



Transparência sobre o emprego de Inteligência Artificial no Judiciário: um modelo de governança¹

Transparency on the use of Artificial Intelligence in the Judiciary: a governance framework

Transparencia sobre el uso de la Inteligencia Artificial en el Poder Judicial: un modelo de gobernanza

Juliano Souza de Albuquerque Maranhão²

Universidade de São Paulo (São Paulo, SP, Brasil)
ORCID: <https://orcid.org/0000-0002-2705-7440>
E-mail: julianomaranhao@usp.br

Tainá Aguiar Junquillo³

Instituto de Direito Público (Brasília, DF, Brasil)
ORCID: <https://orcid.org/0000-0001-8638-0670>
E-mail: taina.aguiarj@gmail.com

Fernando Antonio Tasso⁴

Universidade Nove de Julho (São Paulo, SP, Brasil)
ORCID: <https://orcid.org/0009-0004-0977-5731>
E-mail: ftasso@tjsp.jus.br

¹MARANHÃO, Juliano Souza de Albuquerque; JUNQUILHO, Tainá Aguiar; TASSO, Fernando Antonio. Transparência sobre o emprego de inteligência artificial no Judiciário: um modelo de governança. **Suprema**: revista de estudos constitucionais, Brasília, v. 3, n. 2, p. 145-187, jul./dez. 2023. DOI: <https://doi.org/10.53798/suprema.2023.v3.n2.a231>.

²Professor Associado da Faculdade de Direito da USP, Diretor da Associação *Lawgorithm* de Pesquisa em Inteligência Artificial, Pesquisador Associado do Centro de Inteligência Artificial USP-IBM e da Fundação Alexander von Humboldt – Alemanha, Diretor da *International Association on Artificial Intelligence and Law*. Currículo Lattes: <http://lattes.cnpq.br/3498148304153540>.

³Doutora em Direito com ênfase em Inteligência Artificial pela Universidade de Brasília. Advogada, CEO da Sentelia e Pesquisadora. Professora de Direito, Inovação e Tecnologia no mestrado do IDP. Foi bolsista da FINATEC no Projeto de Pesquisa & Desenvolvimento de aprendizado de máquina (*machine learning*) sobre dados judiciais das repercussões gerais do Supremo Tribunal Federal – STF (Projeto Victor). Vice Líder do Grupo de Pesquisa Observatório de Políticas Públicas (GEOPP) UnB. Foi professora substituta de Direito empresarial na Universidade de Brasília de 2019 a 2021. Mestra em Direito pela Universidade Federal do Espírito Santo (UFES). Currículo Lattes: <http://lattes.cnpq.br/5848504606151120>.

⁴Doutor em Direito Empresarial pela Universidade Nove de Julho. MBA em Gestão e Governança de Tecnologia da Informação pela Faculdade de Informática e Administração Paulista. Membro da Comissão de Proteção de Dados

Resumo

O presente artigo propõe, com base em revisão bibliográfica de matriz de transparência apresentado por Kaminski⁵, um modelo de governança da transparência no Judiciário brasileiro sobre sistemas de Inteligência Artificial (como sistemas sociotécnicos), a partir da segmentação do objeto da transparência em: transparência sobre o uso, a operação e os benefícios e riscos do sistema de Inteligência Artificial (IA). O modelo especifica quais seriam as questões-chave e indicações para definir o conteúdo informativo relevante e adequado para diferentes contextos e interlocutores, tanto participantes internos (servidores e magistrados, servidores de Tecnologia da Informação (TI), etc.) quanto destinatários externos (cidadãos, advogados, etc.). A matriz apresentada, além de possibilitar controle público democrático dos projetos de IA, serve de base para estudos empíricos sobre a transparência dos Tribunais em relação às ferramentas empregadas. Os Tribunais também poderão, a partir do modelo criado neste artigo, desenvolver políticas de governança de transparência internos e aplicáveis aos sistemas de IA que utiliza.

Palavras-chave

Inteligência artificial; transparência; governança; regulação.

Sumário

1. Introdução. 2 Metodologia para construção do modelo de governança de transparência em organizações. 3. Estrutura de gestão da tecnologia no Poder Judiciário: os Tribunais e o CNJ. 4. Aplicações de Inteligência Artificial nos Tribunais. 5. Matriz de transparência para IA no Judiciário. 6. Conclusão.

Abstract

This article proposes, based on a bibliographic review of the transparency matrix presented by Kaminski, a transparency governance model in the Brazilian Judiciary on Artificial Intelligence systems (such as socio-technical systems), based on the segmentation of the transparency object into: transparency about the use, operation and benefits and risks of the Artificial Intelligence (AI) system. The

da Corregedoria Nacional de Justiça. Membro do Comitê Consultivo de Dados e Proteção de Dados Pessoais do Conselho Nacional de Justiça. Membro do Grupo de Trabalho sobre inteligência artificial no Poder Judiciário. Coordenador de Direito Digital da Escola Paulista da Magistratura. Juiz de Direito no Tribunal de Justiça de São Paulo. Currículo Lattes: <https://lattes.cnpq.br/9827320445786010>.

⁵ KAMINSKI, Margot E. Understanding transparency in algorithmic accountability. In: BARFIEL, Woodrow (ed.). **Forthcoming in Cambridge handbook of the law of algorithms**. Cambridge, UK: Cambridge University Press, 2021. p. 121-140. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3622657. Acesso em: 3 out. 2023.

model specifies what the key questions and indications would be to define relevant and appropriate information content for different contexts and interlocutors, both internal participants (servants and magistrates, Information Technology (IT) servers, etc.) and external recipients (citizens, lawyers, etc.). The matrix presented, in addition to enabling democratic public control of AI projects, serves as a basis for empirical studies on the transparency of the Courts in relation to the tools used. Courts will also be able, based on the model created in this article, to develop internal transparency governance policies applicable to the AI systems they use.

Keywords

Artificial Intelligence; Transparency; Governance; Regulation.

Contents

1. Introduction. 2. Methodology for building a transparency governance model in organizations. 3. Technology management structure in the Judiciary: the courts and the CNJ. 4. Applications of Artificial Intelligence in Courts. 5. Transparency matrix for AI in the Judiciary. 6. Conclusion.

Resumen

Este artículo propone, a partir de una revisión bibliográfica de la matriz de transparencia presentada por Kaminski, un modelo de gobernanza de la transparencia en el Poder Judicial brasileño sobre sistemas de Inteligencia Artificial (como los sistemas sociotécnicos), basado en la segmentación del objeto de transparencia en: transparencia sobre el uso, funcionamiento y beneficios y riesgos del sistema de Inteligencia Artificial (IA). El modelo especifica cuáles serían las preguntas e indicaciones clave para definir contenidos informativos relevantes y apropiados para diferentes contextos e interlocutores, tanto participantes internos (funcionarios y magistrados, servidores de Tecnologías de la Información (TI), etc.) como destinatarios externos (ciudadanos, abogados, etc.). La matriz presentada, además de permitir el control público democrático de los proyectos de IA, sirve de base para estudios empíricos sobre la transparencia de los Tribunales en relación con las herramientas utilizadas. Los tribunales también podrán, basándose en el modelo creado en este artículo, desarrollar políticas internas de gobernanza de la transparencia aplicables a los sistemas de IA que utilicen.

Palabras clave

Inteligencia artificial; Transparencia; Gobernanza; Regulación.

Índice

1. Introdução. 2 Metodología para la construcción de un modelo de gobernanza de la transparencia en las organizaciones. 3. Estructura de gestión de la tecnología en el Poder Judicial: los tribunales y el CNJ. 4. Aplicaciones de la Inteligencia Artificial en los Tribunales. 5. Matriz de transparencia de la IA en el Poder Judicial. 6. Conclusión.

1. Introdução

Os riscos provenientes da aplicação de sistemas de inteligência artificial (IA) já foram objeto de uma literatura significativa. A bibliografia produzida traz propostas de soluções técnicas e organizacionais, envolvendo os agentes humanos que participam do desenvolvimento, emprego e monitoramento do sistema, bem como sugestões de regulação, seja exógena, imposta pelo Estado, seja interna, com modelos de autorregulação, seja por meio de arranjos institucionais híbridos, em que há participação dos regulados na formulação das políticas públicas e no debate sobre o conteúdo da regulação, ou em que há a formulação de códigos de conduta reconhecidos pelo Estado por meio de procedimentos de legitimação institucional.⁶

O primeiro movimento no sentido de regulação, para mitigar riscos inerentes a algumas metodologias de IA, centrou-se no debate e elenco de valores éticos que deveriam pautar o desenvolvimento e emprego dessa tecnologia. Nessa linha, foram divulgados diversos documentos por empresas, entidades governamentais e não governamentais, no Brasil e no exterior, propondo ou adotando códigos de conduta, com base em princípios éticos. Embora possa haver algum consenso quanto a valores abstratos, observam-se, naqueles documentos, divergências quanto a seu alcance e implementação prática.⁷

⁶Veja-se por exemplo, DIGNUM, Virginia. **Responsible artificial intelligence**: how to develop and use AI in a responsible way. Cham, Switzerland: Springer Nature, 2019; Artificial intelligence: foundations, theory, and algorithms); FLORIDI, Luciano *et al.* AI4people – an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. **Minds and Machines**, v. 28, n. 4, p. 689 –707, Nov. 2018. Disponível em: <https://link.springer.com/article/10.1007/s11023-018-9482-5>. Acesso em: 3 out. 2023; DIGNUM, Virginia. Ethics in artificial intelligence: introduction to the special issue. **Ethics and Information Technology**, v. 20, n. 1, p. 1 –3, 2018. Disponível em: <https://link.springer.com/article/10.1007/s10676-018-9450-z>. Acesso em: 3 out. 2023; COECKELBERGH, M. **AI ethics**. Cambridge, MA: MIT Press, 2020.

⁷JOBIN, Anna; IENCA, Marcello; VAYENA, Effy. The global landscape of AI ethics guidelines. **Nature Machine Intelligence**, v. 1, p. 389, 2019. Disponível em: <https://www.nature.com/articles/s42256-019-0088-2>. Acesso em: 3 out. 2023.

Esse fator, aliado ao questionamento sobre a legitimidade e efetiva implementação por empresas de códigos abstratos de autorregulação, tem levado, de um lado, a literatura a criticar modelos de *soft law* para a inteligência artificial como inadequados⁸ e, de outro, a elaboração de documentos por entidades técnicas (e.g. *ISO* e *Institute of Electric and Electronic Engineers-IEEE*), reunindo melhores práticas de governança, com procedimentos concretos e auditáveis a serem adotados,⁹ bem como projetos de regulação com obrigações procedimentais efetivas, como é o caso do *AI Act Europeu*.¹⁰

Entre os diversos documentos que enumeram princípios, diretrizes e valores para o desenvolvimento responsável dessa tecnologia, a transparência, além de ser nota comum, tem um papel central em relação aos demais valores elencados. Nas publicações por instituições de referência em governança, a transparência também se pronuncia com destaque, merecendo inclusive documentos exclusivos para sua promoção, tanto do ponto de vista técnico, quanto organizacional. No Brasil, publicou-se, em dezembro de 2022, relatório substitutivo ao Projeto de Lei n. 21/2020, elaborado por Comissão de Juristas do Senado, que prevê, como medida de governança para quaisquer aplicações de IA, a transparência, tanto na

⁸ MITTELSTADT, Brent. Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, v. 1, p. 501–507, Nov. 2019. Disponível em: <https://www.nature.com/articles/s42256-019-0114-4>. Acesso em: 5 out. 2023; FLORIDI, Luciano. The end of an era: from self-regulation to hard law in the digital industry. *Philosophy & Technology*, v. 34, n.4, p. 619–622, Dec. 2021. Disponível em: <https://link.springer.com/article/10.1007/s13347-021-00493-0>. Acesso em: 5 out. 2023.

