edu/sites/hwpi.harvard.edu/files/center-for-ethics/files/811330-ai-data-ethics-committee-report_v10.0.pdf?m=1567167985. Acesso em: 12 fev. 2024.

será acionado, formato, periodicidade e estrutura das reuniões, quais informações são disponibilizadas ao colegiado, como ocorrem as deliberações, se as decisões tomadas serão disponibilizadas ao público, entre outras.

Na prática, pouco se tem notícia das atividades do Conselho Consultivo de Inovação do Poder Judiciário. Sua primeira reunião ocorreu, em 15 de fevereiro de 2022, de maneira pública e muito abreviada, ficando gravada e disponibilizada em vídeo na internet²²⁰. Em maio do mesmo ano, ocorreu a segunda reunião em que foram apontadas como sugestões quanto à ampliação da segurança dos sistemas tecnológicos e a padronização dos produtos oferecidos pelo Poder Judiciário, além da realização de audiências públicas para conhecer as demandas das pessoas que os utilizam²²¹. Não há informação sobre a realização de outras reuniões ou mesmo da efetiva concretização das audiências públicas sugeridas.

Iniciativa semelhante, que permite a colaboração da sociedade civil com projetos e iniciativas do CNJ, são o Observatório dos Direitos Humanos e o Observatório do Meio Ambiente do Poder Judiciário.

O Observatório dos Direitos Humanos do Poder Judiciário foi instituído pela Portaria n. 190, de 17/9/2020, considerando a necessidade de interação com outros órgãos, entidades e organizações, de caráter nacional ou internacional, a fim de “desenvolver boas práticas e o aperfeiçoamento das políticas, dos projetos, das diretrizes e das atividades destinadas à tutela dos direitos humanos e fundamentais”²²².

O colegiado foi instituído como grupo de trabalho e “órgão de caráter consultivo vinculado à Presidência do Conselho Nacional de Justiça” e tem por objetivo fornecer subsídios para a adoção de iniciativas que promovam os direitos humanos e fundamentais no âmbito dos serviços judiciários²²³.

Os membros do Observatório, indicados pelo Presidente do CNJ, desempenham atividades em caráter honorífico e não remunerado e devem possuir experiência ou formação na área de Direitos Humanos, sendo preferencialmente escolhidos dentre profissionais do meio

²²⁰ CONSELHO NACIONAL DE JUSTIÇA. Reunião do Conselho Consultivo de Inovação do Poder Judiciário. Brasília: CNJ. 15 FEV. 2022. Disponível em: <https://www.youtube.com/watch?v=LJuQfX-yb8k>. Acesso em: 15 jul. 2023.

²²¹ CONSELHO NACIONAL DE JUSTIÇA. Conselho Consultivo de Inovação discute segurança de sistemas do Judiciário. Brasília: CNJ. 18 maio. 2022. Disponível em: <https://www.cnj.jus.br/conselho-consultivo-de-inovacao-discute-seguranca-de-sistemas-do-judiciario/>. Acesso em: 23 jan. 2024

²²² CONSELHO NACIONAL DE JUSTIÇA. **Portaria n. 190, de 17 de setembro de 2020**. Institui o Grupo de Trabalho denominado “Observatório dos Direitos Humanos do Poder Judiciário” e dá outras providências. DJe/CNJ nº 307/2020, de 18/09/2020, p. 2-3. Republicada no DJe/CNJ nº 308/2020, de 18/09/2020, p. 2-3. Republicada no DJe/CNJ nº 325/2020, de 06/10/2020, p. 2-3.

²²³ *Ibidem*.

acadêmico, da sociedade civil ou de entidades representativas.

Os objetivos do Observatório de Direitos Humanos do Poder Judiciário estão elencados no art. 3º da Portaria n. 190, de 17/9/2020, da seguinte maneira:

- I – promover a articulação do Poder Judiciário com instituições nacionais ou internacionais que atuem na defesa dos direitos humanos, bem como parcerias para o intercâmbio de informações, de dados, de documentos ou de experiências;
- II – municiar a atuação do Poder Judiciário na formulação de políticas, projetos e diretrizes destinados à tutela dos direitos humanos;
- III – executar iniciativas e projetos relacionados à temática de direitos humanos;
- IV – elaborar estudos e pareceres sobre demandas que envolvam questões estratégicas de direitos humanos;
- V – propor a celebração de acordos de cooperação afetos ao seu escopo de atribuições;
- VI – organizar publicações referentes à atuação do Poder Judiciário na defesa dos direitos humanos, bem como promover seminários, audiências públicas ou outros eventos concernentes a essa área temática; e
- VII – propor ao Plenário do Conselho Nacional de Justiça medidas que considere pertinentes e adequadas ao aprimoramento da tutela dos direitos humanos no âmbito do Poder Judiciário²²⁴.

Interessante a previsão específica de que as atividades do Observatório de Direitos Humanos serão documentadas em relatório circunstanciado, a ser publicado anualmente, podendo ser apresentados relatórios parciais dos trabalhos realizados, antes da consolidação do relatório anual, a juízo do Presidente do CNJ.

O “Guia para Integrantes e Embaixadores do Observatório dos Direitos Humanos do Poder Judiciário” aponta como fluxo de trabalho do referido grupo:

- As reuniões ordinárias ocorrem a cada três meses.
- Todos(as) terão voz e voto nas deliberações do Observatório.
- Cada integrante poderá propor as iniciativas que julgar pertinentes.
- Os debates dos temas e os encaminhamentos ocorrerão por meio do Comitê Executivo.
- As propostas serão avaliadas por equipe técnica do CNJ, que poderá sugerir adequações.
- Os integrantes poderão sugerir a formulação de convites a especialistas para que participem das reuniões.²²⁵

Por sua vez, o Grupo de Trabalho “Observatório do Meio Ambiente do Poder Judiciário”, foi instituído pela Portaria n. 241, de 10 de novembro de 2020, com o objetivo de

²²⁴ CONSELHO NACIONAL DE JUSTIÇA. **Portaria n. 190, de 17 de setembro de 2020**. Institui o Grupo de Trabalho denominado “Observatório dos Direitos Humanos do Poder Judiciário” e dá outras providências. DJe/CNJ nº 307/2020, de 18/09/2020, p. 2-3. Republicada no DJe/CNJ nº 308/2020, de 18/09/2020, p. 2-3. Republicada no DJe/CNJ nº 325/2020, de 06/10/2020, p. 2-3.

²²⁵ CONSELHO NACIONAL DE JUSTIÇA. **Guia para Integrantes e Embaixadores do Observatório dos Direitos Humanos do Poder Judiciário**. 17 set. 2020. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2023/11/guia-para-integrantes-e-embaixadores-web-odh-2023-2025.pdf>. Acesso em: 14 jan. 2024.

“traçar estudo, monitoramento, pesquisas, programas, projetos e ações para a construção de diagnósticos das boas práticas, formulação de políticas e implementação de projetos e iniciativas para a tutela do meio ambiente natural da Amazônia Legal pela atuação do Poder Judiciário e do Sistema de Justiça”²²⁶.

Observada a área específica de atuação, referido grupo adota estrutura bastante similar ao Observatório dos Direitos Humanos, tanto quanto aos objetivos, composição e metodologia de trabalho.

Iniciativa semelhante ao Conselho Consultivo Nacional da Inovação, no âmbito internacional, porém mais estruturada, é o já citado Conselho Consultivo de IA do CEPEJ (*Artificial Intelligence Advisory Body* - AIAB), composto por cinco membros selecionados entre integrantes do sistema de justiça, da academia, da política, da indústria e do terceiro setor, responsável por monitorar o efetivo aparecimento de aplicações de IA nos sistemas de justiça, contribuir para a reflexão contínua sobre o uso de IA no referido domínio e propor estratégias relativas à utilização da IA²²⁷.

No caso do Conselho Consultivo de IA do CEPEJ (AIAB), o ato que criou o referido órgão indicou precisamente que o colegiado deve reunir-se virtualmente de três em três meses e que deve apresentar relatórios duas vezes ao ano aos grupos de trabalho do CEPEJ aos quais é vinculado.

O órgão tem como função fornecer orientações sobre questões envolvendo o uso de IA no sistema de justiça e a operacionalização dos princípios da Carta Ética da CEPEJ, bem como aconselha os grupos de trabalho da CEPEJ sobre possíveis novas estratégias relacionadas ao tema²²⁸.

Considerando o modelo de política de governança elaborada pelo CNJ para a gestão da inovação e de gestão de tecnologia da informação e comunicação, pode-se extrair alguns instrumentos interessantes para a construção de um modelo específico para o desenvolvimento e implementação de ferramentas de IA.

