

Banco de Incidentes de Inteligência Artificial do Judiciário - Dissertação Mestrado ENFAM

EUNICE MARIA BATISTA PRADO <eunice.prado@tjpe.jus.br>

qui 13/10/2022 17:26

Para:CJSUBIA <CJSUBIA@senado.leg.br>; stj.gmrvbc@stj.jus.br <stj.gmrvbc@stj.jus.br>;

✉ 4 anexos

Dissertação Banco de Incidentes de IA do Judiciário - registrada em blockchain.pdf; Ata de Aprovação com Louvor SEI_STJ - 3084912.pdf; Portaria CNJ 197, de 22.11.2019 - Grupo de Trabalho Ética e IA.pdf; Resolução CNJ 332 de 21.08.2020 texto original.pdf;

Você não costuma receber emails de eunice.prado@tjpe.jus.br. [Saiba por que isso é importante](#)

Exmo. Sr. Presidente da Comissão de Juristas - Inteligência Artificial

Ministro Ricardo Villas-Bôas Cueva

No intuito de colaborar com os trabalhos da Comissão de Juristas CJSUBIA no Senado Federal, encaminho em anexo minha dissertação de mestrado com o título “**Banco de Incidentes de Inteligência Artificial do Judiciário**”, defendida em 22/08/2022 e aprovada com louvor, no âmbito do Mestrado Profissional em Direito e Poder Judiciário da ENFAM – Escola Nacional de Formação e Aperfeiçoamento de Magistrados, vinculada ao Superior Tribunal de Justiça. O arquivo em PDF está com a ficha catalográfica e certificado de registro em blockchain.

O trabalho se debruçou sobre o **tratamento dos erros, incidentes e eventos adversos que vierem a ocorrer com o uso da IA** aplicada nos tribunais brasileiros, propondo sua catalogação num banco de dados formado a partir de registros de usuários, de modo a criar uma ferramenta útil ao gerenciamento de riscos dessa tecnologia, atendendo aos princípios da prevenção e da precaução.

Embora a pesquisa tenha ficado restrita ao uso da IA em tribunais, **a proposta de banco de incidentes pode ser utilizada para todas as áreas de aplicação**, e inclusive existe uma iniciativa global nesse sentido, o *AI Incident Database* <https://incidentdatabase.ai/>, que foi apontado na dissertação como um dos parâmetros possíveis para adaptação no Brasil.

Informo ainda que fui uma das magistradas integrantes do Grupo de Trabalho criado no Conselho Nacional de Justiça pela Portaria nº 197/2019, de modo que participei da elaboração dos estudos e propostas que embasaram a Resolução nº 332/2020, que dispõe sobre a ética, a transparência e a governança na produção e no uso da Inteligência Artificial no Poder Judiciário (documentos anexos).

Assim, venho colocar minha contribuição acadêmica e profissional à disposição da CJSUBIA.

Atenciosamente,

Eunice Maria Batista Prado

Juíza de Direito no Tribunal de Justiça de Pernambuco

Mestra em Direito e Poder Judiciário

Pesquisadora do Grupo de Pesquisa "Justiça Digital: Tecnologias Disruptivas e Inteligência Artificial no Direito" (GEPDI7/ENFAM, registrado no CNPq)

Membro do Conselho de Inovação do AMBLab - Laboratório de Inovação e Inteligência da Associação dos Magistrados Brasileiros



Certificado de Autenticidade

PACDigital

Hash do Documento Original

f75ad65046f8ab93fa92d714f14a50965e3b802295812c8f3fb9ddb4d784aaa5

Registrado por:

Nome
EUNICE MARIA BATISTA PRADO

E-mail
euniceprado_@hotmail.com



Redes Blockchain

O hash que autentica o documento original foi registrado na(s) rede(s) blockchain listadas abaixo e pode ser conferido no endereço indicado pelo identificador da transação.



Enviado
02/10/2022 02:30:07 -03:00 UTC

Confirmado
02/10/2022 02:30:02 -03:00 UTC

ETC

Identificador da transação
0x05dd17c5edd0ed0f907f15c52d6ccf3cdcc9992fd38
cad1f1de5d022611f6f26



Enviado
02/10/2022 02:00:00 -03:00 UTC

Confirmado
02/10/2022 03:01:32 -03:00 UTC

DCR

Identificador da transação
3defa7bf2be01d5e51c3dadea149dfecdc8cf33c90313
7bd6dafef65d827d7dc

Merkle:

f75ad65046f8ab93fa92d714f14a50965e3b802295812
c8f3fb9ddb4d784aaa5



Escola Nacional
de Formação
e Aperfeiçoamento
de Magistrados



Programa de Pós-Graduação
Profissional em Direito da Enfam

**ESCOLA NACIONAL DE FORMAÇÃO E APERFEIÇOAMENTO DE
MAGISTRADOS**
PROGRAMA DE PÓS-GRADUAÇÃO PROFISSIONAL EM DIREITO
ÁREA DE CONCENTRAÇÃO: DIREITO E PODER JUDICIÁRIO
CURSO DE MESTRADO PROFISSIONAL

EUNICE MARIA BATISTA PRADO

**BANCO DE INCIDENTES DE INTELIGÊNCIA ARTIFICIAL DO
JUDICIÁRIO**

BRASÍLIA-DF
2022

EUNICE MARIA BATISTA PRADO

**BANCO DE INCIDENTES DE INTELIGÊNCIA ARTIFICIAL DO
JUDICIÁRIO**

Trabalho de conclusão de curso, na modalidade de dissertação, apresentado ao Programa de Pós-Graduação Profissional em Direito da Escola Nacional de Formação e Aperfeiçoamento de Magistrados, como requisito parcial para obtenção do título de Mestre em Direito. Área de concentração: Direito e Poder Judiciário.

Orientador: Prof. Dr. Samuel Meira Brasil Júnior

**BRASÍLIA - DF
2022**

**Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)**

P896b Prado, Eunice Maria Batista.
Banco de incidentes de inteligência artificial do judiciário / Eunice Maria Batista Prado. – Brasília (DF), 2022.
261 f.

Orientador: Prof. Dr. Samuel Meira Brasil Júnior.

Dissertação (Mestrado em Direito) – Escola Nacional de Formação e Aperfeiçoamento de Magistrados.

1. Direito. 2. Inteligência artificial – Judiciário – Incidentes.
3. Administração judiciária. I. Brasil Júnior, Samuel Meira. II. Escola Nacional de Formação e Aperfeiçoamento de Magistrados. III. Título.

CDU 34

EUNICE MARIA BATISTA PRADO

BANCO DE INCIDENTES DE INTELIGÊNCIA ARTIFICIAL DO JUDICIÁRIO

Trabalho de conclusão de curso, na modalidade de dissertação, apresentado ao Programa de Pós-Graduação Profissional em Direito da Escola Nacional de Formação e Aperfeiçoamento de Magistrados, como requisito parcial para obtenção do título de Mestre em Direito. Área de concentração: Direito e Poder Judiciário.

Aprovado em: 22/08/2022.

BANCA EXAMINADORA

Prof. Dr. Samuel Meira Brasil Júnior
Escola Nacional de Formação e Aperfeiçoamento de Magistrados

Prof. Dr. Fabrício Castagna Lunardi
Escola Nacional de Formação e Aperfeiçoamento de Magistrados

Prof. Dr. Richard Paulro Pae Kim
Universidade Santo Amaro

Dedico este trabalho a meu filho Guilherme, que tem 11 anos, é da Geração Alpha (a primeira a ser 100% digital), ama tecnologia e quer ser cientista da computação quando crescer, para que saiba desde cedo que os robôs devem ser desenvolvidos para o bem dos humanos.

AGRADECIMENTOS

A Deus, que não me faltou nesses tempos difíceis.

A minha mãe, Graça Prado, porque sem seu apoio e incentivo teria sido simplesmente impossível fazer esse curso de mestrado e ao mesmo tempo ser mãe de um filho pré-adolescente e juíza em duas unidades, tendo enfrentado a organização das eleições municipais de 2020 e sido chamada a colaborar com o CNJ em dois grupos de trabalho, tudo isso no meio de uma pandemia.

A meu filho, Guilherme, pela paciência e compreensão. O mais difícil de tudo foi lidar com o sentimento de culpa por destinar para o mestrado parte do tempo que era para ser todo seu. Faltou tempo, mas sobrou amor.

A meus irmãos, Francisco e Fábio, que mesmo estando em outro Estado e até outro país me ajudaram à distância com torcida e suporte emocional.

A Matheus, por estar sempre presente, ouvindo, incentivando, apoiando.

Às colegas do mestrado, que se tornaram amigas mais que especiais. Unidas por um sentimento em comum – a certeza de que o desafio seria maior para nós mulheres magistradas, com nossos múltiplos papéis e tripla jornada – soubemos dar amparo e força umas às outras mesmo que virtualmente, com a ideia genial da nossa querida Claudia, *expert* da Justiça Restaurativa, que aplicou a técnica dos círculos de diálogo entre nós mesmas. Assim, pudemos preservar nossa saúde mental apesar da sobrecarga, construindo uma relação de sororidade em vez de competição (que ficou restrita ao período do difícil processo seletivo). Adriana Coningham, Adriana Nóbrega, Audrey Kramy, Carmen Ramajo, Claudia Catafesta, Cristina Vieira, Daniela Bandeira, Daniella Prado, Fabiane Saraiva, Ítala Colnaghi, Lívia Borba, Marcela Lobo, Mariana Yoshida, Michelle Amorim e Raffaela Sousa, muito obrigada a todas.

Aos colegas magistrados da primeira turma, pela amizade, companheirismo e parceria nas diversas atividades e disciplinas.

Aos colegas do Grupo de Pesquisa 7 pelos debates, pela rica troca de conhecimento e pelo compartilhamento de material de pesquisa.

A meu orientador, Desembargador Samuel Brasil Jr., pelo tratamento acolhedor, compreensivo, respeitoso e inspirador, sem deixar de ser exigente e rigoroso, que me encorajou a buscar entregar o melhor resultado possível.

À professora Luciane Münch e sua irmã Márcia Villarroel, pela exitosa parceria no artigo científico que publicamos juntas na revista Direito Público, e às

professoras Luíza Figueiredo e Priscilla Corrêa, pela também exitosa parceria que resultou em artigo publicado na revista do CNJ e em pesquisa nacional com a AMB.

Aos magistrados-professores das disciplinas que cursei, André Bezerra e José Marcos Lunardelli (Pesquisa e Metodologia/Sistema de Justiça e Pesquisa Empírica), Carlos Haddad (Administração Judicial Aplicada), Roger Raupp (Ética e Integridade), Rosimeire Ventura (Comunicação e Transparência no Poder Judiciário) e Taís Ferraz (Prevenção de Conflitos e Gestão de Precedentes nos Centros de Inteligência), além dos já mencionados Samuel Brasil Jr. (O Judiciário na Era Digital: Ciências Cognitivas e Tecnologias Disruptivas no Judiciário), Luciane Münch (Inovação e Design Organizacional do Poder Judiciário) e Luíza Figueiredo (Liderança Organizacional e Atividade Judicial), pelos valiosos ensinamentos, ampliando meus horizontes. Neste trabalho há influência de todos.

A Cíntia Brunetta, Victor Magalhães, Jaqueline Mello e todos que fazem parte da ENFAM. Deixo uma homenagem especial ao saudoso Desembargador Eladio Lecey, que se emocionava quando via o sucesso dos projetos da Escola.

À Associação dos Magistrados Brasileiros (AMB), Associação dos Juízes Federais do Brasil (AJUFE) e Associação Nacional dos Magistrados do Trabalho (ANAMATRA), pela colaboração ao enviar o *survey* para os magistrados de todo o país questionando seu nível de conhecimento em Inteligência Artificial, permitindo alcançar o nível de confiança desejado na amostra para o artigo científico publicado.

Ao Desembargador Jones Figueirêdo Alves, decano do Tribunal de Justiça de Pernambuco, pelas palavras de encorajamento e incentivo desde o primeiro momento em que externei minha vontade de fazer um mestrado. Jamais esquecerei.

Aos amigos Rafael Lemos e Carla Moutinho, pelo material e incentivo.

A João Paulo Vasconcelos e Fernando Macedo, da 4ª Vara Cível de Olinda, por todo o apoio nesse período.

Ao Desembargador aposentado José Eduardo Resende Chaves Jr., o “Pepe”, do Tribunal Regional do Trabalho de Minas Gerais, que em abril de 2019 me convidou para participar do IDEIA (Instituto Direito e Inteligência Artificial), permitindo-me desde então aprender diariamente com grandes especialistas do Brasil e do exterior. Vários temas debatidos foram abordados neste trabalho.

Finalmente, um agradecimento muito especial ao Ministro Herman Benjamin, que tanto se empenhou para tornar realidade o Mestrado Profissional da ENFAM, do qual me orgulho de fazer parte da primeira turma.

RESUMO

Esta é uma pesquisa interdisciplinar sobre gerenciamento dos riscos de ferramentas de Inteligência Artificial utilizadas como apoio à prestação jurisdicional nos tribunais brasileiros, que objetiva encontrar uma estrutura eficaz para o tratamento dos eventos adversos que vierem a ocorrer, à luz do princípio da precaução, de modo que o conhecimento sobre falhas verificadas ajude a evitar incidentes futuros, preservando os direitos fundamentais dos jurisdicionados. Mediante pesquisa bibliográfica e documental em diversas áreas do conhecimento, elaborou-se uma visão holística do ciclo de desenvolvimento da tecnologia, acrescentando contexto histórico-cronológico dos avanços científicos, para proceder à descrição e análise crítica de diretrizes éticas internacionais sobre Inteligência Artificial; da estrutura regulatória brasileira; dos riscos relacionados a ataques cibernéticos, discriminação algorítmica, falta de transparência e controle do usuário; do panorama de aplicação da tecnologia nos tribunais; das formas disponíveis para controle; até chegar na última etapa do ciclo de gerenciamento de riscos, quanto ao tratamento dos erros e consequências negativas. Os principais resultados obtidos foram dois parâmetros de bancos de dados para registro e armazenamento dos eventos adversos, MAUDE/MedWatch e AI Incident Database. A conclusão é que tais parâmetros podem ser adaptados em laboratório de inovação para o aprimoramento do controle da Inteligência Artificial usada pelo Poder Judiciário brasileiro, em sua transformação digital e inovação responsável.

Palavras-chave: banco de dados; incidentes; inteligência artificial; tribunais.

ABSTRACT

This is an interdisciplinary research on risk management for Artificial Intelligence tools used to support the provision of judicial services in Brazilian courts, which aims to find an effective structure for dealing with adverse events that may occur, in light of the precautionary principle, in a way that knowledge about verified failures can help prevent future incidents, while preserving the fundamental rights of the people. Through bibliographic and documentary research in various fields of knowledge, a holistic view of the technology development cycle was elaborated, adding the historical-chronological context of scientific advances, to proceed with the description and critical analysis of international ethical guidelines on Artificial Intelligence; the Brazilian regulatory structure; the risks related to cyber attacks, algorithmic discrimination, lack of transparency and user control; the panorama of technology application in the courts; the methods available for control; until reaching the last stage of the risk management cycle, regarding the treatment of errors and negative consequences. The main results obtained were two database parameters for recording and storing adverse events, MAUDE/MedWatch and AI Incident Database. The conclusion is that such parameters can be adapted in an innovation lab to improve the control of Artificial Intelligence used by the Brazilian Judiciary, in its digital transformation and responsible innovation.

Keywords: database; incidents; artificial intelligence; courts.

SUMÁRIO

1	INTRODUÇÃO.....	9
2	DIRETRIZES INTERNACIONAIS PARA USO ÉTICO DA IA.....	17
2.1	Os Princípios de Asilomar na interseção entre IA e Bioética.....	18
2.2	Os Princípios de IA da OCDE.....	32
2.3	A Carta Europeia de Ética Sobre o Uso da IA em Sistemas Judiciais.....	40
3	REGULAÇÃO DOS RISCOS DA IA NO BRASIL.....	54
3.1	A Constituição Federal de 1988 à luz da “ética para a civilização tecnológica” de Hans Jonas e da “sociedade de risco” de Ulrich Beck.....	54
3.2	O Projeto de Lei nº 21/2020 e o futuro marco legal da IA no Brasil.....	61
3.3	A Estratégia Brasileira de Inteligência Artificial.....	66
3.4	A Resolução nº 332/2020 do Conselho Nacional de Justiça.....	69
3.4.1	<i>Os principais pilares da Resolução CNJ nº 332/2020.....</i>	72
3.4.2	<i>Riscos relacionados a falhas na cibersegurança e proteção de dados.....</i>	75
3.4.3	<i>Riscos de modelos de IA discriminatórios.....</i>	87
3.4.4	<i>Riscos da falta de transparência algorítmica.....</i>	112
3.4.5	<i>Riscos pela falta de controle do usuário.....</i>	133
3.5	A Portaria nº 271/2020 do Conselho Nacional de Justiça.....	144
4	APLICAÇÃO DA IA NOS TRIBUNAIS BRASILEIROS, CONTROLE DOS RISCOS E TRATAMENTO DOS INCIDENTES.....	150
4.1	Panorama de aplicação da IA nos tribunais.....	151
4.1.1	<i>Projetos de IA na Justiça Estadual.....</i>	155
4.1.2	<i>Projetos de IA na Justiça Federal.....</i>	166
4.1.3	<i>Projetos de IA na Justiça Eleitoral.....</i>	170
4.1.4	<i>Projetos de IA na Justiça do Trabalho.....</i>	171
4.1.5	<i>Projetos de IA nos Tribunais Superiores.....</i>	173
4.1.6	<i>Projetos de IA nos Conselhos.....</i>	174
4.2	Controle dos riscos da IA nos tribunais.....	175
4.3	Tratamento dos erros e incidentes de IA nos tribunais.....	200
4.3.1	<i>Critérios de solução pela metodologia de design thinking.....</i>	214
4.3.2	<i>“MAUDE database” e sistema “MedWatch” como parâmetros.....</i>	222
4.3.3	<i>“AI Incident Database” como parâmetro.....</i>	228
5	CONCLUSÃO.....	232
	REFERÊNCIAS.....	238

1 INTRODUÇÃO

“Devido ao grande potencial da IA, é importante pesquisar como colher seus benefícios ao mesmo tempo em que evitamos potenciais imprevistos”¹, disse o físico britânico Stephen Hawking no último livro que estava escrevendo quando faleceu em 2018, baseado em seus arquivos pessoais de palestras e entrevistas sobre temas candentes a que fora instado a opinar durante a vida.

Nessa linha de pensamento, com foco na chamada à precaução que se depreende da frase, esta é uma pesquisa interdisciplinar² que envolve tema concernente à auditoria, controle e gerenciamento de riscos da Inteligência Artificial (IA), especificamente quando utilizada como ferramenta auxiliar pelo Poder Judiciário brasileiro, objetivando contribuir para que seja conferida aplicabilidade prática ao art. 27 da Resolução nº 332/2020 do Conselho Nacional de Justiça (CNJ), que dispõe o seguinte: “Art. 27. Os órgãos do Poder Judiciário informarão ao Conselho Nacional de Justiça todos os registros de eventos adversos no uso da Inteligência Artificial”³.

Tal Resolução, baseada na Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus Ambientes⁴, trata sobre ética, transparência e governança no desenvolvimento e utilização de IA no Judiciário, trazendo entre seus dispositivos o referido art. 27 como uma providência a ser adotada em caso de incidentes.

Acontece que a Resolução, assumindo de antemão que é esperado que falhas aconteçam, estabeleceu esse dever de informar eventos adversos e manter registros, mas não definiu como fazer isso, incorrendo na crítica de Morley et al. quanto ao debate ético sobre IA que se popularizou nos últimos anos, que “focalizou essencialmente os princípios – o ‘o quê’ da ética da IA (beneficência, não maleficência, autonomia, equidade e explicabilidade) – em vez de práticas, o ‘como’”⁵.

¹ HAWKING, Stephen. **Breves respostas para grandes questões**. Rio de Janeiro: Editora Intrínseca, 2018. p. 94.

² KUSTERS, Remy et al. Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities. **Frontiers in Big Data**, v. 3, p. 45, 2020. Disponível em: <https://www.frontiersin.org/articles/10.3389/fdata.2020.577974/full>. Acesso em: 11 out. 2021.

³ CNJ. **Resolução nº 332, de 21 de agosto de 2020**.

Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 10 jul. 2021.

⁴ CEPEJ. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente**. Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-re-vista/168093b7e0>. Acesso em: 10 jul. 2021.

⁵ MORLEY, Jessica et al. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. In: **Ethics, Governance, and Policies in Artificial Intelligence**. Springer, Cham, 2021. p. 153-183.

A lacuna apontada levou à formulação da seguinte pergunta de pesquisa: como criar um método eficaz para informar e registrar as informações sobre os eventos adversos que vierem a ocorrer com o uso da IA pelo Poder Judiciário, de um modo que seja útil ao efetivo controle e gerenciamento de riscos?

A par desse questionamento principal, formularam-se outras perguntas: Existem iniciativas desse tipo em outros países? Em caso positivo, seria possível adaptá-las para uso no Poder Judiciário brasileiro?

Em busca de parâmetros, após estudo exploratório sem nada semelhante encontrar em sistemas judiciais de outros países, a ideia para esta pesquisa interdisciplinar surgiu de reflexões provocadas com a leitura do livro “A Quarta Revolução Industrial”⁶, de Klaus Schwab, fundador e presidente executivo do Fórum Econômico Mundial, que escreveu a obra para servir de fio condutor da Reunião Anual de 2016 em Davos-Klosters na Suíça, cujo tema naquele ano foi “Para dominar a Quarta Revolução Industrial”⁷.

Nesse contexto, após falar das megatendências⁸ para o presente e futuro, nas categorias física e digital – entre elas, a IA –, Schwab adentrou no campo biológico, abordando temas como sequenciamento de genomas, biologia sintética, engenharia genética, medicina de precisão, edição biológica etc., quando então vaticinou: “Na verdade, a ciência avança tão rápido que, no momento, as limitações são mais jurídicas, regulamentares e éticas que técnicas”⁹. E, mais adiante, afirmou: “Acredito que os maiores desafios em relação ao desenvolvimento de normas sociais e regulamentações adequadas estão no domínio biológico”¹⁰.

Essas frases provocaram a seguinte reflexão: se os maiores desafios éticos, segundo Schwab, diante da velocidade vertiginosa das transformações tecnológicas em curso, está no campo das ciências biológicas e da saúde – que lidam com o direito humano mais elementar e ao mesmo tempo mais relevante, que é a vida, então como está sendo tratado o gerenciamento dos riscos da IA aplicada, por exemplo, à Medicina? E como o Direito – que, além da vida, lida com outros valores também importantes como liberdade, igualdade, dignidade, honra, propriedade, entre

⁶ SCHWAB, Klaus. **A quarta revolução industrial**. Tradução Daniel Moreira Miranda. São Paulo: Edipro, 2016. Tradução de: *The Fourth Industrial Revolution*.

⁷ Ibidem. pp. 5-7.

⁸ Ibidem. pp. 23-29.

⁹ Ibidem. p. 30.

¹⁰ Ibidem. p. 31.

outros –, poderá extraír lições e parâmetros a partir do rigor científico das pesquisas na área médica?

Sob outro prisma, bem poderiam ser prospectados modelos de gerenciamento de riscos da IA aplicada a inúmeros setores, como indústria, aeronáutica, mercado financeiro, mercado de trabalho, entre outros, a partir de critérios elaborados em diferentes áreas do conhecimento humano como Engenharia e Economia, por exemplo. Contudo, optou-se inicialmente neste trabalho pelo recorte da área médica, pelos motivos já explicitados acima, sobretudo por lidar diretamente com o mais fundamental de todos os direitos humanos, a vida, “já que se constitui em pré-requisito à existência e exercício de todos os demais”¹¹, conforme Alexandre de Moraes.

Nesse diapasão, esta pesquisa irá investigar a hipótese de saber se é viável adequar, para o Poder Judiciário brasileiro, o método utilizado na Medicina norte-americana para controle de eventos adversos causados por dispositivos médicos e cirurgias robóticas, a partir da experiência dos usuários, tomando como modelo o banco de dados MAUDE – *Manufacturer and User Facility Device Experience*¹², que tem como órgão regulatório a FDA – *U.S. Food & Drug Administration*.

A ideia inicial é que, com as devidas adaptações, um banco de dados em moldes semelhantes talvez possa, no Brasil, servir de suporte ao processo decisório do CNJ quanto às aplicações de IA nos tribunais que devem ser mantidas, encorajadas, aperfeiçoadas, limitadas ou proibidas, conforme seu potencial para causar lesão aos direitos humanos fundamentais dos jurisdicionados.

A propósito, pesquisas preliminares realizadas entre os anos de 2018 e 2019 em trabalho acadêmico¹³ anterior na área de Direito do Consumidor e Responsabilidade Civil, já haviam tangenciado superficialmente o tema quanto à existência do banco de dados MAUDE, ao analisar danos a consumidores-pacientes causados por dispositivo médico de braços mecânicos para cirurgias robóticas, que apesar da adjetivação na verdade sequer utilizavam IA.

¹¹ MORAES, Alexandre de. **Direito Constitucional**. 17 ed. São Paulo: Atlas, 2005. p. 30.

¹² FDA. MAUDE - **Manufacturer and User Facility Device Experience**. Disponível em: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Acesso em: 10 jul. 2021.

¹³ PRADO, Eunice M B. Responsabilidade civil por atos de inteligência artificial nas relações de consumo: uma análise dos primeiros casos concretos de falhas em aplicações ao redor do mundo e dos rumos da legislação brasileira sobre o tema. **Monografia** (Especialização em Direito do Consumidor e Responsabilidade Civil) - Centro Universitário Mauricio de Nassau/ESMAPE. Recife, 2020.

Na ocasião, a monografia focou no tema da responsabilidade civil, fazendo um relato descritivo dos primeiros casos noticiados na imprensa internacional quanto a danos a consumidores possivelmente associados a falhas de aplicações de IA.

Em tal trabalho acadêmico, foram reportados prejuízos financeiros por erro de sistema inteligente de investimentos, acidente com morte de pedestre envolvendo veículo autônomo, ofensas a internautas por assistente de conversação em rede social, erro de reconhecimento facial, entre outras situações.

Porém, como dito, não foram encontrados casos concretos e documentados no banco de dados MAUDE quanto a falhas de IA aplicada à Medicina, pelo fato de que, na época, ainda eram muito incipientes as autorizações da agência reguladora norte-americana para uso da tecnologia em pacientes reais.

Paralelamente à pesquisa para a monografia acima referida no âmbito de uma pós-graduação *lato sensu*, estavam sendo desenvolvidos estudos no Grupo de Trabalho “Ética e Inteligência Artificial”, criado pelo CNJ por meio da Portaria nº 197/2019¹⁴, que elaborou as propostas que resultaram na Resolução CNJ nº 332/2020.

Na divisão de tarefas entre os participantes do grupo, esta pesquisadora colaborou nos estudos a respeito do controle do usuário e da responsabilidade civil, apresentando posteriormente suas conclusões à sociedade e à comunidade jurídica em evento oficial do CNJ realizado por videoconferência no dia 04 de setembro de 2020¹⁵, junto com os demais integrantes.

O Grupo de Trabalho, ao concluir os estudos, apontou ao CNJ a necessidade de manter registros sobre eventos adversos, o que foi destacado no voto¹⁶ do Relator e acabou sendo acolhido à unanimidade pelo Plenário, passando a constar no art. 27 da Resolução. Contudo, o fato é que não houve tempo suficiente para os pesquisadores apontarem *como* isso poderia ser feito – o Grupo teve prazo de apenas 90 dias¹⁷, prorrogado por mais 60¹⁸, para conclusão dos trabalhos, e no

¹⁴ CNJ. **Portaria n. 197, de 22 de novembro de 2019.** Disponível em:

<https://atos.cnj.jus.br/files/original190722201911265ddd77ea7b46e.pdf>. Acesso em: 10 jul. 2021.

¹⁵ CNJ. **Webinar A Resolução CNJ n. 332: ética e governança para produção e uso da Inteligência Artificial no Poder Judiciário.** 04 set. 2020. Disponível em: <https://www.cnj.jus.br/agendas/webinar-a-resolucao-n-332-e-a-etica-na-producao-e-uso-da-i-a-no-poder-judiciario/>. Acesso em: 10 jul. 2021.

¹⁶ CNJ. **Procedimento de Ato Normativo nº 0005432-29.2020.2.00.0000**, Plenário Virtual, 14 ago. 2020, votação unânime.

¹⁷ CNJ. **Portaria n. 197, de 22 de novembro de 2019.** Disponível em:

<https://atos.cnj.jus.br/files/original190722201911265ddd77ea7b46e.pdf>. Acesso em: 10 jul. 2021.

¹⁸ CNJ. **Portaria n. 78, de 19 de maio de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3325>. Acesso em: 10 jul. 2021.

transcurso ainda foi impactado com o advento da pandemia de Covid-19¹⁹, o que fez com que as reuniões fossem reduzidas e seguissem apenas à distância.

Assim, quando em julho de 2020 adveio a seleção para a primeira turma do Mestrado Profissional da ENFAM²⁰, já no fim do prazo do Grupo de Trabalho, surgiu a oportunidade ideal para submeter o projeto desta pesquisa, buscando aprofundar os estudos em gerenciamento de riscos de IA aplicada ao Poder Judiciário com tempo suficiente para tentar completar a lacuna apontada, desta feita no âmbito de uma pós-graduação *stricto sensu*, possibilitando a apresentação, no final, de sugestões de aperfeiçoamento da Resolução e de mecanismos de efetivação.

Tendo o projeto sido selecionado, retomou-se a reflexão provocada pelo livro de Klaus Schwab, de possível abordagem interdisciplinar médico-jurídica do tema, por ter sido constatado que, enquanto se passava, no Brasil, a elaboração da Resolução CNJ nº 332/2020 como primeiro instrumento regulatório de IA oficialmente emitido no país por um Poder da República – conforme será demonstrado mais adiante – prosseguia nos Estados Unidos uma discussão pública proposta pela FDA desde abril de 2019 quanto a uma proposta de regulação de IA aplicada a dispositivos médicos²¹.

Encerrado o período de discussão e como resposta às sugestões apontadas pelas partes interessadas, a FDA emitiu, em janeiro de 2021, um plano de ação²² em cinco etapas que pretendia desenvolver até validar *software* baseado em Inteligência Artificial/Aprendizado de Máquina como dispositivo médico (SaMD – *Software as a Medical Device*) – sujeito, portanto, a registro de eventos adversos no banco de dados MAUDE.

Com tais considerações, o objetivo geral desta pesquisa é apresentar uma proposta para concretizar o comando normativo do art. 27 da Resolução CNJ nº

¹⁹ ONU News. **Organização Mundial da Saúde declara novo coronavírus uma pandemia**. Disponível em: <https://news.un.org/pt/story/2020/03/1706881#:~:text=A%20Organiza%C3%A7%C3%A3o%20Mundial%20da%20Sa%C3%BAde,pre-sente%20agora%20em%20114%20pa%C3%ADses>. Acesso em: 10 jul. 2021.

²⁰ STJ. ENFAM/PPGP. **Edital n. 1, de 08 de julho de 2020**. DJe/STJ, 9 jul. 2020. Disponível em: <https://bdjur.stj.jus.br/jspui/handle/2011/144345>. Acesso em: 10 jul. 2021.

²¹ FDA. **Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD): Discussion Paper and Request for Feedback**. Regulations.gov. 2 abr. 2019. Disponível em: <https://www.regulations.gov/document/FDA-2019-N-1185-0001>. Acesso em: 10 jul. 2021.

²² FDA. **Artificial Intelligence and Machine Learning (AI/ML) Software as a Medical Device Action Plan**. 12 jan. 2021. Disponível em: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>. Acesso em: 10 jul. 2021.

332/2020, no sentido de desenvolver um modelo conceitual de banco de dados que possa ser gerenciado pelo CNJ, voltado para catalogar os eventos adversos que forem detectados no uso de aplicações de IA pelos diversos ramos da Justiça brasileira, criando um mecanismo útil para gestão de riscos, precaução, prevenção e aprimoramento, verificando se seria possível tomar como parâmetro o banco de dados estadunidense MAUDE.

Os objetivos específicos são: (i) verificar parâmetros internacionais para gerenciar riscos da aplicação de IA; (ii) examinar o panorama de aplicação de IA nos tribunais brasileiros, investigando se já ocorreram eventos adversos, em que circunstâncias, e se já existem iniciativas para registrá-los; (iii) examinar de que forma o banco de dados MAUDE é estruturado, como é alimentado, como é consultado, como é gerenciado; (iv) identificar quais dados são relevantes para catalogar e parametrizar.

A justificativa da presente pesquisa, no âmbito de um Mestrado Profissional, reside na sua contribuição teórico-prática para buscar suprir uma lacuna sobre tema relevante, que foi recentemente regulado pelo CNJ porém de uma forma vaga, visto que apenas reconheceu expressamente a necessidade de informação e registro de eventos adversos no uso de IA em tribunais, mas não detalhou de que maneira isso deve ser feito.

Nesse diapasão, a pesquisa irá trabalhar com a ideia de um banco de dados nacional que, ao concentrar o registro de eventos adversos ocorridos em aplicações de IA pelo Poder Judiciário brasileiro, em qualquer tribunal do país, permita identificar padrões, causas e recorrências, tornando possível ao CNJ verificar a aderência das aplicações de IA usadas pelos tribunais a princípios éticos e constitucionais, fazendo-lhes controle.

Dessa maneira, supõe-se que o CNJ poderá ter embasamento em dados para determinar a adequação ou mesmo a descontinuidade de algum sistema de IA, evitando-se ou minorando-se lesões a direitos fundamentais dos usuários do sistema de Justiça e consequente responsabilização civil do Estado.

Supõe-se ainda que, na hipótese de ser implementado, o banco também tenha potencial para, no futuro, à medida que for sendo alimentado de dados, servir como ferramenta de consulta para os desenvolvedores de soluções de IA para a Justiça, porque condensará em um só lugar informações sobre o que já foi tentado, o que deu certo, o que deu errado, como foi consertado, se foi possível consertar, e os

desafios a serem superados; podendo ser um manancial de conhecimento técnico que poderá otimizar o tempo de pesquisas e testes.

Para testar a hipótese, há que se levar em conta que a natureza desta pesquisa é eminentemente aplicada, objetivando contribuir de modo propositivo para a solução do problema identificado, visto ser realizada no âmbito de um Mestrado Profissional da escola oficial de treinamento de juízes de direito e juízes federais brasileiros, com área de concentração em Direito e Poder Judiciário.

O trabalho será desenvolvido na Linha de Pesquisa 2 (LP2), Ética, Integridade e Efetividade da Atividade Jurisdicional, que pressupõe uma atuação estratégica e ativa, “sempre com perspectiva ética e preocupação com qualidade e resultados”²³, em busca de solução para um problema relacionado à tecnologia, no momento em que o sistema judicial brasileiro se encontra em franca transformação digital.

Quanto ao percurso metodológico, considerando que a regulação de IA é um tema novo e desafiador para pesquisadores e formuladores de políticas públicas do mundo inteiro, será desenvolvida pesquisa exploratória, descritiva e analítica, numa abordagem de essência qualitativa, ancorando-se sobretudo em pesquisa bibliográfica e documental junto a fontes de outras áreas do conhecimento além do Direito, no campo das ciências humanas, exatas e biológicas, nos cenários nacional e internacional. A abordagem interdisciplinar será acrescida de contexto histórico-cronológico, indicando-se como e por que há relação com o tema pesquisado.

Com tais premissas, a dissertação será desenvolvida em três capítulos.

O primeiro capítulo buscará embasamento teórico-normativo no cenário internacional, descrevendo aquelas que, entre muitas, foram consideradas as principais diretrizes para uso ético e benéfico da IA, em diálogo com o escopo desta pesquisa aplicada ao Poder Judiciário, mas também observando, quando pertinente, a interdisciplinaridade com ciências biológicas e da saúde.

O capítulo foi organizado partindo das diretrizes mais gerais e distantes no tempo, para as mais específicas e recentes, elaboradas coletivamente nos últimos anos por centenas de especialistas de vários países e áreas de atuação.

²³ ENFAM. **Linhas de pesquisa.** Disponível em: <https://www.enfam.jus.br/programa-de-pos-graduacao-profissional-em-direito/mestrado-profissional/pesquisa/linhas-de-pesquisa/>. Acesso em: 6 jan. 2022.

Desenvolveu-se, assim, o estudo de uma sequência lógica de: (i) princípios gerais para aplicações de IA tanto em áreas privadas como públicas (Princípios de IA de Asilomar); (ii) princípios a serem observados por formuladores de políticas públicas (Princípios de IA da Organização para a Cooperação e Desenvolvimento Econômico – OCDE), (iii) princípios específicos a serem observados para aplicação de IA em sistemas judiciais (Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus Ambientes).

O segundo capítulo, dedicado exclusivamente ao cenário nacional, iniciando com referencial teórico sobre o princípio da precaução, descreverá o estado atual da regulação da IA no Brasil, no âmbito de cada um dos três Poderes.

Será dada ênfase aos atos normativos emitidos pelo Conselho Nacional de Justiça, analisando-os criticamente e contextualizando-os com os avanços nas pesquisas científicas sobre IA, além de notícias sobre fatos recentes, a respeito de riscos relacionados a ataques cibernéticos, à discriminação algorítmica, à falta de transparência e à falta de controle do usuário, que podem acarretar eventos adversos.

O terceiro capítulo descreverá e analisará os projetos oficiais de IA em desenvolvimento nos tribunais brasileiros, sua aderência aos princípios éticos e aos atos normativos tratados nos capítulos anteriores, bem como os métodos atualmente existentes para auditar as ferramentas de IA.

Finalmente, a todos os elementos anteriormente pesquisados serão acrescidos conhecimentos buscados nas áreas de Filosofia, Sociologia, Ciência Política, Administração Pública, Bioética, História, Direito, Design, Tecnologia da Informação e Comunicação, Medicina e até Aviação, para dar endereçamento ao problema de pesquisa, permitindo chegar aos resultados de forma propositiva, atingindo os objetivos do trabalho.

2 DIRETRIZES INTERNACIONAIS PARA USO ÉTICO DA IA

Em novembro de 2017 Stephen Hawking fez uma de suas últimas palestras por ocasião do *Web Summit*²⁴, surgindo nos telões do evento anual realizado em Lisboa, que é tido como a maior conferência de tecnologia da Europa²⁵, para falar por videochamada sobre as potencialidades e riscos da IA. Ele próprio fazia uso de um avanço tecnológico bastante elucidativo da utilização de IA benéfica, sem o qual a Humanidade não saberia o que uma mente brilhante como a dele era capaz de produzir ao longo de sua vida: um *software* sintetizador, atualmente de código aberto gratuito²⁶, que lhe permitiu comunicar-se com voz robótica pelos movimentos da face durante mais de três décadas²⁷, quando quase todo o resto de seu corpo já estava paralisado em razão da esclerose lateral amiotrófica com que conviveu desde a juventude até falecer aos 76 anos em março de 2018²⁸.

Nas palavras a um só tempo otimistas e alarmantes, ora traduzidas para o português²⁹, disse ele:

[...] Há muitos desafios e oportunidades à nossa frente neste momento, e acredito que um dos maiores é o impacto da Inteligência Artificial para a Humanidade. [...] O sucesso na criação de uma Inteligência Artificial eficaz poderá ser o maior evento na História da nossa civilização, ou o pior. [...] A menos que aprendamos a evitar e preparar-nos para os potenciais riscos, a Inteligência Artificial pode ser o pior evento da História da nossa civilização. [...] Simplesmente, precisamos estar alerta para os perigos, identificá-los, empregar as melhores práticas possíveis e a melhor gestão e preparar-nos para suas consequências com antecedência. [...].

²⁴ O'CONNOR, Niamh. Leading from the Centre: Web Summit's hall-of-famers. **Web Summit**. 6 mar. 2019. Disponível em: <https://websummit.com/blog/hall-of-famers>. Acesso em: 14 nov. 2021.

²⁵ REUTERS. Europe's biggest tech conference to be held in-person in November. 11 ago. 2021. Disponível em: <https://www.reuters.com/technology/europe-s-biggest-tech-conference-be-held-in-person-november-2021-08-11/>. Acesso em: 14 nov. 2021.

²⁶ ACAT. Assistive Context-Aware Toolkit (ACAT) is a free, open source software to enable people with severe disabilities to communicate. Disponível em: <https://01.org/acat/>. Acesso em: 21 nov. 2021.

²⁷ FAGONE, Jason. Exclusive: The Silicon Valley quest to preserve Stephen...: How a Silicon Valley team helped rebuild his distinctive robotic sound. **San Francisco Chronicle**, 18 mar. 2018. Disponível em: <https://www.sfchronicle.com/bayarea/article/The-Silicon-Valley-quest-to-preserve-Stephen-12759775.php>. Acesso em: 14 nov. 2021.

²⁸ HAWKING. **Biography**. Disponível em: <https://www.hawking.org.uk/biography>. Acesso em: 14 nov. 2021.

²⁹ OBSERVADOR. **Tudo o que Stephen Hawking disse à Web Summit**. Youtube, 14 mar. 2018. Disponível em: <https://www.youtube.com/watch?v=nVgoWh7Wang>. Acesso em: 14 nov. 2021.

Essa perspectiva de formulação de boas práticas de gestão – no sentido de se antecipar, de se preparar para os riscos da IA, buscando identificá-los para tentar evitá-los ou minorar suas consequências – é justamente a baliza que norteia a presente pesquisa, que se ocupa de como documentar, para gerenciar e prevenir, casos de erros de sistemas de IA aplicados pelo Poder Judiciário brasileiro, que sejam potencialmente prejudiciais em alguma medida a jurisdicionados.

2.1 Os Princípios de Asilomar na interseção entre IA e Bioética

Para iniciar a reflexão sobre ética e IA neste capítulo, optou-se pela referência à palestra de Hawking em Lisboa – que também foi documentada em seu último livro³⁰ – por todo o desenvolvimento que é possível seguir a partir da menção que ele fez ao *AAAI 2008-2009 Presidential Panel on Long-Term AI Futures* (Painel Presidencial AAAI sobre Futuros de IA de Longo Prazo)³¹, um evento científico ocorrido em fevereiro de 2009 no balneário californiano de Asilomar, como resultado de estudos avançados sobre IA que se desenvolveram ao longo dos meses antecedentes, organizado por pesquisadores da *AAAI – Association for the Advancement of Artificial Intelligence* (Associação para o Avanço da Inteligência Artificial).

Trata-se essa organização de uma sociedade científica fundada em Palo Alto, Califórnia, em 1979³², à época denominada *American Association for Artificial Intelligence* (Associação Americana de Inteligência Artificial), que congrega expoentes da pesquisa em IA desde então, e cujos primeiros presidentes inclusive estiveram na conferência seminal de *Dartmouth College* no verão de 1956, a exemplo de Marvin Minsky e John McCarthy³³, a quem se atribui a criação do termo “Inteligência Artificial”³⁴.