⁹ Ver, por exemplo: INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC TR 24027:2021**: information technology – artificial intelligence (AI) – bias in AI systems and AI aided decision making. Geneva, Switzerland: ISO, 2021. Disponível em: <https://www.iso.org/standard/77607.html>. Acesso em: 4 out. 2023; INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC 38507:2022**: information technology – governance of IT – governance implications of the use of artificial intelligence by organizations. Disponível em: <https://www.iso.org/standard/56641.html>. Acesso em: 4 out. 2023; INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE P2863**: recommended practice for organizational governance of artificial intelligence. New York: IEEE, 2020. Disponível em: <https://standards.ieee.org/ieee/2863/10142/>. Acesso em: 4 out. 2023; INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE P2894**: IEEE draft guide for an architectural framework for explainable artificial intelligence. New York: IEEE, 2023. Disponível em: <https://standards.ieee.org/ieee/2894/11296/>. Acesso em: 4 out. 2023; INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE 7000-2021**: IEEE model process for addressing ethical concerns during system design. New York: IEEE, 2021. Disponível em: <https://standards.ieee.org/ieee/24748-7000/11098/7000/6781/>. Acesso em: 4 out. 2023; INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE 7001-2021**: IEEE standard for transparency of autonomous systems. New York: IEEE, 2021. Disponível em: <https://standards.ieee.org/ieee/7001/6929/>. Acesso em: 4 out. 2023; INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE P7003**: algorithmic bias considerations. New York: IEEE, 2023. Disponível em: <https://standards.ieee.org/ieee/7003/11357/>. Acesso em: 4 out. 2023; FLORIDI, Luciano *et al.* **capAI**: a procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act. Oxford, UK: Saïd Business School, University of Oxford, 2022. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4064091. Acesso em: 4 out. 2023.

¹⁰ UNIÃO EUROPEIA. Comissão Europeia. **The Artificial Intelligence Act**. Bruxelas: EU, 2021. Disponível em: <https://artificialintelligenceact.eu/>. Acesso em: 2 abr. 2023.

interface humano-máquina, quanto no que tange ao uso e à clareza e poder informativo das medidas de governança adotadas no desenvolvimento dos projetos.¹¹

Tal ênfase ocorre porque a transparência é elemento fundamental para que terceiros, sejam aqueles que operam o sistema, sejam aqueles que são diretamente afetados pelo sistema, bem como as autoridades envolvidas na regulação e adjudicação de conflitos que abarquem a IA, possam obter elementos críticos para sua avaliação ética, seus impactos sociais e econômicos e suas consequências jurídicas. Ou seja, a transparência é veículo social de prestação de contas (*accountability*) e de autonomia dos humanos envolvidos e afetados pela operação da IA, fatores cruciais para o desenvolvimento e emprego democrático dessa tecnologia.¹² A autonomia dos humanos envolvidos, que interagem ou que são impactados pela IA é central para a construção de confiança da sociedade na tecnologia, dentro do que se convencionou chamar de *IA confiável* ou *responsável*, consubstanciada em uma série de valores como: confiabilidade, não discriminação (equidade/justiça), proteção de dados pessoais e responsabilidade.

Além de valor democrático intrínseco, a transparência é instrumental em relação à promoção daqueles valores que consubstanciam a IA confiável.¹³ Assim, a transparência em relação aos dados coletados e seu processamento pode trazer instrumentos para a avaliação do seu impacto sobre a privacidade e proteção de dados pessoais. Informações detalhadas sobre o domínio de aplicação e populações afetadas, bem como representatividade, trazem elementos para avaliação de possíveis vieses. Informações sobre níveis de acurácia permitem a adoção de medidas para mitigar riscos de erro ou medidas organizacionais para possibilitar revisões. Informações sobre os centros de decisão ao longo do ciclo de vida da IA trazem elementos relevantes para a apuração de responsabilidade em caso de danos.

¹¹ O relatório final da Comissão está disponível em: BRASIL. Congresso Nacional. Senado Federal. Projeto de Lei n° 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. Disponível em: https://legis.senado.leg.br/diarios/BuscaDiario?tipDiario=1&datDiario=09/12/2022&paginaDireta=3&indSuplemento=Sim&codSuplemento=B&desVolumeSuplemento=&desTomoSuplemento=&_gl=1*hu1qbi*_ga*MjgwNzgz1MDUzLjE2MjQ0ODA3Nzk.*_ga_CW3ZH25XMK*MTY5NjUzMDI1MS4yNS4xLjE2OTY1MzA3MDcuMCAwLjA. Acesso em: 2 abr. 2023.

¹² WISCHMEYER, Thomas. Artificial intelligence and transparency: opening the black box. In: WISCHMEYER, Thomas; RADEMACHER, Timo (ed.). *Regulating artificial intelligence*. Cham, Switzerland: Springer Nature, 2020. p. 75-101.

¹³ ARRIETA, Alejandro Barredo *et al.* Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, v. 58, p. 82-115, June 2020. Disponível em: <https://arxiv.org/pdf/1910.10045>. Acesso em: 3 out. 2023.

Quando se trata de IA aplicada ao Poder Judiciário brasileiro, a transparência tem ainda maior relevo, por se tratar de aplicação ao setor público de alta proeminência e impacto social, sujeito ao controle democrático dos cidadãos e à prestação de contas, bem como pautado pela exigência de motivação e justificação de suas decisões. Não por outro motivo, aplicações de IA na administração da justiça, incluindo sistemas que auxiliem autoridades judiciárias na investigação de fatos e na aplicação da lei são previstas pelo art. 17, VII, do PL (Projeto de Lei) n. 2.338/2023, proposto pela Comissão de Juristas, como aplicações de *alto risco*. A transparência sobre a IA nos Tribunais exige não só informações gerais em relação ao uso e seus impactos, benefícios e riscos, como também maior detalhamento da relação humano-máquina quando houver algum processo decisório que impacte cidadãos em particular, o que é o caso dos serviços de adjudicação em relação às partes de um processo.

Levantamento realizado pela Fundação Getúlio Vargas¹⁴ e por Junquillo¹⁵ aponta que existem atualmente cerca de 110 aplicações de IA, com finalidades variadas. O Conselho Nacional de Justiça (CNJ), diante das diversas aplicações de IA na justiça (muito embora a própria definição e classificação de *softwares* como IA seja disputável), publicou a Resolução n. 332/2020 do CNJ, que determina, em seu Capítulo IV, a publicidade e transparência das aplicações.

Mesmo com a publicação da Resolução, pairam dúvidas sobre o que significa determinar que cada aplicação de IA, com suas respectivas particularidades, seja transparente e tem-se criticado a falta de transparência do Poder Judiciário em relação aos sistemas de IA¹⁶.

De fato, não há sistematicidade, seja na atuação fiscalizadora do CNJ ou mesmo consenso na literatura em relação ao que informar, sobre qual aspecto e para qual destinatário trazer que tipo de informação. Tampouco há parâmetros sobre o momento e a forma de veicular as informações relevantes. Ou seja, não há qualquer moldura ou “*framework*” de governança da transparência para uma organização como o Judiciário, muito embora exista regulamentação vigente exigindo o emprego transparente da IA pelos Tribunais.

¹⁴ SALOMÃO, Luis Felipe (coord.). **Inteligência artificial**: tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário brasileiro. 2. ed. Rio de Janeiro: FGV Conhecimento Centro de Inovação, Administração e Pesquisa do Judiciário, 2022. Disponível em: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/relatorio_ia_2fase.pdf. Acesso em: 2 fev. 2023.

¹⁵ JUNQUILHO, Tainá Aguiar. **Inteligência artificial no direito**: limites éticos. São Paulo: Juspodivm, 2022.

¹⁶ JUNQUILHO, Tainá Aguiar; SUGAI, Kin Modesto; VALENTE, Ana Cláudia de Souza. A inteligência artificial no TJDF: estratégias de publicidade e princípio da transparência. In: FARRANHA, Ana Cláudia. **Direito ao acesso à informação**: perspectivas no âmbito da tecnologia e governança digital. Campinas: Pontes, 2022. p. 163-178.

O presente artigo ambiciona preencher essa lacuna, ou ao menos estruturar os elementos principais de uma governança da transparência pelos Tribunais quanto ao emprego de sistemas de inteligência artificial, com base no modelo geral proposto por Kaminski.¹⁷ Para tanto, descreverá, na próxima seção, a metodologia para que cada Tribunal, como organização responsável por empregar de modo transparente sistemas de IA, forneça informações relevantes e adequadas para diferentes interlocutores. A seção seguinte discutirá a estrutura de gestão e de governança da tecnologia em Tribunais, tomando como exemplo o Tribunal de Justiça do Estado de São Paulo, de modo a identificar atores relevantes no sistema de adjudicação que seriam os destinatários das informações. Em seguida, abordaremos alguns tipos mais conhecidos de aplicação de sistemas de IA nos Tribunais, indicando seus principais objetivos, riscos e preocupações decorrentes. Por fim, fundados na metodologia delineada, nos interlocutores relevantes dentro da estrutura de governança dos Tribunais e considerando os principais tipos de sistemas empregados, apresentaremos matrizes de governança da transparência de sistemas de IA para o Judiciário. Ao final, traremos algumas considerações à guisa de conclusão.

2. Metodologia para construção do modelo de governança de transparência em organizações

A transparência, como elemento de governança de sistemas de IA, consiste na *prestação de informações a terceiros sobre o emprego da inteligência artificial na organização*. Ao colocar o escopo com essa abrangência, já assumimos a natureza *sociotécnica* de sistemas de IA.¹⁸ Ou seja, sistemas de IA, nesta acepção, não se resumem ao *software* ou conjuntos de *softwares* empregados para uma determinada tarefa, mas também aos humanos envolvidos no emprego da(s) ferramenta(s) que levam a determinado resultado (preditivo, de recomendação ou decisório), conforme seu *design* no funcionamento de uma organização particular. Com isso, distanciamos-nos de expressões como *transparência algorítmica* (que traz uma conotação de informações estritamente vinculadas ao modelo ou à metodologia

¹⁷ KAMINSKI, Margot E. Understanding transparency in algorithmic accountability. In: BARFIEL, Woodrow (ed.). *Forthcoming in Cambridge handbook of the law of algorithms*. Cambridge, UK: Cambridge University Press, 2021. p. 121-140. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3622657. Acesso em: 3 out. 2023.

¹⁸ Isso é, “[...] Inteligência artificial (IA) é uma família de tecnologias em rápida evolução que pode trazer uma ampla gama de benefícios econômicos e sociais em todo o espectro de indústrias e atividades sociais, [...]” GAETANI, Francisco; ALMEIDA, Virgílio. Inteligência artificial e democracia. *Valor Econômico*, São Paulo, 22 jul. 2021. Disponível em: <https://valor.globo.com/opiniao/coluna/inteligencia-artificial-e-democracia.ghtml>. Acesso em: 23 fev. 2023.

usada no algorítmico) ou transparência quanto ao *agente inteligente* (um pouco mais abrangente que o anterior, mas que ainda se restringe ao elemento tecnológico).

Outro aspecto importante a ressaltar sobre essa escolha inicial está no fato de que o emissor dessa interlocução, aquele que veiculará informações sobre o objeto “sistemas de IA” a terceiros, é a própria organização. Isso permite incluir entre os “terceiros”, destinatários do conteúdo informativo, os próprios atores envolvidos nas decisões relevantes para o desenvolvimento, emprego e monitoramento do sistema de IA. Assim, para usar um exemplo do campo de aplicação escolhido para este artigo, temos que o “tribunal” como “organização” se compromete, dentro do modelo de governança de transparência, a informar o magistrado ou o operador de TI sobre aspectos do sistema que sejam relevantes para o desempenho de suas tarefas, como atores (humanos) do próprio sistema de IA envolvido na prestação de serviços pelo Tribunal. Isto é, os destinatários da política de transparência, em um modelo completo de governança, não podem se limitar a interlocutores “fora” da organização. Deve também abranger os integrantes da própria organização, para que se promova confiabilidade não só *externa*, como também *interna* em relação ao emprego dessa tecnologia. Mais do que isso, a transparência deve abranger, inclusive, como objeto, a relação *humano-máquina* no funcionamento e entregas do sistema. Ou seja, um ator da organização pode ser, ao mesmo tempo, destinatário e objeto da informação sobre o uso da IA na organização.