²²⁶ CONSELHO NACIONAL DE JUSTIÇA. **Portaria n. 241, de 10 de novembro de 2020**. Institui o Grupo de Trabalho “Observatório do Meio Ambiente e das Mudanças Climáticas do Poder Judiciário”. DJe/CNJ nº 358/2020, de 11/11/2020, p. 8-9.

²²⁷ CEPEJ. **Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment**. 9 dez. 2021. Disponível em: <https://rm.coe.int/cepej-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 15 nov. 2023.

²²⁸ COUNCIL OF EUROPE. **CEPEJ Artificial Intelligence Advisory Board (AIAB)**. Disponível em: <https://www.coe.int/en/web/cepej/ai-advisory-board>. Acesso em: 17 mar. 2024.

3.3. Sugestão de criação de comitês de ética de IA no Poder Judiciário brasileiro

Cabe aqui fazer algumas considerações sobre a necessidade de construção de uma política de governança de IA com instrumentos e mecanismos próprios e não apenas sua inclusão dentre os escopos da gestão de inovação ou de tecnologia da informação e comunicação (TIC).

Apesar de sistemas de IA utilizarem tecnologia inovadora e disruptiva, os impactos causados pela sua utilização geram questões éticas a serem dirimidas com cautela, razão pela qual não parece adequada que o tratamento dado ao tema seja realizado no âmbito dos laboratórios de inovação do Judiciário.

De acordo com o art. 2º, IV, da Lei n. 10.973/2004, inovação consiste na “introdução de novidade ou aperfeiçoamento no ambiente produtivo e social que resulte em novos produtos, serviços ou processos ou que compreenda a agregação de novas funcionalidades ou características a produto, serviço ou processo já existente que possa resultar em melhorias e em efetivo ganho de qualidade ou desempenho”²²⁹.

Já a Resolução CNJ n. 395/2020¹ define inovação como “a implementação de ideias que criam uma forma de atuação e geram valor para o Poder Judiciário, seja por meio de novos produtos, serviços, processos de trabalho, ou uma maneira diferente e eficaz de solucionar problemas complexos encontrados no desenvolvimento das atividades que lhe são afetas”²³⁰.

Assim, a inovação não está atrelada ao desenvolvimento ou uso de recursos tecnológicos, muito menos a tecnologias específicas de inteligência artificial. A cultura de inovação está relacionada muito mais à criatividade, podendo acontecer, por exemplo, com a reestruturação de um determinado fluxo procedimental, alterando sequências ou eliminando etapas, sem a necessidade de incorporar elementos tecnológicos para tanto²³¹.

Os laboratórios de inovação devem propiciar um ambiente favorável para a concepção de ideias e experimentação de instrumentos e métodos, exigindo, assim, flexibilidade, adaptabilidade, de modo a promover agilidade, otimização de recursos e ganho de eficiência à

²²⁹ BRASIL. **Lei nº 10.973, de 2 de dezembro de 2004**. Dispõe sobre incentivos à inovação e à pesquisa científica e tecnológica no ambiente produtivo e dá outras providências. Diário Oficial da União de 03 de dez. de 2004, p. 2. Disponível em: <https://legislacao.presidencia.gov.br/atos/?tipo=LEI&numero=10973&ano=2004&ato=884UTRE5UeRpWT79e>. Acesso em: 12 fev. 2024.

²³⁰ CONSELHO NACIONAL DE JUSTIÇA. **Resolução n. 395, de 07 de junho de 2021**. Institui a Política de Gestão da Inovação no âmbito do Poder Judiciário. DJe/CNJ nº 148/2021, de 9 de junho de 2021, p. 3-7. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 12 fev. 2024.

²³¹ LIMA, Caio Moysés. Laboratório de inovação e tecnologia, in LUNARDI, Fabrício Castagna; CLEMENTINO, Marco Bruno Miranda (coord.). Inovação judicial: fundamentos e práticas para uma jurisdição de alto impacto. Brasília: ENFAM, 2021, p. 110-111.

prestação de serviços.

O debate sobre ética da IA, por sua vez, exige, entre outras habilidades, capacidade de reflexão sobre valores humanos em diferentes contextos, antecipação de possíveis impactos e riscos, avaliação sobre medidas de mitigação, familiaridade com as capacidades e limitações da tecnologia, bem como com conceitos de viés, discriminação, privacidade, *accountability* e *fairness*.

O papel de um órgão como um comitê de ética de IA é exatamente enxergar o que a equipe técnica, sozinha, não conseguiria: identificar conflitos de valores; refletir criticamente sobre os impactos sociais e dilemas éticos; contribuir para que sejam observados princípios éticos; avaliar e documentar sobre as soluções dadas aos conflitos apreciados.

Esse aspecto essencialmente multi e interdisciplinar das questões relacionadas ao uso ético de IA justifica que o tema seja tratado de modo apartado da estrutura de governança de tecnologia da informação (TI).

Dada a importância estratégica que o uso de ferramentas tecnológicas possui nas organizações, a governança de TI tem ganhado destaque, assegurando-se estrutura, processos e mecanismos voltados a direcionar as iniciativas aos objetivos definidos para instituição e monitorar os resultados.

A necessidade de que sejam observados princípios éticos no desenvolvimento e implementação de ferramentas de IA, por sua vez, cria novos desafios, exigindo que haja reflexão sobre os diferentes valores envolvidos, bem como sejam ponderados diferentes interesses de muitos atores envolvidos durante todo o ciclo de vida dos sistemas, além de aspectos puramente econômicos ou de eficiência.

Considerando essa necessidade de trabalhar com a complexidade dos valores humanos e avaliar os impactos gerados em todas as partes envolvidas e na sociedade como um todo, importante desenvolver uma estrutura própria de governança de IA, que congregue profissionais habilitados para lidar com a operacionalização de princípios éticos e compreendam os riscos do uso da IA no sistema de justiça.

Com efeito, para profissionais da área de tecnologia a avaliação de um sistema está relacionada apenas ao funcionamento do modelo, se ele cumpre a tarefa para a qual foi desenvolvido, sem que sejam examinados os impactos éticos e sociais da implementação da ferramenta. Ainda que sejam instruídos a realizar esse tipo de avaliação, engenheiros e cientistas de dados não estão preparados para essa tarefa.

Conforme aponta Reid Blackman, frequentemente profissionais de TI tendem a

acreditar que os desafios éticos relacionados aos sistemas por eles desenvolvidos podem ser resolvidos apenas com ajustes técnicos seja na base de dados de treinamento ou em qualquer outro ponto do processo de construção dos modelos²³².

Ocorre, contudo, que nem todos os problemas decorrentes do emprego de sistemas de IA são exatamente da tecnologia, muitas vezes podem ser consequência de abusos, usos indevidos, ausência ou insuficiência das diretrizes de uso ou de deficiências na cultura da instituição. Ademais, a maioria dos profissionais e cientistas da computação não está preparada para lidar com julgamentos morais e dilemas éticos.

Nesse cenário, considerando as razões que justificam a criação de comitês de ética nas organizações, bem como os modelos de gestão implementados pelo CNJ em outras searas, propõe-se a criação de estrutura de governança específica para gestão de sistemas de IA no Judiciário integrada, entre outros mecanismos, por comitês de ética de IA locais dentro de cada tribunal, uma rede nacional de comitês de ética de IA do Poder Judiciário, um Comitê Nacional de Ética de IA e um Conselho Consultivo Nacional de Ética de IA, conforme passa-se a expor²³³.