³⁰ HAWKING, Stephen. **Breves respostas para grandes questões**. Rio de Janeiro: Editora Intrínseca, 2018.

³¹ AAAI. **AAAI Presidential Panel on Long-Term AI Futures: 2008-2009 Study**. Disponível em: <https://www.aaai.org/Organization/presidential-panel.php>. Acesso em: 16 nov. 2021.

³² AAAI. **Association for the Advancement of Artificial Intelligence**. Disponível em: <https://www.aaai.org/>. Acesso em: 16 nov. 2021.

³³ AAAI. **Past AAAI Officials**. Disponível em: <https://aaai.org/Organization/past-officers.php>. Acesso em: 16 nov. 2021.

³⁴ MCCARTHY, J.; MINSKY, M. L.; ROCHESTER, N. **A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence**. 1955. Disponível em: <http://jmc.stanford.edu/articles/dartmouth.html>. Acesso em: 16 nov. 2021.

Para o Painel de 2008-2009 mencionado por Stephen Hawking, cientistas da computação vinculados à AAAI dividiram-se em três subgrupos focais de estudos sobre questões-chave psicossociais, socioeconômicas e ético-legais da IA a longo prazo.

Na ocasião, os painelistas projetaram benefícios, impactos e riscos, bem como discutiram sobre formas de controle e atitudes proativas em termos de regulação, segurança e políticas públicas que a pesquisa científica poderia indicar, visando a uma IA benéfica para a Humanidade – inclusive no tocante a minimizar e gerenciar resultados inesperados, bem como tratar da responsabilidade por comportamentos imprevistos, à medida que a tecnologia for avançando.

Importa ressaltar que, de acordo com o relatório³⁵ de Eric Horvitz e Bart Selman, copresidentes do Painel, a escolha de Asilomar para sediar o evento sobre IA em 2009 não foi nada aleatória: levou-se em conta o simbolismo de ter sido o mesmo local situado em *Pacific Grove*, Califórnia, no qual ocorreu em 1975 a *Asilomar Conference on Recombinant DNA* (Conferência de Asilomar sobre DNA Recombinante), considerada “um marco na história da ética aplicada à pesquisa”, segundo Goldim³⁶, pois “pela primeira vez na história da tecnologia, refletia-se sobre o que poderia dar errado antes que acontecesse o primeiro acidente”, de acordo com Hervé Kempf³⁷.

Dada a dimensão da importância de tal evento para a ciência em geral, a partir do qual muitas lições tiradas reverberam até hoje, cabe aqui tecer breves considerações sobre essa reunião científica de 1975, a fim de melhor ilustrar o motivo de ter sido tomada como parâmetro para a pesquisa em IA no Painel de 2008-2009.

Objetivando tratar dos riscos e dilemas éticos envolvendo a engenharia genética, a Conferência de Asilomar reuniu cerca de 140 participantes e teve como um dos principais organizadores o bioquímico Paul Berg, Professor da Universidade

³⁵ HORVITZ, Eric; SELMAN, Bart. Interim Report from the Panel Chairs: AAAI Presidential Panel on Long-Term AI Futures. In: **Singularity Hypotheses**. Springer, Berlin, Heidelberg, 2012. p. 301-308. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-32560-1_15. Acesso em: 16 nov. 2021.

³⁶ GOLDIM, J.R. **Conferência de Asilomar**. Disponível em: <https://www.ufrgs.br/bioetica/asilomar.htm>. Acesso em: 16 nov. 2021.

³⁷ KEMPF, Hervé. **Asilomar: a ciência aos cientistas**. Transgênicos, terapia genética e células-tronco: questões para a ciência e para a sociedade. Brasília: Núcleo de Estudos Agrários e Desenvolvimento Rural, Instituto Interamericano de Cooperação para a Agricultura, 2004. Disponível em: http://www.ensp.fiocruz.br/portal-ensp/uploads/documentos-pessoais/documento-pessoal_11531.pdf. Acesso em: 16 nov. 2021.

de Stanford que cinco anos depois foi um dos vencedores no Prêmio Nobel de Química de 1980³⁸.

Listada pela revista *Nature* em 2008 entre as “reuniões que mudaram o mundo”, porque segundo Berg “estabeleceu padrões que permitem aos geneticistas levar a pesquisa ao limite, sem colocar a saúde pública em risco”, e assim “marcou o início de uma era excepcional para a ciência e para a discussão pública da política científica”³⁹, a Conferência de 1975 é tida como um significativo exemplo de aplicação prática do princípio da precaução, no campo da biotecnologia.

Isso porque o evento ocorreu após a suspensão global de pesquisas nessa área, pleiteada pelos próprios cientistas que desenvolveram a então nova técnica de DNA Recombinante (que hoje é capaz de avanços tão notáveis quanto o desenvolvimento de vacinas contra a COVID-19 em tempo recorde⁴⁰), enquanto não houvesse ampla discussão da natureza e dimensão dos riscos, e enquanto não fossem estabelecidas diretrizes de biossegurança – as quais vieram a ser instituídas na Declaração⁴¹ elaborada como resultado de tal encontro científico internacional.

Cerca de trinta e três anos depois da Conferência de Asilomar, no já citado artigo publicado na revista *Nature* de 2008, Paul Berg ao recapitular os fatos destacou que na maior parte do tempo as divergências entre os participantes eram quase insuperáveis, até que se chegou ao ponto fulcral que na época permitiu o consenso para o estabelecimento das diretrizes mínimas: “atribuir uma estimativa de risco aos diferentes tipos de experimentos previstos e de aplicar diretrizes de segurança de rigor variável de acordo com o grau de risco”⁴².

Além disso, ele ressaltou que, desde então, incontáveis experimentos de manipulação genética foram realizados no mundo inteiro seguindo os paradigmas estabelecidos em 1975, não tendo sido relatada qualquer intercorrência, numa

³⁸ NOBEL PRIZE. The Nobel Prize in Chemistry 1980. Disponível em: <https://www.nobelprize.org/prizes/chemistry/1980/summary/>. Acesso em: 16 nov. 2021.

³⁹ BERG, Paul. Asilomar 1975: DNA modification secured. *Nature*, v. 455, n. 7211, p. 290-291, 2008. Disponível em: <https://www.nature.com/articles/455290a>. Acesso em: 16 nov. 2021.

⁴⁰ LIMA, Eduardo Jorge da Fonseca; ALMEIDA, Amalia Mapurunga; KFOURI, Renato de Ávila.

Vacinas para COVID-19 - o estado da arte. *Revista Brasileira de Saúde Materno Infantil*, v. 21, p. 13-19, 2021. Disponível em: <https://doi.org/10.1590/1806-9304202100S100002>. Acesso em: 16 nov. 2021.

⁴¹ BERG, Paul et al. Summary statement of the Asilomar conference on recombinant DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America*, v. 72, n. 6, p. 1981, 1975. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC432675/>. Acesso em: 16 nov. 2021.

⁴² Op. cit.

evidência de eficácia das diretrizes. Destacou ainda que muitos avanços projetados e outros tantos sequer imaginados acabaram acontecendo, em campos diversos como na área médica, em termos de diagnósticos, vacinas e medicamentos, bem como na agricultura e na indústria alimentícia; mas ainda assim houve projetos de tamanha complexidade técnica que ainda não foi possível desenvolver.

Adicionalmente ao objetivo principal de estabelecer regras para uma pesquisa científica em DNA Recombinante segura e benéfica para a Humanidade, outro trunfo da Conferência de Asilomar, de acordo com Paul Berg, foi a conquista da confiança da opinião pública pelos cientistas, em relação à nova tecnologia.

Esse fato é atribuído por Berg à composição eclética do grupo de 140 participantes do evento, que além de cientistas geneticistas incluiu não-cientistas como advogados, funcionários públicos e jornalistas. Estes últimos compunham cerca de 15% dos participantes, o que possibilitou divulgar na mídia as discussões e sobretudo a forma como o consenso foi construído, mantendo a opinião pública informada e assim contribuindo para desmistificar receios, ao se passar a mensagem de que houve uma discussão responsável.

Ainda no artigo da revista *Nature*, Paul Berg fez um alerta de uma questão por ele considerada crucial para a análise dos limites éticos da pesquisa científica para uma nova tecnologia: o fato de que, na Conferência de Asilomar, a maior parte dos cientistas era vinculada a instituições públicas, não a organizações privadas com o interesse econômico que lhes é inerente – e que, ao ser levado em consideração, pode comprometer a disposição isenta dos pesquisadores envolvidos para o estabelecimento dos limites adequados.

Tanto é assim, que Berg conclui o artigo da seguinte forma (em tradução livre)⁴³:

Dito isso, há uma lição em Asilomar para toda a ciência: a melhor maneira de responder às preocupações criadas por conhecimentos emergentes ou tecnologias em estágio inicial é fazer com que cientistas de instituições com financiamento público encontrem uma causa comum com o público em geral sobre a melhor maneira para regular - o mais cedo possível. Assim que os cientistas das corporações começarem a dominar o empreendimento de pesquisa, simplesmente será tarde demais.

⁴³ BERG, Paul. Asilomar 1975: DNA modification secured. *Nature*, v. 455, n. 7211, p. 290-291, 2008. Disponível em: <https://www.nature.com/articles/455290a>. Acesso em: 16 nov. 2021.

Traçando um paralelo entre a novidade no campo da biotecnologia em 1975 e o estágio de desenvolvimento da IA no final da primeira década dos anos 2000, foi o tom de cautela e rigor científico de pesquisa atingido pela Conferência de Asilomar, ao tratar de uma nova tecnologia de efeitos futuros cujas dimensões ainda são desconhecidas, que os cientistas da computação da AAAI declaradamente pretendiam imprimir ao Painel de 2008-2009, “em termos de objetivo de alto nível de responsabilidade social para cientistas”⁴⁴, conforme o relatório já mencionado de Horvitz e Selman.

Quanto ao resultado do trabalho dos três grupos focais do Painel Presidencial AAAI sobre Futuros de IA de Longo Prazo, de acordo com o relatório dos copresidentes, o primeiro grupo de cientistas tratou do tema relacionado ao ritmo de evolução da IA e das preocupações com seu controle, tanto do ponto de vista da comunidade científica como da opinião pública.

Enquanto a preocupação em 1975 com o DNA Recombinante era o risco de se criarem agentes patogênicos prejudiciais à saúde pública e ao meio ambiente, que poderiam ficar fora de controle – o que acabou não se confirmado, como já foi dito acima, citando Paul Berg – em se tratando da IA os painelistas da AAAI debateram que há uma certa visão utópica ou mesmo catastrófica por parte de cientistas e da opinião pública em geral, de máquinas e robôs supostamente vindo a suplantar humanos, podendo prejudicá-los e subjugá-los.

Quanto a essas tendências, os painelistas ponderaram que tanto a utopia de uma superinteligência para breve como o receio de uma IA maléfica fora de controle são compreensíveis, porque o imaginário popular vem sendo influenciado ao longo de décadas com obras e filmes de ficção científica, a exemplo dos inspirados nos contos de Isaac Asimov e suas Leis da Robótica⁴⁵ inobservadas por robôs de aparência humana; assim como há cientistas como Ray Kurzweil, com sua tese de que “a singularidade está próxima” e que a IA superará a inteligência humana ainda na primeira metade deste século⁴⁶.

⁴⁴ HORVITZ, Eric; SELMAN, Bart. **Interim Report from the Panel Chairs: AAAI Presidential Panel on Long-Term AI Futures**. Ago. 2009. Disponível em: <https://www.aaai.org/Organization/Panel/panel-note.pdf>. Acesso em: 16 nov. 2021.

⁴⁵ ASIMOV, Isaac. **Eu, robô**. Tradução Aline Storto Pereira. 1. ed. São Paulo: Aleph, 2014. Título original: I, robot.

⁴⁶ KURZWEIL, Ray. **A singularidade está próxima: quando os humanos transcendem a biologia**. Tradução Ana Goldberger. Itaú Cultural, 2019. Título original: The singularity is near.

O primeiro grupo focal de painelistas da AAAI, no entanto, imbuído de um posicionamento cético quanto a visões radicais, dado o fato concreto de que não havia a chamada singularidade tecnológica naquele momento (como ainda não há) e não é possível saber se e quando ocorrerá essa superinteligência artificial, preferiu adotar uma perspectiva pragmática, revisando a literatura, para concluir que dificilmente será vista no horizonte próximo uma profusão de IA avançada com perda de controle de seus sistemas de forma generalizada ou em larga escala.

De todo modo, conforme o relatório, o grupo focal considerou que pesquisas contínuas são importantes e necessárias para compreender melhor o funcionamento dos sistemas inteligentes do momento, fazer projeções realistas de sua evolução, segmentá-los em razão de suas funcionalidades e graus de risco, estudar os resultados esperados e inesperados, bem como o que fazer concretamente quando surgirem intercorrências.

Além disso, tal qual outrora pontuado por Paul Berg para a pesquisa em manipulação genética, o primeiro grupo focal do Painel de 2008-2009 mencionou a importância da divulgação adequada das pesquisas em IA, não só perante a comunidade científica como também perante a opinião pública, quanto às conclusões dos painelistas sobre a remota plausibilidade de acontecimentos catastróficos, e dando ênfase às potencialidades de uso benéfico da IA, bem como aos esforços dos cientistas para encontrar soluções aos desafios daquele tempo presente e do futuro próximo, que já se afiguravam com o advento da nova tecnologia.

O segundo grupo focal de cientistas da computação da AAAI, participantes do Painel de 2008-2009 em Asilomar, na mesma linha pragmática do primeiro grupo, ocupou-se justamente de estudar os desafios da IA no curto prazo, colocando no topo da lista a necessidade de aumentar a proteção à privacidade das pessoas, perante a dificuldade de compatibilizar essa proteção com a possibilidade de que empresas sigam personalizando serviços de maneira cada vez mais sofisticada e invasiva a partir dos dados e rastros digitais que as pessoas fornecem.

Outro desafio de curto prazo identificado pelo segundo grupo focal foi aperfeiçoar a interação entre as pessoas e os sistemas inteligentes, para que atuem em sinergia “na solução conjunta de tarefas e no apoio a transições fluidas entre o

raciocínio automatizado e o controle humano”⁴⁷, conforme consta no relatório dos copresidentes do Painel.

Mais um tópico a merecer pesquisas aprofundadas, apontado pelo segundo grupo, envolve aumentar a transparência de sistemas inteligentes, bem como a explicação de como se desenvolvem as operações até os resultados.

O derradeiro desafio de curto prazo, identificado pelo segundo grupo e mencionado no relatório, também envolve a questão da privacidade e o risco de uso maléfico da IA a partir de dados das pessoas – acessados e manipulados de forma mal intencionada por pessoas, empresas e governos –, fato que recomenda esforços de pesquisa em segurança e criptografia.

Já o terceiro e último grupo focal do Painel 2008-2009 tratou de questões éticas e legais, a exemplo do uso da IA para decisões de alto risco como diagnósticos médicos ou definição de alvos por armas inteligentes, quanto à responsabilidade em caso de resultados inesperados, começando com reflexões sobre a suficiência ou não das regras legais já existentes para dirimir esse tipo de situação. Outra questão ética apontada diz respeito às implicações psicológicas da interação entre pessoas e máquinas quando estas emulam emoções. Perante tais dilemas, os painelistas apontaram a necessidade de pesquisas adicionais interdisciplinares nos campos da Ética, do Direito e da Psicologia.

Ao concluir o relatório com o resultado das discussões e a definição dos principais pontos de atenção pelos três grupos focais de especialistas, os copresidentes do Painel Presidencial AAAI sobre Futuros de IA de Longo Prazo destacaram que, em sua percepção, a pesquisa contínua focada na solução de problemas reais e nos impactos sociais da IA é o caminho indicado para que cientistas e pesquisadores cumpram sua função de proporcionar à sociedade o máximo proveito benéfico de que a IA é capaz em inúmeras searas.

Em janeiro de 2017, quase uma década depois do Painel Presidencial AAAI sobre Futuros de IA de Longo Prazo, mais uma vez Asilomar foi palco de outro importante encontro científico, confirmado sua vocação histórica para eventos desse jaez, tratando-se da BAI 2017⁴⁸ – *Asilomar Conference on Beneficial AI* (Conferência

⁴⁷HORVITZ, Eric; SELMAN, Bart. **Interim Report from the Panel Chairs: AAAI Presidential Panel on Long-Term AI Futures.** Ago. 2009. Disponível em: <https://www.aaai.org/Organization/Panel/panel-note.pdf>. Acesso em: 16 nov. 2021.

⁴⁸FLI. **Beneficial AI 2017.** Future of Life Institute. Disponível em: <https://futureoflife.org/bai-2017/>. Acesso em: 20 dez. 2021.

Asilomar sobre IA Benéfica), da qual resultou o desenvolvimento dos *Asilomar AI Principles* (Princípios de IA de Asilomar), após reunir mais de 100 participantes entre destacados cientistas da computação, especialistas, membros da academia de diversas áreas e empresários interessados em financiar pesquisas para uso benéfico da IA.

A BAI 2017 foi organizada pelo FLI – *Future of Life Institute* (Instituto Futuro da Vida), um instituto de pesquisa sem fins lucrativos sediado nos Estados Unidos, na região de Boston, que “trabalha na redução do risco catastrófico e existencial global de tecnologias poderosas”⁴⁹, com foco nas áreas de IA, armas nucleares, biotecnologia e mudanças climáticas.

Presidido e cofundado por Max Tegmark, professor do MIT – Instituto de Tecnologia de Massachusetts, o FLI foi criado em 2014 e tem entre seus cofundadores⁵⁰ o patrono, investidor filantropo e engenheiro estoniano Jaan Tallinn, fundador do Skype, que também cofundou o CSER⁵¹ – Centro para o Estudo de Risco Existencial, ligado à Universidade de Cambridge, no Reino Unido; além de apoiar financeiramente pesquisas relacionadas à temática das poderosas tecnologias emergentes em instituições como o FHI⁵² – Instituto Futuro da Humanidade (da Universidade de Oxford), o GCRI⁵³ – Instituto Global de Risco Catastrófico e o MIRI⁵⁴ – Instituto de Pesquisa de Inteligência de Máquina (de Berkeley, Califórnia).

O FLI conta ainda com um Conselho Consultivo Científico multidisciplinar de notáveis, ligados a algumas das instituições de maior prestígio no mundo, como o MIT e as Universidades de Harvard, Oxford, Cambridge e UC Berkeley, incluindo dois vencedores do Prêmio Nobel de Física (Frank Wilczek, de 2004; e Saul Perlmutter, de 2011), além do empresário e investidor filantropo Elon Musk (fundador da Tesla Motors e SpaceX). Até falecer em 2018, também fazia parte do Conselho Consultivo Científico do FLI (do qual foi membro fundador⁵⁵) o físico britânico Stephen Hawking. Foi no final

⁴⁹FLI. **Future of Life Institute**. Disponível em: <https://futureoflife.org/>. Acesso em: 20 dez. 2021.

⁵⁰FLI. **About us**. Disponível em: <https://futureoflife.org/team/>. Acesso em: 20 dez. 2021.

⁵¹CSER. **Centre for the Study of Existential Risk**. Disponível em: <https://www.cser.ac.uk/>. Acesso em: 20 dez. 2021.

⁵²FHI. **Future of Humanity Institute**. Disponível em: <https://www.fhi.ox.ac.uk/>. Acesso em: 20 dez. 2021.

⁵³GRCI. **Global Catastrophic Risk Institute**. Disponível em: <https://gcrinstitute.org/>. Acesso em: 20 dez. 2021.

⁵⁴MIRI. **Machine Intelligence Research Institute**. Disponível em: <https://intelligence.org/>. Acesso em: 20 dez. 2021.

⁵⁵TEGMARK, Max. **Stephen Hawking in Memoriam**. 14 mar. 2018. Disponível em: <https://futureoflife.org/2018/03/14/stephen-hawking-memoriam/>. Acesso em: 20 dez. 2021.

do ano em que os Princípios de IA de Asilomar foram desenvolvidos que ele proferiu a palestra mencionada na abertura deste Capítulo.

Toda essa expertise reunida num instituto, com investidores de peso para o financiamento de pesquisas em IA, foram agregados ao conhecimento de uma centena de participantes na conferência BAI 2017, tratando-se de um grupo heterogêneo formado por especialistas em IA e expoentes de diversas áreas como Ética, Filosofia, Economia e Direito, para deliberar sobre governança de IA mediante procedimentos⁵⁶ organizados pelo Instituto, que se iniciaram com a reunião do máximo possível de relatórios e publicações até então existentes sobre o que deveria ser feito para gerenciamento adequado da IA nas décadas vindouras, do ponto de vista acadêmico, da sociedade, de governos, da indústria.

Nessa fase prévia ao evento, a partir das dezenas de relatórios reunidos e analisados, foram compiladas opiniões que pareciam em alguma medida convergentes, mesmo que ainda passíveis de discussão e interpretações diversas, sendo então submetidas à minuciosa análise dos participantes da conferência. Com o retorno dado por estes, o material foi depurado e aprimorado, com o objetivo de levar à conferência uma lista de pontos de vista que pudessem ser traduzidos em princípios gerais.

Por ocasião da conferência propriamente dita, que durou 5 dias no total, o material compilado ainda passou por uma rodada de discussões em grupos, saindo dessa etapa uma nova versão, que foi finalmente levada ao escrutínio da plenária, decidindo-se na BAI 2017 manter apenas os preceitos que alcançassem no mínimo 90% de concordância dos participantes.

Assim transcorreu o processo pelo qual se chegou aos 23 Princípios de IA de Asilomar, que se subdividem em 5 princípios sobre questões de pesquisa; 13 princípios sobre ética e valores; e 5 sobre problemas de longo prazo. Dada sua importância, vão a seguir integralmente reproduzidos⁵⁷, com tradução livre para o português:

⁵⁶ FLI. A principled AI discussion in Asilomar. Disponível em: <https://futureoflife.org/2017/01/17/principled-ai-discussion-asilomar/>. Acesso em: 20 dez. 2021.

⁵⁷ FLI. **Asilomar AI Principles**. Disponível em: <https://futureoflife.org/2017/08/11/ai-principles/>. Acesso em: 20 dez. 2021.

Princípios de IA de Asilomar

A inteligência artificial já forneceu ferramentas benéficas que são usadas diariamente por pessoas de todo o mundo. Seu desenvolvimento contínuo, guiado pelos seguintes princípios, oferecerá oportunidades incríveis para ajudar e capacitar as pessoas nas décadas e séculos à frente.

Questões sobre Pesquisa

- 1) **Objetivo da pesquisa:** O objetivo da pesquisa em IA não deve ser criar inteligência sem objetivo definido, mas criar inteligência benéfica.
- 2) **Financiamento de pesquisa:** Os investimentos em IA devem ser acompanhados de financiamento para pesquisas que garantam seu uso benéfico, incluindo questões espinhosas em ciência da computação, economia, direito, ética e estudos sociais, tais como:
 - Como podemos tornar os sistemas de IA do futuro altamente robustos, de modo que façam o que queremos sem avariar ou ser hackeados?
 - Como podemos aumentar nossa prosperidade por meio da automação e, ao mesmo tempo, manter os recursos e o propósito das pessoas?
 - Como podemos atualizar nossos sistemas jurídicos para ser mais justos e eficientes, para acompanhar o ritmo da IA e para gerenciar os riscos associados à IA?
 - Com qual conjunto de valores a IA deve estar alinhada e que status legal e ético deve ter?
- 3) **Conexão entre Ciência e Política:** Deve haver um intercâmbio construtivo e saudável entre pesquisadores de IA e formuladores de políticas.
- 4) **Cultura de pesquisa:** Uma cultura de cooperação, confiança e transparência deve ser promovida entre pesquisadores e desenvolvedores de IA.
- 5) **Prevenção de competição:** As equipes que desenvolvem sistemas de IA devem cooperar ativamente para evitar atalhos que possa comprometer padrões de segurança.

Ética e Valores

- 6) **Segurança:** Os sistemas de IA devem ser seguros durante toda a sua vida operacional e verificável quando aplicável e viável.
- 7) **Transparência de falha:** Se um sistema de IA causar danos, deve ser possível determinar o porquê.
- 8) **Transparência Judicial:** Qualquer envolvimento de um sistema autônomo na tomada de decisões judiciais deve fornecer uma explicação satisfatória auditável por uma autoridade humana competente.
- 9) **Responsabilidade:** Desenvolvedores de sistemas avançados de IA são partes envolvidas nas implicações morais de seu uso, uso indevido e ações, cabendo-lhes a responsabilidade e a oportunidade de moldar essas implicações.

- 10) **Alinhamento de valores:** Sistemas de IA altamente autônomos devem ser projetados de forma que seus objetivos e comportamentos possam ter a garantia de estar alinhados com os valores humanos durante toda a sua operação.
- 11) **Valores humanos:** Os sistemas de IA devem ser projetados e operados de forma a serem compatíveis com os ideais de dignidade humana, direitos, liberdades e diversidade cultural.
- 12) **Privacidade pessoal:** As pessoas devem ter o direito de acessar, gerenciar e controlar os dados que geram, dado o poder dos sistemas de IA de analisar e utilizar esses dados.
- 13) **Liberdade e privacidade:** A aplicação de IA aos dados pessoais não deve restringir injustificadamente a liberdade real ou percebida das pessoas.
- 14) **Benefício compartilhado:** As tecnologias de IA devem beneficiar e capacitar o maior número possível de pessoas.
- 15) **Prosperidade compartilhada:** A prosperidade econômica criada pela IA deve ser amplamente compartilhada, para beneficiar toda a humanidade.
- 16) **Controle Humano:** Os humanos devem escolher como e quais decisões se delegar aos sistemas de IA, para cumprir os objetivos escolhidos pelos humanos.
- 17) **Não-subversão:** O poder conferido pelo controle de sistemas de IA altamente avançados deve respeitar e melhorar, ao invés de subverter, os processos sociais e cívicos dos quais a saúde da sociedade depende.
- 18) **Corrida armamentista de IA:** Uma corrida armamentista em armas autônomas letais deve ser evitada.

Problemas de longo prazo

- 19) **Precaução de capacidade:** Não havendo consenso, devemos evitar suposições fortes sobre os limites superiores das capacidades futuras de IA.
- 20) **Importância:** IA avançada pode representar uma mudança profunda na história da vida na Terra e deve ser planejada e administrada com cuidado e recursos adequados.
- 21) **Riscos:** Riscos apresentados por sistemas de IA, especialmente riscos catastróficos ou existenciais, devem estar sujeitos a planejamento e esforços de mitigação proporcionais ao seu impacto esperado.
- 22) **Autoaprimoramento recursivo:** os sistemas de IA projetados para se autoaperfeiçoar recursivamente ou se autorreplicar de maneira que possam levar a um rápido aumento da qualidade ou quantidade devem estar sujeitos a medidas rígidas de segurança e controle.
- 23) **Bem comum:** a superinteligência só deve ser desenvolvida a serviço de ideais éticos amplamente compartilhados e para o benefício de toda a humanidade, e não de um estado ou organização.⁵⁸

⁵⁸ FLI. **Asilomar AI Principles.** Disponível em: <https://futureoflife.org/2017/08/11/ai-principles/>. Acesso em: 20 dez. 2021

Passados cinco anos desde que tais princípios foram divulgados em janeiro de 2017, é possível depreender que tiveram ampla aceitação junto à comunidade científica mundial, considerando que em janeiro de 2022 a lista⁵⁹ de signatários já alcançava 5.720 pessoas, entre 1.797 pesquisadores de IA e robótica e 3.923 outros endossantes de várias partes do mundo.

Tornando efetivos os princípios relativos às questões de pesquisa, constata-se que o FLI incrementou o programa de concessão de bolsas para pesquisa de segurança em IA (que iniciara em 2015 logo após sua fundação), com novas concessões em 2018, resultando em dezenas de publicações científicas de ponta para o uso seguro da tecnologia, deixando clara a ideia de precaução como baliza do Instituto, ao destacar as palavras de um de seus principais patronos, Elon Musk: “[...] É melhor tentar evitar que uma circunstância negativa ocorra do que esperar que ela ocorra e então ser reativo”⁶⁰.

Em 2021, evidenciando novamente a importância da precaução, ao afirmar que “a COVID-19 mostrou que nossa civilização é frágil e pode lidar melhor com os riscos ao planejar com antecedência”⁶¹, o FLI lançou programas de concessão de bolsas de doutorado e pós-doutorado para a pesquisa em IA nas áreas de segurança existencial, políticas públicas e ciência comportamental, dessa vez com o dobro do patrocínio até então distribuído, mediante subsídios plurianuais da ordem de 25 milhões de dólares patrocinados por Vitalik Buterin, o inventor da criptomoeda Ethereum⁶².

A propósito, cabe ressaltar a miríade de temas a serem objeto de pesquisa para uma IA benéfica, sendo possível ter a dimensão do desafio ao visualizar as mais de duas centenas de tópicos e suas interconexões no gráfico interativo elaborado pelo FLI (Figura 1), que o intitulou “FLI Cenário de Pesquisa de Alinhamento de Valor” e assim o definiu⁶³ (em tradução livre):

⁵⁹ FLI. **Principles signatories**. Disponível em: <https://futureoflife.org/2017/01/11/principles-signatories/>. Acesso em: 21 dez. 2021.

⁶⁰ FLI. **AI Safety Research**. Disponível em: <https://futureoflife.org/ai-safety-research/>. Acesso em: 21 dez. 2021.

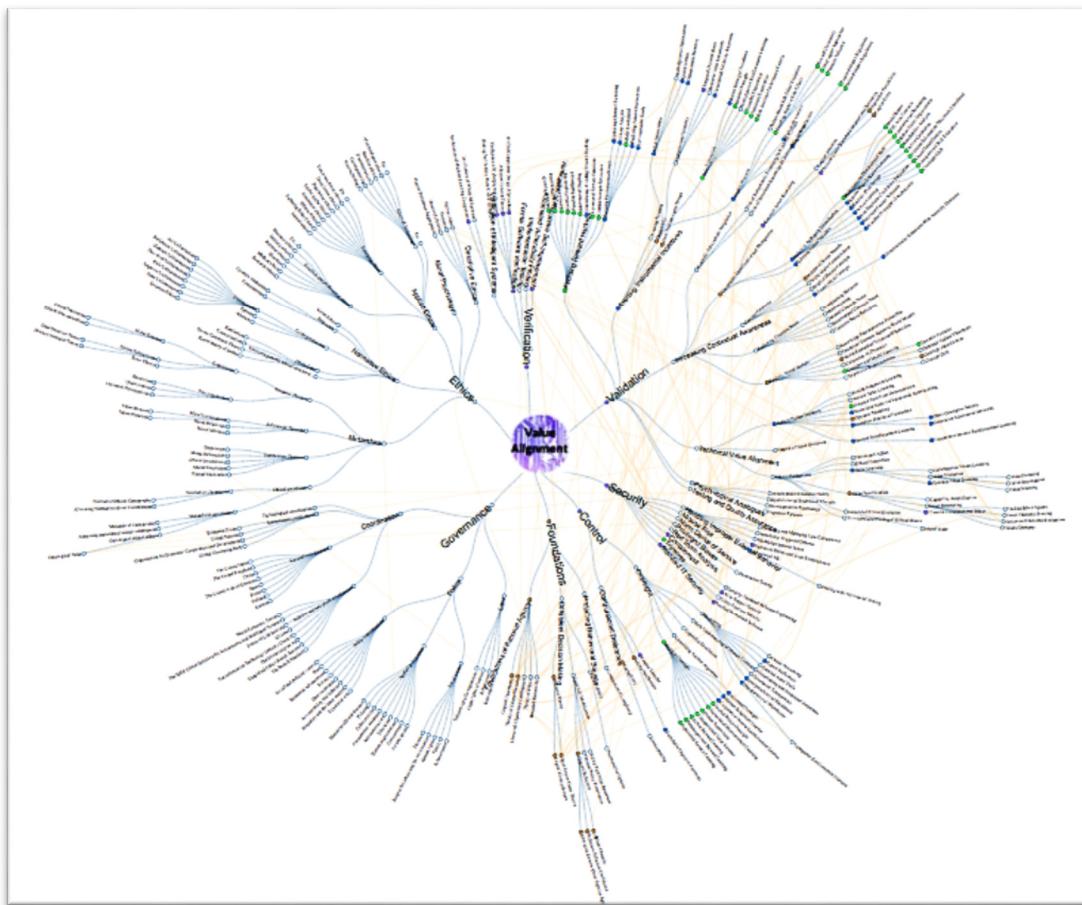
⁶¹ FLI. **Grant Programs**. Disponível em: <https://futureoflife.org/grant-programs/>. Acesso em: 21 dez. 2021.

⁶² PYBUS, Robert. The uncanny mind that build Ethereum. **Wired**. 13 jun. 2016. Disponível em: <https://www.wired.com/2016/06/the-uncanny-mind-that-built-ethereum/#.u88ip6pb1>. Acesso em: 21 dez. 2021.

⁶³ FLI. **FLI Value Alignment Research Landscape**. Disponível em: <https://futureoflife.org/valuealignmentmap/>. Acesso em: 21 dez. 2021.

O projeto de criar IA alinhada a valores é talvez uma das coisas mais importantes que faremos. No entanto, há questões abertas e muitas vezes negligenciadas sobre o que significa realmente 'IA benéfica'. O alinhamento de valor é o projeto de um dia criar IA benéfica e foi expandido para fora de seu contexto técnico usual para refletir e modelar sua natureza verdadeiramente interdisciplinar. Para que a IA alinhada a valores se torne uma realidade, precisamos não apenas resolver a Inteligência, mas também os fins aos quais a Inteligência se destina e o contexto social/político, as regras e as políticas em que e por meios dos quais tudo isso acontece. Este cenário sintetiza uma variedade de agendas de pesquisa de segurança de IA junto com outros artigos sobre IA, aprendizado de máquina, ética, governança e segurança de IA, robustez e pesquisa de beneficência. Ele descreve quais tópicos de pesquisa técnica podem nos ajudar a criar IA benéfica e descreve como esses muitos tópicos se relacionam.

Figura 1 - Mapa de Pesquisa de Alinhamento de Valor



Fonte: <https://futureoflife.org/valuealignmentmap/>

Para além do fomento à pesquisa científica e sobretudo numa atuação efetiva quanto ao princípio nº 3, percebe-se que o FLI assumiu uma posição de liderança na conexão entre Ciência e Política, atuando em consultas públicas sobre regulação, sendo chamado a opinar e sendo convidado a participar de organizações

multilaterais, contribuindo ativamente para o intercâmbio entre pesquisadores e formuladores de políticas públicas nos Estados Unidos e na União Europeia⁶⁴.

Nessa linha de atuação, o Instituto figura entre as entidades da sociedade civil nominadas como campeãs de colaboração sobre IA junto à ONU – Organização das Nações Unidas, segundo o Roteiro para Cooperação Digital da ONU, cujo relatório do Secretário-Geral (quanto ao trabalho sobre o tema iniciado em 2018 e desenvolvido ao longo dos anos seguintes) foi divulgado em junho de 2020⁶⁵.

Tal relatório, na parte que tratou de IA⁶⁶, destacou o risco das armas autônomas letais – a recomendar um esforço mundial pelo compromisso de proibi-las ou restringi-las; bem como destacou o risco da manipulação da opinião pública por *deepfakes*. Também chamou a atenção da comunidade internacional para os avanços nas tecnologias de reconhecimento facial e identificação digital, a recomendar esforços para que o uso da IA não atente contra os direitos humanos ao acentuar desigualdades e causar discriminação.

Além dessas preocupações principais, o relatório apontou três desafios⁶⁷ que ainda persistem, apesar do aumento global das iniciativas, pesquisas e discussões sobre IA nos últimos anos:

(i) A ausência ou sub-representação de países em desenvolvimento nos principais fóruns de discussão sobre IA, que deixam de aproveitar todo o potencial da tecnologia para seu progresso social e econômico;

(ii) A ausência de uma coordenação geral ou de uma plataforma que sirva para agregar todas as iniciativas, visto que já foram mapeados mais de 160 conjuntos de princípios de ética e regras de governança para o uso de IA ao redor do mundo, porém estão dispersos, o que dificulta o compartilhamento e o uso colaborativo;

(iii) A falta de sistematização para que os setores públicos dos países possam se beneficiar do conhecimento e experiência em IA já existentes, os quais seriam úteis para o incremento da supervisão e governança, para mitigar riscos e para agregar conhecimento sobre como a IA pode ser utilizada no sentido de ajudar a atingir os Objetivos de Desenvolvimento Sustentável.

⁶⁴ FLI. **FLI Policy Work**. Disponível em: <https://futureoflife.org/policy-work/>. Acesso em: 21 dez. 2021.

⁶⁵ ONU. Report of the Secretary General. Roadmap for Digital Cooperation. Jun. 2020. Disponível em: https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf#page=32. Acesso em: 21 dez. 2021.

⁶⁶ Ibidem. p. 17.

⁶⁷ Ibidem. p. 18.

Entre esses mais de 160 conjuntos de princípios de IA mapeados pela ONU mundialmente até junho de 2020, elaborados por organizações públicas e privadas diversas, cabe destacar que o relatório do Secretário-Geral da ONU, na nota explicativa número 42, faz menção expressa a 4 iniciativas específicas que estabelecem diretrizes para o uso benéfico da IA, o que indica uma maior relevância destas no cenário internacional quanto ao tema, razão pela qual também serão objeto de estudo neste trabalho. Segue a nota traduzida⁶⁸:

42. Atualmente, há uma série de iniciativas globais que estabelecem normas para o desenvolvimento e uso de inteligência artificial. Alguns, como a Parceria Global em Inteligência Artificial, são lideradas por uma coalizão de Estados-Membros que pensam da mesma maneira; outros, como o Observatório de Políticas sobre Inteligência Artificial da Organização para Cooperação e Desenvolvimento Econômico, buscam apoiar os esforços dos Estados-Membros; iniciativas técnicas como a Iniciativa Global sobre Ética dos Sistemas Autônomos e Inteligentes do Instituto de Engenheiros Elétricos e Eletrônicos estabelecem normas e padrões técnicos, enquanto a Parceria em Inteligência Artificial atua como uma plataforma de parceria entre seus fundadores corporativos e a sociedade civil.

Levando em conta o Princípio de Asilomar nº 3 (Conexão entre Ciência e Política), assumem especial relevo para esta pesquisa os princípios que foram adotados em 2019 pela OCDE – Organização para a Cooperação e Desenvolvimento Econômico, que serão examinados no tópico a seguir e que estão na origem de duas das quatro iniciativas acima mencionadas.

2.2 Os Princípios de IA da OCDE

Foi em novembro de 2016 que a OCDE iniciou formalmente a discussão sobre o tema, a partir de um fórum realizado em Paris para fazer uma previsão das implicações sociais e econômicas da IA, publicando em fevereiro de 2017 um resumo⁶⁹ dos principais pontos abordados nas palestras e debates, cujo foco foi a chamada IA aplicada, ou seja, aquela projetada para realizar tarefas específicas, já

⁶⁸ ONU. Report of the Secretary General. Roadmap for Digital Cooperation. Jun. 2020. Disponível em: https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf#page=39. Acesso em: 21 dez. 2021.

⁶⁹ OCDE. Summary of the CDEP Technology Foresight Forum Economic And Social Implications Of Artificial Intelligence. Paris, 10 fev. 2017. Disponível em: [https://www.oecd.org/digital/ieconomy/DSTI-CDEP\(2016\)17-ENG.pdf](https://www.oecd.org/digital/ieconomy/DSTI-CDEP(2016)17-ENG.pdf) . Acesso em: 02 jan. 2022.

utilizada em larga escala e em franca expansão nos mais diversos setores da economia e da sociedade.

Essa abordagem foi deliberadamente adotada por ser considerada mais pragmática e realista, porém sem deixar de lado, mas em segundo plano, ações necessárias que, por precaução, formuladores de políticas públicas deveriam adotar para uma futura e hipotética IA geral (AGI, na sigla em inglês, aquela que seria capaz de igualar ou até superar a totalidade das habilidades cognitivas humanas), que ninguém sabe exatamente se e quando vai se tornar realidade.

O encaminhamento dado pelos palestrantes do fórum de 2016 foi a elaboração de diretrizes gerais pela OCDE, o que foi endossado após outra conferência de discussão mais ampla (IA: Máquinas Inteligentes, Políticas Inteligentes), também realizada na sede da entidade em Paris no mês de outubro de 2017, cujo relatório⁷⁰ foi publicado em agosto de 2018.

Nesse relatório consta que, entre os diversos parâmetros que seriam adotados como referência, foram incluídos os Princípios de IA de Asilomar, os quais, àquela altura, já haviam sido divulgados pelo FLI. Consta ainda ter ficado definido que a OCDE coordenaria o trabalho de um grupo de especialistas que se debruçaria sobre as questões práticas envolvendo IA aplicada nas áreas de economia digital, ciência/pesquisa, modelos de negócio, proteção ao consumidor, educação e trabalho/emprego.

O desafio era criar diretrizes para a formulação de regras e políticas públicas de IA pelos governos, que fossem suficientemente gerais, que pudessem resistir à passagem do tempo numa área que é altamente dinâmica e que, ao mesmo tempo, fossem flexíveis o bastante para não tolher o avanço tecnológico da IA.

Na conferência de outubro de 2017 a OCDE também foi instada a oferecer suporte no tocante às discussões sobre IA travadas no âmbito do G7 e do G20, respectivamente as reuniões dos 7 e dos 20 países de maior economia do mundo. Ressalte-se que “juntos, os membros do G20 representam mais de 80% do PIB mundial, 75% do comércio internacional e 60% da população mundial”⁷¹.

⁷⁰ OCDE. AI: Intelligent Machines, Smart Policies. Conference Summary. **OECD Digital Economy Papers n. 270**. Paris: OECD Publishing, 2018. Disponível em: <https://doi.org/10.1787/f1a650d9-en>. Acesso em: 02 jan. 2022.

⁷¹ G20. **About the G20**. Disponível em: <https://g20.org/about-the-g20/>. Acesso em: 2 jan. 2022.

O grupo⁷² multidisciplinar de mais de 50 especialistas de diversas nacionalidades e instituições (oriundos de setores variados como academia, área técnica, governo, indústria, sociedade civil e sindicatos) foi formado em maio de 2018 e se reuniu entre setembro de 2018 e fevereiro de 2019, realizando encontros na França, nos Estados Unidos e nos Emirados Árabes Unidos.

O resultado⁷³ do trabalho do grupo de especialistas embasou um documento formal proposto pelo Comitê de Política de Economia Digital da OCDE e finalmente concebido na forma de Recomendação, que foi adotada pelo Conselho da entidade em 21 de maio de 2019, tornando-se “o primeiro padrão intergovernamental sobre IA”⁷⁴.