É importante considerar que a própria política de transparência deve ser também objeto da transparência a ser adotada na organização perante terceiros, ou seja, a organização deve informar sobre a sua política de transparência, de modo que a transparência é autorreferente. Não entraremos aqui nesse elemento de transparência de segunda ordem, mas faremos, ao final, algumas considerações sobre essa dimensão informativa sobre o uso de IA nos Tribunais.

Como componente de uma atividade de comunicação, a transparência é sempre relacional: o emissor (a organização) fornece informações (i) *relevantes* e (ii) *adequadas* ao (iii) *receptor* sobre (iv) o sistema de IA. Esses quatro elementos formam os pontos cardeais para definir o conteúdo informativo nessa relação comunicativa. Primeiro, a *transparência* (não a *informação* específica) deve dizer respeito a determinado *objeto*, que define o universo de discurso, ou seja, delimita o campo no qual devem ser veiculadas as informações. No nosso caso, o objeto, em geral, é o sistema de IA, na acepção ampla acima indicada. Mas o “sistema de IA” comporta

uma série de componentes distintos, que merecem uma segmentação, em diferentes dimensões, o que será feito mais à frente¹⁹.

Uma vez estipulado o objeto da transparência, por exemplo, transparência quanto ao modelo usado pelo algoritmo de IA, deve-se perguntar pelo destinatário da informação. Apenas fixando o tipo de destinatário (*quem?*), é possível delimitar os demais fatores: *relevância* e *adequação*.²⁰ Como o nível de transparência varia conforme aplicações e seu destinatário, documentos de governança recomendam diferentes níveis de transparência, que possam ser mensurados e testados.²¹

Assim, para cada destinatário identificado, exige-se que a informação seja *relevante*. Com isso, queremos dizer que deve haver uma razão determinada (*por que?*) para que aquela informação seja fornecida, dentro daquele objeto, àquele (tipo de) destinatário. Por exemplo, qual a razão para fornecer informações sobre determinado modelo de IA para os advogados das partes de um processo? A resposta a essa pergunta permite indagar qual seria o conteúdo informativo adequado (*o que?*), considerando aquela justificativa particular e o destinatário. Se entendermos que os advogados das partes devem receber informações sobre o modelo para que possam questionar, no sentido de “contraditar”, eventualmente, o *output* de determinado sistema de IA que traz um efeito na movimentação do processo, não faria sentido trazer informações técnicas sobre a função ou funções matemáticas adotadas, ou sobre o total de variáveis e parâmetros utilizados, mas, de modo inteligível para um jurista, quais seriam os fatores determinantes ou mais importantes que influenciaram o resultado do sistema empregado na movimentação particular do processo (ver abaixo as considerações sobre *explicabilidade*). Já para os servidores de TI de um tribunal, que teriam por papel monitorar o funcionamento do sistema, uma simplificação não técnica da relação entre *inputs* e *outputs* não seria suficiente, ou poderia mesmo ser completamente fora de propósito (no caso de modelos de explicação agnósticos, ver abaixo).

A resposta às perguntas “*quem?*”, “*por que?*” e “*o que?*” para cada dimensão do objeto escolhido são cruciais para a política de transparência em governança de

¹⁹ DIAKOPOULOS, Nicholas. Transparency. In: DUBBER, Markus D.; PASQUALE, Frank; DAS, Sunil (ed.). **The Oxford handbook of ethics of AI**. New York: Oxford University Press, 2020. p. 197–213.

²⁰ Embora as matrizes aqui estruturadas tenham metodologia um tanto diversa, por estruturar de modo particular aos nossos propósitos a segmentação do objeto de transparência, vale ver a indicação das questões relevantes em relação à transparência em KAMINSKI, Margot E. Understanding transparency in algorithmic accountability.

²¹ Nesse sentido, por exemplo, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE 7001-2021: IEEE standard for transparency of autonomous systems**. New York: IEEE, 2021. Disponível em: <https://standards.ieee.org/ieee/7001/6929/>. Acesso em: 4 out. 2023.

IA. Mas para que a política seja implementável, é ainda necessário abordar a forma e o momento adequado para que a informação definida na política seja veiculada. A forma (*como?*) está diretamente relacionada à acessibilidade pelo destinatário escolhido e inclui a definição sobre a mídia, o canal (*onde?*) e o formato do conteúdo a ser veiculado (será em documento físico? Em plataforma *online*? Em *e-mails* enviados aos interessados? No andamento do processo? Em qual tipo de arquivo? Acessível ao público? Acessível as partes? Acessível aos servidores do tribunal?, entre outros.). O momento (*quando?*) é bastante sensível à justificativa que delimitou o conteúdo informado, pois está ligado à prerrogativa ou função que o receptor pode exercer ao conhecer ou aplicar a informação recebida.

No campo de aplicação aqui escolhido, as informações sobre o resultado do uso integral ou parcial de IA que resultou em uma movimentação processual, caso seja direcionada aos advogados, deve vir em momento útil para eventual exercício contraditório e é razoável exigir que seja registrado na movimentação processual. Já para uma informação relevante aos cidadãos, que exercem o controle democrático da atividade dos Tribunais, não é razoável que esteja disponível no documento que registra um andamento processual, mormente em um processo sigiloso, mas em local facilmente identificável e acessível, por exemplo, na página oficial do tribunal na rede mundial de computadores.

Como indicamos acima, o objeto da transparência – “o sistema de IA na organização” – comporta diferentes aspectos que merecem segmentação para que o conteúdo informativo seja melhor delimitado, em cada contexto. Para que uma organização seja transparente quanto a determinado sistema de IA empregado, há o pressuposto básico de que a organização dê conhecimento ao destinatário de que usa determinado sistema de IA (*transparência quanto ao uso*). Somente assim pode trazer informações sobre o sistema que sejam inteligíveis ao seu interlocutor.

Outro aspecto é a *operação* do sistema (sócio-técnico) de IA. Essa *transparência quanto à operação* pode ser dividida em três aspectos fundamentais, especialmente quando tratamos de metodologias de *machine learning*, sistemas que aumentam sua eficiência a partir da experiência trazida pelo *conjunto de dados* para treinamento, com base em um *modelo matemático*. Ou seja, quanto à operação, falamos, de um lado, em transparência quanto aos dados, e, de outro, em transparência quanto ao modelo. Soma-se a esses dois aspectos, a interação entre o humano e máquina para que se estabeleça o resultado almejado pela organização (*transparência quanto ao envolvimento humano*).

A última dimensão da transparência não está propriamente no sistema, mas nos desdobramentos ou impactos previstos deste sobre os destinatários do serviço ou do emprego da IA e sobre a sociedade em geral. Trata-se da transparência quanto à própria avaliação feita pela organização em relação aos benefícios e riscos do emprego do sistema de IA. Trataremos abaixo de cada uma das dimensões do objeto “sistemas de IA” indicadas.

Transparência quanto ao uso. A transparência quanto ao uso do sistema requer daqueles que empregam o sistema de IA que informem não só que o usuário interage direta ou indiretamente com um sistema de IA ou que estão sujeitos a um processo de tomada de decisão que é influenciado por um sistema de IA, mas também que indiquem qual o grau de influência da IA no resultado final da decisão. Quando tratamos da interação direta do usuário com o sistema de IA, ou seja, quando ele se relaciona de modo imediato fornecendo os *inputs* e recebendo os *outputs* na comunicação, atenção especial deve ser dada à interface do sistema. Para que a interação ocorra com transparência, a informação sobre o sistema de IA deve ser acompanhada da informação sobre a função exercida pelo sistema, em linguagem simples e com destaque. A interface adequada volta-se para o sucesso da interação, de tal modo que a compreensão pelo usuário de sua finalidade e das ações que deve executar para que atinja o fim almejado na interação permita que o usuário exerça sua autonomia. Tal autonomia no campo informacional pressupõe também que sejam de fácil acesso as informações sobre o uso de dados pessoais, de modo a evitar o efeito inibitório da incerteza quanto ao tratamento e suas consequências. Outro aspecto relevante a considerar é o grau de informação sobre as limitações da ferramenta em relação aos resultados almejados²². Em particular, a interface deve não só trazer as informações sobre a finalidade, como também educar o usuário sobre a possibilidade de erro, margens de incerteza quanto ao resultado, bem como indicar de modo claro como o usuário pode solicitar revisão do resultado, caso o sistema de IA, ainda que não integralmente automatizado, possa influenciar de modo relevante uma decisão que afete de modo significativo o usuário.

Transparência quanto à operação. Como indicado acima, a transparência quanto à operação refere-se a três elementos-chave no desenvolvimento de sistemas de IA, em particular aqueles baseados em aprendizado de máquina: a) transparência quanto aos dados; b) transparência quanto ao modelo e suas inferências; c) transparência quanto ao envolvimento humano.

²² PINHANHEZ, Claudio S. Expose uncertainty, instill distrust, avoid explanations: towards ethical guidelines for AI. In: NEURIPS 2021 WORKSHOP ON HUMAN-CENTERED AI, 2021, online. [Proceedings...]. Disponível em: <https://arxiv.org/abs/2112.01281>. Acesso: 16 abr. 2023.

A *transparência quanto aos dados* refere-se, em primeiro lugar, à indicação sobre o emprego de dados pessoais e uso de *profiling* dos indivíduos sujeitos à decisão. Medidas nesse sentido permitem a verificação de legitimidade e *accountability* em relação à proteção de dados pessoais. Em segundo lugar, a transparência também deve conter elementos para avaliar a qualidade dos dados empregados: sua completude, fidedignidade, confiança quanto à fonte, forma de coleta ou possíveis adulterações, frequência de atualização e representatividade em relação a populações de interesse (grupos menos favorecidos).

Vale lembrar que a construção de bancos de dados para treinamento de algoritmos e desenvolvimento de modelos de IA requerem quantidades expressivas de trabalho de pré-processamento, antes do início da fase operacional, e de pós-processamento, para treinamento e refinamento da acurácia do modelo.

A transparência desses processos requer a divulgação de informações sobre os dados usados para treinamento e aprendizagem supervisionada, bem como quaisquer distorções, riscos ou vieses já detectados no modelo.

A *transparência quanto ao modelo* empregado, em particular sobre os critérios relevantes para o resultado, comporta uma digressão mais alongada, tendo em vista o risco de opacidade inerente aos sistemas mais complexos de inteligência artificial.

Um dos problemas em termos de governança e transparência está no *trade-off* entre a capacidade de compreensão dos resultados do sistema e sua confiabilidade (precisão, cobertura, acurácia),²³ o que torna o esforço de explicação do funcionamento do sistema mais difícil, sob o ponto de vista técnico, em sistemas baseados em modelos mais complexos. Estes, por sua vez, têm poder de previsão bastante elevado, mas são considerados “caixas pretas” em razão da opacidade quanto aos critérios relevantes para seus resultados (decisões, predições, recomendações).

Em estágios iniciais de emprego de sistemas computacionais para decisões, predições ou recomendações, havia uma espécie de “engenharia de fatores”, em que especialistas em determinado domínio, com base em seu conhecimento ou em modelos teóricos, selecionavam as variáveis explicativas relevantes como hipóteses

²³“Here, then, is a core and, for the moment, unavoidable trade-off in designing algorithmic accountability regimes: Interpretability often comes only at the cost of efficacy” (p. 7 do PDF). ENGSTROM, David Freeman; HO, Daniel E. Artificially intelligent government: a review and agenda. In: VOGL, Roland (ed.). **Research handbook on big data law**. Cheltenham, UK: Edward Elgar, 2021. p. 57-86. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3551549. Acesso em: 3 out. 2023.

causais de determinado resultado. Essas hipóteses seriam então testadas com métodos estatísticos para verificar sua capacidade explicativa das observações e preditiva.