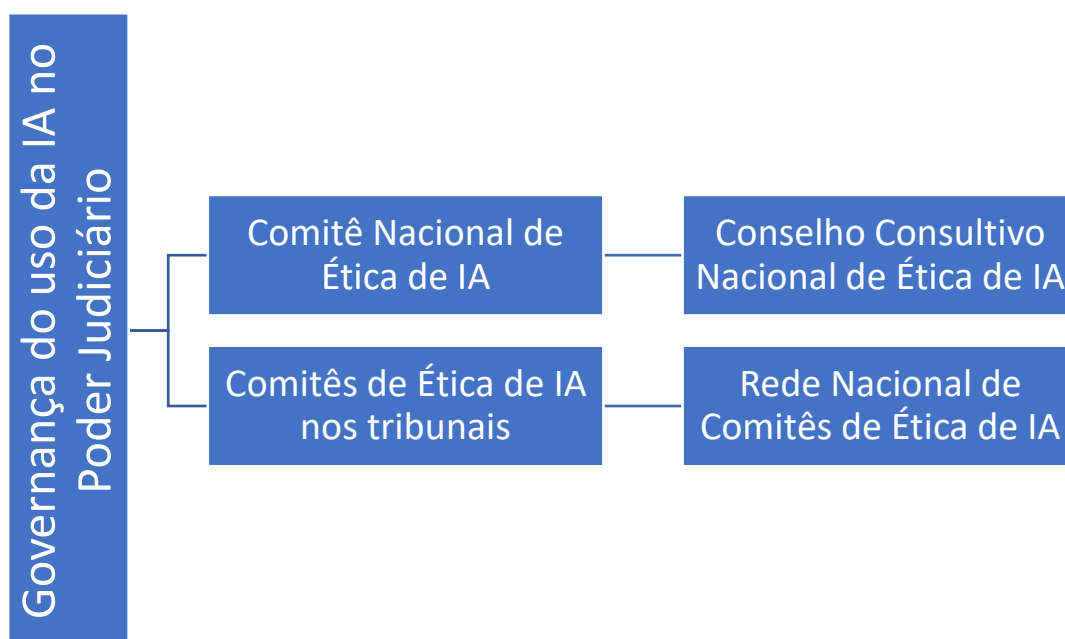


Figura 2: Representação visual da estrutura de governança do uso da IA no Poder Judiciário proposta pela autora
Fonte: Elaborado pela autora.

²³² BLACKMAN, Reid. Ethical Machines: Your Concise Guide to Totally Unbiased, Transparent, and Respectful AI. **Harvard Business Review Press**, 2022.

²³³ A proposta em questão baseia-se nos seguintes atos normativos do CNJ: Resolução Nº 370/2021, que estabelece a Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário (ENTIC-JUD); Portaria Nº 47/2014, que institui o Comitê Nacional de Gestão de Tecnologia da Informação e Comunicação do Poder Judiciário; e Resolução CNJ n. 395/2021, que instituiu a Política de Gestão da Inovação no âmbito do Poder Judiciário.

Os comitês de ética de IA constituídos no âmbito de cada tribunal devem ter caráter permanente, natureza consultiva e deliberativa, e devem ser compostos por representantes de todas as áreas estratégicas do órgão, incluindo usuários finais (internos) dos produtos de IA desenvolvidos ou em utilização na instituição, assegurando-se uma perspectiva multidisciplinar.

O referido colegiado ficará responsável, em essência, pela identificação e mitigação de riscos éticos do uso de IA no âmbito de atuação do tribunal. Também propõe-se incluir dentre as competências do comitê de ética de IA de cada tribunal, as seguintes:

- I – identificar e mitigar os riscos éticos dos produtos de IA que são desenvolvidos internamente ou adquiridos de fornecedores terceirizados;
- II - apoiar e acompanhar o desenvolvimento de projetos de IA, bem como eventuais processos de aquisição de produtos desenvolvidas por terceiros, de modo independente, analisando os riscos éticos relacionados ao uso intencional e não intencional da ferramenta;
- III – apoiar na estruturação e implementação do projeto de governança de IA;
- IV – orientar sobre as diretrizes de uso dos produtos de IA para assegurar que sejam empregados de forma ética e responsável;
- V – recomendar e acompanhar a adoção de boas práticas, assim como a eficácia de seus processos, propondo atualizações e melhorias quando necessário;
- VI – disseminar a cultura do uso da IA de modo ético e responsável, incentivando ações de capacitação na temática;
- VII – recomendar ações de treinamento e de conscientização dos riscos associados ao uso de sistema de inteligência artificial no sistema de justiça;
- VIII – estimular o diálogo entre as diferentes áreas do órgão em assuntos relacionados aos riscos éticos dos sistemas de IA;
- IX – receber e tratar opiniões, reclamações e denúncias de usuários internos e externos ou de partes interessadas externas ao Poder Judiciário que envolvam a utilização de sistemas de IA;
- X – assegurar a convergência dos interesses entre o Poder Judiciário e a sociedade;
- XI – promover relatórios periódicos sobre os resultados obtidos com a utilização de sistemas de IA e os riscos envolvidos que devem ser encaminhados para o Comitê Nacional de Ética de IA²³⁴.

Quanto à composição do comitê de ética de IA de cada tribunal, ela deve ser multidisciplinar e diversa em perfis e experiências. Deve-se buscar congregar pessoas com capacidade de diálogo, de diferentes departamentos e áreas, assim como de diferentes níveis, do mais operacional à alta gestão. Destaca-se que é desejável que o grupo consiga “impulsionar, manter e consolidar o processo de transformação por sua própria autoridade ou por interação

²³⁴ A proposta apresentada é inspirada nas atribuições de órgão previstos nos seguintes atos normativos do CNJ: Resolução Nº 370/2021, que estabelece a Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário (ENTIC-JUD); Portaria Nº 47/2014, que institui o Comitê Nacional de Gestão de Tecnologia da Informação e Comunicação do Poder Judiciário; e Resolução CNJ n. 395/2021, que instituiu a Política de Gestão da Inovação no âmbito do Poder Judiciário.

direta com a cúpula da organização”²³⁵.

É importante que, além da composição, seja definida a forma de atuação do colegiado, de preferência, com a definição prévia de reuniões periódicas para cumprimento de seus objetivos e divulgação sobre as atividades desenvolvidas.

Transparência sobre as operações do comitê de ética é um ponto relevante para que seja construído um ambiente de confiança no uso de sistemas de IA no Judiciário, ainda que eventualmente certas avaliações ou recomendações possam não ser divulgadas para o público, bem como para evitar a prática de *ethics washing*. Todas as atividades desenvolvidas no âmbito do comitê devem ser documentadas, cabendo uma avaliação cuidadosa sobre aquelas que não podem ser publicizadas.

Mostra-se essencial também que, consistindo em colegiado formado apenas por integrantes da instituição, que, em regra, não possuem formação específica para lidar com os aspectos técnicos da tecnologia, bem como com os dilemas éticos envolvidos, seja promovida a necessária capacitação dos seus integrantes tanto no aspecto técnico quanto ético.

No ponto, destaca-se que não obstante as iniciativas do Conselho Nacional de Justiça de disponibilizar cursos de capacitação em IA, não há, dentre as ações desenvolvidas, nenhuma voltada para abordar as questões éticas e conscientizar sobre os potenciais impactos do uso da tecnologia no sistema de Justiça.

É importante que os integrantes dos comitês de ética de IA tenham conhecimento não apenas sobre a lógica, o funcionamento e as limitações da inteligência artificial, mas também sobre os princípios e debates éticos, de modo que estejam preparados para realizar avaliações que envolvam os impactos da implementação da tecnologia para além da questão de eficiência e acurácia, incluindo a antecipação de consequências não intencionais no devido processo legal e na sociedade em geral.

Destaca-se que a capacitação que inclua também conhecimento sobre as más práticas de *ethics shopping* e *ethics washing* contribui para que a existência do colegiado não se torne mera fachada para aparentar um respaldo de comportamento ético sem que seja realizada atividade efetivamente crítica.

Assim, um projeto de qualificação dos membros dos diferentes comitês de ética voltado para promover aprendizagem sobre as questões éticas e técnicas da IA é essencial. Nessa perspectiva, também se mostra importante assegurar oportunidades de compartilhamento de

²³⁵ COELHO, Alexandre Zavaglia et al. **Governança da inteligência artificial em organizações: framework para comitês de ética em IA: versão 1.0**. CEPI FGV Direito SP, 2023.

conhecimento entre os membros de diferentes comitês, possibilitando o aprendizado por meio da troca de experiências e estudo de casos.

Nesse sentido, propõe-se a criação de uma Rede de Comitês de Ética de IA do Poder Judiciário com o propósito de promover interação entre membros de diferentes comitês a fim de criar um ambiente colaborativo, que promova a troca de informações sobre projetos e desafios, além de proporcionar debates sobre temas relevantes relacionados à utilização de IA no sistema de justiça.

Os referidos comitês de ética devem propiciar um ambiente de reflexão, que permita a diversidade de ideias, adotando uma abordagem de precaução, para avaliar a aceitabilidade social, legal e ética dos sistemas de IA. Sugere-se que sejam endereçados questionamentos como: quais são os impactos não intencionais do uso da ferramenta? Os usuários estão adotando boas práticas na utilização das ferramentas? Quais são as hipóteses de usos indevidos do sistema?

Os referidos colegiados, ressalta-se, não devem ser responsáveis pela implementação de todo o programa de governança de IA da instituição, mas complementam os esforços existentes, orientando, elaborando diretrizes, avaliando dilemas éticos e auxiliando a tomada de decisões tanto em relação a sistemas desenvolvidos internamente quanto no que toca à sistemas desenvolvidos por terceiros.