A estrutura da Recomendação sobre IA da OCDE contém (i) um conjunto de cinco conceitos essenciais, objetivando uniformizar o entendimento de expressões a respeito de IA mencionadas no texto; (ii) a Seção 1, descrevendo cinco “Princípios para administração responsável de IA confiável”⁷⁵; (iii) a Seção 2, dispondo sobre cinco diretrizes quanto a “Políticas nacionais e cooperação internacional para IA confiável”⁷⁶; além de (iv) instruções de coordenação e monitoramento periódico quanto à aplicação da Recomendação, endereçadas ao Comitê de Política de Economia Digital da OCDE. As três primeiras partes principais estão reproduzidas no quadro sinóptico abaixo (Figura 2):

⁷² OECDAI. List of participants in the OECD Expert Group on AI (AIGO). Disponível em: <https://oecd.ai/en/list-of-participants-oecd-expert-group-on-ai>. Acesso em: 2 jan. 2022.

⁷³ OCDE. Scoping the OECD AI sources: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO). **OECD Digital Economy Papers** n. 291. Paris: OECD Publishing, 2019. Disponível em: https://www.oecd-ilibrary.org/science-and-technology/scoping-the-oecd-ai-principles_d62f618a-en. Acesso em: 2 jan. 2022.

⁷⁴ OCDE. Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. **OECD Legal Instruments**. 21 maio 2019. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 02 jan. 2022.

⁷⁵ Ibidem.

⁷⁶ Ibidem.

Figura 2 - Quadro sinóptico da Recomendação sobre IA da OCDE

Termos-chave	Princípios para administração responsável de IA confiável	Políticas nacionais e cooperação internacional para IA confiável
<ul style="list-style-type: none"> • Sistema de IA • Ciclo de vida do sistema de IA • Conhecimento de IA • Atores de IA • Partes interessadas 	<ul style="list-style-type: none"> • Crescimento inclusivo, desenvolvimento sustentável e bem-estar • Valores centrados no ser humano e justiça • Transparência e explicabilidade • Robustez, segurança e proteção • Responsabilidade 	<ul style="list-style-type: none"> • Investir em pesquisa e desenvolvimento de IA • Promover um ecossistema digital para IA • Moldar um ambiente de política favorável à IA • Capacitação humana e preparação para a transformação do mercado de trabalho • Cooperação internacional para IA confiável

Fonte: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

Cumpre destacar que o Brasil figura na lista de países aderentes⁷⁷ à Recomendação sobre IA da OCDE, na condição de não-membro.

A propósito, o Brasil é considerado “Parceiro-Chave” da OCDE desde 2007⁷⁸, vindo a firmar em junho de 2015 o Acordo de Cooperação Brasil-OCDE, que entre diversas formas de parceria inclui a “participação do Brasil na revisão e no desenvolvimento de instrumentos e normas de políticas da OCDE” (Seção 3, 3.5)⁷⁹.

Após aprovação do Congresso Nacional, o Acordo foi promulgado pelo Decreto nº 10.109, de 7 de novembro de 2019, passando então a integrar o ordenamento jurídico brasileiro, nos termos do art. 49, I, e do art. 84, VIII, ambos da Constituição Federal⁸⁰.

⁷⁷ OCDE. Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. **OECD Legal Instruments**. 21 maio 2019. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#adherents>. Acesso em: 2 jan. 2022.

⁷⁸ OCDE. A OCDE e o Brasil: uma relação mutuamente benéfica. Disponível em: <https://www.oecd.org/latin-america/paises/brasil-portugues/>. Acesso em: 2 jan. 2022.

⁷⁹ BRASIL. Decreto nº 10.109, de 7 de novembro de 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/D10109.htm#:~:text=DECRETO%20N%C2%BA%2010.109%2C%20DE%207,3%20de%20junho%20de%202015. Acesso em: 2 jan. 2022.

⁸⁰ BRASIL. Constituição da República Federativa do Brasil de 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 2 jan. 2022.

O Brasil ainda se encontra em meio ao seu processo de acessão à OCDE, que é rigoroso⁸¹, tendo sido instituído o Conselho para a Preparação e o Acompanhamento do Processo de Acessão do Brasil à OCDE pelo Decreto nº 9.920, de 18 de julho de 2019⁸².

Fato é que, independentemente da acessão, os tratados internacionais já firmados pelo Brasil com a OCDE permitiram-lhe aderir à Recomendação sobre IA da OCDE na mesma data em que foi adotada, 21 de maio de 2019. Além disso, como membro do G20, o Brasil endossou no Japão a Declaração de Osaka dos Líderes do G20⁸³, em 29 de junho de 2019, que adotou os Princípios sobre IA do G20⁸⁴, incorporando a redação das Seções 1 e 2 da Recomendação sobre IA da OCDE.

Pelo enfoque prático, bem como pela finalidade de embasamento para formulação de políticas públicas alinhadas às pesquisas científicas de ponta e às melhores práticas internacionais na área de IA, trata-se a Recomendação de um instrumento de alta relevância para o escopo deste trabalho.

Afinal, esta pesquisa está sendo realizada no âmbito de um mestrado profissional de instituição pública que é “o órgão oficial de treinamento de juízes de direito e juízes federais brasileiros”⁸⁵, buscando solução para um problema prático de IA aplicada por um dos três Poderes⁸⁶ da República Federativa do Brasil, o Poder Judiciário: encontrar um método adequado para reportar e catalogar eventos resultantes de erros de sistemas de IA em operação nos tribunais, que atentem ou possam atentar contra direitos humanos dos jurisdicionados.

Relembre-se que o objetivo é apresentar proposta para conferir efetividade a uma norma em branco de gerenciamento de risco editada pelo órgão de controle – o Conselho Nacional de Justiça – que no art. 27 da Resolução nº 332/2020 determinou

⁸¹ OCDE. **Where: Global reach.** Disponível em: <https://www.oecd.org/about/members-and-partners/>. Acesso em: 2 jan. 2022.

⁸² BRASIL. **Decreto nº 9.920, de 18 de julho de 2019.** Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/D9920.htm. Acesso em: 2 jan. 2022.

⁸³ BRASIL. **Nota à imprensa nº 170/2019.** Ministério das Relações Exteriores. Disponível em: https://www.gov.br/mre/pt-br/canais_atendimento/imprensa/notas-a-imprensa/2019/declaracao-de-osaka-dos-lideres-do-g20. Acesso em: 2 jan. 2022.

⁸⁴ G20 Insights. **G20 AI Principles.** Disponível em: <https://www.g20-insights.org/wp-content/uploads/2019/07/G20-Japan-AI-Principles.pdf>. Acesso em: 2 jan. 2022.

⁸⁵ ENFAM. **Sobre a Enfam.** Disponível em: <https://www.enfam.jus.br/institucional/sobre-a-escola/>. Acesso em: 2 jan. 2022.

⁸⁶ BRASIL. **Constituição da República Federativa do Brasil de 1988.** Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 2 jan. 2022.

que sejam reportados pelos tribunais todos os eventos adversos no uso de IA, mas não disse como fazer isso.

Da leitura da Recomendação, levando em conta o enfoque desta pesquisa, destacam-se as diretrizes da OCDE quanto à existência de riscos e a necessidade de gerenciá-los, bem como a possibilidade de haver mau uso ou operação de sistemas de IA em condições adversas, que possam afetar direitos das pessoas.

Nesse caso, será preciso haver verificação e explicação sobre como o resultado negativo ocorreu, para que seja possível não só compreendê-lo como também permitir que a parte afetada possa contestá-lo; além de possibilitar que sejam adotadas providências para evitar que o resultado negativo ocorra novamente, podendo até se resolver pela descontinuação do sistema de IA, se for o caso.

A OCDE também destacou a importância da pesquisa e do intercâmbio de informações e boas práticas; a necessidade de regulação constantemente revisada; a necessidade de capacitar os agentes de IA para utilização adequada e eficaz da tecnologia; assim como a necessidade de avaliação periódica de resultados – tudo isso devendo ocorrer simultaneamente, nos termos da Recomendação.

Todos esses aspectos serão levados em consideração no próximo capítulo deste trabalho, que examinará o panorama atual de aplicação e regulação da IA nos tribunais brasileiros, verificando a observância ou não de tais diretrizes, bem como o que está sendo feito e o que falta fazer para uso benéfico da IA no sistema de justiça do Brasil, sobretudo no tocante ao gerenciamento de riscos.

Da mesma forma, nesta pesquisa que envolve a busca de um tratamento adequado para os casos de erros de IA, serão levadas em conta outras diretrizes constantes do item 12 da Declaração de Osaka dos Líderes do G20, como a possibilidade de “regras experimentais (*‘regulatory sandboxes’*)” e a necessidade de “promover a confiança do público em tecnologias de IA”⁸⁷.

Para atingir esse desiderato, será de grande valia examinar o material que vem sendo produzido pelo OECD.AI – *Artificial Intelligence Policy Observatory of the Organization for Economic Cooperation and Development* (Observatório de Políticas de Inteligência Artificial da Organização para Cooperação e Desenvolvimento Econômico).

⁸⁷ BRASIL. **Nota à imprensa nº 170/2019.** Ministério das Relações Exteriores. Disponível em: https://www.gov.br/mre/pt-br/canais_atendimento/imprensa/notas-a-imprensa/2019/declaracao-de-osaka-dos-lideres-do-g20. Acesso em: 2 jan. 2022.

Trata-se desse Observatório de uma plataforma de políticas públicas de IA que foi criada em 2020, na esteira da Recomendação sobre IA da OCDE, que agrupa práticas internacionalmente observadas e comparáveis entre os países, que “pode fornecer análises de políticas públicas multidisciplinares e baseadas em evidências sobre IA em toda a gama de questões de política econômica e social, bem como métricas e medições”⁸⁸.

Também em 2020 a OCDE criou a ONE AI – *OCDE Network of Experts on AI* (Rede de Especialistas em IA da OCDE), tendo em vista o sucesso da experiência realizada com o grupo diversificado de especialistas que forneceu subsídios para a elaboração da Recomendação sobre IA da OCDE. A Rede conta hoje com 164 participantes que se dividem em subgrupos temáticos e é definida como “um grupo de múltiplas partes interessadas e multidisciplinar que fornece consultoria política específica para IA”⁸⁹, contribuindo para o Observatório.

Passados cerca de dois anos da criação dessa rede de especialistas, há vários estudos e ações importantes em andamento, a exemplo da criação de uma força-tarefa⁹⁰ dedicada a estudar e formular orientações de políticas públicas quanto aos investimentos necessários dos países na área da infraestrutura computacional que a IA exige para funcionar de maneira eficaz, incluindo *hardware*, conectividade, serviços em nuvem, *data center* e outros elementos essenciais.

Em junho de 2021 foi divulgado o primeiro relatório⁹¹ produzido pelo subgrupo de especialistas que analisa as políticas e estratégias nacionais de IA, verificando de que maneira os países estão passando dos princípios à prática, ao implementarem as diretrizes da Recomendação sobre IA da OCDE.

Apesar de os países se encontrarem em diferentes estágios quanto à estratégia nacional de IA de cada um, já foi possível encontrar padrões, por exemplo, quanto aos diferentes modelos de governança, coordenação e supervisão ética adotados, que têm sido principalmente os seguintes⁹²:

⁸⁸ OECDAI. **About OECD.AI**. Disponível em: <https://oecd.ai/en/about>. Acesso em: 2 jan. 2022.

⁸⁹ OECDAI. **OCDE Network of Experts on AI (ONE AI)**. Disponível em: <https://oecd.ai/en/network-of-experts>. Acesso em: 2 jan. 2022.

⁹⁰ OECDAI. **Compute & climate**. Disponível em: <https://oecd.ai/en/network-of-experts/working-group/1138>. Acesso em: 2 jan. 2022.

⁹¹ OCDE. State of implementation of the OECD Principles: insights from national AI policies. **OECD Digital Economy Papers** n. 311. Paris: OECD Publishing, 2021. Disponível em: https://www.oecd-ilibrary.org/science-and-technology/state-of-implementation-of-the-oecd-ai-principles_1cd40c44-en. Acesso em: 2 jan. 2022.

⁹² OCDE. State of implementation of the OECD Principles: insights from national AI policies. **OECD Digital Economy Papers** n. 311. p. 21. Paris: OECD Publishing, 2021. Disponível em:

- i. atribuir a supervisão do desenvolvimento e implementação de uma estratégia a um ministério ou agência já existentes;
- ii. criar um novo órgão governamental ou de coordenação para IA;
- iii. estabelecer grupos consultivos de especialistas em IA;
- iv. receber contribuições de órgãos de supervisão e consultoria em ética de IA e dados.

Inclusive o relatório mencionou, ainda que brevemente, o advento da Estratégia Brasileira de Inteligência Artificial, informando apenas que foi lançada em 2021 e foi conduzida pelo Ministério da Ciência, Tecnologia e Inovações⁹³. No próximo capítulo desta dissertação, o cenário regulatório brasileiro no âmbito de cada um dos três Poderes da República, com ênfase no Poder Judiciário, será analisado de forma detalhada.

Além de analisar como os países individualmente considerados estão atuando, o relatório de implementação dos Princípios sobre IA da OCDE também se reporta a diversas organizações intergovernamentais que, em cooperação internacional, estão trabalhando para enfrentar os desafios éticos da aplicação prática da IA.

A propósito, para os fins específicos desta pesquisa, há um ponto muito relevante a destacar do relatório, que é a parte que apresenta “uma seleção de abordagens regulatórias de IA”⁹⁴, quando então surge a **única menção**, em todo o relatório de 92 páginas, à **aplicação de IA em sistemas judiciais**, que foi tratada por uma das organizações intergovernamentais analisadas, qual seja, o Conselho da Europa.

Criado em 1949 e com sede em Estrasburgo, na França, o Conselho da Europa é “a principal organização de direitos humanos do continente”, e “inclui 47 Estados-membros, 27 dos quais são membros da União Europeia”⁹⁵.

Um de seus órgãos de monitoramento⁹⁶ é a CEPEJ – Comissão Europeia para Eficiência da Justiça, criada “no final de 2002, por iniciativa dos Ministros da Justiça europeus”, que pretendiam “estabelecer um organismo inovador para melhorar

https://www.oecd-ilibrary.org/science-and-technology/state-of-implementation-of-the-oecd-ai-principles_1cd40c44-en. Acesso em: 2 jan. 2022.

⁹³ Ibidem, pp. 16 e 22.

⁹⁴ OCDE. State of implementation of the OECD Principles: insights from national AI policies. **OECD Digital Economy Papers** n. 311. p. 30. Paris: OECD Publishing, 2021. Disponível em: https://www.oecd-ilibrary.org/science-and-technology/state-of-implementation-of-the-oecd-ai-principles_1cd40c44-en. Acesso em: 2 jan. 2022.

⁹⁵ COE. **About us**. Disponível em: <https://www.coe.int/en/web/about-us/who-we-are>. Acesso em: 5 jan. 2022.

⁹⁶ COE. **Values**. Disponível em: <https://www.coe.int/en/web/about-us/values>. Acesso em: 5 jan. 2022.

a qualidade e eficiência dos sistemas judiciários europeus e reforçar a confiança dos usuários dos tribunais em tais sistemas”⁹⁷.

Em dezembro de 2018, a CEPEJ publicou a Carta Ética Europeia sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente⁹⁸, que foi justamente o parâmetro adotado pelo Conselho Nacional de Justiça, no Brasil, para o ato normativo que é objeto desta pesquisa – a Resolução nº 332, de 21 de agosto de 2020⁹⁹, que será analisada pormenoradamente no capítulo seguinte desta dissertação, cuja lacuna quanto à efetivação do art. 27 se busca preencher. Por ora, examinam-se na próxima seção a Carta Ética e seus desdobramentos.

2.3 A Carta Europeia de Ética sobre o Uso da IA em Sistemas Judiciais

Adotada pela CEPEJ, um dos órgãos do Conselho da Europa, no início de dezembro de 2018, a Carta Ética Europeia sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente – doravante aqui referida de forma abreviada como Carta Ética, estabeleceu cinco princípios para guiar a aplicação de Inteligência Artificial no âmbito da Justiça, assim definidos¹⁰⁰:

- 1. Princípio do respeito dos direitos fundamentais:** assegurar que a concepção e a aplicação de instrumentos e serviços de inteligência artificial sejam compatíveis com os direitos fundamentais;
- 2. Princípio da não-discriminação:** prevenir especificamente o desenvolvimento ou a intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos;
- 3. Princípio da qualidade e segurança:** em relação ao processamento de decisões e dados judiciais, utilizar fontes certificadas e dados intangíveis com modelos elaborados de forma multidisciplinar, em ambiente tecnológico seguro;
- 4. Princípio da transparência, imparcialidade e equidade:** tornar os métodos de tratamento de dados acessíveis e comprehensíveis, permitindo auditorias externas;
- 5. Princípio “sob o controle do usuário”:** impedir uma abordagem prescritiva e garantir que os usuários sejam atores informados e controlem as escolhas feitas.

⁹⁷ CEPEJ. Council of Europe European Commission for the efficiency of justice (**CEPEJ**). Disponível em: <https://www.coe.int/en/web/cepej/home>. Acesso em: 5 jan. 2022.

⁹⁸ CEPEJ. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente**. Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 5 jan. 2022.

⁹⁹ CNJ. **Resolução n. 332, de 21 de agosto de 2020**. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 10 jul. 2021.

¹⁰⁰ CEPEJ. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente**. Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 5 jan. 2022.

Além dos princípios, a Carta Ética contém quatro anexos. O primeiro apresenta um estudo sobre a aplicação de IA em sistemas judiciais; o segundo trata de categorias de utilizações – as que são encorajadas, as que exigem maior precaução, as que dependem de maior avanço científico e as que exigem extrema cautela; o terceiro anexo traz um glossário e o quarto e último apresenta uma lista de itens para verificar se um sistema de IA aplicado à Justiça observa os cinco princípios éticos.

Para chegar a esses princípios, a CEPEJ se baseou num estudo aprofundado que um grupo de especialistas elaborou sobre a utilização de IA em sistemas judiciais, com ênfase nas aplicações que processam decisões e dados judiciais. O resultado do estudo consta no Anexo I da Carta Ética Europeia e abordou dez tópicos¹⁰¹.

O primeiro tópico foi o panorama de utilização da IA em sistemas judiciais, até 2018, de países membros do Conselho da Europa, verificando-se muito pouca aplicação na esfera pública àquela altura, tanto que no estudo foi mencionada expressamente apenas a ferramenta HART, de análise criminal e risco de reincidência, que estava em operação no Reino Unido¹⁰².

Em contrapartida, foram mencionadas algumas ferramentas comerciais, pois se verificou que estava em franca expansão a aplicação de IA no setor privado, sobretudo em escritórios de advocacia, departamentos jurídicos de empresas e companhias de seguro. Nestas, as aplicações mais difundidas naquele momento eram: motores de pesquisa avançada de jurisprudência; assistência na redação de documentos; análise preditiva e de escalas; classificação e análise de cláusulas contratuais; *chatbots* para atendimento e informação¹⁰³.

O segundo tópico do estudo aprofundado que forneceu subsídios para a elaboração da Carta Ética Europeia analisou o panorama das políticas de dados abertos quanto a decisões judiciais, no âmbito de sistemas judiciais dos países membros do Conselho da Europa. Levou-se em conta que “quanto mais dados disponíveis, mais a IA é capaz de refinar modelos, melhorando a sua capacidade

¹⁰¹ CEPEJ. **European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment**. Disponível em: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 5 jan. 2022.

¹⁰² Ibidem, p.p. 17 e 18.

¹⁰³ Ibidem, p.p. 16 e 17.

preditiva". Contudo, verificou-se que "o tratamento destes dados levanta uma série de questões, tais como alterações na formação da jurisprudência e na proteção dos dados pessoais (incluindo os nomes dos profissionais)"¹⁰⁴.

Entre as muitas questões levantadas, destacaram-se tanto aquelas de ordem prática, que dizem respeito à administração e organização, como a insuficiência de coordenação para coleta e reunião de todas as decisões judiciais publicáveis nos diferentes níveis dos tribunais¹⁰⁵; como também questões mais profundas que dizem respeito à independência e imparcialidade de magistrados.

Nesse sentido, questionou-se sobre "o risco de as decisões judiciais serem feitas de acordo com uma lógica reprodutiva", ou seja, "à qual os juízes seriam incentivados a recorrer sem questionar, com um efeito induzido pela excessiva padronização das decisões judiciais"¹⁰⁶.

Outra preocupação do estudo, em se tratando de política de dados judiciais abertos, disse respeito à proteção de dados pessoais de partes e testemunhas. Mesmo que se ocultem determinadas informações de identificação, ressaltou-se a probabilidade de que tal providência seja o equivalente a "uma simples pseudoanonimização dos dados, e não a uma completa anonimização". Isso porque¹⁰⁷:

O volume e a variedade das informações contidas nas decisões judiciais, juntamente com a crescente facilidade de cruzamento de dados com outras bases de dados, tornam impossível, na prática, garantir que a pessoa em questão não possa ser reidentificada.

Destacou-se ainda que assumia especial relevo a proteção de dados particularmente sensíveis que podem estar contidos nos textos das decisões judiciais, como os que "revelem a origem étnica ou racial, opiniões políticas, filiação sindical, crenças religiosas ou outras, saúde física ou mental ou vida sexual, que são considerados detalhes íntimos"¹⁰⁸.

O estudo aprofundado considerou que todos esses dados sensíveis requerem uma proteção especial, visto que "sua divulgação em massa apresentaria

¹⁰⁴ Ibidem, p. 18.

¹⁰⁵ CEPEJ. **European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment**. p. 23. Disponível em: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 5 jan. 2022.

¹⁰⁶ Ibidem, p. 24.

¹⁰⁷ Ibidem, pp. 25 e 26.

¹⁰⁸ Ibidem, p. 26.

sérios riscos de discriminação, perfilamento e violação da dignidade humana”¹⁰⁹. Por perfilamento, entendeu-se em nota explicativa conforme “definido na seção 4 do Regulamento Geral de Proteção de Dados”¹¹⁰:

É o tratamento dos dados pessoais de um indivíduo para analisar e prever seu comportamento ou situação, tais como a determinação do seu desempenho no trabalho, situação financeira, saúde, preferências, hábitos de vida, etc.

Depois de abordar a questão dos dados, o terceiro tópico do estudo aprofundado tratou das características operacionais da IA aplicada a decisões judiciais, com ênfase no processamento de linguagem natural e no aprendizado de máquina, ressaltando que “na maioria das vezes, o objetivo desses sistemas não é reproduzir o raciocínio jurídico, mas identificar as correlações entre os diferentes parâmetros de uma decisão”, de modo que, aplicando técnicas de aprendizado de máquina, seja possível inferir um ou mais modelos que serviriam para “predizer” ou ‘prever’ uma futura decisão judicial”¹¹¹.

Num enfoque eminentemente prático, deixando claro que se está tratando de “IA fraca”, em contraposição à chamada “IA forte”, que “permanece puramente ficcional”¹¹², ressaltou-se no item 64 do estudo que¹¹³:

64. Para superar quaisquer considerações instintivas ou fantasiosas, estas tecnologias de processamento e análise de informação devem ser compreendidas com base nos três conceitos seguintes.

- **IA não é um objeto único e homogêneo:** ele é na verdade um conjunto de ciências e técnicas (matemática, estatística e ciência da computação) capaz de processar dados para projetar tarefas de processamento computacional muito complexas.

- **Os motores da IA não produzem inteligência por si só, mas usam uma abordagem indutiva:** a ideia é associar de forma quase automatizada um conjunto de observações (entradas) com um conjunto de resultados possíveis (saídas) usando várias propriedades pré-configuradas. Especificamente para a justiça preditiva, o motor constrói ligações entre os diferentes grupos lexicais que compõem as decisões judiciais. Estes grupos estão correlacionados entre os identificados na fase de entrada (fatos e raciocínio) e os identificados na fase de saída (a parte operacional da decisão) e depois classificados.

¹⁰⁹ Ibidem, pp. 25 e 26.

¹¹⁰ Ibidem, p. 26.

¹¹¹ CEPEJ. **European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment.** p. 29. Disponível em: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 5 jan. 2022.

¹¹² Ibidem, p. 31.

¹¹³ Ibidem, pp. 31 e 32.

- A confiabilidade do modelo (ou função) construído depende fortemente da qualidade dos dados utilizados e da escolha da técnica de aprendizagem automática.

Ainda abordando a questão do raciocínio jurídico, o quarto tópico do estudo aprofundado reforça que a utilização da IA em decisões judiciais limita-se objetivamente a categorizar de forma automatizada “os diferentes parâmetros identificados pelos projetistas ou aqueles descobertos pela máquina”¹¹⁴, auxiliando a tomada de decisão ou “antecipando as prováveis decisões de um juiz em casos semelhantes”¹¹⁵. Não há que se falar, assim, de emular o raciocínio jurídico que um magistrado faria, analisando fatos, provas e regras de acordo com o sistema em que atua – *civil law* ou *common law* – porque não é de interpretação que a IA se trata.

O quinto tópico do estudo examinou se seria possível à IA analisar o comportamento dos juízes em retrospectiva, a fim de constatar eventuais vieses em suas decisões, tendo em vista fato ocorrido na França, assim descrito no item 86¹¹⁶:

86. Algumas empresas de tecnologia jurídica foram mais longe e pensaram que poderiam identificar possíveis preconceitos pessoais de juízes e alimentar suspeitas de parcialidade. Os dados abertos dos nomes de certos juízes presidentes dos tribunais administrativos e dos tribunais administrativos recursais na França permitiram desenvolver um indicador da taxa de rejeição de recursos contra as obrigações de sair do território francês tomadas pelas autoridades administrativas. Alguns doutrinadores argumentaram fervorosamente que a alegada imparcialidade dos juízes foi, portanto, posta em dúvida pela inteligência artificial.

Refutando a possibilidade, o estudo concluiu que “para que haja viés pessoal nos processos decisórios dos juízes (diferindo de suas declarações pessoais e públicas no caso em questão)”, seria necessário que a causa determinante e isolada para tais decisões fossem os “traços de personalidade, opiniões ou religião”¹¹⁷ dos juízes, tarefa que requer capacidade analítica e interpretativa de todo o contexto globalmente considerado, que um sistema de IA simplesmente não tem, dado que uma miríade de fatores pode ser levada em conta pelos juízes ao decidirem.

¹¹⁴ Ibidem, p. 35.

¹¹⁵ CEPEJ. **European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment.** p. 37. Disponível em: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 5 jan. 2022.

¹¹⁶ Ibidem, pp. 39 e 40.

¹¹⁷ Ibidem, p. 40.

O sexto tópico do estudo analisou de que maneira a IA deveria ser aplicada no âmbito da justiça civil, comercial e administrativa. Para tanto, já começou alertando que “um discurso instigante que promove um produto de IA pode esconder falhas inaceitáveis de design e resultados de análise totalmente errôneos”¹¹⁸, ao falar sobre uma experiência malsucedida nos tribunais recursais de Douai e Rennes, na França, que em 2017 testaram um *software* de justiça preditiva oferecido por uma empresa.

O objetivo da solução testada era fornecer um instrumento de tomada de decisão que pudesse contribuir para reduzir a discrepância e variabilidade das decisões, colaborando para garantir tratamento mais igualitário aos jurisdicionados. Mas a experiência fracassou porque¹¹⁹:

Foram revelados vieses de raciocínio de software que levaram a resultados aberrantes ou inadequados devido à confusão entre meras ocorrências lexicais do raciocínio jurídico e as causalidades que foram decisivas no raciocínio dos juízes.

Tendo em vista tal exemplo negativo, o estudo aprofundado considerou que o estágio de desenvolvimento tecnológico dos métodos de aprendizado de máquina, naquele momento, ainda não permitia um nível aceitável de confiança para aplicação preditiva de decisões judiciais¹²⁰.

Entretanto, verificou-se enorme potencial de uso da IA para cálculos de escalas, demandas de massa, litígios fiscais e de segurança social, pequenos litígios, resolução *on-line* de disputas, métodos alternativos de solução de conflitos, uniformização de jurisprudência, entre outras aplicações no âmbito cível, comercial e administrativo¹²¹.

Ainda assim, com a ressalva de se atentar para a qualidade dos dados selecionados e desde que observadas as principais garantias processuais, a saber: (i) que as pessoas, mesmo podendo utilizar soluções tecnológicas, conservem seu direito de acesso a um tribunal; (ii) que seja assegurado o princípio do contraditório; (iii) que se busquem soluções para observar a paridade de armas, considerando as dificuldades dos excluídos digitais para o acesso a soluções tecnológicas; (iv)

¹¹⁸ Ibidem, p. 41.

¹¹⁹ Ibidem, p. 42.

¹²⁰ Ibidem, p. 41.

¹²¹ Ibidem, pp. 41-44.

imparcialidade e independência dos juízes, que podem ser indiretamente afetadas por sistemas de controles de precedentes; (v) direito a aconselhamento por advogado¹²².

Especificamente no que diz respeito ao que significa garantir o princípio do contraditório em aplicações de IA a processos judiciais, o estudo considera que¹²³:

Parece imperativo tornar acessível aos cidadãos e, sobretudo, às partes num processo, um certo número de informações quantitativas (por exemplo, o número de decisões processadas para obter a escala) e qualitativas (origem das decisões, representatividade das amostras selecionadas, distribuição das decisões entre diferentes critérios, como o contexto econômico e social), a fim de compreender como foram construídas as escalas, de medir os seus possíveis limites e de as poder debater perante um juiz.

Já no tocante à aplicação de IA no sistema de justiça criminal, abordado no sétimo tópico do estudo aprofundado, são suscitadas questões ainda mais delicadas no uso dessa tecnologia para prevenir infrações e avaliar riscos de perigo e de reincidência¹²⁴.

Tendo em vista experiência negativa verificada nos Estados Unidos com o uso da ferramenta de avaliação de risco de reincidência COMPAS, e críticas à ferramenta de policiamento preditivo HART, do Reino Unido, respectivamente apontadas pelas organizações não-governamentais *ProPublica* e *Big Brother Watch*, quanto a resultados de discriminação de raça e classe, o estudo aprofundado chegou à conclusão que¹²⁵:

Mesmo não sendo especificamente concebidas para serem discriminatórias, a utilização de estatísticas e de IA nos processos penais mostrou um risco de provocar o ressurgimento de doutrinas deterministas em detrimento de doutrinas de individualização da sanção, que foram amplamente adquiridas desde 1945 na maioria dos sistemas judiciais europeus.

O oitavo tópico do estudo retomou a preocupação com a proteção dos dados pessoais dos jurisdicionados, utilizados em sistemas de IA no âmbito dos tribunais, concluindo que “o princípio da precaução deve ser aplicado às políticas de avaliação dos riscos”¹²⁶, desde o design e concepção do sistema como também em

¹²² Ibidem, pp. 47 e 48.

¹²³ Ibidem, p. 47.

¹²⁴ Ibidem, p. 48.

¹²⁵ Ibidem, p.48.

¹²⁶ Ibidem, p. 56.

caso de falha, quando se deve buscar mitigar os riscos que forem identificados. Sobre o tema, ponderou que:¹²⁷

Quando se utiliza a inteligência artificial, os direitos dos titulares dos dados revestem-se de especial importância, e o controle que cada um de nós deve ter sobre a nossas informações pessoais implica que deve ser possível exercer os seguintes direitos: o direito dos titulares de dados de não estarem sujeitos a decisões automatizadas que os afetem significativamente sem que o seu ponto de vista seja levado em conta, o direito a obter informações sobre a fundamentação subjacente ao tratamento de dados efetuado por algoritmos, o direito de oposição a esse tratamento e o direito a um recurso judicial.

O nono tópico do estudo, por sua vez retomando o debate sobre predição travado no terceiro tópico, avaliou que não se pode deixar de ponderar o enorme potencial da aplicação a decisões judiciais, mas não é assunto que se possa deixar a cargo exclusivo de empresas privadas, cabendo o desenvolvimento à esfera pública com participação dos próprios operadores do direito, juntamente com os desenvolvedores técnicos, inclusive em parceria com instituições acadêmicas e de pesquisa científica. Sobre o tema, avaliou-se que¹²⁸:

O termo justiça preditiva deve ser descartado porque é ambíguo e enganoso. Estes instrumentos baseiam-se em métodos de análise da jurisprudência, utilizando métodos estatísticos que não reproduzem de forma alguma o raciocínio jurídico, mas que podem tentar descrevê-lo. Os vieses analíticos, se não puderem ser totalmente eliminados, devem ser identificados. O processo de concepção e a utilização da ferramenta devem ser integrados num quadro ético claro.

O décimo e último tópico do estudo aprofundado justificou a necessidade de estabelecimento dos princípios constantes da Carta Ética Europeia, diante do desafios da aplicação de IA a sistemas judiciais, de modo a se criar “um quadro ético para promover o rápido desenvolvimento de uma forma de IA que inclua mecanismos de prevenção de preconceitos e discriminação nos seus próprios processos de concepção”¹²⁹.

Esse último tópico do estudo tratou também da “importância de debater, testar e rever continuamente a aplicação destas ferramentas antes da implementação de políticas públicas”¹³⁰, com debate público sobre os benefícios e riscos da IA

¹²⁷ Ibidem, p. 57.

¹²⁸ Ibidem, p. 57.

¹²⁹ Ibidem, p. 59.

¹³⁰ Ibidem, p. 60.

aplicada à Justiça, envolvendo não só profissionais da área de tecnologia mas também os operadores do direito, incluindo tribunais, órgãos associativos e representativos de magistrados e de advogados, além de escolas judiciais.

A propósito, sobre as escolas judiciais, o estudo destacou que “podem desempenhar um papel fundamental na sensibilização dos profissionais da justiça para estas questões, para que possam compreender melhor e contribuir na prática para os desenvolvimentos atuais”¹³¹, inclusive sendo capazes de identificar riscos e colaborar em processos de certificação, que são recomendados “a fim de evitar a repetição de erros como o algoritmo COMPAS acima referido”¹³².

O estudo aprofundado finalizou chamando a atenção para a dimensão da responsabilidade de todos os atores de IA, e para a importância de agregar as humanidades à tecnologia, na chamada ciberética. Para tanto, inclusive fez um paralelo com a medicina (assim como a hipótese da presente pesquisa também busca fazer), ao mencionar o Juramento de Hipócrates¹³³:

161. Hoje em dia, todos os especialistas envolvidos no desenvolvimento da IA, incluindo pesquisadores, engenheiros e desenvolvedores de computadores, têm responsabilidades excepcionais e sem precedentes. O seu trabalho poderia ser acompanhado por um reforço ainda maior das humanidades. O exemplo de algumas escolas inovadoras de desenvolvedores de computadores mostra que por trás da vontade de "hackear o sistema" esconde-se, segundo alguns observadores, na realidade um pragmatismo sem qualquer contextualização da responsabilidade que agora paira sobre os técnicos com poderes de quase-demiurgo. O Juramento Hipocrático tem certamente os seus limites na medicina, mas ritualiza a responsabilidade e fornece um quadro ético.

Percebe-se que a Carta Ética, publicada alguns meses antes dos Princípios sobre IA da OCDE (que tanto primaram pelo foco prático, ou seja, na IA aplicada), de fato se deteve nas especificidades do uso real da IA em sistemas judiciais, com seu potencial, limitações e pontos de atenção.

Depreende-se ainda a importância dada ao princípio da precaução, bem como um forte senso de responsabilidade que se pretendeu incutir nos desenvolvedores de tecnologia e formuladores de políticas públicas envolvidos na aplicação de IA nos tribunais; não só no tocante à concepção dos sistemas como

¹³¹ Ibidem, p. 60.

¹³² Ibidem, p. 61.

¹³³ Ibidem, p. 61.

também em relação ao tratamento de erros que porventura venham a acontecer – para que não voltem a ocorrer –, temas que convergem com o escopo deste trabalho.

Será oportuno, aliás, cotejar a aplicação da Resolução brasileira à luz de novos instrumentos que foram posteriormente elaborados pela CEPEJ, sobretudo: (i) “Possível introdução de um mecanismo de certificação de ferramentas de inteligência artificial e serviços na esfera da justiça e do judiciário: estudo de viabilidade”¹³⁴, adotado na 34^a reunião plenária da CEPEJ em 8 de dezembro de 2020; e (ii) “Roteiro revisado para garantir um acompanhamento adequado da Carta Ética da CEPEJ sobre o uso de inteligência artificial em sistemas judiciais e seu ambiente”¹³⁵, aprovado na 37^a reunião plenária da CEPEJ, nos dias 8 e 9 de dezembro de 2021.

Tais instrumentos resultam das atividades de um novo grupo de trabalho que foi criado pela CEPEJ em dezembro de 2019 (um ano depois da elaboração da Carta Ética), especificamente sobre justiça cibernética e IA (CEPEJ-GT-CYBERJUST), “para promover o uso responsável da inteligência artificial, com base no respeito dos direitos fundamentais e dos princípios éticos e atendendo às necessidades dos profissionais do direito”¹³⁶, com o encargo de também desenvolver programas de capacitação. Sobre tudo cabe ao CEPEJ-GT-CYBERJUST¹³⁷:

[...] desenvolver ferramentas com vista a oferecer um enquadramento e garantias aos Estados membros e profissionais do direito que pretendam criar ou utilizar Informação e Tecnologias de Comunicação e/ou mecanismos de Inteligência Artificial nos sistemas judiciais para melhorar a eficiência e a qualidade da justiça.

Salienta-se que, para o estudo de viabilidade de um possível processo de certificação de ferramentas de IA aplicadas a sistemas judiciais, o novo grupo de trabalho da CEPEJ buscou referência inclusive nos padrões de uma instituição técnica que já foi mencionada no final da primeira seção deste capítulo, tendo sido nominada

¹³⁴ CEPEJ. Possible introduction of a mechanism for certifying artificial intelligence tools and services in the sphere of justice and the judiciary: Feasibility Study. Dez. 2020. Disponível em: <https://rm.coe.int/feasability-study-en-cepej-2020-15/1680a0adf4>. Acesso em: 5 jan. 2022.

¹³⁵ CEPEJ. Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment. Dez. 2021. Disponível em: <https://rm.coe.int/cepej-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 5 jan. 2022.

¹³⁶ CEPEJ. Cyberjustice and artificial intelligence used in the field of justice. Disponível em: <https://www.coe.int/en/web/cepej/cepej-working-group-cyber-just>. Acesso em: 6 jan. 2022.

¹³⁷ Ibidem.

pelo Secretário-Geral da ONU no Roteiro de Cooperação Digital¹³⁸ entre as mais relevantes organizações no cenário mundial sobre IA ética: *The IEEE Global Initiative*¹³⁹ – *Global Initiative on Ethics of Autonomous and Intelligent Systems of the Institute of Electrical and Electronic Engineers* (Iniciativa Global sobre Ética de Sistemas Autônomos Inteligentes do Instituto de Engenheiros Elétricos e Eletrônicos).

Trata-se o IEEE, segundo sua própria definição, da “maior organização profissional técnica do mundo dedicada ao avanço da tecnologia para o benefício da humanidade”¹⁴⁰. Criada em Nova York em 1884, congregava em 2020 “mais de 395.000 membros em 160 países”¹⁴¹. Especificamente no tocante à Iniciativa Global IEEE, assim define sua missão¹⁴²:

Garantir que todas as partes interessadas envolvidas no projeto e desenvolvimento de sistemas autônomos e inteligentes sejam educadas, treinadas e capacitadas para priorizar considerações éticas para que essas tecnologias avancem para o benefício da humanidade.

As realizações mais importantes da Iniciativa, e que inclusive serviram de referência para os Princípios de IA de Asilomar elaborados pelo FLI, bem como para os Princípios de IA da OCDE, continuando a embasar políticas públicas ao redor do mundo, são as seguintes (em tradução livre)¹⁴³:

- A criação de um documento seminal criado por centenas de especialistas de todo o mundo ao longo de três anos denominado EAD – *Ethically Aligned Design* (Design Eticamente Alinhado): uma visão para priorizar o bem-estar humano com sistemas autônomos e inteligentes. A versão original foi lançada em 2016, recebeu mais de 500 páginas de comentários e foi lançada novamente em 2017 (EAD / v2). A versão mais recente do EAD foi lançada em 2019.

¹³⁸ ONU. **Report of the Secretary General. Roadmap for Digital Cooperation.** Jun. 2020. Disponível em: https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf#page=32. Acesso em: 6 jan. 2021.

¹³⁹ IEEE. **The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems.** Disponível em: <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>. Acesso em: 6 jan. 2021.

¹⁴⁰ IEEE. **Mission & Vision.** Disponível em: <https://www.ieee.org/about/vision-mission.html>. Acesso em: 6 jan. 2022.

¹⁴¹ IEEE. **History of IEEE.** Disponível em: <https://www.ieee.org/about/ieee-history.html>. Acesso em: 6 jan. 2022.

¹⁴² IEEE. **The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems.** Disponível em: <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>. Acesso em: 6 jan. 2022.

¹⁴³ IEEE. **The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems Key Information, Milestones, and FAQs about The Initiative.** Disponível em: <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/faqs/gieais-faq-11.22.2020.pdf>. Acesso em: 6 jan. 2022.

- A criação de mais de doze grupos de padrões inspirados no EAD – *Ethically Aligned Design* (Design Eticamente Alinhado). A série de padrões IEEE P7000 aborda as principais questões sociotécnicas identificadas pelo EAD de formas pragmáticas e úteis para colocar os princípios em prática para Sistemas de Inteligência Artificial (AIS).

Relevante pontuar que, no design de sistemas de IA alinhados à ética, a Iniciativa Global IEEE afirma que “o mais importante é ser transdisciplinar desde o início, sempre se perguntando: ‘quem deveria estar neste espaço para garantir que o Sistema de IA que criamos é o mais seguro e responsável possível?’”¹⁴⁴

Nesse sentido, a Iniciativa Global informa que seu comitê executivo preza pela diversidade, numa composição em torno de 50% de homens e 50% de mulheres, buscando inclusão de gênero, variedade de localização geográfica e multidisciplinaridade, reunindo especialistas não só na área de engenharia como também de ciências humanas e da área de negócios. Também informa que está envidando esforços para incluir a perspectiva de indígenas e povos originários, bem como de outros grupos da população que são afetados pela aplicação da IA ou poderiam dela beneficiar porém comumente ficam deixados à margem do processo de design¹⁴⁵.

Toda essa perspectiva de design eticamente alinhado da Iniciativa Global IEEE é de suma importância para este estudo, que versa sobre a possibilidade de ocorrerem erros/falhas de sistemas de IA aplicados nos tribunais, que resultem em violações a direitos humanos de jurisdicionados.

A ideia central desta pesquisa é que, se tais erros/falhas e violações fossem reportados de maneira eficaz ao órgão de controle, e catalogados num banco de dados, poderiam ser examinados para investigação da causa – que poderia estar, por exemplo, ligada a má qualidade dos dados de entrada, correlações inexatas, preconceitos e vieses, entre muitos outros fatores.