Em modelos de aprendizado de máquina, porém, a lógica se inverte. A chave para a seleção das variáveis explicativas do modelo passa a ser feita a partir da mensuração de correlações estatísticas entre as variáveis de *input* e a variável de *output*. Tais correlações são calculadas, por vezes, a partir de milhões de observações pelo sistema computacional.²⁴

Assim, não se almeja apenas uma relação causal entre cada variável de *input* e o resultado. Busca-se, a partir de uma grande massa de dados, obter uma regra de decisão ou função objetiva probabilística, conforme o modelo matemático escolhido (regressão linear, árvore de decisão, florestas randômicas, redes neurais, etc.) que atribua pesos relativos aos dados de entrada, de modo a obter a predição mais acurada da variável de saída.²⁵

Ou seja, é o resultado ou sucesso, das múltiplas observações de predições processadas pelo sistema informático, a partir de diferentes conjuntos de variáveis, com diferentes pesos, que determina quais variáveis de entrada e com qual peso seriam relevantes. Assim, para algoritmos mais complexos de aprendizado de máquina, em que lidamos com milhões de observações, de múltiplos pacotes de variáveis, com diferentes sopesamentos, é praticamente impossível conectar um *input* específico a determinado *output*.²⁶ Com isso, há um desencontro entre a otimização matemática das predições típicas dos modelos de *machine learning* com a escala humana de raciocínio e estilo de interpretação semântica.²⁷

Tal característica de alguns sistemas complexos de *machine learning*, porém, não impedem que sejam abertas informações valiosas dentro de um arranjo institucional envolvendo outros atores, como auditores independentes, *board* de experts, que possam apoiar internamente e monitorar a operação do sistema, bem como

²⁴ ANDERSON, Chris. The end of theory: the data deluge makes the scientific method obsolete. *Wired*, 23 June 2000. Disponível em: <https://www.wired.com/2008/06/pb-theory/>. Acesso em: 3 out. 2023.

²⁵ LEHR, David; OHM, Paul. Playing with the data: what legal scholars should learn about machine learning. *UC Davis Law Review*, v. 51, n. 2, p. 653-717, Dec. 2017. Disponível em: https://lawreview.law.ucdavis.edu/issues/51/2/Symposium/51-2_Lehr_Ohm.pdf. Acesso em: 3 out. 2023.

²⁶ WISCHMEYER, Thomas. Artificial intelligence and transparency: opening the black box.

²⁷ BURRELL, Jenna. How the machine thinks: understanding opacity in machine learning algorithms. *Big Data & Society*, v. 3, n. 1, Jan. 2016. Disponível em: <https://doi.org/10.1177/2053951715622512>. Acesso em: 3 out. 2023.

apoiar autoridades em tomadas de decisão sobre os sistemas de inteligência artificial de alto risco ou que tenham provocado danos.²⁸ Daí por que é importante a distinção entre interlocutores e o duplo papel destes, como atores no sistema socio-técnico de IA e como destinatários das informações em graus de complexidade, aspectos e enfoque diversos. A interação entre os mesmos no arranjo institucional da organização permite que o risco de opacidade seja mitigado.

Feitas essas observações quanto às limitações à pretensão de explicar causalmente o papel das variáveis no resultado, algumas distinções são importantes, para determinação do conteúdo informativo adequado para cada interlocutor na governança da transparência.

Em primeiro lugar, cabe distinguir a interpretabilidade ou inteligibilidade da explicabilidade.²⁹ Por inteligibilidade, compreende-se a capacidade de um modelo permitir que um humano compreenda sua função, sem a necessidade de explicar sua estrutura interna ou o processamento algorítmico. Modelos inteligíveis ou interpretáveis têm seus *outputs* previsíveis por humanos, podendo ser esclarecido por interpretações intuitivas de seus resultados.³⁰ Já a explicabilidade diz respeito a uma intervenção necessária, por vezes empregando métodos de inteligência artificial, para que modelos não inteligíveis possam ser compreendidos por humanos. Há diferentes métodos de explicação e sua compreensão é relevante quando pensamos em diferentes interlocutores, por exemplo, quando temos um jurista ou um técnico de TI como destinatário das informações sobre o modelo de IA empregado.

Assim, se, para modelos interpretáveis, mais simples, a informação sobre seu funcionamento já permite ao humano entender os fatores relevantes para o resultado, em modelos não interpretáveis, pode haver uma dissociação entre esses elementos, a depender do seu grau de explicabilidade, que leva em consideração

²⁸ WISCHMEYER, Thomas. Artificial intelligence and transparency: opening the black box. p. 76-97.

²⁹ Em: ARRIETA, Alejandro Barredo *et al.* Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI, o autor propõe uma série de distinções ulteriores entre inteligibilidade, compreensibilidade e interpretabilidade, que são desnecessárias para o nosso propósito. Aqui interessa apenas a distinção entre sistemas que necessitam ou não de intervenções auxiliares para que os resultados possam ser compreendidos.

³⁰ ZHONG, Jinfeng; NEGRE, Elsa. AI: to interpret or to explain? In: CONGRÈS INFORSID 2021: INFormatique des Organisations et Systèmes d'Information et de Décision, 39., 2021, Dijon, France. [Actes...]. Disponível em: <https://hal.science/hal-03529203/document>. Acesso em: 3 out. 2023.

o estado da arte da própria tecnologia. Conforme indicam Maranhão *et al.*,³¹ há modelos de explicação, baseados em inteligência artificial (*Explainable Artificial Intelligence – XAI*), como a *análise de sensibilidade* e o *modelo decomposicional*, que trazem à luz aspectos do funcionamento interno do modelo, o que permite uma compreensão dos seus critérios. Em algumas hipóteses, porém, para modelos opacos, a explicação gerada é chamada *agnóstica*, pois não examina o seu funcionamento interno, mas baseada em um modelo de IA interpretável, que relaciona fatores da entrada do modelo (dados de *input*) com suas saídas (*outputs*).

Essa distinção é particularmente interessante para o *framework* aqui desenvolvido, pois, conforme o destinatário da informação sobre modelos opacos, as ferramentas de explicação podem ser diversas, *e.g.*, ferramentas de sensibilidade ou decompositivas são mais adequadas para técnicos de TI, ao passo que ferramentas agnósticas podem ser mais interessantes para destinatários que não tenham o domínio de conhecimento técnico. É importante, porém, destacar que modelos agnósticos, por não se fiarem no funcionamento interno, podem gerar explicações descoladas daqueles fatores efetivamente relevantes. Nesse caso, a política de transparência deve deixar claro o que será considerado, em última análise, como fator de explicação, o que é particularmente relevante no âmbito de Tribunais, quando pensamos no direito ao contraditório, de supedâneo constitucional.

Há, por outro lado, como também indicam Maranhão *et al.*³², diferentes exigências de explicabilidade, conforme o destinatário. Pode-se exigir a *simulatabilidade* (capacidade de o humano reproduzir os resultados em tempo razoável diante dos fatores relevantes), a *contrastividade* (capacidade de comparar o *output* com alternativas plausíveis que poderiam trazer resultado diverso), ou a *contestabilidade* (trazer elementos suficientes para a contestação do resultado). Explicações simuláveis podem caber em relação a IAs que substituem determinada tarefa humana, em modelos mais simples. Em casos de alta opacidade e complexidade, pode ser inviável. Já a exigência de contrastividade pode caber para destinatários que têm o papel de monitorar o funcionamento do modelo, controlando seu desempenho e eventuais vieses ou falhas de acurácia. Pode também ser relevante em contextos nos quais o destinatário do resultado possa ajustar sua conduta ou as informações relevantes

³¹ MARANHÃO, Juliano Souza de Albuquerque; COZMAN, Fabio Gagliardi; ALMADA, Marco. Concepções de explicação e do direito à explicação de decisões automatizadas. In: VAINZOF, Rony; GUTIERREZ, Andriei Guerrero (coord.). *Inteligência artificial: sociedade, economia e Estado*. São Paulo: Revista dos Tribunais, 2021. p. 137-154.

³² MARANHÃO, Juliano Souza de Albuquerque; COZMAN, Fabio Gagliardi; ALMADA, Marco. Concepções de explicação e do direito à explicação de decisões automatizadas. p. 137-154.

de entrada. Já a exigência de contestabilidade parece ser uma condição geral mais adequada ao funcionamento de um processo, em relação às partes interessadas.

Por fim, a transparência sobre o modelo de IA pode ser *centrada no modelo*, ou seja, como um esclarecimento geral e *ex-ante* sobre o modelo emprego e quais são geralmente os elementos relevantes para o *output*, de tal forma que o indivíduo possa usar as informações para um questionamento genérico do seu emprego. A transparência também pode ser *centrada no indivíduo*, com informações concretas dos elementos relevantes para o *output* particular em relação àquele indivíduo. A transparência centrada no modelo é mais adequada para os destinatários que não são impactados diretamente pelo emprego da IA, como, por exemplo, os cidadãos que têm interesse em realizar o controle democrático das ferramentas empregadas pelo tribunal, ao passo que as partes de um processo estariam melhor atendidas por uma transparência centrada no indivíduo.

A transparência quanto ao envolvimento humano diz respeito à explicitação da participação humana nas diferentes etapas do ciclo de desenvolvimento e emprego do sistema de inteligência artificial. Trata-se de identificar em quais etapas há o envolvimento de decisões humanas relevantes na construção e emprego do sistema, seja na coleta e preparação dos dados, na escolha do modelo, na forma de monitoramento dos resultados, etc. Trata-se de medida relevante para a rastreabilidade das decisões e atribuições de responsabilidade por eventuais consequências negativas advindas do emprego de sistemas de IA. Por outro lado, é de particular interesse a participação humana no resultado de uma aplicação particular da ferramenta, seja ela de predição, recomendação ou decisão automatizada. Para cada tipo de resultado, pode haver maior ou menor envolvimento humano, o que pode contar com diferentes indicações relevantes para os destinatários da informação sobre o sistema de IA.

Quando tratamos de decisões automatizadas, temos algumas categorias principais. Primeiro, a decisão pode ser integralmente automatizada (sem envolvimento humano), ou pode ser parcialmente automatizada (com envolvimento humano). O envolvimento humano pode estar na coleta e alimentação do sistema com os *inputs* relevantes, ou como ator na tomada de decisão a partir do *output* gerado pela máquina. Em sistemas de recomendação, pode haver a recomendação final do resultado, cabendo ao humano a discricionariedade para revisá-la, ou a recomendação pode dizer respeito a uma ação preliminar relevante no processo decisório humano. Por sua vez, sistemas de predição, a não ser quando a entrega na prestação do serviço seja a própria predição, são normalmente usados como informação relevante para que o humano tome a decisão (por exemplo, um sistema

de *credit scoring* que indique a pontuação de risco de determinado solicitante, cabendo ao humano decidir sobre a concessão ou não de crédito). Para os sujeitos impactados pelo processo decisório que envolve a interação humano-máquina, a indicação exata do tipo de envolvimento humano é fundamental para eventual questionamento do resultado.

Transparência quanto a benefícios e riscos. A dimensão de transparência quanto a benefícios e riscos se justifica pelo caráter instrumental da transparência. O fornecimento de informações aos diferentes interlocutores tem a função de permitir que esses não só desempenhem suas funções no sistema (interlocutores internos), como também aproveitem os seus resultados (usuários-interlocutores externos) e avaliem os riscos e os ganhos sociais do emprego do sistema. Todos os envolvidos podem melhor monitorar o alcance dos objetivos e benefícios, bem como mitigar os riscos trazidos caso obtenham preliminarmente uma avaliação e indicação preditiva da própria organização sobre os aspectos relevantes a serem observados quanto ao impacto no desempenho do serviço, benefícios sociais eventualmente produzidos e, sobretudo, riscos a serem mitigados.

A metodologia de construção da matriz baseia-se, portanto, na compreensão de que a transparência é relacional. Dessa forma, para cada dimensão de transparência, será elaborada matriz delimitando os critérios para consecução da governança.

3. Estrutura de gestão da tecnologia no Poder Judiciário: os Tribunais e o CNJ

A estrutura de gestão da transparência para os Tribunais deve partir da própria organização proporcionada pelo Conselho Nacional de Justiça (CNJ) para exercer o controle da atuação administrativa e financeira do Poder Judiciário e do cumprimento dos deveres funcionais dos juízes. À exceção do Supremo Tribunal Federal (STF), todos os 90 Tribunais do país são regulados pelo CNJ, que, por meio de atos normativos, com obrigatoriedade equiparada à legal, tem por função padronizar as estruturas de governança das diversas Cortes e uniformizar as rotinas de trabalho, respeitando a autonomia de cada Tribunal.