No âmbito nacional, defende-se a criação de um Comitê Nacional de Ética de IA no Poder Judiciário, vinculado ao Conselho Nacional de Justiça, de caráter permanente, que tenha como objetivo geral supervisionar e monitorar a identificação e mitigação de riscos éticos dos produtos de IA que são desenvolvidos internamente ou adquiridos de fornecedores independentes, bem como promover estudos e coordenar trabalhos sobre o tema.

A composição do referido comitê deve assegurar diversidade de representação das diferentes esferas do Poder Judiciário (Justiça Federal, Estadual, do Trabalho, Eleitoral e Militar), bem como a participação de atores externos ligados às atividades essenciais à Justiça indicados pelo Conselho Nacional do Ministério Público (CNMP) e a Ordem dos Advogados do Brasil (OAB), ainda que esses possam vir a ter limitação quanto a eventuais deliberações.

A existência de membros externos no comitê garante independência, legitimidade e *accountability* ao colegiado²³⁶.

²³⁶ A propósito: “*The members of such an external committee are presumably more accountable to the public interest than the company convening the ethics committee. These external corporate ethics boards are beneficial for their independence, external accountability, and perceived legitimacy*”. MARCHANT, Gary E.; GUTIERREZ, Carlos Ignacio. “Soft Law 2.0: An Agile and Effective Governance Approach for Artificial Intelligence.” *Minnesota Journal of Law, Science and Technology*, vol. 24, no. 2, 2023, pp. 375-424. P. 404.

Deve-se atentar, contudo, que “comitês com muitos membros podem deixar o processo mais moroso”²³⁷, de modo que se sugere que o referido colegiado de caráter nacional tenha 15 membros, de modo a garantir diversidade e multiplicidade de perspectivas sem comprometer a eficiência da atuação do colegiado.

O Comitê Nacional de Ética de IA no Poder Judiciário sugerido deve ter como objetivos específicos:

- I – apoiar a implementação da Estratégia Nacional de IA do Poder Judiciário;
- II – disseminar a cultura do uso da IA de modo ético e responsável, incentivando pesquisas e estudos sobre as consequências do uso da IA no Poder Judiciário, assim como ações de capacitação sobre ética de IA;
- III – recomendar a adoção de boas práticas, de modo a mitigar os riscos éticos do uso de IA;
- IV – incentivar o desenvolvimento, implantação e o aperfeiçoamento de diretrizes de uso dos produtos de IA para assegurar sejam empregados de forma ética e responsável;
- V – propor ações de capacitação de servidores e magistrados em IA e ética de IA;
- VI – supervisionar e monitorar a identificação e mitigação dos riscos éticos dos produtos de IA implementada no âmbito dos comitês de ética dos tribunais;
- VII – auxiliar os comitês de éticas dos tribunais em questões de interesse relevante quando provocados por usuários internos ou externos;
- VIII – estabelecer comunicação com a Rede de Comitês de Ética do Poder Judiciário;
- IX – propor critérios para realização de auditorias de sistemas de IA;
- X – identificar problemas ou necessidades passíveis de intervenção pelo Conselho Nacional de Justiça, encaminhando recomendações se for o caso;
- XI – estabelecer interlocução com agentes externos visando a convergência de interesses entre o Poder Judiciário e a sociedade;
- XII – propor ao CNJ normas relacionadas à ética de IA e governança de IA no Poder Judiciário²³⁸.

A fim de assegurar imparcialidade e independência dos integrantes do Comitê Nacional de Ética de IA no Poder Judiciário, deve-se garantir que, ao menos, dois terços de seus membros não estejam diretamente envolvidos em projetos de desenvolvimento de sistemas de IA, pois o foco deve ser o debate sobre os valores envolvidos e as consequências do uso da tecnologia, não os desafios operacionais de desenvolvimento e implementação.

De fato, o referido órgão propõe-se a estabelecer uma instância oficial para diálogo sobre os dilemas e desafios éticos enfrentados envolvendo soluções tecnológicas em todas as atividades do Poder Judiciário, bem como preocupações macro com direitos e garantias

²³⁷ COELHO, Alexandre Zavaglia et al. **Governança da inteligência artificial em organizações: framework para comitês de ética em IA: versão 1.0**. CEPI FGV Direito SP, 2023. P. 137.

²³⁸ A proposta apresentada é inspirada nas atribuições de órgão previstos nos seguintes atos normativos do CNJ: Resolução Nº 370/2021, que estabelece a Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário (ENTIC-JUD); Portaria Nº 47/2014, que institui o Comitê Nacional de Gestão de Tecnologia da Informação e Comunicação do Poder Judiciário; e Resolução CNJ n. 395/2021, que instituiu a Política de Gestão da Inovação no âmbito do Poder Judiciário.

fundamentais, independência judicial e Estado de Direito.

Busca-se com a criação deste ambiente de diálogo promover reflexões para além da discussão de resultados, mas também sobre os meios utilizados, os efeitos colaterais, o que fica de fora da análise focada em eficiência, bem como desafios a serem abordados em termos de política judiciária.

Além do Comitê Nacional de Ética de IA no Poder Judiciário, que assegura, em menor proporção, a participação de integrantes externos no colegiado (membros do Ministério Público e advogados), propõe-se a criação de um Conselho Consultivo Nacional de Ética de IA com o objetivo de subsidiar a atuação Comitê Nacional de Ética de IA, garantindo uma maior participação da sociedade no debate por meio de membros especialistas, com renomada experiência e atuação na área, integrantes de entidades do setor privado, da academia e do terceiro setor.

A propósito, ao tratar de desenvolvimento e uso de sistemas de IA, Hartmann aponta que “a formação de espaços de discussão entre especialistas multidisciplinares, funcionários públicos, organizações civis, associações e representantes da iniciativa privada permitirá avançar em visões mais amplas, coletivas e coerentes sobre os desafios éticos da IA (oportunidades e riscos)”²³⁹.

A criação de um colegiado que assegure ampla participação das partes interessadas, incluindo representantes da sociedade civil, pesquisadores e profissionais de renome, contribui para a aceitação e criação de confiança na utilização ética e segura de IA no Judiciário, bem como permite que se aproveite do conhecimento e sugestões de terceiros, que podem contribuir com diferentes perspectivas sobre o tema.

Nessa linha, a recomendação da UNESCO sobre ética de IA indica que:

Os Estados-membros devem encorajar entidades públicas, empresas do setor privado e organizações da sociedade civil a envolverem diferentes partes interessadas em sua governança de IA, bem como a considerar o acréscimo do papel de um oficial de ética da IA independente, ou algum outro mecanismo para supervisionar a avaliação do impacto ético, auditoria e esforços de monitoramento contínuo, e garantir a orientação ética dos sistemas de IA²⁴⁰.

Também no âmbito do Conselho Consultivo Nacional de Ética de IA deve-se considerar a necessidade de efetiva descrição das atividades a serem desenvolvidas, da periodicidade das

²³⁹ PEIXOTO, Fabiano Hartmann. *Op. Cit.* p. 152.

²⁴⁰ UNESCO. **Recomendação sobre a Ética da Inteligência Artificial**. 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 23 nov. 2023.

reuniões, do nível de acesso à informação dado aos membros externos, do processo de funcionamento do colegiado, dos direitos e deveres de seus membros, estando sempre alerta para que a criação de tal iniciativa não caracterize mera prática de *ethics washing*.

No ponto, pode-se tomar como inspiração o Conselho Consultivo de IA do CEPEJ (*Artificial Intelligence Advisory Body* - AIAB), que possui atribuições, composição e forma de trabalho bem definidas.

No que toca à organização dos trabalhos, o referido colegiado europeu deve reunir-se virtualmente de três em três meses, seguindo uma agenda preparada pelos grupos de trabalho do CEPEJ aos quais é vinculado, devendo apresentar relatórios duas vezes ao ano em reuniões com os referidos grupos.

Quanto à sua composição, os cinco especialistas que compõem o grupo são nomeados para um mandato de dois anos, renováveis por decisão do CEPEJ, após seleção realizada por meio de chamamento público de interessados²⁴¹.

Podem concorrer para seleção especialistas provenientes do Judiciário, da academia, formuladores de políticas públicas, da indústria e do terceiro setor, desde que atendam aos seguintes requisitos mínimos:

- a) apresentem estudos em direito, ciências políticas, ciências sociais, matemática, engenharia, ciências técnicas (desenvolvimento de software, etc.) ou outras áreas relacionadas;
- b) forte experiência profissional em relação à justiça ou TI, sendo uma combinação de ambos considerada uma vantagem;
- c) interesse comprovado em ética e/ou direitos fundamentais, especialmente nas áreas de justiça ou aplicações de TI;
- d) conhecimento do funcionamento e do trabalho da CEPEJ;
- e) muito bom conhecimento de inglês (tradução livre)²⁴².