Surgiria então o potencial de um banco de dados desse tipo para servir como ferramenta de gerenciamento de risco, numa aplicação prática do princípio da precaução, pois se fosse descoberta a causa e fosse possível retificar o sistema, talvez se conseguisse evitar que outras violações viessem a ocorrer. Também seria

¹⁴⁴ IEEE. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems Key Information, Milestones, and FAQs about The Initiative. Disponível em: <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/faqs/gieais-faq-11.22.2020.pdf>. Acesso em: 6 jan. 2022.

¹⁴⁵ Ibidem.

possível suspender o funcionamento do sistema de IA para redesign mais amplo e alinhamento/realignamento ético, e em último caso até proibi-lo, caso não fosse possível sanar a causa da falha, ou fossem constatados riscos desproporcionais aos benefícios.

A ideia também envolve condensar essas informações sobre erros/falhas e violações numa plataforma única nacional, no âmbito do órgão geral de controle – que no Brasil é o Conselho Nacional de Justiça – para que não fiquem dispersas, de modo a tornar possível constatar recorrências e padrões a partir da observação do conjunto.

Sob outro prisma, supõe-se que um banco de erros de IA aberto a consultas também teria sua utilidade por uma questão de transparência, que é tão cara em se tratando de IA, inclusive contribuindo para obter a confiança da população e dos operadores jurídicos em relação à tecnologia aplicada.

Uma utilidade adicional de um banco de erros de IA, que a princípio também se afigura, é a possibilidade de servir de fonte permanente de pesquisa e consulta por desenvolvedores, com potencial de contribuir para o aprimoramento de sistemas.

Tais ideias iniciais deste trabalho parecem encontrar ressonância no novo instrumento da CEPEJ, o “Roteiro revisado para garantir um acompanhamento adequado da Carta Ética da CEPEJ sobre o uso de inteligência artificial em sistemas judiciais e seu ambiente”¹⁴⁶, cujo item 17 sugere, ainda que de forma vaga, a criação de um banco de dados onde possa ser registrada, entre outras informações, a experiência do usuário.

É justamente centrada na experiência do usuário a hipótese com a qual se pretende trabalhar nesta pesquisa, que é saber se é possível adaptar, para uso do Poder Judiciário brasileiro, um banco de dados existente nos Estados Unidos para registrar a experiência do usuário quanto a eventos adversos na utilização de dispositivos médicos e cirurgias robóticas.

Trata-se do banco de dados MAUDE – *Manufacturer and User Facility Device Experience* (mantido pelo órgão regulatório FDA – *U.S. Food & Drug Administration*)¹⁴⁷, cuja estrutura e funcionamento serão explicados detalhadamente no último capítulo desta dissertação.

¹⁴⁶ CEPEJ. **Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment.** 9 dez. 2021. Disponível em: <https://rm.coe.int/cepel-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 6 jan. 2022.

¹⁴⁷ FDA. **MAUDE - Manufacturer and User Facility Device Experience.** Disponível em:

Quanto ao paralelo buscado no campo da Medicina, como hipótese da presente pesquisa (cujo projeto foi apresentado em julho de 2020), também parece haver convergência com o outro documento produzido depois da Carta Ética pela CEPEJ (publicado em dezembro de 2020), “Possível introdução de um mecanismo de certificação de ferramentas de inteligência artificial e serviços na esfera da justiça e do judiciário: estudo de viabilidade”¹⁴⁸.

Isso porque tal estudo de viabilidade de certificação, de modo semelhante a esta pesquisa, buscou referência no órgão europeu de controle de medicamentos e dispositivos médicos, a Diretoria Europeia para a Qualidade de Medicamentos e Cuidados de Saúde (EDQM)¹⁴⁹:

A razão para a EDQM operar dentro do Conselho da Europa é que, de acordo com sua declaração de missão, ela “contribui para o direito humano básico de acesso a medicamentos e cuidados de saúde de boa qualidade e promove e protege a saúde humana e animal”. O objetivo da certificação discutido no presente estudo parece ainda mais estreitamente orientado para as missões de proteção dos direitos humanos do Conselho da Europa na tentativa de adaptá-las ao universo digital e, em particular, aos sistemas de inteligência artificial, na esfera judicial.

Embora atue em uma área que é muito diferente da esfera jurídica e judicial, a experiência da *Pharmacopoeia Europeia* pode ser altamente instrutiva para elaborar uma certificação obrigatória antes de os produtos entrarem no mercado e para estruturar um sistema de certificação do ponto de vista institucional e organizacional (ver seção IV abaixo) e em relação aos organismos de normalização e à União Europeia.

É possível perceber que a linha de raciocínio adotada pela CEPEJ é bastante parecida com a que motivou a escolha da hipótese desta pesquisa. A diferença é que o CEPEJ buscou referências no órgão europeu de controle de medicamentos para tentar adaptar o processo de certificação aos sistemas de IA judiciais; enquanto esta pesquisa buscará referências no órgão regulador norte-americano, em análise comparativa médico-judicial que será feita no último capítulo, para verificar a possibilidade de adaptação de outra ferramenta distinta, o banco de dados alimentado pelos usuários quanto a eventos adversos ocorridos pelo uso de dispositivos médicos, para registro de erros de IA aplicada em tribunais.

<https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Acesso em: 6 jan. 2022.

¹⁴⁸ CEPEJ. **Possible introduction of a mechanism for certifying artificial intelligence tools and services in the sphere of justice and the judiciary: Feasibility Study**. Dez. 2020. Disponível em: <https://rm.coe.int/feasability-study-en-cepej-2020-15/1680a0adf4>. Acesso em: 6 jan. 2022.

¹⁴⁹ Ibidem, p. 10.

3 REGULAÇÃO DOS RISCOS DA IA NO BRASIL

Ainda não existe uma lei brasileira sobre IA, não obstante alguns projetos de lei sobre o assunto estejam em tramitação no Congresso Nacional. Diante dessa lacuna, serão analisados no presente capítulo os princípios gerais que podem ser extraídos da Constituição Federal e aplicados à temática, bem como serão abordadas as iniciativas atualmente existentes para a regulação da IA no âmbito de cada um dos três Poderes – com ênfase no Poder Judiciário e enfoque nos riscos que poderão dar ensejo a possíveis erros e incidentes da IA aplicada pelos tribunais, tema que é o objeto de estudo desta pesquisa.

Partindo da Constituição Federal, as subseções seguintes irão tratar, com embasamento teórico e registros histórico-cronológicos, do marco legal da IA, cuja discussão no momento se encontra no Senado, além da Estratégia Brasileira de Inteligência Artificial, elaborada pelo Poder Executivo, finalizando com uma análise crítica detalhada da Resolução CNJ nº 332/2020 e da Portaria CNJ nº 271/2020, em vigor para o Poder Judiciário.

3.1 A Constituição Federal de 1988 à luz da “ética para a civilização tecnológica” de Hans Jonas e da “sociedade de risco” de Ulrich Beck

Em sua publicação original, a Constituição da República Federativa do Brasil de 1988 tratou “Da Ciência e Tecnologia”¹⁵⁰ dedicando ao tema o Capítulo IV, contendo apenas o art. 218 e o art. 219, que permaneceram tais quais redigidos pela Assembleia Constituinte por quase 27 anos, até que sobreveio a Emenda Constitucional nº 85, de 2015, dando nova redação a dispositivos e incluindo outros, com o objetivo de “atualizar o tratamento das atividades de ciência, tecnologia e inovação”¹⁵¹.

De modo resumido, tais modificações disseram respeito ao seguinte:

¹⁵⁰BRASIL. **Constituição Federal de 1988 (publicação original).** Disponível em: <https://www2.camara.leg.br/legin/fed/consti/1988/constituicao-1988-5-outubro-1988-322142-publicacaooriginal-1-pl.html>. Acesso em: 22 maio 2022.

¹⁵¹BRASIL. **Emenda Constitucional nº 85, de 26 de fevereiro de 2015.** Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc85.htm#art1. Acesso em: 22 maio 2022.

- Competência comum dos entes federativos de proporcionar os meios de acesso à ciência, à tecnologia, à pesquisa e à inovação (art. 23, V);
- Competência concorrente da União, Estados e Distrito Federal para legislar sobre o tema (art. 24, IX);
- Permissão para realocação de recursos orçamentários sem necessidade de prévia autorização legislativa, para viabilizar projetos de ciência, tecnologia e inovação (art. 167, § 5º);
- Incremento tecnológico do Sistema Único de Saúde (art. 200, V);
- Apoio financeiro do Poder Público para atividades de pesquisa e incentivo à inovação realizadas por universidades ou instituições de educação profissional e tecnológica (art. 213, § 2º);
- Estímulo à articulação entre entes públicos e privados em todas as esferas de governo para execução de atividades relacionadas à ciência, pesquisa, desenvolvimento e capacitação em tecnologia e inovação (art. 218, § 6º);
- Incentivo à atuação, no exterior, de instituições públicas que executem atividades de ciência, tecnologia e inovação (art. 218, § 7º);
- Estímulo à cultura da inovação nos entes públicos e privados, bem como ao fomento de parques e polos tecnológicos (art. 219, parágrafo único);
- Possibilidade de instrumentos de cooperação firmados entre entes federativos e outros órgãos públicos e entidades privadas, com ou sem contrapartida financeira, para compartilhamento de infraestrutura tecnológica ou recursos humanos especializados (art. 219-A);
- Previsão do Sistema Nacional de Ciência, Tecnologia e Inovação (SNCTI), em regime colaborativo entre entes públicos e privados, a ser criado por lei federal, com possibilidade de competência concorrente para que os Estados, o Distrito Federal e os Municípios possam legislar sobre suas particularidades (art. 219-B).

A propósito, sobre esse último item, passados mais de 7 anos da promulgação da Emenda Constitucional nº 85/2015, o Sistema Nacional de Ciência, Tecnologia e Inovação (SNCTI) ainda não foi criado por lei federal, havendo notícia do

Ministério da Ciência, Tecnologia e Inovações no sentido de que a previsão para remeter o Projeto de Lei sobre o assunto para a Câmara dos Deputados é até o terceiro trimestre de 2022¹⁵².

No tocante ao tema central desta dissertação – Inteligência Artificial (IA) – não há qualquer menção ou regulação de *status* constitucional referente a essa tecnologia específica, como se pode perceber nos dispositivos mencionados, nem mesmo em virtude da atualização efetivada em 2015 pelo poder constituinte derivado.

Falou-se no capítulo anterior deste trabalho sobre tecnologias que têm o potencial de colocar em risco a humanidade, entre elas, algumas aplicações de IA, como armas autônomas letais. Todavia, a única tecnologia que envolve risco existencial que está expressamente prevista desde a publicação original da Constituição de 1988 é a nuclear, então mencionada 8 vezes no texto.

Posteriormente, sobrevieram alterações de redação e inclusões por meio da Emenda Constitucional nº 49, de 2006¹⁵³, e da recente Emenda Constitucional nº 118, de 2022¹⁵⁴. Desse modo, o regramento constitucional da tecnologia nuclear se encontra no art. 21, XXIII, “a”, “b”, “c” e “d”; art. 22, XXVI; art. 49, XIV; art. 177, V; e art. 225, § 6º da Constituição Federal.

Em apertada síntese, tais dispositivos estabelecem a competência da União para legislar sobre o tema e para explorar serviços e instalações nucleares, exercendo monopólio estatal relativo a minérios nucleares, tudo para fins pacíficos, mediante aprovação do Congresso Nacional, com possibilidade de autorização em regime de permissão para usos agrícola, industrial e médico, ficando estabelecida expressamente a responsabilidade civil objetiva para danos nucleares.

Há por fim, no capítulo relativo ao meio ambiente, diante do dever de preservação para as futuras gerações, a regra de que é preciso haver lei federal para definir a localização de instalação de usinas nucleares.

¹⁵²BRASIL. **MCTI tem reunião sobre ciência e tecnologia com representantes do ICTP.Br. Ministério da Ciência, Tecnologia e Inovações.** 12 maio 2022. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2022/05/mcti-tem-reuniao-sobre-ciencia-e-tecnologia-com-representantes-do-ictp.br>. Acesso em: 22 maio 2022.

¹⁵³BRASIL. **Emenda Constitucional nº 49, de 8 de fevereiro de 2006.** Disponível em: <https://www2.camara.leg.br/legin/fed/emecon/2006/emendaconstitucional-49-8-fevereiro-2006-540999-publicacaooriginal-42678-pl.html>. Acesso em: 22 maio 2022.

¹⁵⁴BRASIL. **Emenda Constitucional nº 118, de 26 de abril de 2022.** Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc118.htm. Acesso em: 22 maio 2022.

Mas por que está havendo, nesta pesquisa sobre IA, toda essa referência ao regramento constitucional brasileiro relativo à tecnologia nuclear?

A resposta está diretamente relacionada ao contexto da época da promulgação da Carta Magna, em 1988, a seguir analisada à luz de referencial teórico que ajuda a entender todos esses cuidados que o legislador constituinte demonstrou, ao tentar gerenciar riscos advindos de uma nova tecnologia poderosa que trazia progresso, mas também gerava riscos que poderiam resultar em significativos prejuízos ao povo brasileiro, inclusive a longo prazo.

O objetivo, aqui, é estudar o que ocorreu no passado, a respeito de uma nova tecnologia (nuclear), na mesma linha do que foi feito no primeiro capítulo deste trabalho (sobre o DNA recombinante), a fim de extrair lições úteis para o presente e o futuro em relação a outra nova tecnologia (IA).

A Assembleia Constituinte da qual resultou, em 5 de outubro de 1988, a promulgação da Constituição Federal em vigor, foi instalada em 1º de fevereiro de 1987¹⁵⁵, sendo natural que refletisse os acontecimentos de seu tempo.

Justamente no ano anterior, em 26 de abril de 1986, aconteceu aquele que até hoje, passados 36 anos, é considerado “o pior desastre nuclear da história”¹⁵⁶, quando houve uma explosão na Usina Nuclear de Chernobyl, na Ucrânia, resultando numa nuvem de radioatividade que ultrapassou fronteiras, deixando um rastro de destruição e morte por gerações, e uma cidade fantasma “privada de vida pelos próximos vinte mil anos”¹⁵⁷.

No mesmo ano, 1986, poucos meses após o acidente em Chernobyl, o sociólogo alemão Ulrich Beck publicou o livro “Sociedade de Risco: rumo a uma outra modernidade”¹⁵⁸, considerado em 2001 pela *International Sociological Association* (ISA) como um dos 20 livros mais importantes de sociologia do século XX¹⁵⁹.

¹⁵⁵BRASIL. 30 anos da Constituição. Disponível em: <https://www.camara.leg.br/internet/agencia/infograficos-html5/constituinte/index.html>. Acesso em: 22 maio 2022.

¹⁵⁶ONU. **Chefe de agência atômica da ONU visita Chernobil 36 anos após desastre nuclear.** ONU News. 26 abr. 2022. Disponível em: <https://news.un.org/pt/story/2022/04/1787232>. Acesso em: 22 maio 2022.

¹⁵⁷DUPUY, Jean-Pierre. A catástrofe de Chernobyl vinte anos depois. **Estudos Avançados**, [S. I.], v. 21, n. 59, p. 243-252, 2007. Disponível em: <https://www.revistas.usp.br/eav/article/view/10219>. Acesso em: 23 maio 2022.

¹⁵⁸BECK, Ulrich. **Sociedade de risco: rumo a uma outra modernidade.** Tradução de Sebastião Nascimento. São Paulo: Editora 34, 2011 (2ª edição). 384 p.

¹⁵⁹ISA. **Books of the XX century.** Disponível em: <https://www.isa-sociology.org/en/about-isa/history-of-isa/books-of-the-xx-century>. Acesso em: 23 maio 2022.

Beck teorizou que a chamada “sociedade industrial clássica” advinda da Revolução Industrial se transmudou a partir da segunda metade do século XX para o que denominou de “sociedade (industrial) de risco”¹⁶⁰, permeada por “problemas e conflitos surgidos a partir da produção, definição e distribuição de riscos científico-tecnologicamente produzidos”¹⁶¹.

Nessa nova configuração de sociedade apontada por Beck, “a produção social de riqueza é acompanhada sistematicamente pela produção de riscos”¹⁶².

Desse modo, “a reboque das forças produtivas exponencialmente crescentes no processo de modernização, são desencadeados riscos e potenciais de autoameaça numa medida até então desconhecida”¹⁶³.

Discorrendo sobre diversos exemplos da época sobre danos a consumidores, a populações e ao meio ambiente – que teve em Chernobyl a máxima expressão de gravidade – Ulrich Beck apontou, a respeito das reinantes ausência ou minimização de gerenciamento de riscos, que o poder político e/ou econômico impunha aos cientistas desenvolvedores de novas tecnologias a pressão da “desumana lei da infalibilidade”¹⁶⁴, como se o erro, que é inerente à condição humana, não pudesse sequer ser cogitado, sob pena de abalo de credibilidade e finanças de governos e empresas.

O cerne do problema, segundo Beck, é que “no esforço pelo aumento da produtividade, sempre foram e são deixados de lado os riscos implicados”¹⁶⁵, numa espécie de miopia ou cegueira convenientes e sistemáticas quanto à gestão dos riscos.

Nessa toada, ele fez uma crítica, carregada de ironia, dizendo que as possíveis aplicações econômicas das novas tecnologias são exaustivamente previstas e testadas à luz das melhores técnicas, em estado da arte, ao passo que os riscos são deixados de lado, e então quando surgem problemas sérios ou até tragédias, são recebidos com ares de surpresa e profunda consternação, tidos como fatalidades¹⁶⁶.

¹⁶⁰BECK, Ulrich. **Sociedade de risco: rumo a uma outra modernidade**. Tradução de Sebastião Nascimento. São Paulo: Editora 34, 2011 (2^a edição). p. 12.

¹⁶¹Ibidem, p. 23.

¹⁶²Ibidem, p. 23

¹⁶³Ibidem, p. 23.

¹⁶⁴Ibidem, p. 65.

¹⁶⁵Ibidem, p. 73.

¹⁶⁶Ibidem, p. 73.

A solução apontada por Ulrich Beck é o fomento de uma cultura civil de responsabilidade, informando o desenvolvimento tecnológico com valores éticos e filosóficos¹⁶⁷ e, diante dos riscos, assumindo uma atitude proativa para buscar antecipá-los, dimensioná-los e tratá-los ainda que incalculáveis, pelo princípio da “precaução pela prevenção”¹⁶⁸.

Dito de outro modo, à decisão humana de colocar em funcionamento uma tecnologia que trará progresso mas também produzirá riscos, deve corresponder ao mesmo tempo outra decisão humana “sobre suscetibilidades: sobre alcance, tipo e teor de ameaça, círculo de pessoas afetadas, efeitos retardados, medidas a tomar, responsáveis, demandas de reparação”¹⁶⁹.

Outro aspecto apontado por Beck é que a sociedade de risco é também a sociedade “da mídia e da informação”, havendo “aqueles que são afetados pelos riscos e aqueles que lucram com eles”, “aqueles que produzem definições de risco e aqueles que as consomem”. Surge então uma tensão a partir do que é informado por cada grupo, de acordo com seu ponto de vista e interesses, gerando “disputas definitórias em torno da extensão, do grau e da urgência dos riscos”¹⁷⁰, cujo efeito na opinião pública, diante de uma nova tecnologia, é o medo¹⁷¹.

A propósito, é justamente a “heurística do medo” que está na base do “princípio responsabilidade”, desenvolvido pelo filósofo alemão Hans Jonas no livro lançado em 1979 ao ensaiar “uma ética para a civilização tecnológica”¹⁷², diferente da ética clássica antropocêntrica e mais voltada para a esfera da política pública do que do indivíduo, colocando a responsabilidade no centro da ética como medida de autocontenção de poder.

Segundo ele, dados os riscos associados à tecnologia, é a “previsão do perigo”¹⁷³, buscando antever males que podem acontecer, a bússola para o agir diligente nessa nova teoria ética.

¹⁶⁷Ibidem, p. 34.

¹⁶⁸Ibidem, p. 364.

¹⁶⁹Ibidem, p. 65.

¹⁷⁰Ibidem, p. 56.

¹⁷¹Ibidem, p. 60.

¹⁷²JONAS, Hans. **O princípio responsabilidade: ensaio de uma ética para a civilização tecnológica.** Tradução de Marijane Lisboa, Luiz Barros Montez. Rio de Janeiro: Contraponto: Ed. PUC-Rio, 2006.

¹⁷³Ibidem, p. 21.

Atribui-se a Hans Jonas, pela obra acima mencionada, “o primeiro estudo ético sobre o princípio da precaução”¹⁷⁴, nas palavras da professora Teresa Ancona Lopez, que abordou o tema na tese com a qual conquistou em 2009 o cargo de Professora Titular de Direito Civil da Faculdade de Direito da Universidade de São Paulo. Segundo ela:

Os princípios da prevenção (para riscos determinados) e da precaução (para riscos hipotéticos) poderão gerir os graves riscos oriundos das descobertas tecnológicas e científicas, distribuindo mais segurança para todos, cumprindo, assim, nesse aspecto, a justiça social. Podemos ilustrar o aparecimento desse princípio, primeiro lembrando dos problemas ambientais sérios que o mundo passou a enfrentar a partir dos anos 70, culminando com a positivação do princípio da precaução na Carta da ECO-92, no Rio de Janeiro¹⁷⁵.

As ideias de sociedade de risco de Ulrich Beck e de inovação com prudência ética de Hans Jonas são fundamentais para a compreensão do “princípio da precaução como novo fundamento da responsabilidade civil”¹⁷⁶ no século XXI, cujo lastro jurídico, no Brasil, assim como do princípio da prevenção, de acordo com Tereza Ancona Lopez, “é a **obrigação geral de segurança**, que veio se acrescentar aos princípios da **reparação integral** e da **solidariedade**, com a crescente ‘socialização de riscos’”¹⁷⁷.

Em se tratando nesta dissertação de um instrumento de gerenciamento de riscos quanto ao uso de IA pelo Poder Judiciário, em aplicação prática do princípio da precaução, tudo o que foi visto no presente tópico foi para demonstrar que, embora até o momento não exista qualquer dispositivo na Constituição Federal que trate especificamente de IA, é preciso ter em mente que, como diz Marçal Justen Filho, “aquele que é investido de competências estatais tem o dever objetivo de adotar as providências necessárias e adequadas a evitar danos às pessoas e ao patrimônio”¹⁷⁸.

Para tanto, cumpre observar que na Constituição a **obrigação geral de segurança** está prevista no art. 5º, *caput*, ao mencionar a segurança entre os direitos

¹⁷⁴LOPEZ, Teresa Ancona. **Princípio da precaução e evolução da responsabilidade civil**. São Paulo: Quartier Latin, 2010, p. 110.

¹⁷⁵Ibidem, p. 110.

¹⁷⁶Ibidem, p. 86.

¹⁷⁷Ibidem, p. 240.

¹⁷⁸JUSTEN FILHO, Marçal. **Curso de Direito Administrativo**. 10. ed. São Paulo: Revista dos Tribunais, 2014, p. 1.335.

e garantias fundamentais, no sentido mais amplo possível, assim como consta a segurança entre os direitos sociais previstos no *caput* do art. 6º.

O princípio da **reparação integral** também se insere no rol dos direitos e garantias fundamentais, estabelecido no art. 5º, V e X, enquanto o **princípio da solidariedade** está no art. 3º, I. Especificamente em relação ao dever estatal de reparação, a **responsabilidade civil objetiva do Estado** está prevista no art. 37, § 6º¹⁷⁹.

Assim, todos os princípios constitucionais mencionados informam este trabalho, que busca refletir sobre mecanismos para uso mais seguro da IA nos tribunais.

3.2 O Projeto de Lei nº 21/2020 e o futuro marco legal da IA no Brasil

Encontra-se em tramitação no Congresso Nacional o Projeto de Lei nº 21/2020, iniciado na Câmara dos Deputados, quando foi apresentado pelo deputado Eduardo Bismarck (PDT/CE) em 04/02/2020, cuja ementa atualizada até o momento é a seguinte: “Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil; e dá outras providências”¹⁸⁰.

Na justificativa do texto original do PL 21/2020, fez-se menção expressa aos princípios para o desenvolvimento da IA estabelecidos pela OCDE em 2019 (que foram analisados no capítulo anterior deste trabalho), levando em conta que o Brasil foi um dos países signatários do documento e se comprometeu a promover e implementar tais princípios.

Inicialmente contendo 16 artigos, o PL 21/2020 trazia algumas definições inspiradas nos princípios da OCDE sobre o que seria considerado como sistemas de inteligência artificial, ciclo de vida do sistema de inteligência artificial, conhecimento em inteligência artificial, agentes de inteligência artificial (subdivididos em agentes de desenvolvimento e agentes de operação), partes interessadas e relatório de impacto de inteligência artificial.

¹⁷⁹BRASIL. **Constituição Federal de 1988.** Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 25 maio 2022.

¹⁸⁰BRASIL. **PL 21/2020.** Disponível em: <https://www.camara.leg.br/propostas-legislativas/2236340>. Acesso em: 25 maio 2022.

Também foram estabelecidos fundamentos, objetivos, princípios, direitos das partes interessadas, deveres dos agentes de inteligência artificial e diretrizes para atuação dos entes federativos, entre outras disposições.

Durante a tramitação na Câmara dos Deputados, ao PL 21/2020 apensaram-se outros três projetos que tratavam sobre IA: PL nº 240/2020, PL nº 4120/2020 e o PL nº 1.969/2021. Houve audiência pública realizada em 08/07/2021, passagem por Comissões e aprovação de requerimento para tramitação em regime de urgência, até que foi levado a plenário na sessão do dia 29/09/2021, quando foram apresentadas 10 emendas, chegando-se após votações à Subemenda Substitutiva Global ao Projeto de Lei nº 21/2020, adotada pela relatora designada, deputada Luiza Canziani (PTB/PR), ficando prejudicados o texto original, os projetos apensados e as emendas, ressalvados os destaques¹⁸¹.

O que é importante ressaltar dessa tramitação é que no parecer da relatora¹⁸² foram explicitadas as razões para alteração do texto original em muitos aspectos, motivada sobretudo pelo avanço dos debates sobre a regulação da IA na Europa, o chamado *AI Act*, que vem a ser a proposta de lei do Parlamento Europeu que é considerada a mais relevante do mundo até agora¹⁸³.

Com efeito, em 21/04/2021 foi apresentada em Bruxelas a “Proposta de Regulamento do Parlamento Europeu e do Conselho que Estabelece Regras Harmonizadas em Matéria de Inteligência Artificial (Regulamento Inteligência Artificial”, após anos de estudos e debates, seguidos de uma consulta pública *on-line* no período de 19/02/2020 a 14/06/2020, que recebeu 1.215 contribuições¹⁸⁴.

O ponto alto da proposta do *AI Act* europeu é sua abordagem ao tratar da **regulação com base no risco**, exatamente como ocorreu no passado em relação à então nova tecnologia sobre DNA recombinante, quando na Conferência de Asilomar

¹⁸¹BRASIL. **PL 21/2020**. Disponível em: <https://www.camara.leg.br/propostas-legislativas/2236340>. Acesso em: 25 maio 2022.

¹⁸²BRASIL. **Parecer de Plenário pelas Comissões de Ciência e Tecnologia, Comunicação e Informática, Trabalho, de Administração e Serviço Público e Constituição e Justiça e de Cidadania ao Projeto de Lei nº 21, de 2020**. Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarIntegra?codteor=2082139. Acesso em: 26 maio 2022.

¹⁸³FLI. **The AI Act**. Disponível em: <https://artificialintelligenceact.eu/>. Acesso em: 26 maio 2022.

¹⁸⁴COMISSÃO EUROPEIA. **Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União**. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>. Acesso em: 26 maio 2022.

de 1975 os cientistas decidiram que o grau de risco deveria ser o fator decisivo para permissão, permissão com ressalvas ou proibição de aplicações, conforme analisado no capítulo anterior desta dissertação.

Na proposta do *AI Act* europeu, aplicações de IA que não apresentarem risco ou tiverem risco mínimo serão permitidas sem restrições. Para aquelas que apresentarem risco, poderão estar sujeitas a maiores obrigações de informação e transparência. Em se tratando de aplicações de alto risco, a permissão de uso estará sujeita a prévia avaliação de conformidade com requisitos ético-legais. Finalmente, para aplicações de risco inaceitável, a proibição será imposta.

Para a nova redação do projeto da lei brasileira, a deputada relatora deixou claro em seu parecer que a base seriam os princípios da OCDE de 2019, mas com os avanços da proposta europeia de regulamento de 2021, sobretudo com dois destaques: (i) uma nova definição, mais detalhada, para o que será considerado “sistema de Inteligência Artificial”, contendo um rol exemplificativo de técnicas computacionais, a fim de diferenciar a IA propriamente dita de sistemas menos sofisticados; (ii) o esforço para resguardar as pessoas dos riscos sem regular demais a IA, tudo para não inibir os avanços tecnológicos, tendo como baliza o grau de risco das aplicações.

Entretanto, de forma diferente da proposta de regulamento europeia, passando a inspirar-se em modelos norte-americanos, optou-se por não estabelecer de antemão proibições absolutas nem quais seriam os graus de risco de possíveis aplicações, deixando a tarefa para leis posteriores ou autorregulação setorial e descentralizada – tanto que foi rejeitada a emenda que sugeriu a criação de um órgão central ou uma espécie de agência reguladora nacional exclusiva sobre IA.

Assim, por exemplo, pelo texto aprovado na Câmara dos Deputados, se a IA for aplicada a dispositivos médicos, caberá à ANVISA regular, se for aplicada a um produto financeiro ficará a cargo dos órgãos reguladores (Bacen, CVM, Susep, Previc), e assim por diante. Por essa lógica, para o que interessa à presente pesquisa, a regulação da IA aplicada ao Poder Judiciário caberá ao órgão que exerce seu controle, que é o CNJ, considerando o art. 103-B, § 4º, da Constituição Federal.

Com as premissas principais acima mencionadas, o PL nº 21-A/2020 recebeu uma nova redação ao ser aprovado na Câmara dos Deputados em 29/09/2021, inclusive no que se refere à responsabilidade civil, seguindo mais enxuto para o Senado, com 10 artigos.

No Senado, em 03/02/2022 a Presidência determinou a tramitação conjunta do PL nº 21/2020 com os Projetos de Lei nº 5.051, de 2019, e 872, de 2021, dada a correlação entre os temas¹⁸⁵.

Além disso, por meio do Ato do Presidente do Senado Federal nº 4, de 2022¹⁸⁶, publicado em 17/02/2022, foi instituída uma Comissão de Juristas¹⁸⁷, presidida pelo Ministro Ricardo Villas Bôas Cueva, do Superior Tribunal de Justiça, com a finalidade de subsidiar a elaboração de minuta de substitutivo para instruir a apreciação dos três projetos apensados, com prazo inicial de 120 dias. Em 09/06/2022, a Comissão anunciou a prorrogação dos trabalhos por mais 120 dias¹⁸⁸.

O fato é que o PL nº 21-A/2020 que saiu da Câmara dos Deputados foi bastante criticado pela comunidade acadêmica brasileira¹⁸⁹, que apesar de apontar alguns aspectos positivos, qualificou o texto de extremamente vago, genérico e até inútil, resultado de uma injustificada tramitação acelerada, à margem da contribuição de especialistas, passando muito longe da profundidade e do debate público que se verificou na proposta de regulamento europeia.

Para o professor Anderson Schreiber, o PL nº 21-A/2020 “cria uma fratura no ordenamento jurídico brasileiro”, destoando completamente do sistema pário, ao prever responsabilidade subjetiva para os desenvolvedores de sistemas de IA. Essa foi tida pelo referido professor como “uma passagem especialmente estarrecedora” do projeto, porque “o dispositivo parece ignorar todo o processo de evolução da disciplina da responsabilidade civil em nosso país”¹⁹⁰, referindo-se ao art. 927,

¹⁸⁵BRASIL. **Projeto de Lei nº 21/2020.** Disponível em: <https://www.congressonacional.leg.br/materias/materias-bicamerais/-/ver/pl-21-2020>. Acesso em: 26 maio 2022.

¹⁸⁶BRASIL. **Ato do Presidente do Senado Federal nº 4, de 2022.** Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9087218&ts=1652881168283&disposition=inline>. Acesso em: 26 maio 2022.

¹⁸⁷BRASIL. **CJSUBIA - Comissão de juristas responsável por subsidiar elaboração de substitutivo sobre inteligência artificial.** Senado. Disponível em: <https://legis.senado.leg.br/comissoes/comissao?codcol=2504>. Acesso em: 24 jun. 2022.

¹⁸⁸FRAGOSO, Roberto. **Comissão da inteligência artificial prorroga o prazo trabalhos por 120 dias.** Senado Notícias, 9 jun. 2022. Disponível em: <https://www12.senado.leg.br/noticias/audios/2022/06/comissao-da-inteligencia-artificial-prorroga-trabalhos-por-120-dias>. Acesso em: 26 maio 2022.

¹⁸⁹NOGUEIRA, Pablo. Projeto de marco legal da IA no Brasil é pouco consistente e pode ser inútil, dizem especialistas. **Jornal da Unesp.** 29 jul 2021. Disponível em: <https://jornal.unesp.br/2021/07/29/projeto-de-marco-legal-da-ia-no-brasil-e-pouco-consistente-e-pode-ser-inutil-dizem-especialistas/>. Acesso em: 26 maio 2022.

¹⁹⁰SCHREIBER, Anderson. **PL da Inteligência Artificial cria fratura no ordenamento jurídico brasileiro.** Jota. 2 nov 2021. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/coluna-do-anderson-schreiber/pl-inteligencia-artificial-cria-fratura-no-ordenamento-juridico-02112021>. Acesso em: 26 maio 2022.

parágrafo único, do Código Civil, pelo qual a regra para danos resultantes de atividades de risco é a responsabilidade objetiva¹⁹¹.

Como se percebe, muitas são as questões sobre as quais a Comissão de Juristas instituída no Senado terá de se debruçar ao elaborar o texto substitutivo, parecendo ainda distante o advento do marco legal da IA no Brasil.

Caso o Legislativo brasileiro continue tomando por parâmetro a proposta europeia, ainda poderá haver bastante modificação, porque as notícias de junho de 2022 dão conta que, no Parlamento Europeu, foram apresentadas milhares de emendas por eurodeputados, indicando que as negociações políticas serão complexas e pelo visto ainda vão longe¹⁹².

Até mesmo a própria definição de IA ainda não tem consenso na Europa, havendo emendas para ampliação do alcance do *AI Act* para regular a IA aplicada a metaverso e *blockchain*, ampliar proibições de uso para reconhecimento de emoções e reconhecimento biométrico, ampliar a definição de sistemas de alto risco e aumentar seu controle, incluir o conceito da confiabilidade e dar-lhe clareza jurídica, modificar regras de governança, aumentar a transparência e reduzir opacidade de sistemas, discutir multas e seus valores, entre diversas questões¹⁹³.

Outra notícia de junho de 2022 quanto ao *AI Act*, que interessa a este trabalho por uma similaridade com a hipótese de pesquisa, informa que na governança da IA poderá ser tomado como parâmetro o Regulamento de Dispositivos Médicos europeu para aplicações consideradas de alto risco¹⁹⁴.

No mais, em maio de 2022 especialistas brasileiras já haviam apontado, ao serem entrevistadas¹⁹⁵, que a proposta de regulação europeia, embora traga referenciais importantes, pode não servir de modelo para o Brasil em vários aspectos.

¹⁹¹BRASIL. Lei nº 10.406, de 10 de janeiro de 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm. Acesso em: 26 maio 2022.

¹⁹²BERTUZZI, Luca. **AI regulation filled with thousands of amendments in the European Parliament.** Euractiv, 2 jun. 2022. Disponível em: <https://www.euractiv.com/section/digital/news/ai-regulation-filled-with-thousands-of-amendments-in-the-european-parliament/>. Acesso em: 23 jun. 2022.

¹⁹³BERTUZZI, Luca. Ibidem.

¹⁹⁴BERTUZZI, Luca. **Czech presidency sets out path for AI Act discussions.** Euractiv, 22 jun. 2022. Disponível em: <https://www.euractiv.com/section/digital/news/czech-presidency-sets-out-path-for-ai-act-discussions/>. Acesso em: 24 jun. 2022.

¹⁹⁵SGARIONI, Mariana. **IA: especialistas alertam que proposta de regulação da UE não serve de modelo para o Brasil.** Mobile Time, 30 maio 2022. Disponível em: <https://www.mobiletme.com.br/noticias/30/05/2022/ia-especialistas-alertam-que-proposta-de-regulacao-da-ue-nao-servi-de-modelo-para-o-brasil/>. Acesso em: 24 jun. 2022.

Como principal ponto positivo do *AI Act*, a professora Dora Kaufman, da PUC-SP, indicou a categorização das aplicações de IA por risco. Como ponto negativo de destaque, mencionou-se a ideia de criação de uma agência reguladora central, em vez de uma regulação multissetorial mais próxima do modelo norte-americano, opinião também seguida por Patrícia Peck, membro do Conselho Nacional de Proteção de Dados Pessoais, e por Loren Spíndola, Diretora de Políticas Públicas na Microsoft Brasil¹⁹⁶.

Fato é que, tanto na Europa como no Brasil, seguem os debates nos respectivos parlamentos, de modo que ainda não há qualquer lei sobre IA – um tema ainda muito novo e que gera inúmeras discussões sob diversos prismas.

Pode-se afirmar que as milhares de emendas recentemente apresentadas à proposta do *AI Act*, no Parlamento Europeu, bem demonstram o tamanho do desafio e dão a devida dimensão do equívoco, no Brasil, da tramitação em regime de urgência e do açoitamento da aprovação do PL 21/2020 na Câmara dos Deputados, passando ao largo da ampla discussão necessária.

3.3 A Estratégia Brasileira de Inteligência Artificial

Comparando com o cenário internacional, segundo análise realizada em 2020 por Raphael Cóbé, da Universidade Estadual de São Paulo, e outros pesquisadores, países como Canadá, Japão, Cingapura, China e Emirados Árabes já haviam lançado suas estratégias nacionais de IA desde 2017, e inclusive estes últimos, os Emirados Árabes Unidos, para além de criar uma agência reguladora, “foram a primeira nação a implantar um Ministério da Inteligência Artificial”¹⁹⁷.

Os pesquisadores mencionaram projeções de que até 2025 haverá a substituição de um terço da mão de obra por robôs, e que a diferença entre organizações pioneiras e retardatárias em adotar a IA, no fluxo de caixa (“benefício econômico obtido menos investimento associado e custos de transição”), será em torno de 145% até 2030. Diante desse cenário, a falta de uma política de Estado consistente para IA significa, segundo eles, colocar o Brasil “sob o risco de

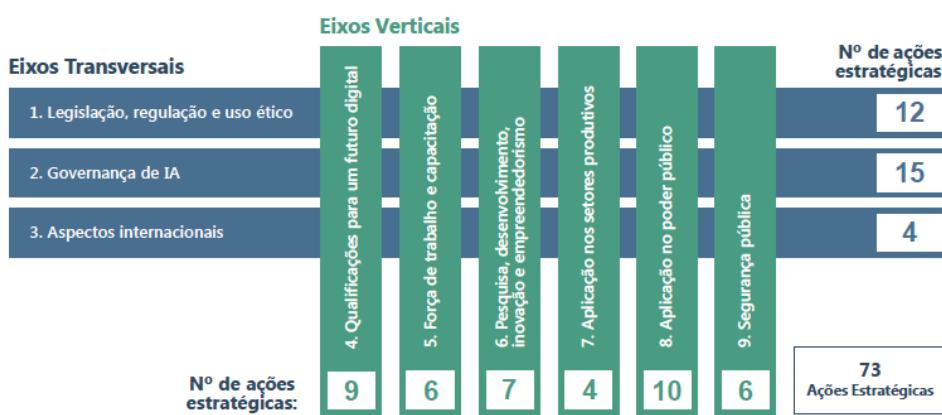
¹⁹⁶SGARIONI, Mariana. *Ibidem*.

¹⁹⁷CÓBE, Raphael M. O. et al. Rumo a uma política de Estado para inteligência artificial. **Revista USP**, n. 124, p. 37-48, 2020. Disponível em: <https://www.revistas.usp.br/revusp/article/view/167914/159995>. Acesso em: 27 maio 2022.

comprometer seriamente seu futuro científico-tecnológico, com importantes reflexos socioeconômicos, na segurança e na própria soberania nacional”¹⁹⁸.

Em 2021, no âmbito do Poder Executivo Federal, por meio do Ministério da Ciência, Tecnologia e Inovações (MCTI), a Portaria MCTI nº 4.167¹⁹⁹, de 06/04/2021, com o anexo explicativo publicado mais de três meses depois na Portaria MCTI nº 4.979²⁰⁰, de 13/07/2021, instituiu a Estratégia Brasileira de Inteligência Artificial (EBIA)²⁰¹ e seus eixos temáticos transversais e verticais, resumidos na Figura 3 a seguir:

Figura 3 – Eixos temáticos da Estratégia Brasileira de Inteligência Artificial



Fonte: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligencia-artificial>

A EBIA buscou se alinhar aos princípios da OCDE aos quais o Brasil aderiu em 2019, e foi precedida de consulta pública²⁰² *on-line* no período de 12/12/2019 a

¹⁹⁸Ibidem.

¹⁹⁹BRASIL. **Portaria MCTI nº 4.167, de 6 de abril de 2021.** Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-portaria_mcti_4-617_2021.pdf. Acesso em: 27 maio 2022.

²⁰⁰BRASIL. **Portaria MCTI nº 4.979, 13 de julho de 2021.** Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-portaria_mcti_4-979_2021_anexo1.pdf. Acesso em: 27 maio 2022.

²⁰¹BRASIL. **Estratégia Brasileira de Inteligência Artificial - EBIA.** Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-diagramacao_4-979_2021.pdf. Acesso em: 27 maio 2022.

²⁰²BRASIL. **Consulta pública – Estratégia Brasileira de Inteligência Artificial.** Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-consulta-publica.pdf>. Acesso em: 27 maio 2022.

03/03/2020, recebendo aproximadamente 1.000 contribuições²⁰³, segundo informado pelo Ministério da Ciência, Tecnologia e Inovações.

Contudo, o fato é que o texto principal da Portaria MCTI nº 4.167/2021 contém apenas 2 artigos, o primeiro laconicamente estabelecendo as finalidades de (i) nortear as ações do Estado e (ii) garantir a inovação no ambiente social e produtivo brasileiro nessa área de IA. O segundo artigo, por seu turno, longe de estabelecer medidas mais concretas, apenas prevê genericamente atribuições ao Ministério no sentido de coordenar ações, criar instâncias de governança, convidar instituições e a academia para colher subsídios, elaborar relatórios e dar-lhes publicidade, e atualizar a Estratégia.