No contexto da Lei do Processo Eletrônico (Lei n. 11.419/2006), artefato base da sua atividade-fim, a padronização da numeração processual (Resolução CNJ n. 65/2008), da unificação taxonômica de classes, assuntos e movimentações

processuais (Resolução CNJ n. 46/2007), o alinhamento dos Planejamentos Estratégicos locais ao Planejamento Estratégico Nacional do Poder Judiciário (Resolução CNJ n. 70/2009) e a criação do modelo de interoperabilidade do Poder Judiciário (Resolução Conjunta CNJ/CNMP n. 03/2013) foram apenas algumas das ações do CNJ que permitiram a evolução tecnológica do Judiciário como um todo. O amadurecimento do processo eletrônico foi o responsável pela digitalização e estruturação dos dados processuais, tornando-os passíveis de manipulação em grande escala, por ferramentas de *business intelligence*, gerando a possibilidade de automatização e, subsequentemente, da automação das rotinas de trabalho, pavimentando a senda evolutiva para o surgimento de aplicações de inteligência artificial em seu âmbito interno.

No recorte específico das tecnologias atreladas à inteligência artificial, a função normativa exercida pelo CNJ editou a Resolução CNJ n. 332, de 21 de agosto de 2020, que dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário, posteriormente regulamentada pela Portaria CNJ n. 271, de 4 de dezembro de 2020. Ambas também preveem o sistema Sinapses para reunir e integrar as informações e adotar providências no sentido de integrar soluções entre Tribunais, evitar a duplicação de custos e avaliar a adequação dos sistemas empregados.

Conforme metodologia acima indicada, para estruturar a matriz de governança da transparência, é preciso definir os interlocutores relevantes do Judiciário que atuam desde a demanda até à entrega do serviço de adjudicação, assim como os que fiscalizam sua execução e uso no dia a dia. Com o fim de compreender os atores envolvidos na gestão da transparência dos projetos de IA no Judiciário, será então realizada análise à luz do modelo de governança no desenvolvimento de aplicações de IA em um órgão do Poder Judiciário, em especial no Tribunal de Justiça de São Paulo.

O TJ/SP estruturou suas áreas temáticas em Secretarias, dotadas de Diretorias, Coordenadorias, Supervisões e, em alguns casos, de Chefias. A Secretaria de Tecnologia da Informação, órgão administrativo vinculado à Presidência do Tribunal de Justiça, passou a ser a responsável por planejar, gerenciar, coordenar e executar ou definir a contratação de prestadores de serviços ou produtos tecnológicos visando atender às necessidades de tecnologia da informação e a modernização dos processos de trabalho do Tribunal de Justiça, por meio da gestão única das atividades de tecnologia da informação, e desempenhar papel estratégico e de apoio técnico à Comissão para Assuntos de Informática, Comitê de Governança de Tecnologia da Informação (CGovTI), Comitê Gestor de Segurança

da Informação (CGESI) e Comitê Gestor de Tecnologia da Informação (CgesTI) – Art. 25, parágrafo único, da Resolução CNJ n. 332/2020.

Uma vez editada a Resolução CNJ n. 332/2020, foi dada ampla publicidade aos públicos interno e externo pela edição do Comunicado n. 120/2020, de 31 de agosto de 2020. O rápido crescimento da demanda e desenvolvimento de modelos de IA e Robotic Process Automation (RPA) deram ensejo a uma reestruturação da Coordenadoria, com a criação de Supervisões que passaram a segmentar integrações, homologações e implantações de automações.

Entretanto, a falta de visibilidade de uma série de projetos pontuais, o retrabalho gerado, a impossibilidade de prestação de suporte técnico eficiente e, sobretudo, a falta de gestão dessas iniciativas que utilizavam os recursos do ambiente tecnológico do Tribunal demonstraram que a falta de transparência dessas iniciativas à equipe técnica da Secretaria de Tecnologia da Informação passou a ser um entrave à boa gestão de recursos e soluções. Nesse contexto, foi normatizada a obrigatoriedade de comunicação à Presidência do Tribunal sobre o desenvolvimento de soluções de automatização, robotização, desenvolvimento de APPs (Aplicativos) e *softwares* em geral, bem como iniciativas relacionadas a desenvolvimento de banco de dados e armazenamento de dados em geral (Art. 3º, V, da Portaria CNJ n. 271/2020).

Uma vez mapeadas as iniciativas, conferindo-se a elas a visibilidade quanto à existência, objetivos e codificação, dotando-as de transparência, foi editada uma nova regulamentação acerca da criação de robôs, aplicações e banco de dados, no âmbito do TJ/SP. Há também grupo de Juízes Assessores da Presidência, cuja função é tática e a missão é de relacionar o nível operacional à gestão estratégica a cargo do Presidente do Tribunal de Justiça. Cada uma das Secretarias temáticas está sob o comando de um ou mais Juízes Assessores da Presidência que, na clássica divisão administrativa, estabelecem a forma de execução das decisões estratégicas da Presidência pela equipe operacional, qual seja, a respectiva Secretaria e seus funcionários.

Dessa forma, ingressando a demanda por solicitação do usuário junto à equipe operacional ou diretamente ao Presidente ou aos demais membros do Conselho Superior, cabe ao Juiz Assessor formalizar a demanda, institucionalizando-a e permitindo que, pelos mecanismos de transparência e na forma prevista na Lei de Acesso à Informação, seja conhecida do público interno e externo. Realizado estudo pela área técnica – Secretaria de Tecnologia da Informação – quanto ao desforço humano e orçamentário necessários para sua realização, faz-se o cotejo do

impacto positivo ao usuário, levando a demanda à consideração do Presidente do Tribunal, trazendo cada qual sua visão, estratégica e tática.

Com base no exemplo do TJ/SP, é possível perceber que, além das figuras comumente associadas a um processo no tribunal, como os advogados, as partes, que são os usuários diretos do sistema de adjudicação, os magistrados e servidores que prestam os serviços, os cidadãos que são beneficiários desse serviço, há uma estrutura interna de tomada de decisões sobre a adoção de ferramentas de TI, que envolve o departamento de TI e um comitê decisório sobre a adoção da tecnologia, formado por juízes e, em regra, integrado pelo Presidente do Tribunal.

Assim, estruturaremos a matriz tendo como destinatários da informação os seguintes interlocutores, aglutinados pelo seu papel como atores relevantes: *usuários* (não vemos razão para distinguir as partes dos advogados que as representam no processo como destinatários de informações sobre o sistema de IA, tendo em vista que eles exercerão as prerrogativas correspondentes às partes representadas, de modo que incluímos aqui todos aqueles que recebem o serviço de adjudicação pelo tribunal, *i.e.*, advogados, partes e funções essenciais à justiça no setor público), *cidadãos* (a sociedade em geral que tem a prerrogativa de exercer o controle democrático das atividades do Tribunal) *magistrados e servidores* da atividade-fim do Tribunal, *servidores de TI do tribunal* (que executam os trabalhos técnicos de implantação e monitoramento da tecnologia), o *Comitê decisório sobre tecnologia/laboratório de TI* e, por fim, o *CNJ* (que, em função da regulamentação vigente, em particular com a criação do sistema Sinapses, deve receber informações sobre os sistemas adotados pelos Tribunais).

4. Aplicações de Inteligência Artificial nos Tribunais

Quando se analisam as aplicações atuais existentes no ecossistema de justiça brasileiro atual, existem algumas cujo objetivo é o agrupamento de processos para julgamento conjunto. Nesses casos, a aplicação da IA busca formar conjuntos de processos com temas similares – são exemplos de projetos que têm esse objetivo: o Athos (STJ); o Bem-Te-Vi (TST) e o Análise Legal Inteligente (ALEI) (TRF/1ª Região). O maior risco dessas aplicações são incorreções: erros de agrupamento podem gerar inclusão de determinados casos processuais em grupos cujo resultado do julgamento pode ser divergente do precedente que deva ser aplicado ao caso, prejudicando as partes envolvidas, por exemplo.

Outra aplicação frequente é a de *Chatbots* – utilizado, por exemplo, no Peticionamento Inteligente (TJ/RO), no *Chatbot* DIGEP (TJ/RS) e no JUDI (TJ/SP), cujo objetivo principal é a interação com servidores, jurisdicionados e cidadãos. Nesse tipo de aplicação de IA, os principais riscos são os vieses discriminatórios e a desinformação na interação com o humano, o que pode afetar interações pré-processuais com partes e cidadãos jurisdicionados.

Também são identificados sistemas preditivos, usados com objetivo de elaboração de alertas, criação de métricas e indicadores. Atualmente são exemplos de projetos que desenvolveram esse tipo de aplicação: o Projeto OMNI (TJ/MT) e o Jurimetria com IA (TJ/MS). A IA usada para predição traz como riscos a possibilidade de generalização inadequada às especificidades dos casos concretos e a inadequação da predição ao caso concreto, o que pode demandar necessidade de revisão humana para conferência da propriedade e conformação do precedente ao caso.

Os maiores riscos estão, entretanto, nas aplicações usadas para sugestão de premissas para tomada de decisões automatizadas com ou sem revisão humana³³. Isso porque essas aplicações vão gerar decisões interlocutórias, despachos e sentenças com ou sem posterior revisão humana, o que pode afetar diretamente o exercício de garantias processuais como contraditório e ampla defesa, dificuldade de contestabilidade e recorribilidade quando da incorreção dessas decisões.

O Quadro 1 lista essas aplicações e riscos. A avaliação prévia de risco do sistema de IA e a análise de seus impactos é necessária para a determinação do grau de detalhamento das informações a serem fornecidas para cada interlocutor e, em particular sobre os próprios riscos envolvidos. As matrizes de transparência apresentadas a seguir, nesse sentido, devem ser lidas, na delimitação do conteúdo informativo, a partir dos níveis de risco de cada aplicação e também de acordo com o tipo e finalidade a que se destinam. Isso é, para se entender a quem se destina a compreensibilidade de cada aplicação, o porquê e o nível de interpretabilidade e explicabilidade relativo a quais questões, o primeiro passo é entender qual a finalidade e objetivo de cada aplicação de IA, quais seus benefícios e as ameaças a serem monitoradas e mitigadas.

³³ Aqui diferenciam-se decisões integralmente automatizadas de recomendações de decisão e de premissas relevantes para a decisão. Quando a máquina entrega premissa importante para decisão, por exemplo, selecionando a jurisprudência a ser consultada, é necessário assegurar se as partes e usuários influenciarão nos argumentos, se a seleção de jurisprudência foi completa, se o juiz procurou mais precedentes ou contentou-se com a sugestão da máquina. Quando a máquina faz recomendações de texto ou decisão e, por exemplo, o magistrado segue raramente as sugestões, a equipe de TI deve avaliar se os resultados são adequados e se a ferramenta deve ser descontinuada. Para cidadãos, nesses casos, é interessante compreender o nível de discricionariedade dos juízes nessas aplicações.

Quadro 1 – Tipos de aplicação, objetivos e principais riscos

TIPOS DE APLICAÇÃO	OBJETIVOS	EXEMPLOS DE PROJETOS NO JUDICIÁRIO BRASILEIRO	PRINCIPAIS RISCOS
Agrupamento	Formação de conjuntos de processos com temas similares	Athos (STJ); · Bem-Te-Vi (TST); · Análise Legal Inteligente (ALEI) (TRF 1ª Região)	Incorreções
<i>Chatbots</i>	Interação com servidores, jurisdicionados e cidadãos	Peticionamento Inteligente (TJ-RO); · <i>Chatbot</i> DIGEP (TJ-RS); · JUDI (TJ-SP)	Discursos discriminatórios ou falsos
Predição	Elaboração de alertas, criação de métricas e indicadores	Projeto OMNI (TJ-MT); · Jurimetria com IA (TJ-MS)	Generalização e inadequação da predição ao caso concreto
Sugestão de premissas para tomada de decisões automatizadas com revisão humana	Geração de decisões interlocutórias, despachos e sentenças com posterior revisão humana	· Victor (STF) · Identificação de Fundamentos de Inadmissão do REsp (STJ)	Contraditório e ampla defesa, Dificuldade de contestabilidade/recorribilidade
Sugestões de premissas para tomada de decisões totalmente automatizadas sem revisão humana	Geração de decisões interlocutórias, despachos e sentenças sem posterior revisão humana	-	Contraditório e ampla defesa, Dificuldade de contestabilidade/recorribilidade / Incorreção

Fonte: elaboração própria com base em JUNQUILHO, Tainá Aguiar. **Inteligência artificial no direito**: limites éticos, e SALOMÃO, Luis Felipe (coord.). **Inteligência artificial**: tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário brasileiro.