No caso do AIAB, os membros recebem honorários por atribuições específicas, de acordo com as regras e práticas habituais do Conselho da Europa e dentro dos limites de sua capacidade orçamentária²⁴³.

O processo de seleção dos membros Conselho Consultivo de IA do CEPEJ muito pode contribuir para a seleção do Conselho Consultivo Nacional de Ética de IA, ora proposto. Com

²⁴¹ COUNCIL OF EUROPE. **CEPEJ Artificial Intelligence Advisory Board (AIAB)**. Disponível em: <https://www.coe.int/en/web/cepej/ai-advisory-board>. Acesso em: 20 mar. 2024.

²⁴² *Ibidem*.

²⁴³ No original: “*Members of the AIAB will receive fees for specific assignments in relation with its terms of reference. The amount of the fees is fixed by the Secretariat of the CEPEJ in accordance with the Council of Europe's usual rules and practices in this field and within the limits of its budgetary capacities*”. **CEPEJ Artificial Intelligence Advisory Board (AIAB)**. Disponível em: <https://www.coe.int/en/web/cepej/ai-advisory-board>. Acesso em: 20 mar. 2024.

efeito, seria interessante uma seleção por meio de chamamento público de interessados que cumpram requisitos mínimos, comprovando a realização de estudos e experiência nas áreas relacionadas (direito, ciências políticas, ciências sociais, ciência da computação, engenharia, matemática, programação), concedendo preferência a pessoas que comprovem formação filosófica ou interesse em ética de IA e/ou sistemas de justiça, bem como aquelas que demonstrem possuir bagagem multidisciplinar. Essa prática reforça o interesse público na seleção, possibilitando a participação de especialistas interessados e atualizados.

A decisão de atribuir mandato de dois anos aos membros do Conselho, podendo ser renovado, é uma escolha estratégica que proporciona estabilidade e independência à equipe, ao mesmo tempo em que permite, com o passar do tempo, a entrada de novos membros para contribuições inovadoras. Essa abordagem assegura um equilíbrio entre a continuidade necessária para a eficácia do grupo e a introdução de novas perspectivas, promovendo assim um ambiente propício ao enriquecimento contínuo das discussões.

Assim, sugere-se que seja garantida maior independência ao Conselho Consultivo Nacional de Ética de IA do que ocorre com os Observatórios dos Direitos Humanos e do Meio Ambiente, que são presididos pelo Presidente do CNJ, que define as metas, prioridades e objetivos dos referidos grupos, e cuja composição pode ser revista a qualquer tempo.

Com efeito, a possibilidade de remover membros do colegiado sem razão específica (*ad nuntum*) compromete a independência do colegiado, que é essencial, inclusive na condução dos trabalhos desenvolvidos, para o adequado cumprimento de seus objetivos, que consistem em: a) municiar a atuação do Poder Judiciário na formulação de políticas, projetos e diretrizes envolvendo o uso de IA; b) elaborar estudos e pareceres sobre o uso ético e responsável da IA no Judiciário; c) propor ao Plenário do Conselho Nacional de Justiça medidas que considere pertinentes e adequadas ao aprimoramento da governança do uso de ferramentas de IA no âmbito do Poder Judiciário.

Nesse contexto, apesar de vinculado ao CNJ, respondendo ao Comitê Nacional de Ética de IA no Poder Judiciário ou a outro grupo de trabalho cujo tema esteja relacionado, bem como ainda que tenha caráter consultivo, o Conselho Consultivo Nacional de Ética de IA deve ser independente.

Para tanto, sugere-se que seus membros, que desempenharão atividades em caráter honorífico e não remunerado, sejam selecionados, por meio de chamamento público de interessados (todos externos ao Judiciário, oriundos da academia, da iniciativa privada e do terceiro setor) que cumpram requisitos mínimos, comprovando a realização de estudos e

experiência nas áreas relacionadas, conforme acima apontado, dando preferência para candidatos com formação filosófica e/ou multidisciplinar, para mandato de dois anos, renovável por igual período.

Recomenda-se que seja bem definida a metodologia de trabalho, com reuniões periódicas virtuais de três em três meses apenas entre os integrantes do colegiado, que devem apresentar relatório público detalhado, ao menos, uma vez por ano. Também poderão ocorrer em reuniões extraordinárias com servidores e/ou magistrados indicados pelo Presidente do Conselho Nacional de Justiça à juízo deste último, que poderá solicitar a apresentação de trabalhos específicos sobre a temática.

Ressalta-se que a eficácia desses colegiados depende em boa parte da escolha de seus membros e de sua disposição e capacidade de cumprir sua missão, de modo que a nomeação de pessoas certas para cada um dos três colegiados sugeridos é fator crucial.

Destaca-se, por fim, que cada um dos órgãos sugeridos possui funções e objetivos próprios.

Os comitês de ética de IA dos tribunais destinam-se a realizar um papel mais operacional de elaborar relatórios de diagnóstico de todas as soluções tecnológicas empregadas no respectivo órgão, bem como de novos projetos em desenvolvimento ou em processo de aquisição; assim como avaliação de impacto pormenorizado de cada sistema. Aponta-se também a necessidade da criação de uma rede entre os diferentes comitês de éticas dos tribunais compartilhamento de conhecimento e experiências.

O Comitê Nacional de Ética de IA no Poder Judiciário, por sua vez, deve realizar uma visão geral sobre as soluções de IA empregadas em todo o país, subsidiando a atividade do Conselho Nacional de Justiça enquanto órgão regulador do Judiciário brasileiro. É o ambiente propício para entender o contexto macro do Poder Judiciário e potenciais riscos envolvidos, bem como avaliar a estrutura de governança e sugerir mudanças e soluções.

Enquanto os comitês de ética de IA de cada tribunal são órgãos internos, formado exclusivamente por membros integrantes do Poder Judiciário, com amplo acesso a todas as informações necessárias para o desempenho da função, o Comitê Nacional de Ética de IA é misto, composto por integrantes internos e externos, assegurando maior legitimidade e independência, combinando assim os benefícios de um comitê interno e externo.

O Conselho Consultivo Nacional de Ética de IA, por sua vez, é formado por membros externos e tem como objetivo assegurar a participação da sociedade na supervisão e no aperfeiçoamento das políticas, dos projetos e das diretrizes envolvendo uso de sistemas de IA

no Poder Judiciário brasileiro.

Pondera-se que a combinação de comitês internos e externos permite que se obtenha as vantagens de cada modelo e a criação de diferentes colegiados contribui para maiores chances de atingir os objetivos almejados, além de assegurar *accountability* e legitimidade.

Deve-se atentar, contudo, em todas as esferas, para o risco de que tais colegiados configurem mera estrutura de supervisão formal desvinculada dos objetivos para os quais foram instituídos, caracterizando apenas mais uma prática de *ethics washing*.

CONCLUSÃO

É incontestável que o Conselho Nacional de Justiça tem adotado a ampla utilização de IA no judiciário como política pública, estimulando a adoção de ferramentas tecnológicas a fim de melhorar a eficiência dos serviços prestados pelo judiciário.

Diante da proliferação de ferramentas em ampla utilização, há necessidade de desenvolvimento e implementação de um modelo de governança de IA no Judiciário para monitoramento dos riscos envolvidos e avaliação das implicações éticas, jurídicas e sociais do uso da tecnologia.

A experiência prática tem demonstrado que, não obstante o intuito da Resolução CNJ n. 332/2020 de assegurar o desenvolvimento colaborativo de modelos de IA e estabelecer mecanismos de governança no desenvolvimento e no uso de IA no Poder Judiciário, o que se verifica é uma atuação fragmentada e desgovernada.

Com o aumento do número de projetos, assim como a proliferação de ferramentas de inteligência artificial generativas disponíveis ao público em geral, os riscos envolvidos tendem a aumentar, o que torna ainda mais relevante a criação de estrutura destinada ao adequado direcionamento dos projetos e gerenciamento dos riscos envolvidos.

Nessa linha, o CNJ, enquanto órgão de coordenação de políticas judiciais, deve intervir e desenvolver um modelo de governança com estratégias adequadas ao gerenciamento e mitigação de riscos, assegurando também a efetiva implementação das medidas estabelecidas a fim de evitar que se configure mera prática de *ethics washing*.