Já o anexo da Portaria MCTI nº 4.979/2021 é um texto de apresentação da EBIA, explicativo da metodologia adotada, dos objetivos, do diagnóstico do ambiente de inovação brasileiro e dos 9 eixos temáticos, sendo 3 qualificados como transversais e 6 como verticais, já demonstrados na Figura 3.

Esse vazio de conteúdo em termos de diretrizes concretas fez com que a EBIA sofresse contundentes críticas de especialistas em tecnologia e inovação, a exemplo de Ronaldo Lemos, diretor do Instituto de Tecnologia e Sociedade do Rio de Janeiro, que ao escrever sobre o assunto em sua coluna na Folha de S. Paulo de 11/04/2021 qualificou a Estratégia como “patética”, parecendo um “trabalho de graduação universitário malfeito”, que sequer considerou ou tabulou as mais de mil contribuições recebidas, resultando o documento numa política de Estado que é um “desastre” e “mostra que o país está à deriva” nessa área²⁰⁴.

Pontuando que o documento parece uma mera carta de intenções, Walter Gaspar e Yasmin de Mendonça, em relatório de análise emitido pelo Centro de Tecnologia e Sociedade da FGV Direito Rio, apontaram que a EBIA é genérica, não define prazos nem metas, não abordou questões essenciais de planejamento, não apontou atores de governança, não estabeleceu “indicadores mensuráveis de referência”, não aproveitou as contribuições de especialistas coletadas na consulta pública, não aprofundou “métodos disponíveis para prover transparência e

²⁰³BRASIL. **Inteligência Artificial.** Ministério da Ciência, Tecnologia e Inovações. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligencia-artificial>. Acesso em: 27 maio. 2022.

²⁰⁴LEMOS, Ronaldo. **Estratégia de IA brasileira é patética.** Folha de S. Paulo. 11 abr. 2021. Disponível em: <https://www1.folha.uol.com.br/columnas/ronaldo-lemos/2021/04/estrategia-de-ia-brasileira-e-patetica.shtml>. Acesso em: 27 maio 2022.

explicabilidade aos sistemas de IA” e incorporou “de forma acrítica as pesquisas sobre o uso de IA na Segurança Pública”²⁰⁵.

Segundo tal relatório da FGV Direito Rio, entre muitos outros aspectos, continuam faltando “diretrizes técnicas e organizacionais para os problemas atrelados à implementação de sistemas de IA”, instrumentos regulatórios e de incentivo, bem como definições de “como se dará a participação da sociedade civil nos processos decisórios acerca do tema”²⁰⁶.

A partir dessa análise que os especialistas brasileiros vêm fazendo da EBIA, é possível dizer que o documento acabou incorrendo no mesmo problema do debate atual sobre IA, já apontado na introdução deste trabalho, quando foi citado artigo de Jessica Morley em coautoria com o filósofo italiano Luciano Floridi e outros, na qualidade de pesquisadores da Universidade de Oxford, quando disseram que a discussão está concentrada nos princípios – “o quê da ética da IA” – deixando de lado as práticas – “o como”, acrescentando que “a capacidade da comunidade de IA de agir para mitigar os riscos associados ainda está em sua infância” (tradução livre)²⁰⁷.

Por enquanto, da forma como está posta, a EBIA continua deixando o Brasil numa espécie de limbo em termos de regulação, métodos, mecanismos e ferramentas para uma política nacional séria e robusta na área de IA.

3.4 A Resolução nº 332/2020 do Conselho Nacional de Justiça

Como mencionado no capítulo anterior deste trabalho, de todas as principais diretrizes internacionais produzidas a respeito da IA nos últimos anos, a única que tratou especificamente da aplicação dessa tecnologia no âmbito dos tribunais foi a Carta Ética Europeia sobre o uso da Inteligência Artificial em sistemas judiciais e seu ambiente, adotada em Estrasburgo, na França, em dezembro de 2018,

²⁰⁵GASPAR, Walter Britto; MENDONÇA, Yasmin Curzi de. A Inteligência Artificial no Brasil ainda precisa de uma estratégia. FGV Direito Rio: CTS: **Papers e textos para discussões**. 2021. Disponível em: <https://bibliotecadigital.fgv.br/dspace;handle/10438/30500>. Acesso em: 27 maio 2022.

²⁰⁶Ibidem.

²⁰⁷MORLEY, Jessica et al. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. In: **Ethics, Governance, and Policies in Artificial Intelligence**. Springer, Cham, 2021. p. 153-183.

por ocasião da 31^a reunião plenária da Comissão Europeia para a Eficiência da Justiça do Conselho da Europa²⁰⁸.

Alguns meses depois, em maio de 2019, ocorreu o primeiro debate público a respeito do assunto por membros do Poder Judiciário no Brasil, até onde se sabe, quando foi apresentada uma versão da Carta Ética Europeia traduzida para o português²⁰⁹, no decorrer da programação científica do VII ENAJE (Encontro Nacional da Justiça Estadual), promovido pela AMB (Associação dos Magistrados Brasileiros) em Foz do Iguaçu, Paraná.

Participaram do painel sobre IA no Judiciário²¹⁰ dois então Conselheiros do Conselho Nacional de Justiça, Márcio Schiefler Fontes e Maria Tereza Uille, e dois então membros da Escola Nacional da Magistratura, os desembargadores da Justiça Estadual de Rondônia e do Espírito Santo, respectivamente Marcos Alaor Grangeia e Samuel Meira Brasil Júnior (este último, a propósito, magistrado brasileiro pioneiro na pesquisa da IA aplicada ao ramo jurídico, que tem estudos publicados internacionalmente desde 2001 sobre argumentação jurídica, raciocínio jurídico e IA²¹¹, é o orientador deste trabalho).

Segundo registro em vídeo²¹², foi a partir das conclusões do painel sobre IA no VII ENAJE que o desembargador Marcos Alaor Grangeia indicou ao CNJ a necessidade de regulação, tendo em vista que na época já existiam as primeiras aplicações de IA em tribunais brasileiros, a exemplo do sistema “Sinapses” em Rondônia.

Foi esse fato que levou o então presidente do CNJ, Ministro Dias Toffoli, no final de 2019, a criar um Grupo de Trabalho “destinado à elaboração de estudos e

²⁰⁸CEPEJ. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente.** Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 31 maio 2022.

²⁰⁹EMERON. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente.** Tradução livre de Teresa Germana Lopes de Azevedo. Disponível em: <https://emeron.tjro.jus.br/images/biblioteca/revistas/Avulso/CartaEticaEuropeia.pdf>. Acesso em: 31 maio 2022.

²¹⁰LANZELLOTTI, Andressa; GUEDES, Maria Rita. A Inteligência Artificial no Judiciário. **AMB Informa** – Especial VII ENAJE, p. 15. Disponível em: https://www.amb.com.br/wp-content/uploads/2019/06/Especial_VII_Enaje_PARA_PROVA-1-1.pdf. Acesso em: 31 maio 2022.

²¹¹BRASIL JR, Samuel Meira. Rules and Principles in Legal Reasoning. A Study of Vagueness and Collisions in Artificial Intelligence and Law. **Information & Communications Technology Law**, v. 10, n. 1, p. 67-77, 2001. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/13600830124910>. Acesso em: 31 maio 2022.

²¹²GRANGEIA, Marcos Alaor. Desafios éticos da inteligência artificial. **Judiciário Exponencial**. Youtube, 15 set. 2020. Disponível em: <https://www.youtube.com/watch?v=HwLGZ69R6cM&t=12s>. Acesso em: 31 maio 2022.

propostas voltadas à ética na produção e uso da inteligência artificial no Poder Judiciário”, por meio da Portaria nº 197, de 22/11/2019²¹³, composto de 9 integrantes, sendo 6 magistrados de primeiro e segundo graus oriundos da Justiça Federal, Justiça Estadual e Justiça do Trabalho, 1 membro da academia (professor da Universidade de Brasília), 1 consultor legislativo do Senado e 1 servidora pública lotada no CNJ.

Na época, existia no Poder Legislativo apenas um projeto de lei no Congresso Nacional sobre o tema, em estágio inicial no Senado, tratando-se do Projeto de Lei nº 5.051, de 2019, que pretendia estabelecer princípios para o uso da IA no Brasil, apresentado em setembro daquele ano pelo Senador Styvenson Valentim (PODEMOS/RN)²¹⁴. No Poder Executivo, ainda estava por iniciar a consulta pública para propostas voltadas à elaboração da Estratégia Nacional de Inteligência Artificial.

Em agosto de 2020, após a conclusão das atividades do Grupo de Trabalho instituído no CNJ, foi publicada a Resolução nº 332, de 21/08/2020, que “dispõe sobre a ética, a transparência e a governança na produção e no uso da Inteligência Artificial no Poder Judiciário e dá outras providências”²¹⁵.

Assim, pode-se afirmar que, dos três Poderes da República, **foi do Poder Judiciário que adveio o primeiro instrumento regulatório da IA no Brasil, a Resolução CNJ nº 332/2020**. O segundo é a EBIA, publicada pelo Poder Executivo em abril de 2021 (Portaria MCTI nº 4.167/2021). No momento, o Poder Legislativo ainda está em fase de elaboração do marco legal brasileiro da IA, com a tramitação conjunta, no Senado, do PL 21/2020 (que foi aprovado na Câmara dos Deputados em setembro de 2021), em apenso ao PL 5.051/2019 e ao PL 872/2021, como visto no item 3.2 deste capítulo.

Sobre a Resolução CNJ nº 332/2020, há no texto indicação expressa de que tomou como parâmetro a Carta Ética Europeia sobre o uso da Inteligência Artificial em sistemas judiciais e seu ambiente, com seus 5 princípios. Entretanto, enquanto a Carta Ética Europeia é um documento de conteúdo apenas principiológico, a Resolução da justiça brasileira buscou ir além, tratando de temas relacionados à

²¹³CNJ. Portaria nº 197, de 22 de novembro de 2019. Disponível em: <https://atos.cnj.jus.br/files/original190722201911265ddd77ea7b46e.pdf>. Acesso em: 31 maio 2022.

²¹⁴BRASIL. Projeto de lei nº 5.051, de 2019. Disponível em: <https://www.congressonacional.leg.br/materias/materias-bicamerais/-/ver/pl-5051-2019>. Acesso em: 31 maio 2022.

²¹⁵CNJ. Resolução nº 332, de 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 31 maio 2022.

governança e aplicação prática da IA na administração da Justiça. Ou seja, nessa seara específica buscou abordar **o como**, em vez de somente **o quê**.

Com efeito, a Resolução CNJ nº 332/2020 encontra-se estruturada em 10 capítulos, assim distribuídos²¹⁶: I. Das disposições gerais; II. Do respeito aos direitos fundamentais; III. Da não discriminação; IV. Da publicidade e transparência; V. Da governança e da qualidade; VI. Da segurança; VII. Do controle do usuário; VIII. Da pesquisa, do desenvolvimento e da implantação de serviços de inteligência artificial; IX. Da prestação de contas e da responsabilização; X. Das disposições finais.

A seguir, serão analisados os pontos-chave da Resolução, bem como seus dispositivos mais relevantes, sob a ótica dos focos de risco que pretendem regular, discutidos com base no estado da arte na pesquisa acadêmica sobre cada tema.

3.4.1 Os principais pilares da Resolução CNJ nº 332/2020

Nas disposições gerais, foram de logo estabelecidos 5 conceitos essenciais para a compreensão da Resolução: algoritmo, modelo de IA, Sinapses, usuário, usuário interno e usuário externo.

A definição de **algoritmo** é “sequência finita de instruções executadas por um programa de computador, com o objetivo de processar informações para um fim específico” (art. 3º, I). **Modelo de Inteligência Artificial** está definido como “conjunto de dados a partir de modelos matemáticos, cujo objetivo é oferecer resultados inteligentes, associados ou comparáveis a determinados aspectos do pensamento, do saber ou da atividade humana” (art. 3º, II)²¹⁷.

Sinapses é uma “solução computacional, mantida pelo Conselho Nacional de Justiça, com o objetivo de armazenar, testar, treinar, distribuir e auditar modelos de Inteligência Artificial” (art. 3º, III)²¹⁸.

Por **usuário**, entende-se a “pessoa que utiliza o sistema inteligente e que tem direito ao seu controle, conforme sua posição endógena ou exógena ao Poder Judiciário, pode ser um usuário interno ou um usuário externo” (art. 3º, IV)²¹⁹.

²¹⁶CNJ. Resolução nº 332, de 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 31 maio 2022.

²¹⁷Ibidem.

²¹⁸Ibidem.

²¹⁹Ibidem.

Já **usuário interno** é tido como “membro, servidor ou colaborador do Poder Judiciário que desenvolva ou utilize o sistema inteligente” (art. 3º, V), enquanto **usuário externo** é aquele não enquadrado na definição de usuário interno que “utiliza ou mantém qualquer espécie de contato com o sistema inteligente”. Essa última conceituação traz um rol exemplificativo de “jurisdicionados, advogados, defensores públicos, procuradores, membros do Ministério Público, peritos, assistentes técnicos, entre outros” (art. 3º, VI)²²⁰.

Dessas definições, é importante tecer algumas considerações sobre a plataforma **Sinapses**, para melhor entendimento, dada a sua importância hoje para a IA aplicada à administração da Justiça no Brasil. A origem é um projeto iniciado pelo Tribunal de Justiça de Rondônia, após 4 analistas participarem de treinamentos da *DataScience Academy* e demonstrarem em janeiro de 2018 um modelo de IA para classificar despachos, ao ler documentos e apresentar sugestões de aplicação, de acordo com as Tabelas Processuais Unificadas, resultando em economia de tempo e minimizando a possibilidade de ocorrer erro humano²²¹.

Em setembro de 2018, o CNJ fez uma visita técnica ao TJRO e logo no mês seguinte resolveu nacionalizar o sistema, por meio do Termo de Cooperação Técnica nº 42/2018. Os analistas de Rondônia passaram a trabalhar em Brasília, nas próprias dependências do CNJ, e desde então o modelo de IA que criaram foi evoluindo e se transformando no que é hoje, uma plataforma que atende “um contexto de governança nacional de modelos de IA”, com o objetivo de que “todos os sistemas do judiciário, que façam uso de modelos de IA, tenham o SINAPSES como um repositório para orquestração e versionamento”²²².

Conforme explicaram os analistas criadores do sistema, trata-se de uma “plataforma para desenvolvimento e disponibilização em larga escala de modelos de Inteligência Artificial, também comumente conhecido como ‘Fábrica de Modelos de IA’”²²³.

²²⁰CNJ. Resolução nº 332, de 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 31 maio 2022.

²²¹TJRO. Sinapses: termo de cooperação técnica com o CNJ completa 1 ano. Disponível em: <https://www.tjro.jus.br/noticias/item/11953-sinapses-termo-de-cooperacao-tecnica-com-o-cnj-completa-1-ano>. Acesso em: 3 jun. 2022.

²²²Ibidem.

²²³Ibidem.

Nesse contexto, a importância de se ter instituído, por Resolução, a plataforma Sinapses, reside no *status* de política pública formal que a ferramenta passou a ter.

Um de seus aspectos mais relevantes é o de servir como medida de combate à descontinuidade de projetos que é característica da Administração Pública judiciária brasileira, conforme apontado em artigo²²⁴ de João Carlos Murta Pereira e Marcus Vinícius Jardim Rodrigues, ao mencionarem que, na raiz do problema, está a regra pela qual a mesa diretora dos tribunais tem mandato de apenas 2 anos, vedada a reeleição, prevista no art. 102 da Lei Orgânica da Magistratura Nacional²²⁵.

Segundo tal artigo, o fato de haver um local para servir de repositório nacional não só garante a continuidade dos modelos de IA já criados nos tribunais, como ainda pode “obstar os efeitos negativos da descontinuidade, sobretudo o desperdício de recursos públicos no abandono ou na paralisação de antigos projetos em decorrência da troca de gestores no Poder Judiciário”²²⁶.

A propósito, no art. 22 da Resolução 332/2020 está expressamente previsto o dever dos tribunais não só de **comunicar o início** da pesquisa, desenvolvimento ou implantação de modelos de IA, como também de **velar por sua continuidade**.

Ao tratar da **governança**, o art. 10 da Resolução 332/2020 da mesma forma estabelece o dever dos tribunais de **informar ao CNJ** a pesquisa, o desenvolvimento, a implantação ou o uso da IA, além do dever de **depositar** os modelos de IA que criarem na plataforma Sinapses, e ainda o dever de **atuar em modelo comunitário**, inclusive sendo vedado desenvolver um projeto em paralelo a outro que já exista ou esteja em fase de desenvolvimento.

Aliás, a par do dever de continuidade, o modelo comunitário ou colaborativo de desenvolver projetos de IA no Judiciário é outro importante pilar da Resolução 332/2020, como medida de racionalização de recursos financeiros, que são limitados em orçamento público, e sobretudo de recursos humanos, tendo em vista que o déficit

²²⁴PEREIRA, João Carlos Murta; RODRIGUES, Marcus Vinícius Jardim. A plataforma Sinapses e a continuidade dos modelos de IA no judiciário. In: ARAÚJO, Valter Shuenquener de; GOMES, Marcus Lívio (Coord.); CANEN, Doris (Org.) **Inteligência Artificial e aplicabilidade prática no Direito**. CNJ. Brasília, 2022. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/04/inteligencia-artificial-e-a-aplicabilidade-pratica-web-2022-03-11.pdf>. Acesso em: 3 jun. 2022.

²²⁵BRASIL. **Lei Complementar nº 35, de 14 de março de 1979**. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp35.htm. Acesso em: 3 jun. 2022.

²²⁶PEREIRA, João Carlos Murta; RODRIGUES, Marcus Vinícius Jardim. Ibidem.

de profissionais na área de Tecnologia da Informação e Comunicação é uma realidade preocupante que se verifica também no Brasil.

De acordo com relatório divulgado em dezembro de 2021 pela BRASSCOM (Associação das Empresas de Tecnologia da Informação e Comunicação e de Tecnologias Digitais), “a oferta anual de 53 mil formandos em TIC não supre a demanda estimada em 159 mil talentos por ano”²²⁷, no cenário projetado até 2025, levando em conta o crescimento exponencial da necessidade de profissionais, por força da transformação digital que se acelerou em todos os setores com o advento da pandemia de Covid-19.

Como será explanado a seguir, a lacuna de profissionais de TI especializados em determinadas áreas por si só já representa uma fonte de risco a gerenciar.

3.4.2 Riscos relacionados a falhas na cibersegurança e proteção de dados

Na área de **cibersegurança** – intimamente ligada à IA – o déficit de pessoal especializado é ainda mais alarmante: segundo a ANATI (Associação Nacional dos Analistas em Tecnologia da Informação), a “falta de profissionais de TI é prelúdio para a catástrofe iminente de Cibersegurança”²²⁸.

A afirmação se baseou em relatório divulgado em 2021 pelo (ISC)² – *International Information System Security Certification Consortium*, um consórcio de instituições que se define líder mundial em segurança cibernética, criador de um programa de certificações que estabelece os padrões globais de melhores práticas na área²²⁹.

Analizando a força de trabalho em 14 países, o (ISC)² identificou no relatório que **o Brasil está na pior situação quando comparado aos demais**,

²²⁷BRASSCOM. **Demanda de talentos em TIC e estratégia Σ TECEM 2021**. São Paulo, dez. 2021. Disponível em: <https://brasscom.org.br/demanda-de-talentos-em-tic-e-estrategia-%cf%83-tcem/>. Acesso em: 4 jun. 2022.

²²⁸ALMEIDA, Romário César de. **Falta de profissionais de TI é prelúdio para a catástrofe iminente de Cibersegurança**. ANATI Notícias, 30 nov. 2021. Disponível em: <https://www.anati.org.br/noticias/falta-de-profissionais-de-ti-e-preludio-para-a-catastrofe-iminente-de-ciberseguranca>. Acesso em: 4 jun. 2022.

²²⁹(ISC)². **(ISC)²: The world's leading cybersecurity professional organization**. Disponível em: <https://www.isc2.org/about>. Acesso em: 4 jun. 2022.

havendo uma lacuna de 441 mil profissionais de segurança cibernética no país, conforme demonstra a Figura 3²³⁰.

Figura 4 - Lacuna global de profissionais de cibersegurança por país



Fonte: <https://www.isc2.org/-/media/ISC2/Research/2021/ISC2-Cybersecurity-Workforce-Study-2021.ashx>

Apenas para se ter uma ideia da dimensão dessa lacuna de 441.000 profissionais de segurança cibernética no Brasil, comparando com os profissionais responsáveis pela segurança e policiamento ostensivo do mundo real, que são os policiais militares, nos termos do art. 144, § 5º, da Constituição Federal²³¹, somando todo o efetivo da Polícia Militar de todas as Unidades da Federação, resulta em 406.384 profissionais, de acordo com o Anuário Brasileiro de Segurança Pública 2022²³².

²³⁰(ISC)². (ISC)² Cybersecurity workforce study 2021. Disponível em: <https://www.isc2.org/-/media/ISC2/Research/2021/ISC2-Cybersecurity-Workforce-Study-2021.ashx>. Acesso em: 1 jul. 2022.

²³¹BRASIL. Constituição da República Federativa do Brasil de 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 1 jul. 2022.

²³²FBSP. Anuário Brasileiro de Segurança Pública. 2022. p. 456. Disponível em: <https://forumseguranca.org.br/wp-content/uploads/2022/06/anuario-2022.pdf?v=4>. Acesso em: 1 jul. 2022.

Acontece que esse tema – cibersegurança – que já é vital para a governança da IA em todos os setores, no âmbito do Poder Judiciário exige redobrada cautela sob todos os prismas, dada a utilização massiva de dados para o treinamento dos modelos, muitos dos quais são dados sob segredo de justiça, bem como dados pessoais sensíveis, protegidos pela Lei nº 13.709/18, a Lei Geral de Proteção de Dados (LGPD)²³³, nos termos do art. 6º da Resolução 332/2020, que está inserido no capítulo II, que trata do respeito aos direitos fundamentais.

Além disso, em 10/02/2022 foi promulgada a Emenda Constitucional nº 115/2022, que incluiu a proteção de dados pessoais no rol dos direitos e garantias fundamentais, acrescentando o inciso LXXIX ao art. 5º da Constituição Federal²³⁴.

Dada a importância do tema, o capítulo VI da Resolução 332/2020 é dedicado à segurança, estabelecendo nos artigos 13 a 15 que os dados devem: (i) ser procedentes de fontes seguras, de preferência governamentais; (ii) ter cópia de segurança; (iii) ser protegidos de riscos, adotando-se providências para evitar que sejam destruídos, modificados, extraviados ou acessados e transmitidos sem autorização. O art. 16, quanto às práticas de armazenamento e execução dos modelos de IA, determina que estejam alinhadas a “padrões consolidados de segurança da informação”²³⁵.

Eis aqui uma seara em que a necessidade de gerenciamento de riscos assume especial relevo, principalmente levando-se em conta que os ataques a sistemas do Poder Judiciário vêm se intensificando nos últimos anos. No período de 18 meses compreendido entre novembro de 2020 e abril de 2022, foram contabilizados 13 ataques cibernéticos a diversos tribunais brasileiros, resultando numa média de 1 a cada 41 dias, conforme a Figura 5²³⁶:

²³³BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/L13709.htm. Acesso em: 4 jun. 2022.

²³⁴BRASIL. Emenda Constitucional nº 115, de 10 de fevereiro de 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc115.htm. Acesso em: 2 jul. 2022.

²³⁵CNJ. Resolução nº 332, de 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 4 jun. 2022.

²³⁶REINA, Eduardo. Em 18 meses, hackers violaram sistemas de tribunais no Brasil a cada 41 dias. Revista Consultor Jurídico, 15 abr. 2022. Disponível em: <https://www.conjur.com.br/2022-abr-15/onda-invasoes-hackers-estruturas-tecnologicas-tribunais>. Acesso em: 4 jun. 2022.

Figura 5 – Ataques cibernéticos a tribunais brasileiros

Ataques <i>hackers</i> aos tribunais brasileiros	
<i>De NOV/20 a ABR/22</i>	
nov/20	TRF-1
nov/20	STJ
nov/20	TSE
jan/21	TRF-3
jul/21	7 ^a Vara Criminal Federal SP
mai/21	STF
abr/21	TJ-RS
ago/21	TSE
out/21	TRT-RS
fev/22	TRT-ES
mar/22	TRF-3
abr/22	JF-PE

Fonte: <https://www.conjur.com.br/2022-abr-15/onda-invasoes-hackers-estruturas-tecnologicas-tribunais>

A investida ocorrida contra o Superior Tribunal de Justiça (STJ) em 03/11/2020 foi tida naquele momento como “o pior ataque cibernético já empreendido contra uma instituição pública brasileira, em termos de dimensão e complexidade”²³⁷.

Na época, pesquisadores da FGV Direito SP apontaram que “o ataque ao STJ é mais um grito de socorro da segurança cibernética no Brasil”, e que “os custos de prevenção, por maiores que possam parecer, são ínfimos frente aos danos que podem decorrer da paralisação desses sistemas”. Para os pesquisadores, ao se descuidar de estratégias de prevenção, “a marcha pela digitalização,

²³⁷STJ. Comunicado da presidência do STJ. 19 nov. 2020. Disponível em: <https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/19112020-Comunicado-da-Presidencia-do-STJ.aspx>. Acesso em: 4 jun. 2020.

desburocratização e eficiência dos serviços públicos só pode resultar em um abismo de vulnerabilidade”²³⁸.

Segundo o relatório *X-Force Threat Intelligence Index 2022*²³⁹, da IBM, o **Brasil é o país que mais sofreu ataques cibernéticos na América Latina em 2021**. No cenário global, o setor governamental figurou empatado em oitavo lugar com o setor de educação na lista das áreas mais visadas pelos *hackers*. Das 3 modalidades de investidas mais cometidas, a principal foi ataque de acesso ao servidor principal dos órgãos públicos, seguida de roubo de dados e fraude. Quanto ao vetor de ataque mais utilizado, consta no relatório que foi a exploração de vulnerabilidade.

Teria sido esse, aliás, o vetor do ataque ao STJ pelo *ransomware* RansomEXX, cuja sofisticação consistiu no fato de que, ao que tudo indica, os invasores visavam chegar até mesmo aos sistemas de cópias de segurança (*backup*), de modo a não restar absolutamente qualquer saída que não o pagamento de resgate, na linha do que parece ser “a nova tendência entre as gangues de *ransomware*”, conforme noticiado pelo CISO Advisor, website brasileiro especializado em cibersegurança²⁴⁰.

As investigações oficiais ainda não foram concluídas pela Polícia Federal, mas analisando as evidências disponíveis, especialistas publicaram o relatório “RansomEXX – Análise do Ransomware Utilizado no Ataque ao STJ”²⁴¹, apontando que essa modalidade de *ransomware*, da forma como se deu, foi 99% similar à de outro grande ataque cibernético ocorrido em 2020, atingindo o Departamento de Transporte do Texas (TxDOT), nos Estados Unidos.

Ao demonstrar, no citado relatório, os detalhes técnicos e a engenharia reversa do *ransomware* usado contra o STJ, os especialistas destacaram que a lição que poderia ser tirada do episódio é “a importância do uso da inteligência gerada a

²³⁸SALVADOR, João Pedro Favaretto; GUIMARÃES, Tatiane. **O ataque ao STJ é mais um grito de socorro da segurança cibernética no Brasil**. FVG Direito. 9 nov. 2020. Disponível em: <https://portal.fgv.br/artigos/ataque-ao-stj-e-mais-grito-socorro-seguranca-cibernetica-brasil>. Acesso em: 4 jun. 2022.

²³⁹IBM. **X-Force Threat Intelligence Index 2022**. IBM Security. Disponível em: <https://www.ibm.com/downloads/cas/ADLMLAZ>. Acesso em: 4 jun. 2022.

²⁴⁰TADEU, Erivelto. **Ataque ao STJ visava servidores de backup, nova tendência entre hackers**. CISO Advisor, 5 fev. 2021. Disponível em: <https://www.cisoadvisor.com.br/ataque-ao-stj-visou-servidores-de-backup-nova-tendencia-entre-hackers/>. Acesso em: 4 jun. 2022.

²⁴¹PALAZOLO, Gustavo; TEIXEIRA, Ialle; DUARTE, Felipe. **RansomEXX: Análise do Ransomware utilizado no ataque ao STJ**. Medium, 15 nov. 2020. Disponível em: <https://gustavopalazolo.medium.com/ransomexx-an%C3%A1lise-do-ransomware-utilizado-no-ataque-ao-stj-918001ec8195>. Acesso em: 4 jun. 2022.

partir de outros incidentes, como possibilidade e melhoria de novos mecanismos de proteção". A título de exemplo do que poderia ser feito concretamente, apontaram que, entre outras tecnologias, a aplicação de "uma simples regra via YARA"²⁴² poderia ser útil. YARA vem a ser uma ferramenta que, em tradução livre, é descrita como "o canivete suíço de correspondência de padrões para pesquisadores de *malware*"²⁴³.

A propósito, esse **aprendizado a partir de incidentes, extraindo deles informações úteis para o aprimoramento de sistemas, buscando evitar que ocorram novos incidentes**, é justamente a ideia central da solução que se busca estudar neste trabalho para a catalogação de erros e incidentes de IA nos tribunais, seguindo a linha proposta por Ulrich Beck de "precaução pela prevenção"²⁴⁴ para o gerenciamento de riscos, mencionada no início deste capítulo, no item 3.1. Parte-se da prevenção relacionada a riscos conhecidos para um horizonte mais amplo de precaução perante riscos incertos e desconhecidos.

Pelas projeções divulgadas, depreende-se que é preciso pensar seriamente sobre o assunto da cibersegurança transversal aos projetos de IA nos tribunais brasileiros, aperfeiçoando mecanismos de cumprimento da Resolução 332/2020 nesse ponto, porque o cenário é alarmante: o Relatório de Riscos Globais 2022, divulgado pelo Fórum Econômico Mundial, informa quanto a ataques cibernéticos que houve aumento de 358% de *malware* e de 435% de *ransomware*, e que a tendência é de ainda mais crescimento²⁴⁵.

Como a pandemia de Covid-19 transformou as sociedades ao redor do mundo, tornando-as cada vez mais dependentes de sistemas digitais, aumentando o uso de dispositivos conectados à Internet, fala-se que está em curso outra pandemia, só que cibernética, porque os métodos dos *hackers* estão mais sofisticados e agressivos, além de haver déficit global de profissionais de cibersegurança, e ainda insuficiência de investimentos e regras eficazes de governança sobre o assunto²⁴⁶.

O relatório do Fórum Econômico Mundial prevê ataques massivos, sistemáticos e estratégicos, e que essa falta de segurança digital também poderá

²⁴²PALAZOLO, Gustavo; TEIXEIRA, Ialle; DUARTE, Felipe. Ibidem.

²⁴³YARA. **Yara: the pattern matching swiss knife for malware researchers (and everyone else).** Disponível em: <https://virustotal.github.io/yara/>. Acesso em: 5 jun. 2022.

²⁴⁴BECK, Ulrich. Op.cit., p. 364.

²⁴⁵WEF. **The Global Risks Report 2022, 17th edition.** World Economic Forum. Disponível em: https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2022.pdf. Acesso em: 5 jun. 2022.

²⁴⁶Ibidem.

acarretar riscos ditos intangíveis, como a possibilidade de **afetar a confiança da população em sistemas tecnológicos**²⁴⁷.

Nesse aspecto em particular, pelo que se depreende do relatório, é possível afirmar também em relação ao Poder Judiciário brasileiro que, do ponto de vista da Ciência Política, é muito importante conquistar e manter a confiança pública nas novas tecnologias, que são cada vez mais utilizadas para a administração da Justiça em tempos de Juízo 100% Digital e Justiça 4.0, sobretudo para legitimá-las.

Cabe aqui lembrar que, no Estudo da Imagem do Judiciário Brasileiro divulgado em 2019, de modo geral o atributo mais importante, indicado pela população, para a atuação considerada ideal dos membros do Judiciário, foi justamente a **confiança** (42%), à frente de atributos como imparcialidade (31%) e transparência (23%)²⁴⁸.

Ataques cibernéticos como o ocorrido contra o STJ e outros tribunais, amplamente divulgados na imprensa, dando ensejo a manchetes como “A impunidade dos hackers que colocaram o Judiciário de joelhos”²⁴⁹, dificultam sobremaneira a conquista e manutenção da confiança da população brasileira na Justiça Digital.

Como disse o professor Paulo Alexandre Batista de Castro, do Instituto Direito Público (IDP), “não apenas a força do Judiciário, mas também a estabilidade da própria democracia depende, em alguma medida, da confiança dos indivíduos no sistema judicial”²⁵⁰. O professor sustenta que estudar a relação entre confiança pública e legitimidade do Poder Judiciário ainda é uma agenda de pesquisa que está pendente no Brasil. Nada foi dito a respeito da confiança perante um sistema judicial cada vez mais digital, com uso de novas tecnologias como a IA, o que indica que a lacuna de pesquisa sob esse prisma possui contornos ainda mais largos.

Ainda sobre a questão da segurança cibernética do sistema judicial brasileiro – tópico que está sendo analisado mais detidamente por tratar de uma das

²⁴⁷Ibidem.

²⁴⁸AMB; FGV; IPESPE. **Estudo da Imagem do Judiciário Brasileiro**. Dez. 2019. Disponível em: https://www.amb.com.br/wp-content/uploads/2020/04/ESTUDO_DA_IMAGEM_DO_JUDICIARIO_BRASILEIRO_COMPLETO.pdf. Acesso em: 5 jun. 2022.

²⁴⁹MOURA, Rafael Moraes; BORGES, Laryssa. **A impunidade dos hackers que colocaram o Judiciário de joelhos**. Veja, 27 mar. 2022. Disponível em: <https://veja.abril.com.br/politica/a-impunidade-dos-hackers-que-colocaram-o-judiciario-de-joelhos/>. Acesso em: 5 jun. 2022.

²⁵⁰CASTRO, Paulo Alexandre Batista de. Confiança pública e legitimidade do Poder Judiciário: a agenda de pesquisa pendente no Brasil. **Revista da CGU**, v. 11, p. 1284-1292, 2019. Disponível em: https://repositorio.cgu.gov.br/bitstream/1/44358/11/V11.N20_Confian%a7a%20p%c3%bablica.pdf. Acesso em: 5 jun. 2022.

maiores áreas de risco, não sendo os ataques uma questão de **se** mas de **quando** ocorrerão, como apontam todas as evidências – cabe observar que o CNJ somente passou a enfrentar o tema, de modo mais consistente, após a investida contra o STJ ocorrida em 03/11/2020.

Com efeito, alguns dias depois do ataque, o CNJ instituiu o Comitê de Segurança Cibernética do Poder Judiciário, por meio da Portaria nº 242, de 10/11/2020²⁵¹, revogando a antiga norma até então existente, que era uma portaria de 2013 que havia instituído o Comitê Gestor de Segurança da Informação (CGSI) do CNJ²⁵².

Em dezembro de 2020, o CNJ instituiu 3 protocolos: o primeiro para gerenciamento de crises cibernéticas, o segundo para prevenção a incidentes cibernéticos e o terceiro para investigação de ilícitos cibernéticos no âmbito do Poder Judiciário, respectivamente pelas Resoluções nº 360²⁵³, nº 361²⁵⁴ e nº 362²⁵⁵, todas do dia 17/12/2020.

Posteriormente, esses atos normativos foram expressamente revogados pela Resolução nº 396, de 07/06/2021, que instituiu a **Estratégia Nacional de Segurança Cibernética do Poder Judiciário (ENSEC-PJ)**, atualmente em vigor²⁵⁶.

Quanto à composição atual do **Comitê Gestor de Segurança da Informação do Poder Judiciário (CGSI-PJ)**, verifica-se pelo ato normativo que está vigente, a Portaria nº 46, de 10/02/2022²⁵⁷, que há 13 especialistas, oriundos do CNJ, de todos os tribunais superiores, dos Conselhos da Justiça Federal e do Trabalho, e do TJRJ, TJDFT e TJPE, sob a coordenação de um juiz auxiliar da Presidência do CNJ.

O que se percebe dessa profusão de atos normativos em curto espaço de tempo é que têm sido muitas idas e vindas, com mudanças e retornos de

²⁵¹CNJ. **Portaria nº 242, de 10 de novembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3566>. Acesso em: 5 jun. 2022.

²⁵²CNJ. **Portaria nº 112, de 11 de julho de 2013.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/1806>. Acesso em: 5 jun. 2022.

²⁵³CNJ. **Resolução nº 360, de 17 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3633>. Acesso em: 5 jun. 2022.

²⁵⁴CNJ. **Resolução nº 361, de 17 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3634>. Acesso em: 5 jun. 2022.

²⁵⁵CNJ. **Resolução nº 362, de 17 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3635>. Acesso em: 5 jun. 2022.

²⁵⁶CNJ. **Resolução nº 396, de 7 de junho de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3975>. Acesso em: 5 jun. 2022.

²⁵⁷CNJ. **Portaria nº 46, de 10 de fevereiro de 2022.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4366>. Acesso em: 5 jun. 2022.

nomenclatura de órgãos e tentativas de aperfeiçoamento da governança em segurança cibernética. Já é um avanço haver essa preocupação mais organizada e sistematizada, mas o futuro dirá se as medidas adotadas estão de fato se traduzindo em maior **resiliência cibernética** do Judiciário brasileiro.

A propósito, um relatório produzido pelo Fórum Econômico Mundial em parceria com a Universidade de Oxford em 2020, *Future Series: cybersecurity, emerging technology and systemic risk*, mencionou que as **melhores práticas de governança de riscos cibernéticos associados à IA** envolvem *design* seguro do modelo de IA, gerenciamento do seu ciclo de vida e gerenciamento dos incidentes verificados²⁵⁸.

Quanto ao **design seguro**, entendeu-se que os sistemas precisam resistir a problemas técnicos, investidas e manipulações, o que também exige que sejam explicáveis, porque só assim sua integridade pode ser verificada.

No tocante ao **gerenciamento do ciclo de vida**, foi dito que deve haver uma verificação rigorosa, constante e dinâmica dos algoritmos, principalmente quando novas versões forem desenvolvidas com uso de novos dados. Ou seja, é preciso fazer o controle de versões em sistemas de IA dinâmicos.

Com relação ao **gerenciamento de incidentes**, o relatório aponta que os desenvolvedores devem ser capazes de identificar a manipulação de algoritmos, ter um plano de ação para mitigar eventuais danos e conseguir restabelecer a integridade algorítmica do sistema.

Já a organização ISACA²⁵⁹, que congrega 150 mil profissionais em 188 países, sendo uma associação internacional dedicada à pesquisa e ao desenvolvimento de métodos de auditoria, controle e certificação para sistemas de informação, divulgou em 2021 um estudo sobre *Cyberrisk Quantification (CRQ)*²⁶⁰, tido como importante ferramenta para o gerenciamento moderno de risco cibernético.

Trata-se de um painel ou relatório de **Quantificação de Risco Cibernético**, que segundo o estudo funciona como “a pedra de Roseta”²⁶¹ porque serve para

²⁵⁸CREESE, Sadie et al. **Future Series: cybersecurity, emerging technology and systemic risk**. In: World Economic Forum. 2020. Disponível em: https://www3.weforum.org/docs/WEF_Future_Series_Cybersecurity_emerging_technology_and_systemic_risk_2020.pdf. Acesso em: 7 jun. 2022.

²⁵⁹ISACA. **About us**. Disponível em: <https://www.isaca.org/why-isaca/about-us>. Acesso em: 24 jun. 2022.

²⁶⁰ISACA. **Cyberrisk quantification**. 2021. Disponível em: <https://store.isaca.org/s/store#/store/browse/detail/a2S4w000004Ko4jEAC>. Acesso em: 24 jun. 2022.

²⁶¹Ibidem.

traduzir, aos gestores de organizações públicas e privadas, a dimensão dos riscos cibernéticos relatados pelos profissionais de TI especializados em cibersegurança, de modo que ambos os setores (de negócio e de TI) possam trabalhar juntos para identificar as áreas críticas a serem objeto de intervenção.

Usando números para medir o risco cibernético, o CRQ “reflete a aplicação de métodos estatísticos rigorosos para quantificar o impacto e a frequência de incidentes de segurança cibernética”²⁶².

De acordo com o estudo da ISACA, o CRQ não substitui outras ferramentas tradicionalmente utilizadas, mas é um importante complemento, uma abordagem mais atual, pois prevê constantes mapeamentos dos prováveis cenários de risco, do ator, do tipo de controle (controle de acesso ou vulnerabilidades) e do impacto na organização, em termos de interrupção de serviços, prejuízos financeiros ou até mesmo risco existencial, que é o impacto que chega a inviabilizar a existência e continuidade da própria organização em si.

Aliás, como exemplo de um caso real desse tipo de impacto de grau máximo para uma organização, pode-se mencionar o fechamento do Lincoln College, uma universidade do Estado norte-americano de Illinois, que tinha 157 anos de existência e fechou as portas em maio de 2022, anunciando que resistiu a eventos como “a gripe espanhola de 1918, a Grande Depressão, a Segunda Guerra Mundial, a crise financeira global de 2008”, mas não ao efeitos de um ataque cibernético que sofreu em dezembro de 2021, deixando inoperantes todos os seus sistemas informatizados e tornando irremediável a dificuldade financeira da instituição, que tentava se recuperar dos efeitos da pandemia de Covid-19²⁶³.

Nem mesmo as próprias empresas de segurança cibernética, até as que são consideradas as maiores do mundo, estão totalmente à prova de ataques, a exemplo do que ocorreu em dezembro de 2020 com a norte-americana FireEye. Segundo apurado pelo jornal britânico The Guardian, o fato “está entre as violações mais significativas da memória recente”, por ser oriundo “de uma nação com capacidades ofensivas de primeira linha”, a Rússia²⁶⁴.

²⁶²Ibidem.

²⁶³LINCOLN COLLEGE. **Abraham Lincoln's namesake college set to close after 157 years.** Disponível em: <https://lincolncollege.edu/home>. Acesso em: 25 jun. 2022.

²⁶⁴THE GUARDIAN. **US cybersecurity firm FireEye says it was hacked by foreign government.** Guardian staff and agencies, 9 dez. 2020. Disponível em: <https://www.theguardian.com/technology/2020/dec/08/fireeye-hack-cybersecurity-theft>. Acesso em: 25 jun. 2022.

Consta que os *hackers* subtraíram ferramentas usadas pelas empresa para testar a defesa dos sistemas de seus clientes, “que incluem governos federais, estaduais, municipais e grandes corporações globais”, e que além das ferramentas os invasores “pareciam interessados em um subconjunto de clientes da FireEye: agências governamentais”²⁶⁵.

A empresa FireEye, por sua vez, em comunicado oficial²⁶⁶ forneceu os detalhes do ataque e informou as providências adotadas, que consistiram em 300 contramedidas para minimizar o impacto da subtração de suas ferramentas de segurança cibernética pelos *hackers*.