Nota-se que a análise realizada no Quadro 1 aponta ferramentas de gestão e ferramentas processuais. Nas ferramentas de gestão, em geral, os mecanismos de transparência correspondem àqueles que são destinados aos cidadãos e à sociedade, pois os riscos relacionados à aplicação dessas ferramentas atingem especialmente o controle democrático da atividade pública do exercício de jurisdição pelo Poder Judiciário. Outras ferramentas voltam-se para atividade judicante, auxiliando juízes e servidores em tarefas que podem influenciar direta ou indiretamente nos processos e, nesse sentido, demandam mecanismos de transparência que atendam não só aos cidadãos, mas também às partes e à própria estrutura de gestão dos Tribunais, uma vez que podem gerar riscos que afetem o exercício apropriado das garantias processuais, por exemplo.

5. Matriz de transparência para IA no Judiciário

Nas seções anteriores, foram delineados os elementos fundamentais para estruturação de um modelo de governança de transparência no Judiciário, tomando o Tribunal como emissor da comunicação com terceiros, que inclui tanto participantes internos do Tribunal, como atores relevantes em decisões sobre desenvolvimento e emprego da IA, como destinatários externos dos serviços de adjudicação. Como destacado, as informações devem ser relevantes e adequadas conforme o receptor, o que permite a organização do conteúdo informativo a ser preenchido, para cada aplicação de IA empregada no Tribunal em matrizes que estruturam a governança de transparência.

Toda a governança de transparência poderia ser organizada em matriz única, segmentada pelos diferentes objetos da transparência, conforme indicado na Seção 1. Porém, por limitações de espaço e didática, elaboraremos cinco matrizes, uma para cada objeto relativo ao sistema de IA: *transparência quanto ao uso de IA*; *transparência quanto à operação* (que se subdivide em *transparência quanto aos dados*, *transparência quanto ao modelo* e *transparência quanto à relação humano-máquina*); e *transparência quanto aos riscos e benefícios* do sistema.

Cada quadro é estruturado com linhas correspondentes aos interlocutores relevantes identificados na Seção 2, ao passo que as colunas colocam, para cada interlocutor, as questões relevantes a serem respondidas para definição do conteúdo informativo sobre o sistema de IA: porque informar, o que informar, quando informar e como informar. Obviamente, não cabe aqui delimitar o conteúdo

específico sobre o que informar, o que somente pode ser definido considerando o tipo de aplicação e os riscos envolvidos, conforme indicado brevemente na Seção 3.

Ou seja, cabe a estudos empíricos posteriores, que analisem criticamente os sistemas de IA empregados pelos diferentes Tribunais, estabelecer qual seria o conteúdo informativo adequado, ou mesmo os próprios Tribunais podem implementar sua política de governança da transparência, buscando responder, em contextos específicos, as questões aqui indicadas.

Quadro 2 – Objeto de transparência: quanto ao uso de IA

QUEM	POR QUE	O QUE	QUANDO	COMO
Usuários	Os usuários são afetados pelo processo e devem ter plena informação sobre aquilo que possa influir em suas prerrogativas e garantias processuais.	Informar sobre o tipo de aplicação e função do sistema de IA e seus possíveis resultados específicos na condução e tramitação do processo.	a) No início do processo, devem ser informadas todas as ferramentas utilizadas que possam ser empregadas ao longo do processo. ³⁴ b) Assim que o resultado (integralmente ou parcialmente automatizado) de uma ferramenta de IA seja agregada a uma movimentação do processo.	a) Na primeira abertura do sistema operacional de tramitação do processo no tribunal. b) Informado no documento que registra a movimentação processual.
Cidadãos	Para o controle democrático da atividade do sistema de adjudicação.	Objetivo da ferramenta e seu efeito na qualidade dos serviços judiciais.	a) De modo contínuo. b) No momento de contratação ou lançamento da ferramenta.	a) No <i>site</i> oficial do tribunal sobre as ferramentas já em uso. b) Em informes previamente à contratação e lançamento (início de uso).

³⁴ O objetivo aqui é influenciar em estratégias processuais. Exemplo: aplicações de IA para agrupamento por semelhança, pode-se destacar as diferenças ou agilidade (e escolher arbitragem ou processo).

QUEM	POR QUE	O QUE	QUANDO	COMO
Magistrados e servidores da atividade-fim	Os operadores das ferramentas computacionais ou de aplicativos usados na tramitação do processo devem entender adequadamente suas utilidades e como seu trabalho pode ser aperfeiçoado.	a) Quais as funcionalidades da ferramenta de IA. b) Quais seus resultados na tramitação ou geração de documentos relevantes para o processo.	Continuamente em oportunidades de treinamento de servidores.	Por meio de treinamentos e <i>workshops</i> .
Servidores de TI	Monitorar o funcionamento da ferramenta e acompanhar seu desempenho.	Função e medidas técnicas de desempenho da ferramenta.	Durante todo ciclo de emprego da ferramenta.	Manuais e documentos técnicos dos sistemas empregados.
Comitê decisório sobre tecnologia/laboratório de TI	Deliberar sobre a concepção e adequação do uso da ferramenta de IA.	A aplicação, objetivo e possíveis resultados da ferramenta, bem como seus custos financeiros e organizacionais.	a) No momento de concepção ou contratação da ferramenta. b) Nos testes e validação das ferramentas antes de seu emprego.	a) Apresentações pela equipe técnica de TI. b) Depósito do modelo no sistema Sinapses. ³⁵
CNJ	a) Exercer a fiscalização das atividades dos Tribunais. b) Promover a integração de soluções técnicas nos Tribunais.	Função e objetivo da ferramenta, possíveis resultados, e momento de emprego na tramitação do processo.	Previamente ao emprego efetivo da ferramenta.	Gestão do modelo na Plataforma Sinapses. ³⁶

Fonte: elaborado pelos autores.

³⁵ De acordo com o que determinam os artigos 5º-7º da Portaria n.271/2020 CNJ. BRASIL. Conselho Nacional de Justiça. Portaria nº 271, de 04/12/2020. **DJe/CNJ**, n. 393, p. 2-4, 14 dez. 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 out. 2023.

³⁶ De acordo com o que determinam os artigos 5º-7º da Portaria n.271/2020 CNJ. BRASIL. Conselho Nacional de Justiça. Portaria nº 271, de 04/12/2020. **DJe/CNJ**, n. 393, p. 2-4, 14 dez. 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 out. 2023.

Quadro 3 – Objeto de transparência: Operação da IA (i) quanto aos dados

QUEM	POR QUE	O QUE	QUANDO	COMO
Usuários	Informações sobre a base de dados e sua gestão, bem como decisões sobre coleta e desenvolvimento da base podem trazer elementos para questionar o resultado de decisões automatizadas que afetem direitos e oportunidades processuais das partes.	a) Quais os critérios de seleção da base de dados. b) Qual a representatividade, relevância e atualidade dos mesmos considerando os objetivos da aplicação.	Assim que o resultado (integralmente ou parcialmente automatizado) de uma ferramenta de IA seja agregado a uma fase de movimentação do processo.	Informado no documento que registra a movimentação processual.
Cidadãos	Informações relevantes sobre os dados para aferir isonomia e potencial de eficácia da ferramenta como forma de controle democrático.	a) Quais os critérios de seleção da base de dados; b) Qual a representatividade, relevância e atualidade dos mesmos considerando os objetivos da aplicação.	a) Permanente incluindo atualização quando houver ampliação ou alteração da base de dados em ferramentas com treinamento prévio. b) Indicação de aspectos da incorporação de novos dados em ferramentas com treinamento continuado durante o emprego.	a) <i>Site</i> oficial do Tribunal. b) Comunicações informativas sobre novas ferramentas incorporadas nos serviços dos Tribunais.

QUEM	POR QUE	O QUE	QUANDO	COMO
Magistrados e servidores da atividade-fim	Informações para que entendam com mais profundidade o funcionamento da ferramenta de modo a gerar confiabilidade (para qualquer tipo de ferramenta) e viabilizar o exercício de discricionariedade em caso de discordância do resultado. ³⁷	Quais os critérios de seleção da base de dados, qual a representatividade, relevância e atualidade dos mesmos considerando os objetivos da aplicação.	Em oportunidades de treinamento interno dos servidores.	Apresentações ou reuniões com servidores de TI do Tribunal.
Servidores de TI	Informações para desenvolvimento e treinamento adequado da ferramenta e monitoramento dos resultados para aferição da eficácia e aspectos éticos da mesma.	a) Quais os critérios de seleção da base de dados, qual a representatividade, relevância e atualidade dos mesmos. b) Forma de partição para treinamento e testes de validação, confiabilidade e descrição dos metadados que revelem os critérios de categorização dos dados pertencentes à base.	De modo permanente desde a adoção até a operação e descontinuação do uso.	a) Manuais técnicos de funcionamento da ferramenta. b) Relatórios que apresentem os indicadores principais de operação.

³⁷ Este último, no caso de sistemas voltados para recomendação bem como monitorar continuamente a eficácia da ferramenta.

QUEM	POR QUE	O QUE	QUANDO	COMO
Comitê decisório sobre tecnologia/laboratório de TI	Para deliberação sobre a adequação do emprego da ferramenta no Tribunal.	a) Quais os critérios de seleção da base de dados, qual a representatividade, relevância e atualidade dos mesmos. b) Forma de partição para treinamento e testes de validação e confiabilidade.	No momento de concepção ou previamente à contratação da ferramenta, bem como periodicamente para avaliação de resultados.	Relatórios técnicos informativos pelo departamento de TI.
CNJ	Para a fiscalização da ferramenta dentro da atividade do tribunal e esforço de integração de soluções entre diferentes Tribunais.	a) Quais os critérios de seleção da base de dados, qual a representatividade, relevância e atualidade dos mesmos. b) Forma de partição para treinamento e testes de validação e confiabilidade e descrição dos metadados ³⁸ que revelem os critérios de categorização dos dados pertencentes à base.	Previamente à contratação de uma nova ferramenta ou logo após a concepção da mesma e tomada de decisão sobre a formação da base de dados para treinamento.	Informação dentro do sistema Sinapses.

Fonte: elaborado pelos autores.

³⁸ Metadados são importantes para o CNJ a fim de que ele possa indicar ou recomendar elementos de uniformização das categorizações de dados que podem ser usados para o desenvolvimento ou integração de ferramentas diversas por outros Tribunais.

Quadro 4 – Objeto de transparência: Operação da IA (ii) quanto ao modelo empregado

QUEM	POR QUE	O QUÊ	QUANDO	COMO
Usuários	Entender os critérios determinantes ou relevantes para o <i>output</i> do sistema de IA ou para interpretar seus resultados, de modo a ter elementos suficientes para questioná-los.	a) Informações gerais <i>ex-ante</i> sobre a metodologia empregada e quais são os critérios relevantes para a tomada de decisão. b) Informações <i>ex-post</i> sobre os fatores relevantes para um <i>output</i> particular ou os elementos levados em consideração para a apresentação ou justificação do resultado que possa afetar o andamento processual.	a) No início do processo, informações gerais sobre as ferramentas que podem ser empregadas no curso do processo. b) Assim que o resultado (integral ou parcialmente automatizado) de uma ferramenta de IA seja agregado/ registrado a uma movimentação do processo.	a) No <i>site</i> do Tribunal, no primeiro contato do usuário com o processo eletrônico na plataforma de tramitação usada pelo Tribunal (informações gerais). b) No documento que registra a movimentação pessoal.
Cidadãos	Informações para que o cidadão possa avaliar o impacto da aplicação no sistema de justiça.	Informações gerais <i>ex-ante</i> sobre a metodologia empregada e quais são os critérios relevantes para a tomada de decisão.	De modo permanente, desde o início de emprego da ferramenta até sua descontinuação.	No <i>site</i> do tribunal e/ou em informes para a comunidade ou sociedade em geral.