O desenvolvimento de um modelo de governança também se mostra importante em um cenário em que são objeto de debate diversas propostas legislativas para regular o desenvolvimento e uso de IA no país, pois é importante no apoio à posterior adequação das instituições à futura regulação emanada do Poder Legislativo.

Ao longo do presente estudo, foram apontados diversos pontos de melhorias na Resolução n. 332/2020, bem como em experiências implementadas nos diferentes órgãos do Poder Judiciário, para construir uma estrutura de governança de IA adequada para direcionar e monitorar o uso de sistemas de IA no âmbito judicial brasileiro.

Promoveu-se breve panorama sobre o cenário regulatório acerca do tema, bem como sobre os princípios éticos da IA e o papel da governança em um contexto de desenvolvimento crescente de normas sobre o tema e aumento dos riscos do uso da tecnologia em diferentes contextos.

Realizou-se também análise sobre os potenciais impactos do uso da IA no sistema de justiça, concluindo para necessidade de se atentar que eventuais falhas ou abusos apresentam perigo não apenas no âmbito de processos ou indivíduos específicos, mas no próprio sistema e na ideia de Estado de Direito e do devido processo legal.

Foram examinados diferentes documentos internacionais que tratam do uso de IA nos sistemas de Justiça, com foco especial na Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes.

Constatou-se que a preocupação da CEPEJ com a avaliação e monitoramento contínuo da implementação dos princípios enunciados na referida Carta, o que se evidencia com a criação do Conselho Consultivo de IA do CEPEJ (*Artificial Intelligence Advisory Body - AIAB*), em 2022, bem como desenvolvimento de estudos acerca de um instrumento de avaliação da conformidade com a sua Carta Ética e um mecanismo de certificação de soluções de IA.

Apontou-se que apesar de toda a estrutura de governança em desenvolvimento para garantir observância dos princípios estabelecidos na Carta Ética do CEPEJ, os dados disponíveis indicam que não há, no âmbito dos países da União Europeia, ferramentas de IA empregadas efetivamente na atividade jurisdicional além de projetos em experimentação ou automação de atividades simples.

Destacou-se a importância da capacitação dos atores envolvidos no uso e implementação de ferramentas de IA no sistema de justiça, de modo a oferecer uma visão sobre os valores, princípios e orientações que fundamentam o uso ético e responsável da tecnologia, o que fica evidenciado com a iniciativa da UNESCO de elaborar o “*Global Toolkit on AI and the Rule of Law for the Judiciary*” (2023).

Ao se debruçar sobre a Resolução CNJ n. 332/2020, verificou-se que o documento não tratou da preocupação com as garantias processuais do contraditório, ampla defesa, independência e imparcialidade dos juízes, pontos abordados em detalhes no Apêndice I da Carta Ética do CEPEJ, assim como com o treinamento de magistrados e servidores acerca das capacidades e limitações da IA, bem como dos dilemas éticos envolvidos.

Destacou-se que, apesar da previsão abstrata de alguns mecanismos de governança como a realização de auditoria e certificação de boas práticas, não houve preocupação com a implementação de tais instrumentos, assim como com a previsão de prévia avaliação de impacto das diferentes ferramentas de IA no sistema de justiça antes do seu desenvolvimento e efetiva utilização. Também foi indicado que, na Resolução CNJ n. 332/2020, não foi endereçada a necessidade de monitoramento contínuo dos riscos envolvidos durante a fase de

operacionalização das ferramentas ou de controle acerca da efetiva implementação de suas regras e/ou avaliação da conformidade com as diretrizes estabelecidas.

Nessa perspectiva, apontou-se que, na prática, não obstante a importância da Resolução CNJ n. 332/2020, o referido documento assemelha-se mais a carta de intenções do que ato normativo cogente suficiente para endereçar os riscos envolvidos.

Ressaltou-se a necessidade de uma abordagem mais proativa em relação à transparência. Foi apresentada como sugestão a divulgação pelo CNJ da lista de todos os modelos hospedados na plataforma Sinapses disponíveis para utilização, destacando os objetivos alcançados, assim como publicação de sumários executivos ou versões resumidas das avaliações realizadas por meio da plataforma. Essa prática proporcionaria transparência, reforçando a prestação de contas à sociedade (*accountability*).

Apontou-se a insuficiência da política de transparência implementada pelo “Painel de Projetos com Inteligência Artificial no Poder Judiciário” em sua atualização de maio de 2022, bem como a necessidade de regulamentação para que seja implementada a previsão do art. 27 da Resolução n. 332/2020, que trata da comunicação de eventos adversos.

Indicou-se que a iniciativa de compilar relatos de eventos adversos pode contribuir na identificação de padrões, prevenção de acontecimentos semelhantes no futuro e melhoria contínua dos sistemas de IA.

Sugeriu-se a criação de um banco de dados de monitoramento de incidentes de IA, de comunicação obrigatória para usuários internos (servidores e magistrados) e facultativa para usuários externos (advogados, membros das funções essenciais à Justiça e jurisdicionados em geral), a fim de compilar ocorrência de incidentes para tratamento e análise por parte do Conselho Nacional de Justiça.

Salientou-se a importância de adotar medidas concretas para implementar a certificação de boas práticas, ferramenta prevista na Resolução n. 332/2020, que promove transparência e confiança nos sistemas de IA.

Apontou-se que a ferramenta de avaliação do impacto, voltada a identificar e avaliar os benefícios, as preocupações e os riscos dos sistemas de IA, bem como medidas adequadas de prevenção, mitigação e monitoramento de riscos, é de suma importância no contexto do Poder Judiciário. Assim, sugeriu-se a inclusão de tal estratégia de governança no âmbito do sistema de justiça brasileiro.

Observou-se que as questões éticas envolvidas na utilização de sistemas de IA no Judiciário não estão na pauta dos debates do Conselho Nacional de Justiça ou dos Tribunais.

Demonstrou-se que a criação de comitês de ética de IA ou órgão semelhante permite: promover debates sobre considerações éticas sobre o uso da tecnologia nas organizações, identificar problemas; desenvolver diretrizes a serem observadas; promover transparência e *accountability*, assim como promover comunicação entre as partes interessadas; desenvolver práticas e treinamento para tratar de questões éticas; promover uma cultura organizacional que valorize o uso da IA de maneira ética e responsável.

Defendeu-se a criação de comitês de ética de IA no âmbito dos tribunais como estratégia de governança para contribuir na identificação e mitigação dos riscos envolvidos no uso e desenvolvimento de ferramentas de IA no Judiciário brasileiro.

Foi apresentada uma proposta de criação não apenas de comitês de ética de IA no âmbito de cada tribunal, mas também de uma rede nacional de comitês de ética de IA do Poder Judiciário, um comitê nacional de ética de IA e um conselho consultivo nacional de ética de IA, com detalhes sobre sua composição, objetivos e atribuições.

Alertou-se para o risco de que a criação e implementação de tais colegiados caracterize apenas mais uma prática de *ethics washing*. Ressaltou-se a necessidade de realizar a adequada capacitação dos membros do comitê de ética dos tribunais para que desenvolvam habilidades para lidar com os conflitos de valores envolvidos na implementação de tecnologia na atividade jurisdicional.

Apontou-se também a necessidade de transparência sobre as operações do comitê de ética, ainda que eventualmente certas avaliações ou recomendações possam não ser divulgadas para o público, sendo essencial que haja, ao menos, transparência sobre a política de transparência adotada quanto às atividades dos referidos colegiados.

Nesse contexto, longe de esgotar o tema de elevada relevância, o presente trabalho contribuiu para o estudo das estratégias de governança do uso de IA no Poder Judiciário brasileiro, apresentando sugestões de melhorias em diferentes mecanismos previstos na Resolução n. 332/2020, bem como apontado a implementação de outros mecanismos capazes de contribuir para uma estrutura de governança sólida.