Apenas uma semana depois desse grande ataque à FireEye, diversos órgãos importantes do governo norte-americano foram atacados, como o Tesouro e os departamentos de segurança interna, estado, defesa e comércio. O cientista-chefe da empresa de segurança cibernética McAfee informou haver evidências de que as ferramentas subtraídas já haviam sido utilizadas em 19 países, incluindo Estados Unidos, Reino Unido, Irlanda, Holanda, Austrália, entre outros, sendo noticiado que as investigações levariam meses, e os impactos podem levar anos para serem percebidos²⁶⁷.

Esses fatos recentes dão a dimensão da relevância do tema da segurança cibernética atualmente, para qualquer organização, principalmente governamental. Até agora os tribunais brasileiros estão conseguindo reagir aos ataques cibernéticos que vêm sofrendo, sendo de vital importância intensificar as medidas de cautela, porque os métodos dos *hackers* estão cada vez mais sofisticados.

É imensurável o prejuízo que haverá para a população e para a própria democracia brasileira se um ataque cibernético conseguir alcançar grau máximo de impacto para o Poder Judiciário, caso uma de suas unidades venha a sofrer impedimento total e irremediável de acesso a todos os seus sistemas e processos. Mesmo ataques em menor grau podem acarretar danos significativos.

Quando a Resolução 332/2020 foi publicada, ainda não havia acontecido a invasão aos sistemas do STJ, bem como a divulgação da série de ataques frequentes

²⁶⁵Ibidem.

²⁶⁶MANDIA, Kevin. **FireEye shares details of recent cyber attack, actions to protect community.** FireEye Stories Blog, 8 dez. 2020. Disponível em: <https://www.fireeye.com/blog/products-and-services/2020/12/fireeye-shares-details-of-recent-cyber-attack-actions-to-protect-community.html>.

Acesso em: 25 jun. 2022.

²⁶⁷BBC News. **SolarWinds Orion: more US government agencies hacked.** 15 dez. 2020. Disponível em: <https://www.bbc.com/news/technology-55318815>. Acesso em: 25 jun. 2022.

a outros tribunais brasileiros, que agora possibilita a percepção de frequência média de 1 a cada 41 dias desde então.

Diante disso, para os modelos de IA em desenvolvimento e produção nos tribunais, talvez seja preciso repensar se são suficientes as medidas de segurança que são aplicadas atualmente para proteger os dados utilizados.

Cabe notar que na Estratégia Nacional de Segurança Cibernética do Poder Judiciário (ENSEC-PJ) não foi feita qualquer menção à Resolução 332/2020. Em nenhum momento do texto da Estratégia, instituída pela Resolução nº 396/2021²⁶⁸, sequer aparece a expressão “inteligência artificial”.

Da mesma forma, também não aparece menção à Resolução 332/2020 e à expressão “inteligência artificial” na Resolução nº 370/2021²⁶⁹, que estabeleceu a ENTIC-JUD, a Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário. Nada foi dito em relação a métodos de quantificação de risco cibernético em geral, e muito menos do risco envolvendo especificamente modelos de IA em desenvolvimento e produção nos tribunais.

Esse cenário de lacunas em ambas as Estratégias permite perceber quão incipiente é o enfrentamento do tema no Poder Judiciário brasileiro, depreendendo-se a utilidade de ser instituída uma ferramenta como o CRQ, ou seja, o Relatório de Quantificação de Risco Cibernético nos moldes técnicos propostos pela ISACA.

Isso porque um relatório dessa natureza, com parâmetros estatísticos, elaborado periodicamente, poderá ser uma importante medida não só para o gerenciamento do risco cibernético no Judiciário, como também para colaborar na continuidade das políticas e estratégias de enfrentamento ao risco, de modo que sobrevivam às mudanças de gestão nos tribunais a cada dois anos.

Para os fins deste trabalho, a parte que se destaca do estudo sobre *Cyberrisk Quantification* (CRQ) da ISACA, por vir ao encontro da hipótese desta pesquisa, é uma das medidas que auxiliam as etapas do mapeamento de risco cibernético: a formação de um banco de dados para documentar eventos adversos e incidentes, reunindo informações que “são inestimáveis para os analistas de risco”²⁷⁰.

²⁶⁸CNJ. Resolução nº 396, de 7 de junho de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3975>. Acesso em: 25 jun. 2022.

²⁶⁹CNJ. Resolução nº 370, de 28 de janeiro de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3706>. Acesso em: 25 jun. 2022.

²⁷⁰ISACA. *Cyberrisk quantification*. 2021. Disponível em: <https://store.isaca.org/s/store#/store/browse/detail/a2S4w000004Ko4jEAC>. Acesso em: 25 jun. 2022.

Adicionalmente, a tendência atual é de que a IA irá “permear todos os aspectos da cibersegurança”, de modo que suas diversas técnicas poderão ser utilizadas **não só em ferramentas de ataque, como também de defesa**, ajudando a tentar equilibrar a luta que hoje está desigual, segundo o professor Bruce Schneier, da Universidade de Harvard. Ele foi o especialista ouvido pelo Fórum Econômico Mundial para opinar na conclusão do relatório *Global Cybersecurity Outlook 2022*, sobre como uma organização pode construir sua resiliência cibernética, à luz da constatação de que “automação e aprendizado de máquina irão introduzir a maior transformação em segurança cibernética no futuro a curto prazo”²⁷¹.

Analizado o capítulo da Resolução 332/2020 relacionado à segurança e proteção de dados, junto com as tendências atuais para controle de riscos cibernéticos, cabe abordar outro tema da máxima importância, que, se não for bem trabalhado no *design* de modelos de IA, também pode ser fonte de risco de acarretar danos, não só aos órgãos do Poder Judiciário, sua imagem, credibilidade e patrimônio, mas sobretudo às pessoas jurisdicionadas: o respeito aos direitos fundamentais, principalmente no que se refere à igualdade.

3.4.3 Riscos de modelos de IA discriminatórios

No Capítulo III da Resolução, houve um especial cuidado com a **não discriminação** (art. 7º), impondo desde o *design* do modelo de IA a cautela de verificar vieses discriminatórios, preconceitos e generalizações que eventualmente pudessem acarretar opressão e marginalização de pessoas, ou ainda erros de julgamento.

Nesse particular aspecto, já foi dito aqui que a Resolução 332/2020 teve como parâmetro a Carta Ética Europeia, que por sua vez foi analisada em detalhes no capítulo anterior deste trabalho, inclusive quanto ao fato de que os estudos de base para o documento europeu levaram em conta a percepção negativa de algoritmo usado nos Estados Unidos, no âmbito da justiça criminal: a ferramenta de avaliação de risco de reincidência denominada COMPAS – acrônimo de *Corretional Offender Management Profiling for Alternative Sanctions*. Trata-se de um modelo de IA que

²⁷¹WEF. **Global Cybersecurity Outlook 2022: insight report.** Jan. 2022. Disponível em: https://www3.weforum.org/docs/WEF_Global_Cybersecurity_Outlook_2022.pdf. Acesso em: 6 jun. 2022.

verificava o perfil de réus, tendo sido apontado que valorava negativamente os negros, como tendo maior propensão do que os brancos a reincidir, segundo a análise de Kehl e Kessler²⁷².

Em 2020, quando a Resolução 332/2020 foi elaborada, muito já havia sido escrito e pesquisado sobre o COMPAS, que talvez seja o exemplo mundialmente mais conhecido dos riscos associados à IA na administração da Justiça, fomentando as discussões sobre vieses algorítmicos discriminatórios.

Entre as estratégias que podiam ser adotadas para mitigação de vieses discriminatórios, pesquisas apontavam que a diversidade da força de trabalho, nas equipes envolvidas com o desenvolvimento de modelos de IA, desde a fase de ideação, poderiam colaborar “para garantir que preconceitos e estereótipos culturais sejam abordados antecipadamente e de forma adequada”²⁷³, como mencionado em relatório de 2019, elaborado por Lee et al., do instituto norte-americano de pesquisa Brookings, de Washington, D.C.

Percebe-se que esse tipo de diretriz foi encampada pela Resolução 332/2020, visto que o art. 20 estabelece, quanto à composição das equipes envolvidas em pesquisa, desenvolvimento e implantação de modelos de IA na Justiça brasileira, que deverá haver **diversidade**, tida “em seu mais amplo espectro, incluindo gênero, raça, etnia, cor, orientação sexual, pessoas com deficiência, geração e demais características individuais”²⁷⁴.

Nesse esforço de evitar vieses discriminatórios, o dispositivo mencionado prevê, no § 1º, que deve haver diversidade nas equipes tanto na área técnica, que é o pessoal de TI, como negocial, ou seja, magistrados, servidores e outros atores envolvidos em todo o ciclo de desenvolvimento do modelo de IA.

Inclusive se estabeleceu no § 4º que as equipes devem ser interdisciplinares, o que significa que, quando recomendável a depender do modelo de IA, deve-se envolver pessoal de outros campos do conhecimento científico. A título

²⁷²KEHL, Danielle Leah; KESSLER, Samuel Ari. Algorithms in the criminal justice system: Assessing the use of risk assessments in sentencing. **Responsive Communities Initiative, Berkman Klein Center for Internet & Society**, Harvard Law School. 2017. Disponível em: <https://dash.harvard.edu/handle/1/33746041>. Acesso em: 7 jun. 2022.

²⁷³LEE, Nicol Turner; RESNICK, Paul; BARTON, Genie. **Algorithmic bias detection and mitigation: best practices and policies to reduce consumer harms**. Brookings. 22 maio 2019. Disponível em: <https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/>. Acesso em: 7 jun. 2022.

²⁷⁴CNJ. **Resolução nº 332, de 21 de agosto de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 7 jun. 2022.

de exemplo, podem ser mencionados profissionais das áreas de Psicologia, Sociologia, Filosofia, Antropologia, Economia, Ciência Política, Ciências Contábeis e muitos outros.

A Resolução 332/2020 também se preocupou em disciplinar, no § 3º do art. 20, que a distribuição de vagas para capacitação em IA, nos tribunais, deve observar o critério de diversidade.

É possível observar que todas essas diretrizes estão de acordo com os princípios do chamado “design eticamente alinhado” (*Ethically Aligned Design*) desenvolvidos pelo IEEE, como analisado no capítulo anterior deste trabalho, no item 3.3, cuja pergunta de partida vale repetir agora (em tradução livre): “quem deveria estar neste espaço para garantir que o Sistema de IA que criamos é o mais seguro e responsável possível?”²⁷⁵.

Vale também repetir que, em tais princípios do IEEE, busca-se a composição paritária entre homens e mulheres, multidisciplinaridade, variedade de localização geográfica e sobretudo a inclusão da perspectiva de indígenas e povos originários, bem como de outros grupos da população que são afetados pela aplicação da IA ou poderiam dela de beneficiar, porém comumente ficam deixados à margem do processo de design.

Acontece que, na Resolução 332/2020, ficou faltando abordar de modo mais efetivo essa perspectiva de grupos da sociedade minoritários ou marginalizados, como será visto mais adiante nesta seção, sendo então um ponto da Resolução a aperfeiçoar.

Para compreensão do problema, outro aspecto importante a destacar é que, pelo § 2º do art. 20 da Resolução 332/2020, somente se permite a **dispensa de diversidade** nas equipes por decisão fundamentada, entre outras razões, **por falta de força de trabalho** especializada nos quadros dos tribunais.

Acontece que, como já foi dito antes, com demonstração de números, a lacuna de profissionais de TI é uma realidade no Brasil, que tende a ficar em dramática situação nos próximos anos, com a crescente digitalização de todos os setores e atividades, públicos e privados.

²⁷⁵IEEE. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems Key Information, Milestones, and FAQs about The Initiative. Disponível em: <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/faqs/gieais-faq-11.22.2020.pdf>. Acesso em: 9 jun. 2022.

Diante disso, é possível perceber que a Resolução 332/2020 até buscou ser avançada, estando de acordo com o estado da arte na pesquisa científica sobre o tema, mas a realidade se impõe e requer uma **atitude proativa dos gestores dos tribunais** para garantir **real diversidade** nas equipes de desenvolvimento de IA.

Do contrário, será letra morta a exigência normativa de diversidade, o que resultará em permanente flanco aberto para o risco de desenvolvimento de modelos de IA com vieses discriminatórios, com potencial de gerar danos a jurisdicionados, ferindo seus direitos humanos.

Disso decorre a possibilidade de acarretar responsabilização civil do Estado, além do potencial prejuízo a um bem intangível da máxima importância, que é a confiança da sociedade nas novas tecnologias adotadas pelo Poder Judiciário, como já foi visto linhas antes. É preciso cuidar para que não venha a ser apontado, no Poder Judiciário brasileiro, um caso como o do norte-americano COMPAS.

A tarefa dos gestores dos tribunais para obter diversidade nas equipes de TI não será fácil. Para começar, faltam dados constantemente coletados e traduzidos em necessidades atuais e projeções consistentes, com foco específico nesse problema. Na pesquisa exploratória realizada para abordagem do tema neste trabalho, não foi encontrado qualquer estudo sobre quão diversas são as equipes responsáveis pelo desenvolvimento de modelos de IA no Poder Judiciário brasileiro.

No entanto, verificou-se um relatório pioneiro divulgado em abril de 2022 pelo Tribunal de Justiça do Distrito Federal e Territórios (TJDFT), “Relatório diagnóstico organizacional de equidade e inclusão da diversidade”²⁷⁶, que examinou os marcadores sociais de gênero, raça, pessoa com deficiência e idade nos quadros de pessoal do referido tribunal.

O documento foi elaborado por uma Comissão Multidisciplinar de Inclusão, presidida pelo juiz Fabio Esteves, que vem a ser uma destacada liderança negra da magistratura brasileira, cofundador de duas iniciativas criadas em 2017, o Fórum Nacional de Juízas e Juízes contra o Racismo e Todas as Formas de Discriminação

²⁷⁶TJDFT. **Diagnóstico organizacional de equidade e inclusão da diversidade**. Disponível em: https://www.tjdft.jus.br/institucional/governanca/integridade-e-etica/programa-pro-equidade-e-diversidade/relatorio_diversidade.pdf. Acesso em: 10 jun. 2022.

(FONAJURD) e o Encontro Nacional de Juízes e Juízas Negros (ENAJUN) – este último, vencedor do Prêmio Desafio Lideranças Públicas Negras 2021²⁷⁷.

A análise da Comissão Multidisciplinar de Inclusão do TJDFT foi feita sob a perspectiva dos Objetivos de Desenvolvimento Sustentável nº 4 (Educação de Qualidade), nº 8 (Trabalho Decente) e nº 10 (Redução das Desigualdades) da Agenda 2030 da ONU²⁷⁸, no âmbito do Programa Pró-Equidade e Diversidade do TJDFT, alinhado à Resolução CNJ nº 325, de 29/06/2020²⁷⁹, que instituiu a Estratégia Nacional do Poder Judiciário 2021-2026. Entre outros valores, buscou “atenuar as desigualdades sociais, garantir os direitos de minorias e a inclusão e acessibilidade a todos”²⁸⁰.

Não obstante a importância e o ineditismo do trabalho, com metodologia que incluiu a coleta de dados com formação de grupos focais inclusivos, de modo a lançar luzes sobre temas tão complexos e multifacetados para resultar em recomendações de ações práticas, fato é que o relatório analisou o quadro de pessoal do TJDFT numa visão geral e mais voltada ao clima organizacional.

Por essa razão, pode-se afirmar que o estudo tem inegável valor para nortear o debate ainda tão incipiente no Poder Judiciário, mas não chegou a adentrar especificamente no tema da diversidade das equipes de TI desenvolvedoras de modelos de IA no âmbito do tribunal analisado, como providência cautelar de controle de risco de viés algorítmico discriminatório.

A respeito da força de trabalho de Tecnologia da Informação no setor público brasileiro, o estudo mais completo encontrado foi o “Levantamento de Pessoal de TI”²⁸¹, realizado pelo Tribunal de Contas da União (TCU), quanto aos órgãos federais dos três Poderes da República, incluindo o Conselho Nacional de Justiça.

²⁷⁷TJDFT. **Juiz do TJDFT vence Prêmio Desafio Lideranças Públicas Negras.** Disponível em: <https://www.tjdft.jus.br/institucional/imprensa/noticias/2021/fevereiro/juiz-do-tjdft-vence-premio-desafio-liderancas-publicas-negras>. Acesso em: 10 jun. 2022.

²⁷⁸ONU. **Os objetivos de desenvolvimento sustentável no Brasil.** Nações Unidas - Brasil. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 10 jun. 2022.

²⁷⁹CNJ. **Resolução nº 325, de 29 de junho de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3365>. Acesso em: 10 jun. 2022.

²⁸⁰TJDFT. **Diagnóstico organizacional de equidade e inclusão da diversidade.** Disponível em: https://www.tjdft.jus.br/institucional/governanca/integridade-e-etica/programa-pro-equidade-e-diversidade/relatorio_diversidade.pdf. Acesso em: 10 jun. 2022.

²⁸¹BRASIL. **Levantamento de Pessoal de TI/Tribunal de Contas da União;** Relator Ministro Raimundo Carreiro. – Brasília: TCU, Secretaria de Fiscalização de Tecnologia da Informação (Sefti), 2015. 52 p. – (Sumário Executivo. Tecnologia da Informação). Disponível em: <https://portal.tcu.gov.br/biblioteca-digital/levantamento-de-pessoal-de-ti-sumario.htm>. Acesso em: 10 jun. 2022.

Porém, tal levantamento foi divulgado em ano que já está longínquo, 2015, com base em dados relativos a um período ainda mais distante, 2012 – portanto, em cenário muito diferente do atual, que é de acelerada transformação digital.

Apesar desse lapso temporal, é possível verificar que já naquela época foram detectados problemas como falta de profissionais de TI no mercado de trabalho, principalmente quanto a funções mais especializadas, e dificuldade de retenção dos quadros, em razão sobretudo da ausência de plano de carreira, aliada à remuneração menos atrativa na comparação com a iniciativa privada. A rotatividade do pessoal de TI nos órgãos do Poder Judiciário da União Federal só fez crescer a cada ano do recorte temporal considerado no estudo: foi de 42% em 2010, 46% em 2011 e 52% em 2012.

Também foi detectada a ausência ou insuficiência de planejamento para manter o contínuo preenchimento dos cargos de TI, além de omissão, demora ou atuação insuficiente dos gestores para identificar e resolver tais problemas.

Consta que o Levantamento foi elaborado por determinação do TCU, no Acórdão 2.189/2013-TCU-Plenário, para verificar a existência e a dimensão da lacuna de profissionais de TI na administração pública federal, bem como para averiguar se os gestores estavam adotando as medidas adequadas para obter e manter quadro de pessoal suficiente, tratando-se de uma providência que se fez necessária porque:

O esvaziamento de recursos humanos na área de TI é considerado uma das causas das recorrentes irregularidades encontradas em contratações públicas concernentes à TI nos últimos anos em várias fiscalizações realizadas pelo TCU, o que tem gerado desperdício de recursos públicos, ineficiência e não atendimento do interesse público²⁸².

A partir dos achados com o levantamento, buscando solucionar o problema, o TCU indicou aos gestores dos órgãos da administração pública federal – inclusive do Poder Judiciário da União – as seguintes medidas:

- a) criação de cargos específicos da área de TI, distribuídos em carreira, de forma a propiciar a oportunidade de crescimento profissional;
- b) atribuição das funções gerenciais exclusivamente a servidores ou empregados públicos ocupantes de cargos efetivos de TI;
- c) estipulação de remuneração coerente com a relevância das atribuições desenvolvidas;

²⁸²Ibidem, p 23.

d) permanente capacitação dos servidores, incluindo nessas ações o conteúdo multidisciplinar necessário ao exercício das atribuições inerentes a essas funções, cujas competências vão além dos conhecimentos de TI²⁸³.

Observe-se que, se já estava difícil resolver o problema de **quantidade** suficiente de profissionais de TI para atender à demanda crescente, depreende-se que, diante dessa necessidade quantitativa mais premente, a **diversidade** acabou ficando em segundo plano.

A rigor, não ficou em plano algum, pois sequer foi cogitada. Os termos “diversidade”, “gênero” e “raça” não apareceram no levantamento realizado pelo TCU, que se limitou ao aspecto quantitativo em geral da força de trabalho em TI no universo delimitado, que foi o da administração pública federal dos três Poderes.

A propósito, no Guia da Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário (ENTIC-JUD) 2021/2026²⁸⁴, na versão 1.1 publicada pelo CNJ em agosto de 2021, atualmente em vigor, os termos “diversidade”, “gênero” e “raça”, também não aparecem em momento algum, e muito menos os outros marcadores de etnia, cor, orientação sexual, pessoas com deficiência e geração, só para citar os que foram expressamente mencionados no art. 20, *caput*, da Resolução 332/2020.

Além disso, no referido Guia, dedicou-se à força de trabalho uma única página, também exclusivamente voltada para a questão quantitativa, trazendo um quadro referencial mínimo para a força de trabalho de TI que todos os tribunais do país devem observar, com base no total de usuários de recursos de TI, para calcular os percentuais aceitáveis de efetivos, comissionados, terceirizados e servidores do quadro permanente.

Destaque-se que o Guia foi instituído pelo art. 5º da Resolução CNJ nº 370, de 28/01/2021²⁸⁵, como o instrumento que serviria de diretriz para execução da Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário (ENTIC-JUD) até 2026.

²⁸³Ibidem, p. 28.

²⁸⁴CNJ. **Guia da Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário ENTIC-JUD 2021-2026.** Versão 1.1. Ago. 2021. Disponível em: <https://atos.cnj.jus.br/files/compilado1841452021102661784be9efedd.pdf>. Acesso em: 10 jun. 2022.

²⁸⁵CNJ. **Resolução nº 370, de 28 de janeiro de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3706>. Acesso em: 10 jun. 2022.

Logo, é no mínimo um **contrassenso** que nada tenha sido mencionado, nem no Guia nem na Resolução 370/2021, a respeito de diretrizes para buscar a diversidade de equipes desenvolvedoras de modelos de IA, que fora considerada obrigatória pelo próprio CNJ na Resolução nº 332/2020.

Isso reforça a percepção de que, sem **atitudes proativas** dos gestores do Poder Judiciário para concretizar a diversidade nas equipes de TI, a norma do art. 20 da Resolução 332/2020 de fato não terá qualquer efetividade, permanecendo latente o risco de viés discriminatório nos modelos de IA desenvolvidos nos tribunais brasileiros.

Acontece que é alarmante essa inércia diante de uma fonte de risco relevante, haja vista a possibilidade de um modelo de IA atentar contra direitos e garantias fundamentais de jurisdicionados, se for enviesado desde o design.

Exemplos não faltam, sistematicamente em desfavor de negros e mulheres. A esse respeito, o pesquisador brasileiro Tarcízio Silva organizou e mantém atualizada uma “Linha do Tempo do Racismo Algorítmico”²⁸⁶, que pode ser consultada de forma on-line e interativa. Em 12/05/2022, ele inclusive participou do painel realizado no Senado, intitulado “Inteligência artificial e riscos: vieses e discriminação”, figurando entre os especialistas ouvidos pela Comissão de Juristas instituída para subsidiar o marco legal da IA no Brasil²⁸⁷.

Entre os exemplos catalogados na linha do tempo organizada por Tarcízio Silva, podem-se mencionar: ferramentas de visão computacional do Google que identificou pessoas negras como gorilas em fotografias, confundiu cabelos afros com perucas, bem como entendeu que instrumentos que estavam nas mãos de negros eram armas; resultados de busca que associavam mulheres negras à prostituição, pornografia e termos sexuais; cerceamento ao uso da expressão *Black Lives Matter* em diversas plataformas de redes sociais; análise de crédito prejudicial a negros; aumento de preço em aplicativos de transporte com destino a bairros pobres de maioria negra; resultado prejudicial a negros e mulheres em buscas de vagas de trabalho e de imóveis etc. A lista é extensa, esses são apenas alguns casos.

²⁸⁶SILVA, Tarcízio. **Linha do tempo do racismo algorítmico: casos, dados e reações.** 2019. Disponível em: <https://tarciziosilva.com.br/blog/destaques/posts/racismo-algoritmico-linha-do-tempo/>. Acesso em: 10 jun.2022.

²⁸⁷SILVA, Tarcízio. **Audiência pública – inteligência artificial e riscos: vieses e discriminação.** 30 maio 2022. Disponível em: <https://tarciziosilva.com.br/blog/audiencia-publica-inteligencia-artificial-e-riscos-vieses-e-discriminacao/>. Acesso em: 10 jun. 2022.

Na raiz do problema, está o fato de que na área de TI existe um “domínio esmagador de homens brancos”²⁸⁸, conforme apontado em relatório especial sobre diversidade no ramo da tecnologia, elaborado pela Comissão de Oportunidades Iguais de Emprego dos Estados Unidos, ao quantificar a medida da sub-representação de negros, mulheres, hispânicos e outros segmentos da sociedade norte-americana, para investigar as causas e servir de direcionamento a políticas públicas que aumentem a diversidade.

Tais políticas são necessárias diante dessa constatação de que há um grupo homogêneo que produz modelos de IA para uma sociedade que, no entanto, é plural.

Para o desenvolvimento dos modelos, várias decisões humanas são tomadas, como a escolha dos dados para treinamento, das técnicas que serão utilizadas, dos índices aceitáveis de acurácia, entre muitas outras. Assim, o desempenho da máquina será tanto melhor quanto forem as decisões humanas que resultarem na sua criação.

Ocorre que os estudos atuais na área de psicologia cognitiva e economia comportamental, a exemplo da obra de Kahneman, Sibony e Sunstein, demonstram que os julgamentos humanos são passíveis de **heurísticas, vieses e ruídos**, o que conduz a **erros** – que podem ser reduzidos, desde que avaliados e medidos²⁸⁹.

Para minorar riscos de erros, precisa-se justamente “de múltiplos julgamentos sobre um mesmo problema”²⁹⁰, o que não acontece se existir apenas a percepção e visão de mundo de um único segmento da sociedade – formado por homens brancos. E é por isso que diversidade nas equipes de TI importa tanto, ainda mais num país desigual como o Brasil, principalmente em se tratando de IA aplicada pelo Poder Judiciário.

No Brasil, as **desigualdades sociais**, o **machismo** e o **racismo** estruturais, de raízes históricas, podem ser demonstrados a partir de vários marcadores, entre os quais os índices de violência e criminalidade, que se relacionam ao sistema de justiça criminal.

²⁸⁸EEOC. **Diversity in high tech. U.S. Equal Employment Opportunity Comission.** Disponível em: <https://www.eeoc.gov/special-report/diversity-high-tech>. Acesso em: 10 jun. 2022.

²⁸⁹KAHNEMAN, Daniel et. al. **Ruído: uma falha no julgamento humano.** Trad. Cássio de Arantes Leite. – 1^a ed. – Rio de Janeiro: Objetiva, 2021.

²⁹⁰Ibidem, p. 56.

Conforme dados compilados em 2021 pelo Fórum Brasileiro de Segurança Pública, 76,2% das pessoas assassinadas no país em 2020 eram negras. Entre menores de idade, 63% das crianças e 81% dos adolescentes assassinados em negros. Das mulheres vítimas de feminicídio, 61,8% eram negras²⁹¹.

Esses são os dados quanto às vítimas. Do outro lado da moeda, a respeito de quem são os infratores, pretos e pardos são 67,34% da população carcerária nos Estados, segundo dados do Departamento Penitenciário Nacional apurados de julho a dezembro de 2021²⁹².

Sob o prisma dos **excluídos digitais** – que o Poder Judiciário precisa levar em consideração ao desenvolver serviços acessíveis pela Internet, com ou sem uso de IA – a pesquisa “O abismo digital no Brasil”, divulgada em março de 2022 pela PwC em parceria com o Instituto Locomotiva, apurou que “menos de um terço da população pode ser considerada plenamente conectada (sobretudo brancos das classes A e B)”, ao passo que “os outros cidadãos (principalmente negros das classes C, D e E) ficam sem conexão quase metade do mês”²⁹³.

Além disso, de acordo com a pesquisa, “58% dos brasileiros acessam a internet apenas via smartphone”²⁹⁴, o que é apontado como um fator limitante do aproveitamento das possibilidades tecnológicas de acesso digital a serviços públicos, ao mercado de trabalho e à educação.

Aliás, no quesito educação, os estudantes brasileiros estão entre os 10 menores desempenhos do mundo em Matemática. Em torno de 67% dos estudantes adolescentes têm dificuldade em interpretar textos e diferenciar fatos de opiniões. Somente 9% dos brasileiros afirmam saber falar Inglês²⁹⁵.

Tais fatores projetam para o futuro próximo um cenário de persistência da lacuna de talentos em TI, indicando que, se nada for feito, muito provavelmente permanecerá o problema não só da quantidade, mas também da falta de diversidade no setor.

²⁹¹FBSP. **A violência contra pessoas negras no Brasil 2021.** 19 nov. 2021. Disponível em: https://forumseguranca.org.br/publicacoes_posts/a-violencia-contra-pessoas-negras-no-brasil-2021/. Acesso em: 15 jun. 2022.

²⁹²BRASIL. **Levantamento nacional de informações penitenciárias – período de julho a dezembro de 2021.** Disponível em: <https://www.gov.br/depes/pt-br/servicos/sisdepen/sisdepen>. Acesso em: 15 jun. 2022.

²⁹³PWC. **O abismo digital no Brasil.** PwC/Instituto Locomotiva. 2022. Disponível em: https://www.pwc.com.br/pt/estudos/preocupacoes-ceos/mais-temas/2022/O_Abismo_Digital.pdf.

Acesso em: 15 jun. 2022.

²⁹⁴Ibidem.

²⁹⁵Ibidem.

Na falta de dados oficiais e atuais do setor público, e especificamente do Poder Judiciário, quanto à diversidade das equipes de TI envolvidas em projetos de IA nos tribunais, é possível ter uma visão geral a partir do trabalho da iniciativa brasileira PretaLab, que entre 2018 e 2019 realizou a pesquisa pioneira “Quem coda o Brasil?”²⁹⁶, em parceria com a consultoria global ThoughtWorks, com a finalidade de verificar o perfil do pessoal de TI no país, e qual era a composição das equipes.

Os principais achados da pesquisa demonstraram que, das pessoas que estavam produzindo tecnologia no Brasil naquele momento, 77% dos respondentes estavam na faixa etária de 18 a 34 anos e 65% moravam na capital dos Estados, com prevalência na Região Sudeste. Mais de 60% tinham renda mensal domiciliar acima de 5 salários mínimos. 58,3% se declararam brancos, ao passo que 36,9% declararam estar entre os pretos e pardos – enquanto na população em geral esse grupo era majoritário, de 54%²⁹⁷.

Além disso, os homens foram 67,1% dos respondentes. No que se refere às equipes de trabalho em TI, 53% eram formadas por até 10 pessoas. Na maioria das equipes, em 64,9% delas as mulheres representavam no máximo 20% da força de trabalho. Não havia qualquer mulher que fosse mãe em 62,1% das equipes²⁹⁸.

Não havia qualquer pessoa oriunda da classe E (com renda menor de 2 salários mínimos) em 69% das equipes. Não havia qualquer pessoa negra em 32,7% dos times. Nos 68,5% das equipes em que havia a presença de pessoa negra, a representatividade não passava de 10%. Nos 67,4% das equipes nas quais havia alguém morando em bairros periféricos e vulneráveis, a representatividade também não passava de 10%²⁹⁹.

Também em torno de 10% foi a representatividade de pessoas com orientação não heterossexual, verificada em 77,1% das equipes. Não havia qualquer pessoa com deficiência em 85,4% dos times. Não havia qualquer indígena em 95,9% das equipes³⁰⁰.

Cumpre destacar que esses números da pesquisa “Quem coda o Brasil?”, embora não advenham de um órgão estatal, apontam na mesma direção do relatório

²⁹⁶PRETALAB. **Quem coda o Brasil?** - #quemcodab. PretaLab/ThoughtWorks. Disponível em: https://www.thoughtworks.com/content/dam/thoughtworks/documents/report/tw_report_quem_coda_br.pdf. Acesso em: 15 jun. 2022.

²⁹⁷Ibidem.

²⁹⁸Ibidem.

²⁹⁹Ibidem.

³⁰⁰Ibidem.

norte-americano antes mencionado, no sentido de que também no Brasil há forte predominância de homens brancos na força de trabalho de TI, com índices reduzidos de diversidade nas equipes de trabalho.

Isso significa que, para a “busca da diversidade em seu mais amplo espectro”, na composição de equipes desenvolvedoras de modelos de IA no Poder Judiciário, tal qual foi mencionada no art. 20, *caput*, da Resolução 332/2020, há que se traçar e executar estratégias consistentes, sem solução de continuidade, porque o cenário é desafiador e requer providências de curto, médio e longo prazos.

Como possíveis táticas, o Poder Judiciário brasileiro poderia tomar por parâmetro as que constam no relatório sobre como construir equipes de IA de alto desempenho, *Building Successful AI Teams*³⁰¹, elaborado em 2021 pelo *Deloitte AI Institute*, que é um braço da Deloitte, empresa multinacional de auditoria e consultoria.

Levando em conta o panorama global de escassez de especialistas, o relatório aponta que é necessária uma combinação de providências, principalmente mudança de cultura e reestruturação das carreiras de TI no âmbito da organização, com novas contratações e requalificação da força de trabalho já existente.

A **mudança de cultura** terá de vir do topo, partindo dos gestores, a quem cabe liderar essa mudança pelo exemplo; criar incentivos como promoções, concursos e prêmios; bem como investir pesadamente em contínua capacitação.

Fala-se em **capacitação** não só para que as equipes de TI envolvidas em projetos de IA estejam sintonizadas com o que há de mais novo nessa área tão dinâmica, mas também para que os profissionais da área de negócio tenham os conhecimentos mínimos necessários em IA, o que permitirá aumentar a sinergia entre as áreas técnica e de negócio, permitindo melhor entendimento dos desafios e otimizando o desenvolvimento conjunto de soluções.

A mudança de cultura inclui o **redesign organizacional**, precedido de um exame minucioso dos atuais times que desenvolvem modelos de IA para a organização, verificando as descrições dos cargos, o organograma, as habilidades técnicas necessárias para o ramo de atuação da organização, o plano de carreira e as remunerações na comparação com o mercado, as atribuições, as características pessoais dos componentes, suas habilidades e aptidões, a medida da presença ou

³⁰¹DELOITTE. **Building successful AI teams.** 2021. Disponível em: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-ai-institute-building-successful-AI-teams.pdf>. Acesso em: 16 jun. 2022.

ausência de diversidade, a projeção de demanda, a necessidade ou não de recrutar talentos externos.

O relatório da Deloitte exemplifica que, nos Estados Unidos, a profissão nº 1 em 2020 foi a de cientista de dados. Como não há profissionais suficientes para suprir a procura, a remuneração está em alta e a rotatividade média é de 2 anos, o que gera para as organizações o desafio extra da **retenção de talentos**.

A Deloitte ainda destaca que não se pode pretender recrutar profissionais com muitos anos de experiência em IA, até porque simplesmente não existem, em se tratando de uma tecnologia que é muito nova.

O mais indicado é recrutar pessoal de TI em geral para depois treiná-lo, sob a mentoria dos profissionais mais experientes. Outra tática, fortemente recomendada, é **prospectar pessoal da própria instituição**, que demonstre capacidade e interesse em mudar de carreira, dispondo-se a receber treinamento para adquirir habilidades no desenvolvimento de modelos de IA.

Assim, trará o ganho de já vir inteirado do funcionamento e necessidades da organização, podendo atuar como uma ponte bastante útil entre a área de negócios e a área de TI, com economia de tempo.

Essas estratégias também são válidas para alcançar a **diversidade** nas equipes de IA, que o relatório inclusive aponta como aquele que deve ser um **item prioritário** da agenda dos gestores, por ser a **melhor maneira de reduzir o risco de vieses discriminatórios** nos modelos de IA desenvolvidos, ao proporcionar diferentes perspectivas na análise de situações.

Para retenção dessa força de trabalho especializada, que está tão escassa e difícil de obter, que leva tempo para treinar e está fácil de perder, o relatório sugere, além da perspectiva de ascensão profissional e remuneração compatível, atenção à descrição dos cargos e funções, oportunidades de mentoria, treinamento contínuo, organização de eventos como *hackathons* que permitam o intercâmbio de experiências, premiações e reconhecimento, flexibilidade de condições para o exercício das funções (a exemplo de trabalho remoto e híbrido), além da possibilidade de trabalho em projetos variados para criar e manter o estímulo ao permitir experimentação.

No tocante ao tamanho das equipes de IA, ao local em que devem figurar no organograma das organizações e à descrição de funções e habilidades, o relatório

da Deloitte menciona que cada instituição deve customizá-los de acordo com suas necessidades.

Porém, consta que de modo geral as três estruturas mais comuns são: (i) um setor específico para IA, submetido a uma supervisão de gestão e de governança; (ii) descentralização e pulverização de equipes de IA em diferentes áreas de negócio que utilizem modelos de IA; (iii) um modelo híbrido, combinando as duas estruturas anteriores.

A fim de apontar parâmetros, a Deloitte também indica as possíveis funções para equipes de IA de alto desempenho, entre as quais se destaca a de “Eticista de IA”, conforme a Figura 6 a seguir (com tradução livre adaptada dos cargos sugeridos, mantendo-se em inglês algumas expressões próprias de uso corrente na área de TI, e destaque em vermelho para a função mencionada)³⁰²:

Figura 6 – Possíveis funções para equipes de IA de alto desempenho



Fonte: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-ai-institute-building-successful-AI-teams.pdf>

³⁰²Ibidem.

Como medida adicional, o relatório a Deloitte sugere **parcerias com universidades e instituições de pesquisa**³⁰³, para planejamento de oportunidades de recrutamento e estágio, de modo a atrair talentos desde o início de suas carreiras.

Além disso, é interessante pontuar que a Deloitte, no artigo de apresentação do relatório sobre como construir equipes de IA bem sucedidas, encorajou a **interdisciplinaridade** entre a TI e outros ramos do conhecimento, a exemplo da Medicina, como medida útil para prospectar soluções.

Com efeito, a Deloitte afirmou que a mudança de cultura, rumo a um programa de IA realmente vantajoso, consistente e sustentável, requer ampliar os horizontes para ir além do departamento de TI, buscando ideias para resolver gargalos, pois “ao combinar a tecnologia de IA com os campos da medicina e da ciência, as organizações buscam oportunidades para transformar alguns dos seus processos mais críticos”³⁰⁴.

Essa é justamente a linha de raciocínio interdisciplinar adotada na presente pesquisa desde o início, considerando que uma das hipóteses trabalhadas é sobre a possibilidade de adaptação de uma ferramenta de controle de risco observada na Medicina norte-americana, a ser analisada no próximo capítulo.

Por ora, busca-se nesta seção fazer uma análise crítica da Resolução 332/2020, à luz das possíveis fontes de risco de erros e incidentes de modelos de IA. No que se refere à diversidade das equipes, demonstrou-se que é medida de alta relevância, por ser tida como uma das principais estratégias apontadas pelos especialistas para mitigar o risco de vieses discriminatórios dos sistemas criados.

Acontece que, ao abordar a questão da diversidade, constatou-se que há um desafio dentro de outro: o primeiro é a lacuna em geral de profissionais especializados em IA, enquanto o segundo é a lacuna ainda maior de especialistas negros e mulheres – embora sejam maioria na população brasileira – além de outros representativos de segmentos minoritários, como pessoas com deficiência.

Como se não bastassem esses dois desafios de **recrutamento**, que por si sós já são difíceis de resolver, adiciona-se um terceiro, que é o desafio da **retenção** em geral, e ainda um quarto, que é o desafio da retenção especificamente da força de

³⁰³Ibidem.

³⁰⁴DELOITTE. Building successful AI teams. 2021. <https://www2.deloitte.com/us/en/pages/consulting/articles/building-diverse-teams-in-tech-organizations.html>. Acesso em: 16 jun. 2022.

trabalho feminina na área de TI, com saídas das mulheres em taxas bem mais elevadas que os homens – e por motivos negativos, muito diferentes de terem recebido ofertas de melhores oportunidades de carreira ou salários mais elevados.

Não havendo dados oficiais desse movimento de saída das mulheres de profissões de TI no Brasil, é possível observar o fenômeno no relatório governamental norte-americano antes referido³⁰⁵, que apontou dados de 2008 em que mais da metade das mulheres em ocupações de TI deixaram seus cargos. Em 2013, 32% das mulheres queixaram-se do ambiente de trabalho e afirmaram que provavelmente desistiriam dentro de um ano. Outro conjunto de dados demonstrou que as mulheres ocupavam 41% dos cargos de TI nos degraus mais baixos da carreira, porém no período analisado mais da metade veio a sair³⁰⁶.

O relatório do governo norte-americano apurou que o ambiente predominantemente masculino leva a uma cultura organizacional inóspita, que gera níveis insuportáveis de desgaste, sendo esse o principal fator que acaba empurrando as mulheres para fora da carreira de TI.

Entre as experiências negativas mais recorrentes que o estudo constatou estão: isolamento; capacidade das mulheres frequentemente desmerecida, questionada ou posta à prova; pressão para desempenhar papéis tradicionalmente reservados às mulheres; críticas às mulheres que se portam com franqueza e assertividade ou se recusam a assumir funções estereotipadas; julgamentos negativos caso socializem com colegas; diminuição de oportunidades após terem filhos; assédio moral e sexual³⁰⁷.

Os números foram ainda mais negativos para as mulheres negras na TI, que são mais impactadas em virtude da intersecção do machismo e do racismo. Esse fato foi reforçado pelo relatório de 2021 do Google, sobre diversidade de suas equipes, em que o índice de desgaste das mulheres negras foi superior ao dos demais grupos, só ficando abaixo do desgaste das mulheres nativas norte-americanas³⁰⁸.

³⁰⁵EEOC. **Diversity in high tech.** U.S. Equal Employment Opportunity Comission. Disponível em: <https://www.eeoc.gov/special-report/diversity-high-tech>. Acesso em: 10 jun. 2022.

³⁰⁶Ibidem.

³⁰⁷Ibidem.

³⁰⁸GOOGLE. **2021 Diversity Annual Report.** Disponível em: https://static.googleusercontent.com/media/diversity.google/pt-BR//annual-report/static/pdfs/google_2021_diversity_annual_report.pdf?cachebust=2e13d07. Acesso em: 18 jun. 2022.

Como é possível perceber, o desafio de obter (e manter) diversidade nas equipes de IA apresenta várias nuances, que exigem providências efetivas e adequadas para cada uma delas, não sendo suficiente que o CNJ apenas tenha feito uma previsão puramente normativa no art. 20 da Resolução 332/2020, ainda por cima prevendo de antemão a possibilidade de dispensa de diversidade nas equipes de IA por falta de profissionais, sem que depois tenha adotado atitudes concretas para que essa dispensa seja realmente uma exceção, de fato e de direito.