QUEM	POR QUE	O QUÊ	QUANDO	COMO
Magistrados e servidores da atividade-fim	Entender os critérios determinantes ou relevantes para o <i>output</i> do sistema de IA ou para interpretar seus resultados, de modo a ter elementos suficientes para avaliar sua adequação e confiar no seu uso.	a) Informações gerais <i>ex-ante</i> sobre a metodologia empregada e quais são os critérios relevantes para a tomada de decisão. b) Informações <i>ex-post</i> sobre os fatores relevantes para um <i>output</i> particular ou os elementos levados em consideração para a apresentação ou justificação do resultado que possa afetar o andamento processual.	Antes de seu emprego nas atividades do tribunal e de modo permanente até sua descontinuação. Antes da formalização do <i>output</i> no registro processual para as informações sobre uma decisão concreta.	Nas atividades de treinamento (informações gerais), nas interfaces das plataformas por meio das quais os servidores desenvolvem suas atividades (informações concretas).
Servidores de TI	Entender os critérios determinantes ou relevantes para o <i>output</i> do sistema de IA e dados relevantes para a criação do modelo, além de elementos para avaliar a eficácia do modelo de modo a balizar a tomada de decisão sobre seu emprego ou descontinuidade.	Informações <i>ex-ante</i> sobre o modelo e sua operação suficientes para entender a forma de geração de <i>outputs</i> a partir dos dados, bem como índices de acurácia ou de dados de entrada mais relevantes conforme o uso do sistema, por meio de relatórios de monitoramento.	Anteriormente à contratação ou adoção do modelo pelo Tribunal e durante o seu emprego.	Por meio de descrição detalhada e técnica para os representantes de TI, e didática para os membros do Tribunal antes da contratação e por meio de relatórios técnicos periódicos de acompanhamento.

QUEM	POR QUE	O QUÊ	QUANDO	COMO
Comitê decisório sobre tecnologia/laboratório de TI	Entender os critérios determinantes ou relevantes para o <i>output</i> do sistema de IA e dados relevantes para a criação do modelo, além de elementos para avaliar a eficácia do modelo, de modo a balizar a tomada de decisão sobre seu emprego ou descontinuidade.	Informações <i>ex-ante</i> sobre o modelo e sua operação suficientes para entender a forma de geração de <i>outputs</i> a partir dos dados, bem como índices de acurácia ou de dados de entrada mais relevantes conforme o uso do sistema, por meio de relatórios de monitoramento.	Anteriormente à contratação ou adoção do modelo pelo Tribunal e durante o seu emprego.	Por meio de descrição detalhada e técnica para os representantes de TI e didática para os membros do Tribunal antes da contratação e por meio de relatórios técnicos periódicos de acompanhamento.
CNJ	Informações adequadas para exercer a fiscalização e integração das soluções nos diferentes Tribunais do Poder Judiciário.	Informações gerais sobre o modelo e critérios relevantes para o resultado, bem como justificativas da escolha desse modelo em detrimento de outros possíveis a serem empregados (destacando interpretabilidade e potencial acurácia/confiabilidade).	Antes do início de emprego da ferramenta pelo tribunal.	Informação dentro do sistema Sinapses.

Fonte: elaborado pelos autores.

Quadro 5 – Objeto de Transparência: Operação da IA (iii) quanto à relação humano-máquina

QUEM	POR QUE	O QUE	QUANDO	COMO
Usuários	Informações sobre o processo de tomada de decisão e momento / grau / discricionariedade de envolvimento humano para que o usuário possa solicitar informações precisas sobre esse processo de modo a ter melhores condições para seu questionamento.	Qual a forma de relação homem/máquina adotada para o emprego da ferramenta, o resultado entregue pelo sistema e se houve ou não revisão ou acatada a recomendação por parte do operador humano e justificativa.	Informação geral sobre a relação homem/máquina no sistema de IA disponível desde o início do processo e por provocação do interessado, após registro da decisão na tramitação do processo.	Informação geral no <i>site</i> do tribunal e relatório sobre a tomada de decisão mediante solicitação específica.
Cidadãos	Para o controle de transparência da atividade jurisdicional e dos servidores que operam os sistemas ou atuam na sua função de assessoria ou jurisdicional.	Tipo de relação homem/máquina e forma de produção do resultado, bem como índices gerais de aproveitamento dos resultados gerados por máquina para se apurar grau de discricionariedade dos humanos envolvidos no <i>loop</i> .	De modo permanente.	No <i>site</i> do Tribunal sobre o funcionamento e periodicamente com relatórios acessíveis no <i>site</i> sobre os índices de aproveitamento e discricionariedade.
Magistrados e servidores da atividade-fim	Para a compreensão adequada da forma de interação e avaliação dos resultados apresentados pela máquina, de modo a aproveitar as informações relevantes geradas ou as recomendações pertinentes.	Elementos sobre o objetivo, a interface e momentos adequados para a intervenção no processo de tomada de decisão.	Periodicamente em oportunidades de treinamento da equipe de servidores.	Por meio de manuais técnicos informativos pelo departamento de TI, tutoriais e treinamentos.

QUEM	POR QUE	O QUE	QUANDO	COMO
Servidores de TI	Para a verificação e controle do desempenho da ferramenta considerando o uso adequado das mesmas pelos operadores humanos envolvidos.	Índices de produtividade e de discricionarieidade ou deferência dos humanos em relação aos resultados entregues pelo sistema de IA.	De modo contínuo durante o ciclo de vida do sistema de IA no tribunal.	Por meio de interfaces alimentadas por informações dos próprios servidores e magistrados e por relatórios periódicos.
Comitê decisório sobre tecnologia/laboratório de TI	Para a avaliação da gestão da equipe de servidores e seu aproveitamento/ treinamento <i>versus</i> produtividade almejada.	Índices atualizados de produtividade e ociosidade da equipe, tempo e recursos para treinamento, custos para emprego da ferramenta e tempo para curva de aprendizado e operação eficiente do sistema homem/máquina.	Antes da contratação ou no momento de concepção da ferramenta e durante o emprego da mesma para avaliação dos resultados.	Por meio de descrição técnica e reuniões de apresentação pela equipe de TI aos membros do tribunal na comissão.
CNJ	Para avaliação da eficácia e grau de produtividade com o emprego da ferramenta, de modo a integrar experiências bem-sucedidas entre diferentes Tribunais.	Informações sobre o grau de produtividade anterior e após o emprego da ferramenta, índices de aproveitamento dos resultados do sistema de IA.	No momento de adoção da ferramenta sobre previsão de produtividade e durante seu emprego.	Apresentação técnica do modelo e tipo de relação homem/máquina e por relatórios periódicos (semestrais ou anuais).

Fonte: elaborado pelos autores.

Quadro 6 – Objeto de transparência: riscos e benefícios do sistema de IA

QUEM	POR QUE	O QUE	QUANDO	COMO
Usuários	Para que possam compreender e melhor cooperar com o uso da ferramenta pelo Tribunal e monitorar o funcionamento das mesmas em casos específicos para que possam tanto proteger suas prerrogativas processuais quanto informar o tribunal sobre possibilidades de aperfeiçoamento e mitigação de riscos.	Informações sobre o grau de confiabilidade, incluindo indicadores de precisão, cobertura e acurácia, dados pessoais das partes eventualmente empregados e categorização em perfis, bem como potenciais impactos do uso da ferramenta em relação a prerrogativas processuais, incluindo possíveis efeitos quanto à paridade entre as partes.	Em local destacado no documento que registra a movimentação processual nos autos.	No momento em que for registrado na movimentação processual o resultado da decisão integral ou parcialmente automatizada.
Cidadãos	Para que possam avaliar e exercer o controle democrático quanto à adequação das ferramentas para a prestação do serviço jurisdicional, sopesando seus benefícios em relação aos riscos.	<p>a) Informações sobre o aumento de produtividade ou melhoria de qualidade dos resultados derivado do uso da ferramenta, sobre riscos de erro, discriminação, privacidade e opacidade, aliado às medidas adotadas para sua mitigação.</p> <p>b) Informações sobre os índices de acerto ou revisão de <i>outputs</i> da ferramenta.</p> <p>c) Justificativas para a decisão de adoção da ferramenta considerando esses fatores;</p> <p>d) Informações e justificativas para a descontinuação do uso de determinada ferramenta e sobre medidas de cibersegurança adotadas.</p>	Permanente durante todo o ciclo de desenvolvimento e emprego da ferramenta até sua descontinuação.	No <i>site</i> oficial do tribunal em <i>link</i> de fácil visualização e acesso.

QUEM	POR QUE	O QUE	QUANDO	COMO
Magistrados e servidores da atividade-fim	Para que possam potencializar os benefícios, mitigar os riscos e informar sobre resultados que possam indicar aspectos negativos do emprego da ferramenta e compreender o desempenho da máquina e sua relação com a mesma na sua atuação.	a) Grau de confiabilidade do <i>output</i> específico da ferramenta. b) Margens de incerteza em determinados tipos de <i>output</i> . c) Informações sobre como serão tratados, pela ferramenta e pelo Tribunal, aqueles dados pessoais ou dados sobre desempenho específico do magistrado ou servidor. d) Informações sobre riscos de discriminação ³⁹ e sobre ganhos de produtividade do servidor com o uso da ferramenta.	Informações gerais de modo permanente e indicadores específicos sobre a atuação do servidor por relatórios periódicos.	a) Informações específicas, relatórios individualizados de desempenho do servidor acessíveis na intranet somente pelo servidor e seus gestores/ superiores. b) Informações gerais acessíveis a todos os servidores e transmitidas em oportunidades de treinamento.
Servidores de TI	Para que possam monitorar a produção dos benefícios, bem como os riscos e adotar ações para sua mitigação.	Índices de desempenho técnico do sistema em relação às variáveis relevantes e a base de dados empregada, parâmetros relevantes conforme o relatório de impacto realizado para cada sistema de IA e resultados de testes de confiabilidade, resiliência e presença de vieses.	Durante todo o ciclo de vida do sistema de IA no tribunal.	Por meio de relatórios técnicos periódicos.

³⁹ Se houver possibilidade de discriminação racial, o servidor pode acompanhar se seu emprego da ferramenta apresenta resultados tendentes à discriminação racial.

QUEM	POR QUE	O QUE	QUANDO	COMO
Comitê decisório sobre tecnologia/laboratório de TI	Para que possa deliberar sobre a contratação ou desenvolvimento interno da ferramenta, ponderando benefícios e riscos, bem como deliberar sobre sua descontinuação considerando os desempenhos concretos dos servidores após o emprego da mesma.	a) Informações técnicas detalhadas sobre previsão de aumento de produtividade previamente ao emprego e relatórios técnicos sobre produtividade durante seu uso. b) Informações técnicas detalhadas sobre os riscos de erro, discriminação, privacidade, opacidade e cibersegurança, bem como medidas adotadas para sua mitigação. c) Relatórios sobre desempenho concreto em relação aos riscos.	a) Previamente à contratação e no momento de concepção da ferramenta quando se tratar de desenvolvimento interno. b) Periodicamente sobre desempenho da ferramenta e para subsidiar reuniões para deliberação da descontinuação do seu uso.	Por meio de apresentações técnicas e relatórios técnicos com a participação da equipe de TI.
CNJ	Para que possa ter uma visão abrangente sobre os riscos e benefícios das ferramentas utilizadas pelos Tribunais brasileiros e recomendar melhores práticas conforme as experiências relatadas por cada tribunal.	a) Informações sobre benefícios e índices de aumento de produtividade do Tribunal e dos magistrados, bem como sobre riscos da ferramenta previstos antes da adoção, medidas de mitigação adotadas e efeitos concretos após o uso da ferramenta.	Periodicamente (semestral).	Por meio de relatórios encaminhados pelos Tribunais ao CNJ.

Fonte: elaborado pelos autores.

6. Conclusão

O presente artigo delineou uma metodologia para a governança de transparência quanto ao emprego de sistemas de IA no Poder Judiciário. A metodologia estrutura-se a partir da concepção de transparência como uma relação comunicacional que envolve o provimento de informações relevantes e adequadas para cada tipo de interlocutor, interno ou externo à organização.