REFERÊNCIAS

- ADA LOVELACE INSTITUTE; AI NOW INSTITUTE; OPEN GOVERNMENT PARTNERSHIP. [2021]. **Algorithmic Accountability for the Public Sector**. Disponível em: <https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/>. Acesso em: 13 ago. 2023.
- AI VERIFY FOUNDATION. **Model Artificial Intelligence Governance Framework**. Second Edition. 21 jan. 2020. Disponível em: <http://go.gov.sg/ai-gov-mf-2>. Acesso em: 10 fev. 2024.
- AI VERIFY FOUNDATION. **Proposed Model AI Governance Framework for Generative AI: Fostering a Trusted Ecosystem**. 16 jan. 2024. Disponível em: <https://aiverifyfoundation.sg/news-and-resources/#proposed-model-governance-framework-for-generative-ai>. Acesso em: 13 jan. 2024.
- ALA-PIETILÄ, Pekka et al. **The assessment list for trustworthy artificial intelligence (ALTAI)**. European Commission, 2020.
- ARÁNGUIZ VILLAGRÁN, Matías. **Auditoria algorítmica para sistemas de tomada de decisão ou suporte à decisão**. Apr. 2022. Disponível em: <http://dx.doi.org/10.18235/0004154>. Acesso em: 15 jul. 2023.
- ARANHA, Márcio Iorio. **Manual de Direito Regulatório: Fundamentos de Direito Regulatório**. 3ª ed., London: Laccademia Publishing, 2015.
- ARKOUDAS, Konstantine. GPT-4 Can't Reason. **ArXiv Preprints**, 2023. Disponível em: <https://doi.org/10.20944/preprints202308.0148.v2>. Acesso em: 13 ago. 2023.
- BAHIA, Alexandre; NUNES, Dierle; PEDRON, Flávio. **Teoria Geral do Processo**. Salvador: JusPodivm, p. 144, 2021.
- BENDER, Emily M. et al. On the dangers of stochastic parrots: Can language models be too big?. In: **Proceedings of the 2021 ACM conference on fairness, accountability, and transparency**. 2021. p. 610-623.
- BLACKMAN, Reid. **Ethical Machines: Your Concise Guide to Totally Unbiased, Transparent, and Respectful AI**. Harvard Business Press, 2022.
- BRADFORD, Anu. **The Brussels effect: How the European Union rules the world**. Oxford University Press, USA, 2020.
- BRASIL. Senado Federal. **Projeto de Lei nº 21, de 2020**: Relatório final da Comissão de Juristas responsável por subsidiar a elaboração de substitutivo sobre inteligência artificial. Diário do Senado Federal, ano 77, n. 204, suplemento n. B, p. 3-915, 9 dez. 2022.
- BRASIL. Senado Federal. **Projeto de Lei nº 2338, de 4 de maio de 2023**. Dispõe sobre o uso da Inteligência Artificial. Brasília: Senado Federal, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 17 jan. 2024.

BRASIL. Ministério da Ciência, Tecnologia e Inovações. **Estratégia Brasileira de Inteligência Artificial**. 6 abr. 2021. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligencia-artificial-estrategia-repositorio>. Acesso em: 17 nov. 2023.

BRASIL. Tribunal de Contas da União. **Referencial básico de governança aplicável a organizações públicas e outros entes jurisdicionados ao TCU**. 3 ed. Brasília: TCU, Secretaria de Controle Externo da Administração do Estado, 2020.

BROWN, Shea; DAVIDOVIC, Jovana; HASAN, Ali. The algorithm audit: Scoring the algorithms that score us. **Big Data & Society**, v. 8, n. 1, p. 2053951720983865, 2021. Acesso em: 20 jul. 2023.

COUNCIL OF EUROPE. CEPEJ. **European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment**. Estrasburgo, 3 dec. 2018. Disponível em <https://rm.coe.int/ethicalcharter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 12 ago. 2023.

COUNCIL OF EUROPE. CEPEJ. **Possible introduction of a mechanism for certifying artificial intelligence tools and services in the sphere of justice and the judiciary: Feasibility Study**. 8 dec. 2020. Disponível em: <https://rm.coe.int/feasability-study-en-cepej-2020-15/1680a0adf4>. Acesso em: 15 nov. 2023.

COUNCIL OF EUROPE. CEPEJ. **Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment**. 9 dez. 2021. Disponível em: <https://rm.coe.int/cepej-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 15 nov. 2023.

CHEN, Lingjiao; ZAHARIA, Matei; ZOU, James. **How is ChatGPT's behavior changing over time?**. 31 oct. 2023. Disponível em: arXiv preprint arXiv:2307.09009, 2023. Acesso em: 5 ago. 2023.

COELHO, Alexandre Zavaglia et al. **Governança da inteligência artificial em organizações: framework para comitês de ética em IA: versão 1.0**. CEPI FGV Direito SP, 2023.

COMISSÃO EUROPEIA. Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões: Inteligência artificial para a Europa. Bruxelas: 2018. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A52018DC0237>. Acesso em: 2 abr. 2022.

CONSELHO NACIONAL DE JUSTIÇA. Delegação da União Europeia no Brasil. Intercâmbio de experiências entre a União Europeia e o Brasil sobre e-Justice: relatório final. Brasília: CNJ, 2022. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/09/seminario-e-justice-v6.pdf>. Acesso em: 24 jan. 2023.

CONSELHO NACIONAL DE JUSTIÇA. **Resolução n. 332, de 21 de agosto de 2020**. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. DJe/CNJ, nº 274, de 25/08/2020, p. 4-8. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 12 fev. 2024.

CONSELHO NACIONAL DE JUSTIÇA. **Resolução n. 370, de 28 de janeiro de 2021.** Estabelece a Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário (ENTIC-JUD). DJe/CNJ nº 158/2021, de 22 de junho de 2021, p. 2-11 (republicação). Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3706>. Acesso em: 12 fev. 2024.

CONSELHO NACIONAL DE JUSTIÇA. **Resolução n. 395, de 07 de junho de 2021.** Institui a Política de Gestão da Inovação no âmbito do Poder Judiciário. DJe/CNJ nº 148/2021, de 9 de junho de 2021, p. 3-7. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 12 fev. 2024.

CORRÊA, Nicholas Kluge; DE OLIVEIRA, Nythamar Fernandes; MASSMANN, Diogo Fernando. Sobre a eficiência da ética como ferramenta de governança da inteligência artificial. **Veritas (Porto Alegre)**, v. 67, n. 1, p. e42584-e42584, 2022.

COURT OF JUSTICE OF THE EUROPEAN UNION. **Artificial Intelligence Strategy.** Nov, 2023. Disponível em: https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf. Acesso em: 23 jan. 2024.

CUKIER, Kenneth; MAYER-SCHÖNBERGER, Viktor; DE-VERICOURT, Francis. **Framers: A vantagem humana em uma era de tecnologia e instabilidade.** 2023. Ebook.

DAHL, Matthew et al. Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models. Disponível em: arXiv preprint arXiv:2401.01301, 2024.

DAVIES, Matt; JONES, Elliot. **Foundation models in the public sector.** Ada Lovelace Institute. 2023. Disponível em: <https://www.adalovelaceinstitute.org/evidence-review/foundation-models-public-sector/>. Acesso em: 17 mar. 2024.

DE ALBUQUERQUE MARANHÃO, Juliano Souza; JUNQUILHO, Tainá Aguiar; TASSO, Fernando Antonio. Transparência sobre o emprego de Inteligência Artificial no Judiciário: um modelo de governança. **Suprema-Revista de Estudos Constitucionais**, v. 3, n. 2, p. 145-187, 2023.

DE ALBUQUERQUE, Adriana Reis. **Poder Artificial de Tributar: Limites e Requisitos à Utilização (Adequada) da Inteligência Artificial pela Administração Tributária.** Editora Thoth, 2022.

FJELD, Jessica et al. Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. **Berkman Klein Center Research Publication**, n. 2020-1, 2020.

FLORIDI, Luciano. Soft Ethics and the Governance of the Digital. **Philosophy & Technology**, v. 31, n. 1, p. 1–8, 2018. Disponível em: <https://doi.org/10.1007/s13347-018-0303-9>. Acesso em: 12 ago. 2023.

FLORIDI, Luciano. The end of an era: from self-regulation to hard law for the digital industry. **Philosophy & Technology**, v. 34, n. 4, p. 619-622, 2021.

FLORIDI, Luciano. Translating principles into practices of digital ethics: Five risks of being unethical. **Philosophy & Technology**, v. 32, n. 2, p. 185-193, 2019.

FLORIDI, Luciano; COWLS, Josh. A Unified framework of five principles for AI In Society. **Harvard Data Science Review**, Issue 1.1, Summer 2019.

FRAZÃO, Ana. Discriminação algorítmica: resgatando os aspectos positivos dos julgamentos humanos. A importância das narrativas, das contextualizações e das molduras cognitivas. **JOTA**, 1 out. 2021. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica-resgatando-os-aspectos-positivos-dos-julgamentos-humanos-01092021>. Acesso em: 17 jan. 2024.

GENTILE, Giulia. **AI in the courtroom and judicial independence: An EU perspective**. 22 aug. 2022. Disponível em: <https://eideas.eui.eu/2022/08/22/ai-in-the-courtroom-and-judicial-independence-an-eu-perspective/> Acesso em: 15 maio. 2023.