A diversidade precisa ser buscada, não virá naturalmente. Embora a TI seja área-meio e não área-fim dos tribunais, é área estratégica e da máxima importância em plena era digital. Especificamente a respeito da subárea de IA, a falta de diversidade é fonte permanente de risco, podendo levar o Poder Judiciário a dar ensejo à injustiça, que é o exato oposto de sua missão e razão de existir.

Em acréscimo às diretrizes do relatório da Deloitte, que foram mencionadas para servirem como um parâmetro de providências que podem ser adotadas, é preciso dizer que adaptações serão necessárias para aplicação pelo Poder Judiciário, submetido que está a limites constitucionais e legais próprios da Administração Pública, sobretudo quanto ao recrutamento de sua força de trabalho, em que a regra é o concurso público.

Como táticas de curto prazo para obter mais diversidade nas equipes de IA, o Judiciário pode se valer, por exemplo, de quadro diferenciado da força de trabalho de TI específica para a subárea de IA, talvez com percentuais diferentes dos que constam atualmente no Guia da ENTIC-JUD 2021/2026 para distribuição dos cargos efetivos, comissionados e terceirizados.

Quanto mais estratégica for a função do(a) servidor(a) no design e ciclo de vida do modelo de IA, mais parece recomendável, a princípio, que seja atribuída a quem tenha vínculo permanente. A propósito, já foi analisada a recomendação nesse sentido do TCU ao CNJ, constante no Levantamento de Pessoal de TI³⁰⁹ sobre o qual já se tratou linhas antes, dado o entendimento de que a presença de pessoal não efetivo em funções estratégicas da Administração Pública representa risco ao interesse público.

³⁰⁹BRASIL. **Levantamento de Pessoal de TI**/Tribunal de Contas da União; Relator Ministro Raimundo Carreiro. – Brasília: TCU, Secretaria de Fiscalização de Tecnologia da Informação (Sefti), 2015. 52 p. – (Sumário Executivo. Tecnologia da Informação). Disponível em: <https://portal.tcu.gov.br/biblioteca-digital/levantamento-de-pessoal-de-ti-sumario.htm>. Acesso em: 10 jun. 2022.

Contudo, enquanto não se alcança, a médio e longo prazos, a diversidade ideal ou pelo menos satisfatória com pessoal efetivo (cujos parâmetros também precisam ser definidos quanto ao que seria considerado ideal e satisfatório), pode-se no curto prazo agregar às equipes de IA o valor da diversidade com pessoal comissionado e terceirizado, recrutados por suas habilidades e seu potencial, proporcionando-lhes capacitação e atribuindo-lhes funções não estratégicas.

Teoricamente, a importância da diversidade nas equipes de IA como uma das medidas mais importantes de gerenciamento de risco de vieses discriminatórios, mundialmente atribuída por diversos pesquisadores e especialistas, justificaria a inclusão de um tópico próprio no Guia da ENTIC-JUD 2021/2026. Afinal, é tema tido como prioritário para as organizações que desenvolvem aplicações de IA.

É até surpreendente que nenhuma linha tenha sido dedicada, num guia de estratégia plurianual de TI que vai até 2026, para tratar de assunto que fora considerado tão relevante pelo próprio CNJ em resolução anterior.

Inclusive abre margem para perquirir se não seria um exemplo concreto do tipo de lapso que pode advir da descontinuidade de projetos pela troca bienal de gestão, que é característica da Administração Pública judiciária brasileira, um vício que a própria Resolução 332/2020 buscou combater, como já foi apontado neste trabalho.

A propósito, outro exemplo concreto de lapso verificado é a completa ausência de menção à Resolução nº 332/2020 na Portaria CNJ nº 170, de 20/05/2022³¹⁰, que instituiu o regulamento do Prêmio CNJ de Qualidade 2022.

Foi feita referência à Resolução nº 331/2020³¹¹, de 20/08/2020, para incentivar os tribunais a providenciarem a alimentação correta dos dados na Base Nacional de Dados do Poder Judiciário – DATAJUD, valendo alto peso de 200 pontos, muito além de todos os outros pesos sobre os demais assuntos considerados importantes, que variam entre 10, 15, 20, 25, 30, 35, 40, 45, 50, 70, 80, 90 e 100 pontos.

Todavia, absolutamente nada foi dito a respeito da Resolução nº 332/2020 no regulamento do Prêmio CNJ de Qualidade 2022.

³¹⁰CNJ. Portaria nº 170, de 20 de maio de 2022. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4549>. Acesso em: 20 jun. 2022.

³¹¹CNJ. Resolução nº 331, de 20 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3428>. Acesso em: 20 jun. 2022.

Note-se que as Resoluções nº 331 e 332 foram publicadas no mesmo dia, 20/08/2020, a primeira tratando sobre dados e a segunda sobre ética na aplicação de IA nos tribunais, indicando que a alta administração do CNJ, na época, deu um tratamento conjunto a tais matérias correlatas.

Ora, tratando-se de temas estratégicos para uma Justiça digital, é curioso que no atual regulamento para o prêmio apenas um deles tenha sido incentivado, tenha merecido um *nudge*³¹², na acepção de Thaler e Sustein, enquanto o outro sequer tenha sido citado, mesmo em plena expansão de projetos de IA nos tribunais.

A ênfase dada à alimentação correta de dados no DATAJUD pareceu privilegiar o modelo gerencialista para aferição de desempenho e produtividade, analisado por Jorge Machado e Daniel Lino, que compararam as limitações desse modelo de gestão com o conceito de justiça aberta, que inclui “participação social, colaboração e cocriação com a sociedade civil, linguagem acessível, promoção de dados em formatos abertos para reuso e a criação de canais de diálogo”, como importantes ações “que permitam tornar a Justiça mais permeável à sociedade – algo importante e necessário num país com grandes desigualdades sociais como o Brasil

³¹³.

Ocorre que já houve omissão sobre a Resolução 332/2020 no Guia da ENTIC-JUD 2021/2026, quanto ao encaminhamento do problema da falta de diversidade nas equipes de IA, e essa outra omissão no regulamento do Prêmio CNJ de Qualidade 2022, quando analisadas em conjunto, fazem parecer que se tornou desimportante ou menos importante para o Judiciário brasileiro produzir uma IA ética – logo quanto a uma atividade de risco, ainda mais na instituição cuja missão é justamente distribuir justiça e corrigir injustiças.

Até mesmo por aplicação prática do princípio da precaução, no sentido de adotar atitudes proativas de mitigação de riscos, seria interessante ter incluído a observância à Resolução 332/2020 como item a ser cumprido pelos tribunais para alcançar o Prêmio CNJ de Qualidade, que tem as categorias diamante, ouro e prata, a depender da soma de pontos.

³¹²THALER, Richard H.; SUSTEIN, Cass R. **Nudge: como tomar melhores decisões sobre saúde, dinheiro e felicidade**. Tradução Ângelo Lessa. 1^a ed. Rio de Janeiro: Objetiva, 2019.

³¹³MACHADO, Jorge; LINO, Daniel. Gerencialismo e justiça aberta. **Acervo - Revista do Arquivo Nacional**, v. 34, n. 3, p. 1-18, 2021. Disponível em: <https://www.brapci.inf.br/index.php/res/v/163601>. Acesso em: 20 jun. 2022.

O relatório da Deloitte, como já analisado, mencionou a importância da promoção de incentivos e prêmios como estratégia. Contudo, ao contrário de instigar os tribunais a incluírem a busca da diversidade nas equipes de IA, o regulamento do prêmio, além de tornar essa uma não-meta, pela omissão, acabou deixando o assunto talvez ainda mais distante de ser abordado, porque é natural que todos os tribunais estejam empenhados e focados somente na já extensa lista de metas a cumprir, que foram definidas como prioritárias pelo CNJ.

A importância da diversidade para minorar riscos de vieses discriminatórios também tornaria justificável que o CNJ viesse a baixar atos normativos específicos para a composição das equipes de IA nos tribunais, contemplando políticas de cotas para negros, mulheres e pessoas com deficiência.

Fundamentos constitucionais e legais não faltam, para aplicação direta ou por analogia. Desde a própria Constituição Federal, com o princípio da igualdade e vedação a tratamento discriminatório, passando pela Lei nº 12.288/2010³¹⁴ (Estatuto da Igualdade Racial, prevendo políticas públicas para redução das desigualdades e enfrentamento da discriminação), pela Lei nº 12.711/2012³¹⁵ (sobre cotas sociais e raciais, estabelecendo a reserva de vagas em universidades e instituições de ensino federais para estudantes de baixa renda, pretos, pardos, indígenas e pessoas com deficiência) e pela Lei nº 12.990/2014³¹⁶ (reserva aos negros de 20% das vagas em concursos federais).

O próprio CNJ baixou atos normativos sobre ações afirmativas, com o objetivo de reduzir até eliminar desigualdades históricas, tratando-se da Resolução CNJ nº 203/2015³¹⁷ (reserva aos negros de 20% das vagas em concursos para a magistratura), a Resolução CNJ nº 336/2020³¹⁸ (reserva aos negros de 30% das vagas de estágio no Poder Judiciário).

³¹⁴BRASIL. Lei nº 12.288, de 20 de julho de 2010. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/lei/l12288.htm. Acesso em: 20 jun. 2022.

³¹⁵BRASIL. Lei nº 12.711, de 29 de agosto de 2012. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12711.htm. Acesso em: 20 jun. 2022.

³¹⁶BRASIL. Lei nº 12.990, de 9 de junho de 2014. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12990.htm. Acesso em: 20 jun. 2022.

³¹⁷CNJ. Resolução nº 203, de 23 de junho de 2015. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2203>. Acesso em: 20 jun. 2022.

³¹⁸CNJ. Resolução nº 336, de 29 de setembro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3497>. Acesso em: 20 jun. 2022.

A Lei nº 8.213/1991³¹⁹ (cotas para emprego de pessoas com deficiência nas empresas) já tem mais de 30 anos, enquanto as leis que tratam de cotas para mulheres na política também são dos anos 1990, tratando-se da Lei nº 9.100/1995³²⁰ (reserva às mulheres de 20% das vagas para vereador), e da Lei nº 9.504/1997³²¹ (reserva de 30% em candidaturas proporcionais). Mais recentemente, a Emenda Constitucional nº 117/2022³²² prevê 30% do Fundo Eleitoral para candidaturas femininas.

Além dessas balizas legislativas e administrativas, existem ainda importantes e históricos entendimentos jurisprudenciais, visto que em 2012 o Supremo Tribunal Federal (STF), por unanimidade, no julgamento da Arguição de Descumprimento de Preceito Fundamental nº 186³²³, considerou constitucional a política de cotas que já era adotada de forma pioneira pela Universidade de Brasília (UnB) desde 2009, antes mesmo da Lei nº 12.711/2012.

Da mesma forma, em 2017 o STF por unanimidade considerou constitucional a lei de cotas raciais para concursos no serviço público federal (Lei nº 12.990/2014), no julgamento da Ação Direta de Constitucionalidade nº 41, que “se funda na necessidade de superar o racismo estrutural e institucional ainda existente na sociedade brasileira”³²⁴.

Em 17/12/2020, ato do Presidente da Câmara dos Deputados instituiu uma Comissão de Juristas para estudos e propostas de aperfeiçoamento da legislação brasileira de enfrentamento ao racismo estrutural e institucional³²⁵.

³¹⁹BRASIL. **Lei nº 8.213, de 24 de julho de 1991.** Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8213cons.htm. Acesso em: 20 jun. 2022.

³²⁰BRASIL. **Lei nº 9.100, de 29 de setembro de 1995.** Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9100.htm. Acesso em: 20 jun. 2022.

³²¹BRASIL. **Lei nº 9.504, de 30 de setembro de 1997.** Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9504.htm. Acesso em: 20 jun. 2022.

³²²BRASIL. **Emenda Constitucional nº 117, de 5 de abril de 2022.** Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc117.htm. Acesso em: 20 jun. 2022.

³²³STF. **Arguição de Descumprimento de Preceito Fundamental nº 186 – Distrito Federal.** Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=6984693>. Acesso em: 20 jun. 2022.

³²⁴STF. **Ação Direta de Inconstitucionalidade nº 41 – Distrito Federal.** Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=13375729>. Acesso em: 20 jun. 2022.

³²⁵BRASIL. **Ato do Presidente de 17/12/2020 – Câmara dos Deputados.** Disponível em: <https://www2.camara.leg.br/atividade-legislativa/comissoes/grupos-de-trabalho/56a-legislatura/comissao-de-juristas-combate-ao-racismo-no-brasil/conheca-a-comissao/criacao-e-constituicao/ato-de-criacao>. Acesso em: 20 jun. 2022.

A Comissão de Juristas entregou em 30/11/2021 um relatório final³²⁶ extenso, de 610 páginas, com propostas para diversas áreas como economia, saúde, assistência social, educação, trabalho, legislação penal e processual penal, segurança pública e sistema de justiça criminal, entre outras.

Contudo, é interessante notar que o relatório não faz qualquer referência a expressões como “inteligência artificial”, “algoritmo”, “discriminação algorítmica”, “viés algorítmico”, não obstante tenha havido ao menos uma discussão pública sobre tecnologia em 26/04/2021 com o tema “Tecnologias emergentes, criminalidade e combate ao racismo”³²⁷, e tenha participado da Comissão de Juristas o professor Adilson José Moreira, que na obra Tratado de Direito Antidiscriminatório reservou um capítulo inteiro para o tema Discriminação e Tecnologia.

Na citada obra, segundo o jurista, “podemos afirmar que a inteligência artificial é afligida pelos mesmos males que afetam uma mente discriminatória”. Um dos fatores para tanto, apontados pelo professor, é que a criação dos algoritmos sofre influência da formação cultural dos programadores, que tende a vir de um grupo demográfico homogêneo. Assim, o viés algorítmico discriminatório acontece “porque grande parte de seus programadores não possuem contato social constante com minorias, o que os faz pensar que todos os membros de certos grupos possuem as mesmas características”³²⁸.

Fato é que, no relatório da Comissão de Juristas, não há qualquer seção específica ou mínima menção para tratar especificamente de IA e diversidade nas profissões de TI relacionadas ao desenvolvimento dessa tecnologia.

Há apenas uma sugestão genérica de inclusão de inciso ao art. 3º da Lei nº 14.129/2021 (que trata sobre os princípios e diretrizes do Governo Digital no Brasil, aplicáveis a todos os órgãos da administração direta e indireta dos três Poderes de todos os entes da Federação), em que o acréscimo seria do inciso XXVII, tornando-se princípio/diretriz do Governo Digital brasileiro “a não discriminação por raça/cor, em

³²⁶BRASIL. Relatório final – Comissão de juristas destinada a avaliar e propor estratégias normativas com vistas ao aperfeiçoamento da legislação de combate ao racismo estrutural e institucional no país. 30 nov. 2021. Disponível em: <https://www2.camara.leg.br/atividade-legislativa/comissoes/grupos-de-trabalho/56a-legislatura/comissao-de-juristas-combate-ao-racismo-no-brasil/documentos/outras-documentos/relatorio-final>. Acesso em: 20 jun. 2022.

³²⁷Ibidem, p. 412.

³²⁸MOREIRA, Adilson José. **Tratado de Direito Antidiscriminatório**. São Paulo: Editora Contracorrente, 2020, p. 513 e 514.

conformidade com a Lei nº 12.288, de 20 de julho de 2010 – Estatuto da Igualdade Racial, identidade de gênero e orientação sexual”.

Se a sugestão for acolhida pelo Poder Legislativo, pode vir a caber interpretação, no tema que toca a esta pesquisa, tanto no sentido de não-discriminação no recrutamento de pessoal de TI para desenvolvimento de IA, como no sentido de não-discriminação dos serviços públicos desenvolvidos com aplicação de IA, inclusive na atividade judiciária.

Atualmente, a discussão mais recente sobre diversidade no Brasil ainda está sendo travada no Congresso Nacional, pois está prevista para 2022 a revisão obrigatória da chamada Lei de Cotas (Lei nº 12.711/2012), cujo texto contempla expressamente a revisão em dez anos.

Buscou-se sintetizar, aqui, todo o arcabouço legal que pode servir de base para fundamentar eventual criação, pelo CNJ, de ações afirmativas que resultem em cotas sociais, raciais, de gênero e por deficiência, no âmbito das equipes de IA nos tribunais brasileiros, para tornar efetiva a regra do art. 20 da Resolução 332/2020, a serem incluídas na seção sobre a força de trabalho do plano plurianual de sua estratégia de TI.

Trata-se de medida que seria não só de observância ao princípio constitucional da igualdade e vedação à discriminação, como também de aplicação do princípio da precaução, por meio de uma ação concreta para mitigação de risco de viés algorítmico discriminatório em modelos de IA aplicados nos tribunais.

Entretanto, ainda que medidas desse tipo venham a ser implementadas continuamente quanto aos cargos efetivos, é preciso ter em mente que levam muito tempo para surtirem o efeito desejado, que é a representatividade em percentual similar à da população em geral.

Para exemplificar, fazendo um paralelo com os resultados e projeções da política de cotas para negros nos concursos para a magistratura, existentes desde 2015, “apenas no ano 2049 haverá o atingimento de pelo menos 22% de magistrados negros em todos os tribunais brasileiros”³²⁹, ou seja, mais de 3 décadas para pouco

³²⁹CNJ. **Igualdade racial no judiciário - Relatório de atividade. Grupo de trabalho Políticas Judiciárias sobre a Igualdade Racial no âmbito do Poder Judiciário** (Portaria n. 108, de 8/7/2020). Out. 2020. Disponível em: https://www.cnj.jus.br/wp-content/uploads/2020/10/Relatorio_Igualdade-Racial_2020-10-02_v3-2.pdf. Acesso em: 20 jun. 2022, p. 14.

mais de 20%, de acordo com o relatório sobre igualdade racial elaborado em 2020 por grupo de trabalho criado sobre o tema no CNJ.

Essa constatação corrobora o entendimento já externado linhas antes, de que para alcançar mais diversidade nas equipes de IA desde já, pode-se apontar a distribuição dos cargos efetivos, comissionados e terceirizados em percentuais diferentes dos que constam no Guia da ENTIC-JUD 2021/2026.

Dessa forma, como tática de curto pode-se buscar a diversidade com o recrutamento de comissionados(as), sem descuidar da estratégia de médio e longo prazos quanto aos efetivos, mediante política de cotas em concursos.

Outra estratégia para mitigar riscos de vieses algorítmicos discriminatórios, que pode ser aplicada desde já, mesmo com as atuais equipes de IA ainda não-diversas, enquanto não se obtém paulatinamente mais diversidade, é a **capacitação continuada** em questões raciais e de gênero, e em enfrentamento a discriminações como capacitismo, etarismo, homofobia.

A respeito do ambiente de trabalho inóspito para mulheres e mais ainda para mulheres negras, apontado nas pesquisas antes analisadas como um fator que dificulta a retenção da força de trabalho feminina e pode prejudicar a diversidade nas equipes de IA, foi mencionada a mudança da cultura organizacional como providência importante para resolver o problema.

Nesse ponto, surge o valor de replicar iniciativas como a do TJDF e seu Relatório Diagnóstico Organizacional de Equidade e Inclusão da Diversidade, antes mencionado, bem como de reforçar a divulgação e aplicação do Protocolo para Julgamento com Perspectiva de Gênero³³⁰, conforme a Recomendação CNJ nº 128/2022³³¹.

Outra medida recente determinada pelo CNJ para melhorar a cultura organizacional é a implementação da Política de Prevenção e Combate do Assédio Moral, do Assédio Sexual e de Todas as Formas de Discriminação, que foi instituída pela Resolução CNJ nº 351/2020³³², que determina a criação, em cada tribunal do

³³⁰CNJ. **Protocolo para Julgamento com Perspectiva de Gênero** [recurso eletrônico] — Brasília: Conselho Nacional de Justiça – CNJ; Escola Nacional de Formação e Aperfeiçoamento de Magistrados — Enfam, 2021. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2021/10/protocolo-18-10-2021-final.pdf>. Acesso em: 21 jun. 2022.

³³¹CNJ. **Recomendação nº 128, de 15 de fevereiro de 2022**. Disponível em: <https://atos.cnj.jus.br/files/original18063720220217620e8ead8fae2.pdf>. Acesso em: 21 jun. 2022.

³³²CNJ. **Resolução nº 351, de 28 de outubro de 2020**. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3557>. Acesso em: 21 jun. 2022.

país, de uma Comissão de Prevenção e Enfrentamento do Assédio Moral e do Assédio Sexual. Também foi reforçada a criação de Comissões Permanentes de Acessibilidade e Inclusão, instituídas desde a Resolução CNJ nº 230/2016³³³, atualmente reguladas pela Resolução nº 401/2021³³⁴.

Adicionalmente, a Resolução CNJ nº 343/2020³³⁵ trata sobre condições especiais de trabalho, não só para magistrados(a) como também para servidores(as), que sejam pessoas com deficiência ou necessidades especiais, ou que tenham dependentes nessa mesma condição.

Todos esses diplomas legais e administrativos encampam as diretrizes de diversos instrumentos internacionais de proteção aos direitos humanos, tais como: (i) a Convenção Interamericana Sobre Toda Forma de Discriminação e Intolerância³³⁶; (ii); a Convenção Sobre a Eliminação de Todas as Formas de Discriminação Contra a Mulher³³⁷; (iii) a Convenção Sobre os Direitos da Pessoa com Deficiência³³⁸; (iv) a Convenção nº 111 Sobre Discriminação em Matéria de Emprego e Profissão da Organização Internacional do Trabalho³³⁹; (v) os Princípios de Yogyakarta em relação à orientação sexual e identidade de gênero³⁴⁰.

Como é possível perceber, por meio de vários atos administrativos o CNJ já vem lançando muitas bases no sentido de fomentar uma cultura organizacional mais saudável, inclusiva e com respeito à diversidade, o que pode contribuir para a retenção da força de trabalho negra e feminina de TI.

³³³CNJ. **Resolução nº 230, de 22 de junho de 2016.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2301>. Acesso em: 21 jun. 2021.

³³⁴CNJ. **Resolução nº 401, de 16 de junho de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3987>. Acesso em: 21 jun. 2022.

³³⁵CNJ. **Resolução nº 343, de 9 de setembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3459>. Acesso em: 21 jun. 2020.

³³⁶BRASIL. **Decreto 10.932, de 10 de janeiro de 2022.** Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2022/Decreto/D10932.htm. Acesso em: 21 jun. 2022.

³³⁷BRASIL. **Decreto nº 4.377, de 13 de setembro de 2002.** Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/2002/d4377.htm. Acesso em: 21 jun. 2022.

³³⁸BRASIL. **Decreto nº 6.949, de 25 de agosto de 2009.** Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/decreto/d6949.htm. Acesso em: 21 jun. 2022.

³³⁹BRASIL. **Decreto nº 10.088, de 5 de novembro de 2019.** Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Decreto/D10088.htm#art5. Acesso em: 21 jun. 2022.

³⁴⁰**PRINCÍPIOS de Yogyakarta: princípios sobre a aplicação da legislação internacional de direitos humanos em relação à orientação sexual e identidade de gênero.** Tradução Jones de Freitas. Disponível em: http://www.dhnet.org.br/direitos/sos/gays/principios_de_yogyakarta.pdf. Acesso em: 21 jun. 2022.

Talvez esteja faltando concatenar essas políticas com um enfoque diferenciado e atento às particularidades das equipes de IA, sob o prisma da importância da diversidade para mitigação de riscos de vieses discriminatórios, objetivando desenvolver modelos de IA de qualidade, pois os algoritmos serão tanto melhores quanto forem os humanos que os programarem.

As normas, por si, não bastam. É necessário fazer um diagnóstico da força de trabalho que lida especificamente com IA nos tribunais, traçar metas de contratação para diversificar as equipes e acompanhar continuamente os resultados das estratégias adotadas.

Para tanto, é preciso que os gestores dos tribunais estejam cientes de toda a problemática que envolve a falta de profissionais de TI no mercado de trabalho, as dificuldades de retenção e a diversidade das equipes de IA como ponto crítico, que merece tanta atenção (ou mais, porque sua falta significa risco) que a busca exclusiva por números e produtividade.

3.4.4 Riscos da falta de transparência algorítmica

O Capítulo IV da Resolução 332/2020 é reservado à publicidade e transparência, que de acordo com o art. 8º consiste em um conjunto de providências que levem à explicação de como um modelo de IA chegou aos resultados que apresentou.

É preciso haver indicação dos objetivos e resultados que se pretende alcançar com cada modelo de IA, havendo informação sobre que tipos de dados foram utilizados para desenvolvimento de cada modelo, tudo envolvendo a divulgação responsável das informações necessárias, levando em consideração a existência de dados sensíveis entre os dados judiciais.

Para haver transparência algorítmica dos modelos de IA aplicados nos tribunais, também é necessário que os riscos identificados sejam documentados, que seja possível aferir os motivos pelos quais um dano ocorreu e quais são as ferramentas de segurança da informação utilizadas, bem como os mecanismos de auditoria e certificação.

Além disso, qualquer proposta de decisão algorítmica, principalmente se for de natureza judicial, deve fornecer explicação suficiente e ser passível de auditoria e controle humanos.

Para a compreensão dessas regras estabelecidas na Resolução 332/2020, é necessário lembrar mais uma vez que o parâmetro para esse ato normativo do Judiciário brasileiro foi a Carta Ética Europeia³⁴¹, cujos princípios foram estabelecidos a partir do Estudo Aprofundado Sobre o uso da IA em Sistemas Judiciais, constantes no Apêndice 1 da Carta, conforme se discorreu detalhadamente no item 3.3 do capítulo anterior deste trabalho.

Também é preciso relembrar que a Carta Ética Europeia foi elaborada em 2018, e que foi fortemente influenciada pelos estudos que existiam na época a respeito de duas aplicações de IA, as ferramentas COMPAS e HART, utilizadas respectivamente por tribunais norte-americanos e pela polícia de Durham, no Reino Unido.

A respeito do COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), destacou-se no item 7.2 do Apêndice 1 da Carta Ética Europeia que se trata de um *software* desenvolvido por uma empresa privada, utilizado em alguns Estados norte-americanos para que juízes possam utilizá-lo como ferramenta auxiliar para avaliar o risco de reincidência de um réu, podendo levar em conta o resultado fornecido pelo algoritmo no momento da sentença.

Consta no Estudo Aprofundado que o algoritmo se baseia em dados de 137 itens, entre questões respondidas pelo réu e informações extraídas de históricos e registros criminais, para classificá-lo numa escala que vai de 1 a 10, de baixo a alto risco de reincidência.

Também consta no Estudo que uma investigação realizada pela ProPublica³⁴², uma organização não-governamental, revelou que havia viés discriminatório no algoritmo, porque atribuiu uma taxa de alto risco de reincidência duas vezes maior às populações afrodescendentes, ao passo que foi atribuído risco mais baixo a outras populações.

Consta que a ONG, ao se dirigir à empresa que criou o software, então denominada Northpointe, não teve acesso aos dados, de modo que “teve de apelar

³⁴¹CEPEJ. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente.** Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 26 jun. 2022.

³⁴²LARSON, Jeff et al. **How we analyzed the COMPAS recidivism algorithm.** ProPublica, 23 maio de 2016. Disponível em: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Acesso em: 26 jun. 2022.

às autoridades públicas para acessar os dados e contratar o seu próprio cientista para examinar o algoritmo”³⁴³.

De forma resumida, a nota 76 do Estudo Aprofundado explica que, de acordo com a ProPublica, quando se verificou anos depois quem de fato reincidiu, “os ‘falsos positivos’ eram majoritariamente negros, enquanto os ‘falsos negativos’ eram majoritariamente brancos”³⁴⁴.

Os peritos que elaboraram Estudo Aprofundado destacaram que “a questão de como conciliar a precisão do algoritmo na detecção de reincidência e a necessidade de evitar efeitos discriminatórios contra populações negras” estava sendo objeto de “intenso debate na literatura”³⁴⁵.

Também destacaram que outro motivo de preocupação é “a falta de transparência nos processos de operação dos algoritmos concebidos por empresas privadas (que reivindicam propriedade intelectual)”, abrindo-se outro debate, sobre a legitimidade de uma empresa privada, “sem qualquer controle institucional, para arbitrar entre dois requisitos opostos: o de defender a sociedade, por um lado, e o de respeitar os direitos dos indivíduos, por outro”.

Na nota 77³⁴⁶ do Estudo Aprofundado para a Carta Ética, os peritos ainda fizeram expressa referência ao caso *Wisconsin v. Loomis*, que se tornou conhecido por fazer chegar, à Suprema Corte dos Estados Unidos, o inconformismo do réu Eric Loomis quanto ao julgamento que recebeu, levando em conta as conclusões que o software COMPAS chegou a respeito do risco de reincidência dele.

Loomis submeteu duas questões à Suprema Corte dos EUA³⁴⁷: (i) se viola o direito constitucional ao devido processo legal o fato de um tribunal de primeira instância confiar na avaliação de risco de um software ao qual o réu não tem acesso por pertencer a uma empresa privada, de modo que o réu não tem como contestar a precisão e validade científica do resultado da avaliação de risco de reincidência realizado pelo COMPAS; (ii) se viola o direito constitucional ao devido processo legal

³⁴³CEPEJ. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente.** Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 26 jun. 2022.

³⁴⁴Ibidem.

³⁴⁵Ibidem.

³⁴⁶Ibidem.

³⁴⁷SCOTUS BLOG. **Loomis v. Wisconsin.** Disponível em: <https://www.scotusblog.com/case-files/cases/loomis-v-wisconsin/>. Acesso em: 26 jun. 2022.

o fato de serem levadas em consideração questões de gênero e raça, pelo *software*, ao formular os critérios de avaliação de risco.

Em 2017, a Suprema Corte dos EUA se negou a conhecer o *writ of certiorari* de Eric Loomis³⁴⁸, no seu poder de discricionariedade absoluta. Como explica Vitor Mello a respeito de tal instituto do Direito norte-americano, a Suprema Corte dos EUA “historicamente tem uma preocupação de firmar-se como um tribunal de superposição e só analisar causas de grande relevância e repercussão político-institucional”³⁴⁹.

Ficou valendo, então, a decisão da instância inferior, a Suprema Corte de Wisconsin, de 13/07/2016, que considerou válido o uso de avaliação de risco algorítmica pelo tribunal de primeira instância, sem que isso significasse ofensa ao devido processo legal, ainda que a metodologia utilizada para produzir o resultado não tivesse sido disponibilizada nem ao tribunal nem ao réu³⁵⁰.

A Corte entendeu que bastava emitir um alerta aos juízes sobre 5 aspectos das avaliações de risco algorítmicas pelo COMPAS:

- A empresa criadora do *software* invocou a proteção à propriedade intelectual para manter em sigilo as informações sobre como são ponderados os fatores de risco e como as pontuações são determinadas;
- As pontuações utilizam dados de grupos, que identificam grupos de infratores de alto risco;
- Havia estudos indicando a possibilidade de haver risco desproporcional em relação a certos grupos, tidos como de alto risco de reincidência;
- As amostras levavam em conta dados nacionais, não havendo cruzamento de dados com amostras de Wisconsin, de modo que a ferramenta necessitava ser continuamente atualizada para atingir maior precisão em relação à população local;

³⁴⁸SCOTUS BLOG. **Loomis v. Wisconsin**. Disponível em: <https://www.scotusblog.com/case-files/cases/loomis-v-wisconsin/>. Acesso em: 26 jun. 2022.

³⁴⁹MELLO, Vitor Tadeu Carramão. A repercussão geral e o *writ of certiorari*: breve diferenciação. **Revista da SJRJ**, Rio de Janeiro, n. 26, p. 139-146. Disponível em: <https://www.ifri.jus.br/sites/default/files/revista-sjrj/arquivo/32-149-1-pb.pdf>. Acesso em: 26 jun. 2022.

³⁵⁰FINDLAW. **State v. Loomis**. Disponível em: <https://caselaw.findlaw.com/wi-supreme-court/1742124.html>. Acesso em: 26 jun. 2022.

- O COMPAS não foi desenvolvido para ser usado em sentenças, e sim para o âmbito da execução penal, útil à supervisão e análise de liberdade condicional.

Mesmo antes de vir a público a investigação da ProPublica em 2018, as premissas desse julgamento de 2016 e as cinco recomendações da Suprema Corte de Wisconsin foram criticadas pela comunidade acadêmica.

Um artigo de 2017 publicado na *Harvard Law Review* apontou o contrassenso da pressão exercida sobre os juízes para usar ferramentas tecnológicas objetivas, que em tese deveriam reduzir erros e garantir mais precisão, para depois ser admitido por uma corte superior que havia sim problemas no COMPAS, de modo que a recomendação “sugere que os juízes devem fazer uma verificação de viés em uma ferramenta projetada para corrigir vieses dos juízes”³⁵¹.

Segundo o artigo, em vez de proibir avaliações de risco algorítmicas de metodologia secreta, exigir mais informações da empresa que criou o *software* ou restringir o uso da ferramenta até que houvesse estudos mais conclusivos, simplesmente se permitiu o uso, priorizando a quantidade e não a qualidade das informações³⁵².

Também em 2017, foi publicado um artigo na revista da ACM – *Association for Computing Machinery*, em que o cientista de dados Keith Kirkpatrick apontou que o problema não era propriamente o algoritmo, mas os dados utilizados no desenvolvimento do COMPAS. Para alimentar *softwares* desse tipo, são usados “dados de crimes relatados, que geralmente se concentram em áreas fortemente policiadas, distorcendo as estatísticas para super-representar os pobres e as minorias”, resultando em ferramentas que “são inherentemente tendenciosas”³⁵³.

Todas essas discussões eram cientes em 2018 quando a Carta Ética Europeia foi formulada, mas a posição foi de otimismo em relação ao futuro, tanto que

³⁵¹HARVARD LAW REVIEW. State vs. Loomis. **Harvard Law Review**, v. 130, n. 5, p. 1530-7, 2017. Disponível em: <https://harvardlawreview.org/2017/03/state-v-loomis/>. Acesso em: 26 jun. 2022.

³⁵²Ibidem.

³⁵³KIRKPATRICK, Keith. It's not the algorithm, it's the data. **Communications of the ACM**, v. 60, n. 2, p. 21-23, 2017. Disponível em: <https://cacm.acm.org/magazines/2017/2/212422-its-not-the-algorithm-its-the-data/fulltext>. Acesso em: 26 jun. 2022.

na nota 77 consta que um excerto da decisão da Suprema Corte de Wisconsin “poderia servir de inspiração”³⁵⁴:

É importante considerar que ferramentas como o COMPAS continuam a mudar e a evoluir. As preocupações que hoje abordamos podem muito bem ser atenuadas no futuro. Cabe ao sistema de justiça criminal reconhecer que, nos próximos meses e anos, dados adicionais de pesquisa estarão disponíveis. Podem ser desenvolvidas ferramentas diferentes e melhores. À medida que os dados mudam, o nosso uso de ferramentas baseadas em evidências também terá de mudar. O sistema de justiça deve acompanhar a investigação e avaliar continuamente a utilização destes instrumentos.

Além do debate que existia a respeito do COMPAS, uma outra ferramenta de IA utilizada no sistema criminal, também apontada como possivelmente discriminatória nos estudos que embasaram a Carta Ética Europeia, foi a HART, acrônimo de *Harm Assessment Risk Tool*.

Em breve síntese, trata-se de uma ferramenta de avaliação de risco desenvolvida por pesquisadores da Universidade de Cambridge para ser usada pela Polícia de Durham, Reino Unido, com a finalidade de classificar em baixo, moderado ou alto o risco de reincidência de infratores, de modo que aqueles que fossem enquadrados na categoria de risco moderado teriam uma alternativa à acusação, que era o encaminhamento a um programa de reabilitação³⁵⁵.

Acontece que, conforme apurado em 2018 pela organização não-governamental britânica *Big Brother Watch* (formada por pesquisadores que se dedicam à defesa das liberdades civis, sobretudo a privacidade, diante da vigilância estatal), a ferramenta HART, para fazer a análise preditiva, levava em conta um conjunto de 34 categorias de dados, incluindo não só informações como idade, sexo e histórico criminal do infrator, mas também dois tipos de código postal³⁵⁶.

Segundo apurado por pesquisadores da referida ONG, um desses códigos se originava de dados de outra ferramenta, Mosaic, fornecida pela empresa de análise

³⁵⁴CEPEJ. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente.** Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 26 jun. 2022.

³⁵⁵OSWALD, Marion et al. Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality. **Information & Communications Technology Law**, v. 27, n. 2, p. 223-250, 2018. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/13600834.2018.1458455>. Acesso em: 11 jun. 2022.

³⁵⁶BIG BROTHER WATCH. **A closer look at Experian big data and artificial intelligence in Durham Police.** 6 abr. 2018. Disponível em: <https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/>. Acesso em: 11 jun. 2022.

de crédito Experian, com base em segmentação sociogeográfica. Ou seja, cruzava dados sobre gênero, sexualidade, raça, etnia e classe social com dados sobre localização e moradia, para aumentar as inferências, baseando-se na suposta percepção de que as pessoas se comportam de modo diferente a depender de onde residem – fato que foi apontado como prejudicial aos moradores de bairros mais pobres³⁵⁷.

Destarte, além de questionamentos sobre a justiça, razoabilidade e proporcionalidade desse tipo de inferência à luz de princípios éticos e direitos humanos fundamentais³⁵⁸, também foi apontado que a ferramenta HART podia levar a um enviesamento com risco de repetição infinita, ao se retroalimentar com novos dados produzidos a partir de conclusões anteriores possivelmente enviesadas.

Dado o contexto acima delineado, do que se sabia em 2018 a respeito das ferramentas COMPAS e HART quando a Carta Ética Europeia foi elaborada, é possível perceber que as questões suscitadas influenciaram o posicionamento brasileiro de cautela, adotado na Resolução 332/2020, em diversos dispositivos.

Por exemplo, depreende-se que a Resolução 332/2020 buscou contornar os problemas de sigilo em razão da proteção à propriedade intelectual ocorridos com o COMPAS, visto que no art. 24 ficou estabelecido o uso preferencial de *software* de código aberto para os modelos de IA do Judiciário, de modo a permitir maior transparência e também possibilitar a integração e interoperabilidade entre os diversos sistemas utilizados pelos tribunais, bem como o desenvolvimento colaborativo e a cooperação com outras áreas do setor público e com a sociedade civil.

Considerando que as ferramentas controvertidas eram usadas na seara da persecução criminal, houve expresso desestímulo ao uso de modelos de IA em matéria penal, conforme o art. 23 da Resolução 332/2020, principalmente no tocante à sugestão de modelos preditivos de decisões.

Contudo, no § 1º do art. 23 permitiu-se a utilização de ferramentas de IA para tarefas mais burocráticas de cálculo de penas, de prescrição, análise de reincidência, elaboração de mapeamentos, classificações e triagem para fins de

³⁵⁷Ibidem.

³⁵⁸BURGESS, Matt. **UK police are using AI to inform custodial decisions – but it could be discriminating against the poor.** Wired. 1 mar. 2017. Disponível em: <https://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit>. Acesso em: 11 jun. 2022.

gestão de acervo processual. Ainda assim, quanto à análise de reincidência, ficou estabelecido no § 2º do art. 23 que modelos de IA sobre o assunto não devem levar a uma conclusão mais prejudicial ao réu do que a que um magistrado teria sem auxílio da ferramenta.

Além disso, como regra geral, a Resolução 332/2020 não exige que os tribunais peçam autorização prévia ao CNJ para iniciar o desenvolvimento de modelos de IA – basta apenas comunicar, de acordo com o art. 22 – à exceção de modelos que usem técnicas de reconhecimento facial, que só podem ser implementados após expressa autorização do CNJ, nos termos do § 2º do mesmo dispositivo.

Cabe pontuar que a Carta Ética Europeia não fez qualquer referência às discussões científicas em torno das técnicas de reconhecimento facial, porém esse tema foi tratado na Resolução brasileira tendo em vista que, ao tempo de sua elaboração, já eram conhecidas pesquisas acadêmicas importantes sobre sérios problemas de acurácia nessa tecnologia.

Um dos referenciais teóricos mais citados sobre o tema é o estudo *Gender shades*³⁵⁹, de 2018, das cientistas de computação Joy Buolamwini e Timnit Gebru, em desdobramento às descobertas de Buolamwini na sua tese³⁶⁰ de doutorado no MIT - *Massachusetts Institute of Technology*, que avaliou a precisão de ferramentas de reconhecimento facial de grandes empresas como IBM e Microsoft, detectando significativos problemas de acurácia quanto a gênero, tons de pele e interseção de tons de pele e gênero, em prejuízo de faces negras e/ou femininas.

Cerca de dois anos depois, o estudo se tornou conhecido do grande público pelo documentário *Coded Bias*³⁶¹, lançado pela Netflix em 2020, no qual a cientista Joy Buolamwini, que é uma mulher negra de pele muito escura, narra quando percebeu que os sistemas simplesmente não classificavam seu rosto, a não ser que usasse uma máscara branca.

³⁵⁹BUOLAMWINI, Joy; GEBRU, Timnit Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In: **Conference on fairness, accountability and transparency**. PMLR, 2018. p. 77-91. Disponível em: http://proceedings.mlr.press/v81/buolamwini18a.html?mod=article_inline. Acesso em: 12 jul. 2022.

³⁶⁰BUOLAMWINI, Joy Adowaa. **Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers**. 2017. Tese de Doutorado. Massachusetts Institute of Technology. Disponível em: <https://dspace.mit.edu/handle/1721.1/114068>. Acesso em: 12 jul. 2022.

³⁶¹NETFLIX. **Coded bias**. 2020. Disponível em: <https://www.netflix.com/br/title/81328723>. Acesso em: 12 jul. 2022.

Na raiz do problema, além das questões relacionadas à transparência algorítmica, apontou-se a falta de diversidade das equipes desenvolvedoras, com a preponderância de homens brancos na área de TI, conforme foi visto na seção anterior deste capítulo, como um fator ligado à discriminação algorítmica. Reproduzindo sua visão de mundo na máquina, treinaram o modelo com muito mais imagens à sua própria semelhança.

Dada a importância da pesquisa de Buolamwini e Gebru, ao apresentaram um método para verificar viés algorítmico de reconhecimento facial e conjuntos de dados de grupos fenotípicos diversos, o *MIT Media Lab* mantém uma página na Internet dedicada ao estudo *Gender shades*, onde é possível consultar um vídeo explicativo, o artigo científico, o conjunto de dados analisados e os resultados, incluindo a constatação de que os sistemas de classificação pesquisados retornaram taxas de erro de até 34,7% para mulheres de pele mais escura, quando o máximo de taxa de erro para homens de pele clara foi muito menor, apenas 0,8%³⁶².