A abordagem foi abstrata no sentido de que não responde às questões-chave para sistemas de IA em particular, que podem ser objeto de estudos empíricos ou ações de implementação por parte dos Tribunais quando da aplicação, uso e/ou desenvolvimento de seus sistemas próprios. Aqui, restringimo-nos a indicar o tipo de informação a ser fornecida como resposta a cada questão-chave. Nesse sentido, usando como base o modelo trazido por Kaminski, foi criada matriz de transparência para avaliação de modelos de IA existentes no Poder Judiciário brasileiro, em relação aos aspectos “quem”, “por que”, “o que”, “quando” e “como”.

Por limitação de espaço, não foi discutida aqui a transparência quanto à própria política de governança de IA pelos Tribunais, uma espécie de transparência de segunda ordem. Trata-se de aspecto igualmente relevante, tendo em vista o objetivo de controle democrático das ações do tribunal e de organizações em geral. Para se proporcionar confiança em sistemas de IA, além da transparência quanto ao uso e sua operação, há uma série de medidas de tratamento dos dados, mensuração de acurácia, controle de vieses, documentação e rastreabilidade, que constituem boas práticas de governança. A informação sobre quais práticas são adotadas por aqueles que empregam o sistema de IA promove a autonomia daqueles impactados pelos sistemas de IA, que, no caso do Judiciário, abrange a sociedade em geral.

Nesse aspecto, o CNJ tem papel fundamental, como indicado na Seção 3. Embora a regulação preveja o registro das aplicações de IA no sistema Sinapses, essa ação é limitada. Isso porque o potencial de integração das tecnologias entre os diferentes Tribunais pode ser perdido quando a semântica de categorização dos dados relevantes for diversa ou mesmo incompatível. Assim, desde a concepção ou contratação do sistema de IA, seria recomendável algum grau de envolvimento do CNJ para aumentar o potencial de benefícios de uma nova ferramenta no sistema judicial brasileiro como um todo e também para evitar duplicação e gastos desnecessários por Tribunais quando já houver ferramenta semelhante disponível e que possa ser adaptada por outros Tribunais.

Para que essa missão tenha resultados efetivos, é importante um papel ativo do CNJ em produzir relatórios de análise sobre o potencial de integração entre diferentes ferramentas e definir *standards* para a uniformização de parâmetros, bem como a divulgação da política de governança de IA perante os Tribunais e avaliações gerais de risco e benefícios alcançados, como forma de pautar a política de implementação de IA no Poder Judiciário e, assim, viabilizar o controle democrático dessas iniciativas na instituição.

Referências

ANDERSON, Chris. The end of theory: the data deluge makes the scientific method obsolete. **Wired**, 23 June 2000. Disponível em: <https://www.wired.com/2008/06/pb-theory/>. Acesso em: 3 out. 2023.

ARRIETA, Alejandro Barredo *et al.* Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. **Information Fusion**, v. 58, p. 82-115, June 2020. Disponível em: <https://arxiv.org/pdf/1910.10045>. Acesso em: 3 out. 2023.

BARFIELD, Woodrow (ed.). **The Cambridge handbook of the law of algorithms**. Cambridge: Cambridge University Press, 2021.

BURRELL, Jenna. How the machine thinks: understanding opacity in machine learning algorithms. **Big Data & Society**, v. 3, n. 1, Jan. 2016. Disponível em: <https://doi.org/10.1177/2053951715622512>. Acesso em: 3 out. 2023.

BUSUIOC, Madalina. Accountable artificial intelligence: holding algorithms to account. **Public Administration Review**, v. 81, n. 5, p. 825-836, Sept./Oct. 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/puar.13293>. Acesso em: 3 out. 2023.

COECKELBERGH, Mark. **AI ethics**. Cambridge, MA: MIT Press, 2020.

DIAKOPOULOS, Nicholas. Transparency. *In*: DUBBER, Markus D.; PASQUALE, Frank; DAS, Sunit (ed.). **The Oxford handbook of ethics of AI**. New York: Oxford University Press, 2020. p. 197-213.

DIGNUM, Virginia. Ethics in artificial intelligence: introduction to the special issue. **Ethics and Information Technology**, v. 20, n. 1, p. 1-3, 2018. Disponível em: <https://link.springer.com/article/10.1007/s10676-018-9450-z>. Acesso em: 3 out. 2023.

DIGNUM, Virginia. **Responsible artificial intelligence: how to develop and use AI in a responsible way.** Cham, Switzerland: Springer Nature, 2019. (Artificial intelligence: foundations, theory, and algorithms).

ENGSTROM, David Freeman; HO, Daniel E. Artificially intelligent government: a review and agenda. *In: VOGL, Roland (ed.). Research handbook on big data law.* Cheltenham, UK: Edward Elgar, 2021. p. 57-86. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3551549. Acesso em: 3 out. 2023.

FLORIDI, Luciano. The end of an era: from self-regulation to hard law in the digital industry. **Philosophy & Technology**, v. 34, n.4, p. 619-622, Dec. 2021. Disponível em: <https://link.springer.com/article/10.1007/s13347-021-00493-0>. Acesso em: 5 out. 2023.

FLORIDI, Luciano *et al.* AI4people – an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. **Minds and Machines**, v. 28, n. 4, p. 689-707, Nov. 2018. Disponível em: <https://link.springer.com/article/10.1007/s11023-018-9482-5>. Acesso em: 3 out. 2023.

FLORIDI, Luciano *et al.* **capAI**: a procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act. Oxford, UK: Saïd Business School, University of Oxford, 2022. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4064091. Acesso em: 4 out. 2023.

GAETANI, Francisco; ALMEIDA, Virgílio. Inteligência artificial e democracia. **Valor Econômico**, São Paulo, 22 jul. 2021. Disponível em: <https://valor.globo.com/opinia0/coluna/inteligencia-artificial-e-democracia.ghtml>. Acesso em: 23 fev. 2023.

JOBIN, Anna; IENCA, Marcello; VAYENA, Effy. The global landscape of AI ethics guidelines. **Nature Machine Intelligence**, v. 1, p. 389-399, 2019. Disponível em: <https://www.nature.com/articles/s42256-019-0088-2>. Acesso em: 3 out. 2023.

JUNQUILHO, Tainá Aguiar. **Inteligência artificial no direito: limites éticos.** São Paulo: Juspodivm, 2022.

JUNQUILHO, Tainá Aguiar; SUGAI, Kin Modesto; VALENTE, Ana Cláudia de Souza. A inteligência artificial no TJDF: estratégias de publicidade e princípio da transparência. *In: FARRANHA, Ana Cláudia. Direito ao acesso à informação: perspectivas no âmbito da tecnologia e governança digital.* Campinas: Pontes, 2022. p. 163-178.

KAMINSKI, Margot E. Understanding transparency in algorithmic accountability. *In: BARFIEL, Woodrow (ed.). **Forthcoming in Cambridge handbook of the law of algorithms***. Cambridge, UK: Cambridge University Press, 2021. p. 121-140. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3622657. Acesso em: 3 out. 2023.

LEHR, David; OHM, Paul. Playing with the data: what legal scholars should learn about machine learning. **UC Davis Law Review**, v. 51, n. 2, p. 653-717, Dec. 2017. Disponível em: https://lawreview.law.ucdavis.edu/issues/51/2/Symposium/51-2_Lehr_Ohm.pdf. Acesso em: 3 out. 2023.

MARANHÃO, Juliano Souza de Albuquerque; COZMAN, Fabio Gagliardi; ALMADA, Marco. Conceções de explicação e do direito à explicação de decisões automatizadas. *In: VAINZOF, Rony; GUTIERREZ, Andriei Guerrero (coord.). **Inteligência artificial: sociedade, economia e Estado***. São Paulo: Revista dos Tribunais, 2021. p. 137-154.

MITTELSTADT, Brent. Principles alone cannot guarantee ethical AI. **Nature Machine Intelligence**, v. 1, 501-507, Nov. 2019. Disponível em: <https://www.nature.com/articles/s42256-019-0114-4>. Acesso em: 5 out. 2023.

PINHANHEZ, Claudio S. Expose uncertainty, instill distrust, avoid explanations: towards ethical guidelines for AI. *In: NEURIPS 2021 WORKSHOP ON HUMAN-CENTERED AI, 2021, online. [Proceedings...]*. Disponível em: <https://arxiv.org/abs/2112.01281>. Acesso em: 16 abr. 2023.

SALOMÃO, Luis Felipe (coord.). **Inteligência artificial: tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário brasileiro**. 2. ed. Rio de Janeiro: FGV Conhecimento Centro de Inovação, Administração e Pesquisa do Judiciário, 2022. Disponível em: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/relatorio_ia_2fase.pdf. Acesso em: 2 fev. 2023.

WISCHMEYER, Thomas. Artificial intelligence and transparency: opening the black box. *In: WISCHMEYER, Thomas; RADEMACHER, Timo (ed.). **Regulating artificial intelligence***. Cham, Switzerland: Springer Nature, 2020. p. 75-101.

WISCHMEYER, Thomas; RADEMACHER, Timo (ed.). **Regulating artificial intelligence**. Cham, Switzerland: Springer Nature, 2020.

ZHONG, Jinfeng; NEGRE, Elsa. AI: to interpret or to explain? *In: CONGRÈS INFORSID 2021: INFormatique des Organisations et Systèmes d'Information et de Décision*, 39., 2021, Dijon, France. [Actes...]. Disponível em: <https://hal.science/hal-03529203/document>. Acesso em: 3 out. 2023.

Legislação citada

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. Disponível em: https://legis.senado.leg.br/diarios/BuscaDiario?tipDiario=1&datDiario=09/12/2022&paginaDireta=3&indSuplemento=Sim&codSuplemento=B&desVolumeSuplemento=&desTomoSuplemento=&_gl=1*hulqbi*_ga*MjgwNzg1MDUzLjE2MjQ0ODA3Nzk.*_ga_CW3ZH25XMK*MTY5NjUzMDI1MS4yNS4xLjE2OTY1MzA3MDcuMC4wLjA. Acesso em: 2 abr. 2023.

BRASIL. Conselho Nacional de Justiça. Portaria nº 271, de 04/12/2020. Regulamenta o uso de Inteligência Artificial no âmbito do Poder Judiciário. **DJe/CNJ**, n. 393, p. 2-4, 14 dez. 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 7 mar. 2023.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE 7000-2021**: IEEE model process for addressing ethical concerns during system design. New York: IEEE, 2021. Disponível em: <https://standards.ieee.org/ieee/24748-7000/11098/7000/6781/>. Acesso em: 4 out. 2023.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE 7001-2021**: IEEE standard for transparency of autonomous systems. New York: IEEE, 2021. Disponível em: <https://standards.ieee.org/ieee/7001/6929/>. Acesso em: 4 out. 2023.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE P2863**: recommended practice for organizational governance of artificial intelligence. New York: IEEE, 2020. Disponível em: <https://standards.ieee.org/ieee/2863/10142/>. Acesso em: 4 out. 2023.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE P2894**: IEEE draft guide for an architectural framework for explainable artificial intelligence. New York: IEEE, 2023. Disponível em: <https://standards.ieee.org/ieee/2894/11296/>. Acesso em: 4 out. 2023.

INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. **IEEE P7003**: algorithmic bias considerations. New York: IEEE, 2023. Disponível em: <https://standards.ieee.org/ieee/7003/11357/>. Acesso em: 4 out. 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC TR 24027:2021**: information technology – artificial intelligence (AI) – bias in AI systems and AI aided decision making. Geneva, Switzerland: ISO, 2021. Disponível em: <https://www.iso.org/standard/77607.html>. Acesso em: 4 out. 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO/IEC 38507:2022**: information technology – governance of IT – governance implications of the use of artificial intelligence by organizations. Disponível em: <https://www.iso.org/standard/56641.html>. Acesso em: 4 out. 2023.

UNIÃO EUROPEIA. Comissão Europeia. **The Artificial Intelligence Act**. Bruxelas: EU, 2021. Disponível em: <https://artificialintelligenceact.eu/>. Acesso em: 2 abr. 2023.