GREEN, Ben; CHEN, Yiling. Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments. In: **Proceedings of the conference on fairness, accountability, and transparency**. 2019. p. 90-99. Disponível em: <https://scholar.harvard.edu/bgreen/publications/disparate-interactions-algorithm-loop-analysis-fairness-risk-assessments>. Acesso em: 13 maio 2023.

GUALTIERI, André; LÓPEZ, Nuria. Decisões automatizadas: auditoria sobre aspectos discriminatórios do art. 20 da LGPD. In: PALHARES, Felipe; FRANCOSKI, Denise de Souza Luiz. Coord. Temas Atuais de Direito Digital. São Paulo: **Revista dos Tribunais**, 2023.

IEEE. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. **Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems**, Version 2. IEEE, 2017. Disponível em: http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html. Acesso em: 17 maio 2023.

JUDICIARY OF UK. Artificial Intelligence (AI) – Judicial Guidance. **Courts and Tribunals Judiciary**, 12 dec. 2023. Disponível em: https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance/#related_content. Acesso em: 17 fev. 2024.

KAHNEMAN, Daniel; SIBONY, Olivier; SUNSTEIN, Cass R. **Ruído: uma falha no julgamento humano**. Rio de Janeiro: Objetiva, 2021.

KAUFMAN, Dora; JUNQUILHO, Tainá; REIS, Priscila. Externalidades negativas da inteligência artificial: conflitos entre limites da técnica e direitos humanos. **Revista de Direitos e Garantias Fundamentais**, v. 24, n. 3, p. 43-71, 2023. Disponível em: <https://doi.org/10.18759/rdgf.v24i3.2198>. Acesso em: 7 jan. 2024.

LABORATÓRIO DE INTELIGÊNCIA ARTIFICIAL APLICADA DA 3ª REGIÃO (LIAA-3R). **Diretrizes de auditabilidade e conformidade no desenvolvimento e testes de soluções de IA no âmbito do LIAA-3R**. Grupo de Validação Ético-Jurídica (GVEJ) do LIAA-3R. 2. ed., rev. e atual. São Paulo: LIAA-3R, 2022.

LAM, Khoa et al. A Framework for Assurance Audits of Algorithmic Systems. Disponível em: **arXiv preprint arXiv:2401.14908**, 2024. Acesso em: 20 mar. 2024.

LARSON, Jeft et al. How we analyzed the COMPAS recidivism algorithm. [S. L.]: **ProPublica**, 23 may. 2016. Disponível em: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Acesso em: 7 de jan. 2024.

LEITÃO, Emiliano Zapata. “Painel de Projetos com Inteligência Artificial no Poder Judiciário” do Conselho Nacional de Justiça (CNJ): avaliação qualitativa da transparência governamental. 2022. 129 f. Dissertação (Mestrado em Direito). **Escola Nacional de Formação e Aperfeiçoamento de Magistrados**: Brasília, 2022.

LESLIE, David. Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. **The Alan Turing Institute**, 2019. Disponível em: <https://doi.org/10.5281/zenodo.3240529>. Acesso em: 5 mar. 2024.

LUNARDI, Fabrício Castagna; CLEMENTINO, Marco Bruno Miranda. **Inovação Judicial: fundamentos e práticas para uma jurisdição de alto impacto**. ENFAM, Escola Nacional de Formação e Aperfeiçoamento de Magistrados, 2021.

LUPO, Giampiero. The Ethics of Artificial Intelligence: An Analysis of Ethical Frameworks Disciplining AI in Justice and Other Contexts of Application. **Oñati Socio-Legal Series**, vol. 12, n. 3, Junho 2022, pp. 614-53. Disponível em: <https://opo.iisj.net/index.php/osls/article/view/1366>. Acesso em: 10 mar. 2024.

MARCHANT, Gary E.; GUTIERREZ, Carlos Ignacio. Soft law 2.0: An agile and effective governance approach for artificial intelligence. **Minn. JL Sci. & Tech.**, v. 24, p. 375, 2022.

MCGREGOR, Sean. Preventing repeated real world AI failures by cataloging incidents: The AI incident database. In: **Proceedings of the AAAI Conference on Artificial Intelligence**. 2021. p. 15458-15463. Disponível em: <https://ojs.aaai.org/index.php/AAAI/article/view/17817>. Acesso em: 18 jul. 2022.

MCLENNAN, Marsh. **The Global Risks Report**. Wolrd Economic Forum, 15 jan. 2024. Disponível em: <https://www.weforum.org/publications/global-risks-report-2024/>.

MELO, Ana Karolina Acris et al. Regulação da Inteligência Artificial: benchmarking de países selecionados. **Enap**, dez. 2022. Disponível em: <http://repositorio.enap.gov.br/handle/1/7419>. Acesso em: 10 mar. 2024.

MITTELSTADT, Brent. Principles alone cannot guarantee ethical AI. **Nature machine intelligence**, v. 1, n. 11, p. 501-507, 2019.

MORAIS, Fausto Santos de. O Uso da Inteligência Artificial na Repercussão Geral: Desafios Teóricos e Éticos. **Revista Direito Público**, Brasília, vol. 18, n. 100, p. 306-326, out./dez., 2021.

MORLEY, Jessica, et al. Ethics as a service: A pragmatic operationalization of AI ethics. **Minds and Machines**, v. 31, n. 2, p. 239-256, 2021. Disponível em: <https://doi.org/10.1007/s11023-021-09563-w>. Acesso em: 10 mar. 2024.

MULHOLLAND, Caitlin Sampaio. Mercado, Pessoa Humana e Tecnologias: a Internet das Coisas e a proteção do direito à privacidade. In: EHRHARDT JÚNIOR Marcos; CORTIANO

JUNIOR Eroulths (coord.). **Transformações no direito privado nos 30 anos da Constituição:** estudos em homenagem a Luiz Edson Fachin. Belo Horizonte: Fórum, 2019, p. 103-116.

NUNES, Dierle; LUD, Nathanael; PEDRON, Flávio. **Desconfiando da imparcialidade dos sujeitos processuais.** Salvador: JusPodivm, 2018.

NUNES, Dierle; MARQUES, Ana Luiza Pinto Coelho. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. In: **Revista de Processo.** 2018. p. 421-447.

OECD. Framework for the classification of AI systems. **OECD Digital Economy Papers**, n. 323, Paris, 2022. Disponível em: <https://doi.org/10.1787/cb6d9eca-en>. Acesso em: 11 nov. 2023.

OECD. Recommendation of the Council on Artificial Intelligence. **OECD Legal Instruments**, 7 nov. 2023. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#relatedInstruments>. Acesso em: 11 nov. 2023.

PACHECO, Alexandre et al. **Ética, Governança e Inteligência Artificial.** Almedina Brasil, 2023.

PAULINO, Ana Flávia Borges. **CNJ: o regulador da atividade judiciária.** Rio de Janeiro: Lumen Iuris, 2021.

PEIXOTO, Fabiano Hartmann. **Inteligência artificial e direito: convergência ética e estratégica.** Alteridade Editora, 2020.

PETERMANN, M. et al. Looking before we leap: Expanding ethical review processes for AI and data science research. **Ada Lovelace Institute**, 13 dec. 2022. Disponível em: <https://www.adalovelaceinstitute.org/report/looking-before-we-leap> (2022). Acesso em: 8 jan. 2024.

PRADO, Eunice Maria Batista. Banco de incidentes de inteligência artificial do judiciário. 2022. 262 f. Dissertação (Mestrado em Direito). **Escola Nacional de Formação e Aperfeiçoamento de Magistrados:** Brasília, 2022.

RUSSEL, Stuart. et al. Updates to the OECD's definition of an AI system explained. **OECD AI Policy Observatory**, 29 nov. 2023. Disponível em: <https://oecd.ai/en/wonk/ai-system-definition-update>. Acesso em: 7 dez. 2024.

SANDLER, Ronald; BASL, John; TIELL, Steven. Building data and AI ethics committees. **Accenture & Northeastern University**, 2019. Disponível em: https://ethics.harvard.edu/sites/hwpi.harvard.edu/files/center-for-ethics/files/811330-ai-data-ethics-committee-report_v10.0.pdf?m=1567167985. Acesso em: 12 fev. 2024.

UNESCO. **Recomendação sobre a Ética da Inteligência Artificial.** 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 23 nov. 2023.

WORLD ECONOMIC FORUM. **Empowering AI Leadership: AI C-Suite Toolkit.** World Economic Forum, 12 jan. 2022. Disponível em: <https://www.weforum.org/publications/empowering-ai-leadership-ai-c-suite-toolkit/>. Acesso em: 15 jul. 2023.