Joy Buolamwini fundou uma organização não governamental intitulada *Algorithmic Justice League*³⁶³ (Liga da Justiça Algorítmica), dedicada à pesquisa e conscientização da sociedade civil, empresas e governos quanto ao uso responsável da IA. Em maio de 2019, junto com diversos especialistas, Buolamwini foi ouvida no Congresso dos Estados Unidos em audiência pública³⁶⁴ específica sobre reconhecimento facial, na qual houve posicionamentos tão duros diante dos riscos que se chegou até a pedir moratória da tecnologia enquanto não houver uma regulação sobre o assunto – ou seja, tal qual aconteceu em 1975 com a tecnologia sobre DNA recombinante, mencionada no primeiro capítulo deste trabalho.

Em junho de 2020, a empresa IBM, uma das que foram pesquisadas por Buolamwini, depois de tentar sem sucesso corrigir o problema do sistema (mesmo depois de treiná-lo com dados de mais de 1 milhão de rostos diferentes), anunciou em carta formal a senadores do Congresso norte-americano que decidiu interromper o fornecimento de *software* de reconhecimento facial de uso geral, o que foi considerado

³⁶²MIT MEDIA LAB. **Gender shades**. 2018. Disponível em: <http://gendershades.org/>. Acesso em: 12 jul. 2022.

³⁶³AJL. **Algorithmic Justice League**. Disponível em: <https://www.ajl.org/>. Acesso em: 13 jul. 2022.

³⁶⁴QUACH, Katyanna. **We listened to more than 3 hours of US Congress testimony on facial recognition so you didn't have to go through it.** The Register. 22 maio 2019. https://www.theregister.com/2019/05/22/congress_facial_recognition/. Acesso em: 13 jul. 2022.

“o movimento mais ousado já feito por uma grande empresa de tecnologia para repudiar o uso discriminatório”³⁶⁵.

A carta, com data de 08/06/2020 e assinada pelo CEO da IBM, Arvind Krishna, faz referência ao contexto da trágica morte de George Floyd³⁶⁶ para o gesto de ação concreta por políticas de tecnologia responsável, dizendo que “a tecnologia pode aumentar a transparência e ajudar a polícia a proteger as comunidades, mas não deve promover discriminação ou injustiça racial”. Destaca-se também o trecho da carta que resume os riscos³⁶⁷:

A IBM se opõe firmemente e não tolerará o uso de qualquer tecnologia, incluindo tecnologia de reconhecimento facial oferecida por outros fornecedores, para vigilância em massa, perfil racial, violações de direitos humanos e liberdades básicas ou qualquer finalidade que não seja consistente com nossos valores e Princípios de Confiança e Transparência. Acreditamos que agora é a hora de iniciar um diálogo nacional sobre se e como a tecnologia de reconhecimento facial deve ser empregada pelas agências policiais domésticas. A Inteligência Artificial é uma ferramenta poderosa que pode ajudar a aplicação da lei a manter os cidadãos seguros. Mas os fornecedores e usuários dos sistemas de IA têm a responsabilidade compartilhada de garantir que a IA seja testada quanto ao viés, particularmente quando usada na aplicação da lei, e que esse teste de viés seja auditado e relatado.

O impacto real do trabalho de pesquisa científica e conscientização sobre os riscos da IA de Joy Buolamwini, por meio da *Algorithmic Justice League*, faz dela uma referência importante, a ponto de a cientista figurar entre os 50 maiores líderes do mundo na lista de 2019 organizada pela revista *Fortune*, segundo a qual “ninguém fez mais que a cientista da computação Joy Buolamwini para chamar a atenção para o viés da IA”³⁶⁸.

Em 2021, a revista *Fast Company*, especializada em inovação, incluiu a *Algorithmic Justice League* entre as 10 empresas mais inovadoras do mundo na área de IA, por seu trabalho em “auditar sistemas para garantir que estejam livres de

³⁶⁵MEYER, David. **IBM pulls out of facial recognition, fearing racial profiling and mass surveillance.** Fortune. 9 jun. 2020. Disponível em: <https://fortune.com/2020/06/09/george-floyd-ibm-exits-facial-recognition-bias-human-rights/>. Acesso em: 13 jul. 2022.

³⁶⁶BBC. **Caso George Floyd: morte de homem negro filmado com policial branco com joelhos em seu pescoço causa indignação nos EUA.** 27 maio 2020. Disponível em: <https://g1.globo.com/mundo/noticia/2020/05/27/caso-george-floyd-morte-de-homem-negro-filmado-com-policial-branco-com-joelhos-em-seu-pescoco-causa-indignacao-nos-eua.ghtml>. Acesso em: 13 jul. 2022.

³⁶⁷IBM. **IBM CEO's Letter to Congress on Racial Justice Reform.** 8 jun. 2020. Disponível em: <https://www.ibm.com/policy/facial-recognition-sunset-racial-justice-reforms/>. Acesso em 13 jul. 2022.

³⁶⁸FORTUNE. **50 world's greatest leaders.** 2019. Disponível em: <https://fortune.com/worlds-greatest-leaders/2019/joy-buolamwini/>. Acesso em: 13 jul. 2022.

preconceitos raciais, de gênero e outros”, de tal maneira que convenceu “gigantes como Amazon, IBM e Microsoft a conter o desenvolvimento de algoritmos de reconhecimento facial”³⁶⁹.

A esse movimento de organizações independentes para iniciativas e pesquisas científica de ponta em IA, por mais transparência algorítmica e combate aos vieses discriminatórios, juntou-se a cientista da computação Timnit Gebru, que é cofundadora da *Black in AI*³⁷⁰, organização não governamental dedicada a fomentar a presença de pessoas negras na área de IA.

Gebru foi coautora do estudo *Gender Shades* e também figura entre os 50 maiores líderes do mundo na lista da revista *Fortune*, edição 2021, descrita como “uma pesquisadora respeitada e uma das mulheres negras mais proeminentes que trabalham em inteligência artificial”, que “incansavelmente fez alertas para os riscos da tecnologia”, até que seus avisos pareceram “se tornar verdades inconvenientes para seu empregador, o Google”³⁷¹.

Segundo reportagem³⁷² da revista *Time* sobre o imbróglio, Gebru era pesquisadora da Microsoft ao tempo em que participou do estudo *Gender Shades* em 2017 e, logo depois da publicação dessa pesquisa pioneira, foi contratada em 2018 pelo Google para ser colíder de sua equipe de ética em IA.

Nessa condição, junto com outros pesquisadores da equipe de ética em IA do Google, participou de diversas pesquisas científicas e publicou vários artigos acadêmicos, alguns de grande importância para a área de transparência algorítmica, como o artigo “*Model cards for model reporting*”³⁷³, que descreve estruturas de cartões para documentar e organizar as informações e as etapas de desenvolvimento de modelos de IA, facilitando a compreensão e criando melhores condições para auditoria³⁷⁴.

³⁶⁹FAST COMPANY. **The 10 most innovative companies in Artificial Intelligence**. 2021. Disponível em: <https://www.fastcompany.com/90600124/artificial-intelligence-most-innovative-companies-2021>. Acesso em: 13 jul. 2022.

³⁷⁰BLACK IN AI. Black in AI. Disponível em: <https://blackinai.github.io/#/about>. Acesso em: 13 jul. 2022.

³⁷¹FORTUNE. **50 world's greatest leaders**. 2021. Disponível em: <https://fortune.com/worlds-greatest-leaders/2021/timnit-gebru/>. Acesso em: 13 jul. 2022.

³⁷²PERRIGO, Billy. **Why Timnit Gebru isn't waiting for Big Tech to fix AI's problems**. Time. 18 jan. 2022. Disponível em: <https://time.com/6132399/timnit-gebru-ai-google/>. Acesso em: 13 jul. 2022.

³⁷³MITCHELL, Margaret et al. Model cards for model reporting. In: **Proceedings of the conference on fairness, accountability, and transparency**. 2019. p. 220-229. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3287560.3287596>. Acesso em: 13 jul. 2022.

³⁷⁴GOOGLE. Model cards. Disponível em: <https://modelcards.withgoogle.com/about>. Acesso em: 13 jul. 2022.

Acontece que, no final de 2020, de acordo com a reportagem da revista *Time*, Timnit Gebru e membros da equipe escreveram um artigo em que teceram diversos questionamentos a respeito da ética de modelos de IA mais potentes que o Google estava desenvolvendo para reprodução da linguagem humana, tanto em termos de riscos a pessoas como quanto à sustentabilidade e impactos ambientais.

Então se estabeleceu, segundo consta na reportagem, um conflito de interesses entre o lucro, por um lado, e o uso ético e seguro da IA, por outro, visto que segundo Gebru o artigo foi censurado pelo chefe, havendo uma troca de *e-mails* que documentou o impasse. Gebru teria sido pressionada a desistir de publicar o artigo ou retirar do texto seu nome e o dos colegas da equipe de ética em IA, para que o trabalho não fosse ligado ao Google, e a resistência dela em se submeter à exigência culminou em demissão.

De acordo com a reportagem, a forma como tudo aconteceu “desencadeou uma tempestade no mundo da IA”, pois o Google “parecia ter expulsado uma das mais respeitadas pesquisadoras de IA ética do mundo depois que ela criticou alguns de seus trabalhos mais lucrativos”³⁷⁵, e “a reação foi feroz”, a exemplo de uma carta pública de protesto assinada por 2.695 funcionários do Google e 4.302 apoiadores acadêmicos, da indústria de tecnologia e da sociedade civil³⁷⁶.

Além dessa carta de protesto, segundo publicado na *MIT Technology Review*, membros do Congresso dos EUA enviaram uma carta formal ao Google cobrando esclarecimentos sobre as circunstâncias da demissão, visto que “a IA ética é o campo de batalha para o futuro dos direitos civis” e que “não podemos mitigar o viés algorítmico se impedirmos aqueles que procuram pesquisá-lo e estudá-lo”³⁷⁷.

Dias após o impasse entre o Google e Gebru se tornar público, a *MIT Technology Review* noticiou³⁷⁸ que, sob a condição de não publicar o texto do artigo

³⁷⁵PERRIGO, Billy. **Why Timnit Gebru isn't waiting for Big Tech to fix AI's problems**. *Time*. 18 jan. 2022. Disponível em: <https://time.com/6132399/timnit-gebru-ai-google/>. Acesso em: 13 jul. 2022.

³⁷⁶GOOGLE WALKOUT FOR REAL CHANGE. **Standing with Dr. Timnit Gebru - #ISupportTimnit #BelieveBlackWomen**. Medium. 3 dez. 2020. Disponível em: <https://googlewalkout.medium.com/standing-with-dr-timnit-gebru-isupporttimnit-believeblackwomen-6dadcd300d382>. Acesso em: 13 jul. 2022.

³⁷⁷HAO, Karen. **Congress wants answers from Google about Timnit Gebru's firing**. *MIT Technology Review*. 17 dez. 2020. Disponível em: <https://www.technologyreview.com/2020/12/17/1014994/congress-wants-answers-from-google-about-timnit-gebrus-firing/>. Acesso em: 13 jul. 2022.

³⁷⁸HAO, Karen. **We read the paper that forced Timnit Gebru out of Google. Here's what it says**. *MIT Technology Review*. 4 dez. 2020. Disponível em: <https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/>. Acesso em: 13 jul. 2022.

questionado, teve acesso a uma cópia que obteve junto a uma das coautoras, Emily Bender, a qual também forneceu algumas declarações gerais sobre o artigo que podiam ser noticiadas, segundo seu entendimento.

E, ao fazê-lo, destacou o valor da liberdade acadêmica que possuía por ser uma pesquisadora ligada a uma instituição de ensino, não a uma empresa, sendo professora titular de linguística computacional da Universidade de Washington. Em contraste, 4 dos 6 coautores do artigo questionado eram pesquisadores do Google e, portanto, funcionários da empresa – cujos nomes Bender não quis revelar para não haver repercussões.

Conforme a reportagem da *MIT Technology Review* a partir do material fornecido por Emily Bender, o título do artigo (aqui em tradução livre) era “Sobre os perigos dos papagaios estocásticos: os modelos de linguagem podem ser tão grandes?”, que questionava os riscos de modelos de IA treinados em massivas quantidades de dados de texto.

Consta que o artigo apresentava a história do processamento de linguagem natural, um panorama dos 4 principais riscos de grandes modelos de linguagem e sugestões para mitigação, com indicação de pesquisas futuras.

De acordo com a reportagem, os riscos mencionados eram: (i) custos ambientais e financeiros, em termos de emissão de carbono e consumo de energia, para atender ao poder de processamento computacional necessário; (ii) riscos de homogeneização da linguagem da IA gerada, refletindo as práticas culturais dos países e populações mais ricas, marginalizando outras, com risco de aumentar a reprodução de linguagem racista, sexista e abusiva em geral, além dos riscos pela dificuldade de documentar e auditar grandes modelos de linguagem, (iii) esforço de pesquisa mal direcionado, que busca aumentar a precisão da manipulação da linguagem em vez de aumentar a precisão da compreensão e interpretação da linguagem, (iv) riscos relacionados ao potencial de confundir e enganar as pessoas, ao ser possível imitar a linguagem humana com cada vez mais precisão, o que pode ter grande impacto negativo em uso para desinformação³⁷⁹.

³⁷⁹ HAO, Karen. **We read the paper that forced Timnit Gebru out of Google. Here's what it says.** MIT Technology Review. 4 dez. 2020. Disponível em: <https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/>. Acesso em: 13 jul. 2022.

A reportagem da *MIT Technology Review* sobre o artigo e a demissão de Gebru finaliza com uma reflexão da professora Emily Bender, de que atualmente a maior parte dos melhores especialistas em ética da IA está trabalhando nas grandes empresas de tecnologia, onde estão os recursos financeiros em abundância para pesquisa – que apesar de beneficiarem o progresso da ciência, nem sempre dão o direcionamento mais benéfico para a humanidade quando há subjacente o conflito de interesse quanto ao lucro.

Em março de 2021, o artigo “*On the dangers of stochastic parrots: can language models be too big?*”³⁸⁰ acabou sendo publicado em conferência da *Association for Computing Machinery* (ACM), e todas as polêmicas a seu respeito têm tudo para torná-lo um marco nas discussões sobre ética em IA. De forma inusitada para um texto científico, seu título oficial vem acompanhado de um emoji de papagaio (🦜), e na lista de autores traz só quatro nomes aparentes: Emily M. Bender e Angelina McMillan-Major, ambas vinculadas à Universidade de Washington, Timnit Gebru, que visto ter saído do Google aparece vinculada à ONG que cofundou, *Black in AI*, e uma última indicação de coautora aparentemente carregada de ironia, Shmargaret Shmitchell, que aparece vinculada à *The Aether*.

A ironia consistiria no fato de que foram acrescidas as letas “Sh” no início do nome e sobrenome de Margaret Mitchell, a outra líder e fundadora da equipe de ética em IA do Google, que foi demitida em fevereiro de 2021, apenas dois meses depois da saída de Timnit Gebru, sob a alegação de que teria retirado documentos da empresa³⁸¹.

Nesse contexto, cabe a interpretação de que é possível que as letras “Sh”, cujo som é comumente associado a pedido de silêncio, tenham sido usadas como forma de protesto. Quanto à vinculação a *The Aether*, é possível que tenha sido uma referência jocosa, pois se trata de uma dimensão virtual, uma espécie de mundo com biomas, existente no jogo eletrônico *Minecraft*³⁸².

³⁸⁰BENDER, Emily M. et al. *On the dangers of stochastic parrots: can language models be too big?*  In: **Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency**. 2021. p. 610-623. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3442188.3445922>. Acesso em: 13 jul. 2022.

³⁸¹BBC. **Margaret Mitchell: Google fires AI ethics founder**. BBC News. 20 fev. 2021. Disponível em: <https://www.bbc.com/news/technology-56135817>. Acesso em: 13 jul. 2022.

³⁸²MICROSOFT. **Aether dimension**. Disponível em: <https://www.minecraft.net/pt-br/marketplace/pdp?id=bedf5379-7c46-4f26-8428-475179761d46>. Acesso em: 13 jul. 2022.

Na parte final do artigo, nos agradecimentos, consta expressamente que nele trabalharam 7 autores, porém 3 ficaram anônimos porque o empregador exigiu que retirasse os nomes.

Pode-se dizer pela conclusão do artigo *Stochastic Parrots*³⁸³ que, depois de saber sobre os riscos sérios e importantes de grandes modelos de linguagem, de ordem financeira, ambiental e para as pessoas, listados por alguns dos maiores especialistas do mundo, levando à reflexão sobre o que é benéfico ou não para a humanidade, fica a pergunta essencial para nortear o desenvolvimento de qualquer modelo de IA:

Pode ser feito, mas deve ser feito?

A pergunta faz lembrar a reflexão de Klaus Schwab, presidente do Fórum Econômico Mundial, mencionada na introdução deste trabalho, que ao falar sobre a 4^a Revolução Industrial em 2016 disse: “Na verdade, a ciência avança tão rápido que, no momento, as limitações são mais jurídicas, regulamentares e éticas que técnicas”³⁸⁴.

Sobre a liberdade acadêmica para falar sobre ética e transparência da IA, em dezembro de 2021, exatamente um ano depois da demissão do Google, Timnit Gebru inaugurou seu instituto de pesquisa científica independente em IA, *Distributed AI Research Institute* – DAIR³⁸⁵, no qual já se encontra disponível a sequência de seu estudo sobre documentação para fins de transparência algorítmica, tratando-se do artigo “*Datasheets for datasets*”³⁸⁶, um instrumento bastante completo para registro e organização dos dados usados no desenvolvimento de modelos de IA, que vai além dos cartões de modelo antes mencionados.

Diante desses fatos, em mais um paralelo com o que ocorreu no passado com a tecnologia de DNA recombinante, mencionada no item 2.1 deste trabalho, o caso sobre Timnit Gebru e o Google é emblemático de quão lúcidas foram as palavras de Paul Berg, o cientista vencedor do prêmio Nobel que liderou o esforço para

³⁸³BENDER, Emily M. et al. On the dangers of stochastic parrots: can language models be too big?  In: **Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency**. 2021. p. 610-623. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3442188.3445922>. Acesso em: 13 jul. 2022.

³⁸⁴SCHWAB, Klaus. **A quarta revolução industrial**. Tradução Daniel Moreira Miranda. São Paulo: Edipro, 2016. Tradução de: *The Fourth Industrial Revolution*. p. 30.

³⁸⁵DAIR. **Distributed AI Research Institute**. Disponível em: <https://www.dair-institute.org/>. Acesso em: 13 jul. 2022.

³⁸⁶GEBRU, Timnit et. al. Datasheets for datasets. **Communications of the ACM**, v. 64, n. 12, p. 86-92, 2021. Disponível em: <https://dl.acm.org/doi/10.1145/3458723>. Acesso em: 13 jul. 2022.

regulação dos riscos da então nova tecnologia em 1975, ao dizer em 2008 que, para as novas tecnologias do presente e futuro, sua lição principal como veterano era alertar para os conflitos de interesse nas pesquisas científicas, quando são bancadas por empresas, em que o lucro invariavelmente falará mais alto³⁸⁷.

Também cabe resgatar, do primeiro capítulo deste trabalho, a análise feita no item 2.3, em relação ao nono tópico do estudo aprofundado que serviu de base para a Carta Ética Europeia, onde foi dito que o desenvolvimento de IA para uso em sistemas judiciais não é assunto que se possa deixar a cargo exclusivo de empresas privadas, cabendo o desenvolvimento à esfera pública com participação dos próprios operadores do direito, juntamente com os desenvolvedores técnicos, inclusive em parceria com instituições acadêmicas e de pesquisa científica.

Para o Poder Judiciário brasileiro, que não é uma empresa que visa ao lucro mas um Poder essencial ao Estado Democrático de Direito, algumas lições podem ser tiradas dos desdobramentos do conflito ocorrido entre Timnit Gebru e o Google, como a importância de:

- Verificar a aderência a padrões éticos em soluções computacionais adquiridas de empresas privadas;
- (Re)examinar continuamente o sentido e razão de ser das aplicações de IA usadas pelos tribunais;
- Ter um comitê permanente de ética em IA, que seja plural e diverso, e que garanta a liberdade de seus membros de dizer “verdades inconvenientes” quando detectarem problemas em aplicações de IA, sem sofrer represálias;
- Investir em pesquisas judiciais sobre IA aplicada ao sistema de Justiça;
- Firmar parcerias com instituições acadêmicas, para pesquisa científica independente, livre de conflitos de interesses, com financiamento público ou oriundo de organizações sem fins lucrativos;
- Acompanhar o trabalho de organizações não governamentais que se dedicam à promoção da IA benéfica para a humanidade.

³⁸⁷BERG, Paul. Asilomar 1975: DNA modification secured. **Nature**, v. 455, n. 7211, p. 290-291, 2008. Disponível em: <https://www.nature.com/articles/455290a>. Acesso em: 16 nov. 2021.

Além da cautela a respeito de reconhecimento facial, a Resolução 332/2020 em seu art. 19 deixou expressamente consignado que qualquer modelo de IA só pode ser usado em decisões judiciais como uma ferramenta auxiliar, sempre sob a supervisão do magistrado.

Isso significa que no Brasil não há que se falar em robô-juiz, pois não é permitido pela Resolução 332/2020 haver decisões judiciais totalmente autônomas por máquinas, sem que haja análise, correção e um comando expresso de confirmação por parte de um magistrado.

Além disso, de acordo com o art. 19, para o desenvolvimento do modelo de IA auxiliar à decisão judicial, o principal critério para escolha da técnica a ser utilizada é a possibilidade de explicação dos caminhos que levaram ao resultado. Assim, para essa aplicação específica, por mais avançada que seja uma técnica, se resultar em opacidade e não-explicabilidade, não será adequada para os fins pretendidos.

Dado o cenário altamente dinâmico no campo da IA, passados mais de três anos da elaboração da Carta Ética Europeia e cerca de dois anos da Resolução 332/2020, é oportuno verificar o que ocorreu em termos de avanços da pesquisa científica desde então, no que se refere à transparência algorítmica, além das contribuições pioneiras de Margaret Mitchell et al. e Timnit Gebru et al. quanto à documentação das etapas de desenvolvimento de modelos de IA em cartões de modelos³⁸⁸ e registro dos conjuntos de dados³⁸⁹, já referidos linhas acima.

O trabalho de maior destaque encontrado é o de Cynthia Rudin, da Universidade Duke, onde lidera o Laboratório de Aprendizado de Máquina Interpretável, como professora de ciência da computação, engenharia elétrica, engenharia da computação, estatística, matemática, bioestatística e bioinformática³⁹⁰.

Ela venceu o *AAAI Squirrel AI Award for Artificial Intelligence for the Benefit of Humanity 2022*, um prêmio que reconhece o trabalho mais destacado do mundo em termos de IA responsável, ética e benéfica para a humanidade, considerado uma espécie de “Nobel da IA”, porque a recompensa financeira no valor de 1 milhão de

³⁸⁸MITCHELL, Margaret et al. Model cards for model reporting. In: **Proceedings of the conference on fairness, accountability, and transparency**. 2019. p. 220-229. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3287560.3287596>. Acesso em: 13 jul. 2022.

³⁸⁹GEBRU, Timnit et al. Datasheets for datasets. **Communications of the ACM**, v. 64, n. 12, p. 86-92, 2021. Disponível em: <https://dl.acm.org/doi/10.1145/3458723>. Acesso em: 13 jul. 2022.

³⁹⁰DUKE UNIVERSITY. Cynthia Rudin. **Interpretable ML Lab.** Disponível em: <https://users.cs.duke.edu/~cynthia/home.html>. Acesso em: 27 jun. 2022.

dólares só é comparável à do Prêmio Nobel e à do Prêmio Turing, da *Association Computing Machinery*³⁹¹.

O prêmio foi concedido a Cynthia Rudin pela Associação para o Avanço da Inteligência Artificial (a prestigiada sociedade científica dedicada à IA que foi mencionada no início do capítulo anterior desta dissertação), “pelo trabalho científico pioneiro na área de sistemas e IA interpretáveis e transparentes em implantações no mundo real”. A AAAI considerou que a abordagem da pesquisadora “em áreas altamente sensíveis, como justiça social e diagnóstico médico” servia como “modelo para pesquisadores e profissionais”³⁹².

O que torna o trabalho de Cynthia Rudin notável é o fato de ter criado “modelos preditivos que se explicam de maneiras que os humanos podem entender. Embora o código para projetar essas fórmulas seja complexo e sofisticado, elas são pequenas o suficiente para serem escritas em poucas linhas de um cartão de índice”³⁹³.

Há mais de 15 anos, a professora vem se dedicando ao estudo de algoritmos de aprendizado de máquina que possam ser interpretáveis, sobretudo em aplicações de IA consideradas de alto risco, por envolverem questões éticas e por estarem relacionadas a decisões que impactam a vida das pessoas.

Nesse contexto, uma das áreas de pesquisa lideradas por Cynthia Rudin envolve justamente diversas análises sobre o COMPAS.

Em artigo publicado em 2019 na revista *Nature Machine Learning*, em contraponto ao modelo do COMPAS, tido como opaco ou de caixa-preta, devido à impossibilidade de acesso ao código e ao conhecimento dos mecanismos pelos quais chega aos resultados, Cynthia Rudin defende que modelos de IA desse tipo não devem ser usados para apoiar decisões em aplicações de alto risco³⁹⁴.

Acrescenta ainda que de nada adianta criar modelos “explicáveis”, para tentar explicar o que acontece dentro da caixa-preta, porque tais modelos são imprecisos por definição, não havendo como saber se os parâmetros são os mesmos,

³⁹¹AAAI. **Duke computer scientist wins \$1 million artificial intelligence prize, a “new Nobel”.** Disponível em: <https://aaai.org/Pressroom/Releases/release-21-1012.php>. Acesso em: 27 jun. 2022.

³⁹²AAAI. **AAAI Squirrel AI Award for Artificial Intelligence for the Benefit of Humanity.** Disponível em: <https://aaai.org/Awards/squirrel-ai-award.php>. Acesso em: 27 jun. 2022.

³⁹³AAAI. **Duke computer scientist wins \$1 million artificial intelligence prize, a “new Nobel”.** Disponível em: <https://aaai.org/Pressroom/Releases/release-21-1012.php>. Acesso em: 27 jun. 2022.

³⁹⁴RUDIN, Cynthia. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. **Nature Machine Intelligence**, v. 1, n. 5, p. 206-215, 2019.

de modo que “modelos de caixa preta com explicações podem levar a um caminho de decisão excessivamente complicado, propício ao erro humano”³⁹⁵.

A título de exemplo, ela argumenta que o cálculo do COMPAS envolve mais de 130 fatores, e basta 1% de erros tipográficos humanos no momento de inserir esses dados para produzir injustiça na análise do caso de réus. Pelo caminho diferente proposto por Rudin, usando apenas 3 informações – idade, antecedentes e sexo (opcional) – foi possível chegar a resultados precisos, interpretáveis e ainda por cima de código livre, ao contrário do *software* protegido por sigilo comercial³⁹⁶.

Em outro artigo, publicado em 2020 na *Harvard Data Science Review*, ao fazer a engenharia reversa parcial do COMPAS em cotejo com as acusações da ProPublica, a pesquisadora constatou que os resultados desiguais e injustos poderiam vir de diversos fatores, não necessariamente apenas em razão de viés discriminatório racial. Assim, tratamentos injustos pareciam estar ligados de forma importante à inconsistência dos dados coletados em registros criminais³⁹⁷.

Diante de tudo o que vem apurando ao longo de suas pesquisas científicas, Cynthia Rudin propõe, em vez de modelos de caixa-preta e de modelos supostamente explicáveis, a adoção de modelos completamente interpretáveis, totalmente transparentes e simples de demonstrar aos usuários – e provou que é possível desenvolvê-los.

No artigo da revista *Nature Machine Learning* já mencionado, a pesquisadora chega a encorajar formuladores de políticas públicas “a não aceitar modelos de caixa preta sem tentativas significativas de modelos interpretáveis (em vez de explicáveis)”. Ela acrescenta que, apesar dos avanços, ainda há muitos desafios no campo dos modelos interpretáveis, mas conclama os formuladores de políticas públicas à conscientização para que exijam mais esforços dos desenvolvedores no sentido de superar tais desafios, para garantir a segurança e a confiança em modelos de IA de alto risco, evitando o uso de modelos de caixa-preta “quando não for seguro usá-los”³⁹⁸.

³⁹⁵Ibidem.

³⁹⁶Ibidem.

³⁹⁷RUDIN, Cynthia et. al. The age of secrecy and unfairness in recidivism prediction. **Harvard Data Science Review**, 2 (1), 3 2020.

³⁹⁸RUDIN, Cynthia. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. **Nature Machine Intelligence**, v. 1, n. 5, p. 206-215, 2019.

Cynthia Rudin também faz um alerta de que pouco adianta haver regulamentos prevendo “direito à explicação” quando isso não é tecnicamente possível por completo em modelos de caixa-preta. Adverte, ainda, que continuar permitindo modelos desse tipo em áreas de alto risco significa permitir que erros venham a acontecer, em prejuízo da sociedade, porque “continuará a haver problemas ao combinar as previsões do modelo de caixa preta com informações externas ao banco de dados, e erros contínuos de cálculo das entradas do modelo de caixa preta”³⁹⁹.

Mais recentemente, em extenso artigo publicado em 2022, Cynthia Rudin et al. apresentam 5 princípios e 10 desafios técnicos atuais para a IA interpretável, ao mesmo tempo em que criticam uma nova terminologia que entendem equivocada mas que vem se difundindo. Trata-se da “XAI” (*Explanable Artificial Intelligence*), ou IA explicável, que difere e muito da IA interpretável, mas que têm sido confundidas pela comunidade científica⁴⁰⁰.

De forma resumida e adaptada, os 5 princípios gerais da IA interpretável podem ser explicados da seguinte maneira⁴⁰¹:

- **Princípio 1:** A IA interpretável obedece a restrições para permitir que suas previsões ou os dados que utilizou sejam facilmente compreendidos pelos humanos; e, como essas restrições têm relação com o campo de aplicação do modelo, podem ser totalmente diferentes a depender da área a que o modelo se destina (suporte a diagnóstico médico, justiça criminal etc.);
- **Princípio 2:** A IA interpretável, ao demonstrar como funciona, permite que os humanos tomem uma decisão sobre o modelo ser confiável ou não, em vez de simplesmente se resignarem a confiar no resultado que o modelo apresenta, como ocorre com os modelos de caixa-preta;
- **Princípio 3:** Ao contrário do que se pensa, não é preciso sacrificar a precisão para ganhar interpretabilidade da IA, pois na realidade é

³⁹⁹Ibidem.

⁴⁰⁰RUDIN et al. Interpretable machine learning: fundamental principles and 10 grand challenges. **Statistics Surveys**, v. 16, p. 1-85, 2022.

⁴⁰¹Ibidem.

justamente o fato de ser possível compreender a IA que permite alcançar mais precisão;

- **Princípio 4:** Quanto aos dados utilizados na IA interpretável, deve haver contínuo refinamento das métricas de desempenho e de interpretabilidade;
- **Princípio 5:** Para aplicações que envolvam decisões de alto risco, devem ser utilizados sempre que possível modelos de IA interpretáveis e não modelos de IA ditos “explicados” mas que continuam sendo de caixa-preta.

A respeito dos 10 desafios técnicos atuais para a IA interpretável, Cynthia Rudin et al. assim resumem (em tradução livre):

(1) Otimização de modelos lógicos esparsos, como árvores de decisão; (2) Otimização dos sistemas de pontuação; (3) Colocação de restrições em modelos aditivos generalizados para encorajar a dispersão e melhor interpretabilidade; (4) Raciocínio moderno baseado em casos, incluindo redes neurais e correspondência para inferência causal; (5) Desembaraçamento supervisionado completo de redes neurais; (6) Desembaraçamento total ou mesmo parcial não supervisionado de redes neurais; (7) Redução de dimensionalidade para visualização de dados; (8) Modelos de aprendizado de máquina que podem incorporar física e outras restrições generativas ou causais; (9) Caracterização do “conjunto Rashomon” de bons modelos; e (10) Aprendizagem por reforço interpretável⁴⁰².

Ao fazer essa breve menção ao estado da arte na pesquisa científica sobre transparência em IA, o que se pretendeu foi destacar os alertas de uma pesquisadora experiente, renomada e premiada como Cynthia Rudin quanto à diferença entre IA interpretável e IA explicável, quanto à probabilidade de ocorrerem erros no uso de modelos de IA opacos, e quanto à relevância do assunto em se tratando de aplicações de alto risco – sobretudo no campo ora tratado, que é para uso pelo Poder Judiciário.

Note-se que na Resolução 332/2020 em momento algum constou esse tipo de abordagem interpretável, e sim “explicação satisfatória” (art. 8º, VI). No inciso IV do mesmo artigo, consta “possibilidade de identificação do motivo em caso de dano”, de maneira talvez dissociada das reais condições técnicas de se identificarem motivos em modelos de caixa-preta. Contudo, tendo em vista o avanço de pesquisas recentes,

⁴⁰²Ibidem.

que demonstram que há caminhos possíveis para o desenvolvimento de modelos de IA mais seguros, confiáveis e transparentes, pode-se afirmar que será oportuno se o CNJ resolver reavaliar a forma como regulou o tema nesse ponto da transparência.

3.4.5 Riscos pela falta de controle do usuário

Durante o desenvolvimento deste trabalho, elaborou-se uma pesquisa empírica no mês de setembro de 2021, mediante questionário intitulado “Nível de conhecimento da magistratura sobre Inteligência Artificial”, que foi submetido a magistrados dos ramos federal, estadual, trabalhista e militar.

A pesquisa teve apoio da Associação dos Magistrados Brasileiros (AMB), Associação dos Juízes Federais do Brasil (AJUFE) e Associação Nacional dos Magistrados da Justiça do Trabalho (ANAMATRA), sendo solicitado que tais entidades de classe encaminhassem o formulário a seus associados.

O número de 688 respostas coletadas, no universo de 17.988 magistrados do Brasil, de acordo com o número de cargos providos conforme o relatório Justiça em Números 2021⁴⁰³, possibilitou atingir uma amostra com nível de confiança de 99%, e margem de erro de 5%, de acordo com fórmula seguinte:

Figura 7: Fórmula estatística para cálculo da amostra da pesquisa⁴⁰⁴.

$$n = \frac{k^2 qpN}{e^2(N - 1) + k^2 pq}$$

Fonte: <https://www.questionpro.com/pt-br/mobile-diaries.html>

Para o cálculo, (N) representa o número total de magistrados que poderiam responder ao formulário; (e) significa a margem de erro aceitável; (k) representa o nível de confiança; (p) significa a probabilidade de sucesso; e (q) representa a probabilidade de fracasso⁴⁰⁵.

Os resultados da pesquisa foram divulgados em janeiro de 2022, no artigo “Sob controle do usuário: formação dos juízes brasileiros para o uso ético da IA no

⁴⁰³CNJ. **Justiça em Números 2021**. Brasília: CNJ, 2021. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2021/09/relatorio-justica-em-numeros2021-12.pdf>. Acesso em: 28 jun. 2022.

⁴⁰⁴QUESTIONPRO. **Calculadora de amostra de pesquisa**. Disponível em: <https://www.questionpro.com/pt-br/mobile-diaries.html>. Acesso em: 28 jun. 2022.

⁴⁰⁵Ibidem.

Judiciário”, em dossier temático sobre IA, Ética e Epistemologia da Revista Direito Público⁴⁰⁶.

Tendo como premissa que “somente é possível controlar o que se conhece”⁴⁰⁷, o artigo buscou investigar qual é o nível de conhecimento dos magistrados para exercerem o papel de usuário interno controlador da IA, que lhes foi reservado no capítulo VII da Resolução 332/2020, dedicado ao controle do usuário.

Isso porque, de acordo com o art. 17 da Resolução 332/2020, embora os magistrados brasileiros não estejam obrigados a utilizar sistemas de IA nem tampouco precisem ficar vinculados a resultados apresentados por algoritmos, é importante que estejam capacitados para terem condições de supervisionar a ferramenta de IA, caso a utilizem, sabendo que o modelo de IA deve permitir tanto a revisão dos resultados como também dos dados que foram utilizados para se chegar aos resultados.

Além disso, pelo parágrafo único do art. 19 da Resolução 332/2020, ferramentas de IA não podem produzir decisões judiciais autônomas, devendo ser usadas apenas como ferramentas auxiliares, que estão sempre sujeitas à supervisão do magistrado competente.

É preciso ter em mente que há magistrados mais iniciados em IA que participam das equipes desenvolvedoras dos modelos, como membros da área de negócio, enquanto há magistrados que apenas utilizam as ferramentas colocadas em produção. Mas estes, por terem contato direto com os resultados produzidos pelo algoritmo, em seu cotidiano forense, têm papel fundamental para perceber erros e incidentes, sinalizando o fato aos desenvolvedores e ao próprio tribunal.

Com efeito, o art. 27 da Resolução 332/2020 estabelece que os órgãos do Poder Judiciário devem informar ao CNJ todos os registros de eventos adversos no uso da IA. O dispositivo em tela, aliás, como dito na introdução deste trabalho, é justamente o que se relaciona diretamente com o problema da presente pesquisa.

Até aqui a dissertação procurou fazer um levantamento dos riscos mais importantes envolvidos no uso da IA no sistema de Justiça brasileiro, verificando de onde podem surgir erros e incidentes. No próximo e último capítulo, buscará

⁴⁰⁶PRADO, Eunice M.B.; MÜNCH, L. A. C.; VILLARROEL, M. A. C. “Sob controle do usuário”: formação dos juízes brasileiros para o uso ético da IA no judiciário. *Direito Público*, v. 18, n. 100, 2022. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/6021>. Acesso em: 28 jun. 2022.

⁴⁰⁷Ibidem.

responder de que forma o registro e a informação dos chamados eventos adversos podem ser mais eficientes e úteis.

Como se demonstrou ao longo deste capítulo, há uma miríade de intrincadas questões envolvendo a utilização de IA no Judiciário, com muitos temas novos e controvertidos que necessitam de maior aprofundamento em pesquisas científicas, e que são complexos até para os especialistas.

Percebe-se que a Resolução 332/2020 atribuiu grandes responsabilidades de controladores aos magistrados, como se fossem bons convededores dos meandros da IA, mas a pesquisa empírica demonstrou que, na realidade, a esmagadora maioria ainda não está preparada para exercer tão difícil mister.

Por exemplo, só 9,3% dos magistrados responderam que sabiam explicar todos os conceitos básicos de “algoritmo”, “aprendizado de máquina”, “aprendizado profundo” e “redes neurais artificiais”⁴⁰⁸.

Apenas 6,7% dos magistrados disseram que sabiam que a Resolução 332/2020 existia e que conheciam bem seu teor⁴⁰⁹.

Quando foi perguntado quão preparados os magistrados se sentiam para exercerem o controle e a supervisão de sistemas que usem IA como ferramentas auxiliares na elaboração de decisões judiciais, somente 8,7% disseram que se sentiam bem preparados⁴¹⁰.

Para julgar litígios envolvendo aplicação de IA, apenas 9,3% responderam que se sentiam preparados⁴¹¹.

No tocante à capacitação em Escolas Judiciais, apenas 11,5% disseram que haviam feito algum curso específico de IA aplicada à função judicante oferecido pela Escola Judicial do tribunal a que estavam vinculados⁴¹².

Como se pode perceber pelos principais achados da pesquisa, poucos magistrados têm noção de seu próprio papel de usuários internos controladores, a quem cabe exercer a supervisão de modelos de IA utilizados em auxílio a decisões

⁴⁰⁸PRADO, Eunice M.B.; MÜNCH, L. A. C.; VILLARROEL, M. A. C. “Sob controle do usuário”: formação dos juízes brasileiros para o uso ético da IA no judiciário. **Direito Público**, v. 18, n. 100, 2022. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/6021>. Acesso em: 28 jun. 2022.

⁴⁰⁹Ibidem.

⁴¹⁰Ibidem.

⁴¹¹Ibidem.

⁴¹²Ibidem.

judiciais. A grande maioria não tem conhecimento dos princípios éticos para a IA responsável nem sequer sabe conceitos básicos sobre IA.

Nesse cenário, dificilmente o magistrado-controlador terá a necessária dimensão dos riscos envolvidos no uso da IA, como também não estará preparado para identificar, corrigir e informar, por exemplo, possíveis discriminações algorítmicas.

Por tais razões, no referido artigo publicado na Revista Direito Público se ressaltou que, além da necessária capacitação em IA, outros treinamentos como “capacitação em julgamento com perspectiva de gênero, em letramento racial e em proteção de dados são de fundamental importância, cabendo também aos Tribunais proverem essa necessidade”⁴¹³.

Inclusive quanto a esse ponto se mencionou o art. 6.3 dos Princípios de Bangalore de Conduta Judicial⁴¹⁴, quanto ao dever que os magistrados têm de se aprimorar, e que os tribunais têm de proporcionar esse aprimoramento.

Cabe mencionar que a Resolução CNJ nº 423⁴¹⁵, de 05/10/2021, determinou a inclusão de novos temas no conteúdo a ser cobrado em concursos para a magistratura, a exemplo de Direito Digital (incluindo IA) e Direito da Antidiscriminação, mas é preciso que essas disciplinas também sejam fartamente disponibilizadas pelas Escolas Judiciais para os quase 18 mil magistrados que já estão na carreira, em aperfeiçoamento encorajado e talvez até mesmo exigido.

Além disso, é importante chamar a atenção para outro tipo de risco ligado ao magistrado, como usuário interno, supervisor e controlador da IA, que está relacionado à possibilidade de uso de IA para auxiliar na observância a precedentes, visto que o art. 5º da Resolução 332/2020 menciona a aplicação da tecnologia para

⁴¹³PRADO, Eunice M.B.; MÜNCH, L. A. C.; VILLARROEL, M. A. C. “Sob controle do usuário”: formação dos juízes brasileiros para o uso ético da IA no judiciário. **Direito Público**, v. 18, n. 100, 2022. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/6021>. Acesso em: 28 jun. 2022.

⁴¹⁴ONU. **Comentários aos Princípios de Bangalore de Conduta Judicial**/Escritório Contra Drogas e Crime. Tradução Marlon da Silva Malha, Ariane Emílio Kloth. Brasília: Conselho da Justiça Federal, 2008. Tradução de: Commentary on The Bangalore Principles of Judicial Conduct. Disponível em: https://www.unodc.org/documents/lpo-brazil/Topics_corruption/Publicacoes/2008_Comentarios_aos_Principios_de_Bangalore.pdf. Acesso em: 29 jun. 2022.

⁴¹⁵CNJ. **Resolução nº 423, de 5 de outubro de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4147>. Acesso em: 29 jun. 2022.

“buscar garantir a segurança jurídica e colaborar para que o Poder Judiciário respeite a igualdade de tratamento aos casos absolutamente iguais”⁴¹⁶.

É possível tirar o máximo proveito da IA para triagem e classificação de processos judiciais, com verificação comparável de adequação entre casos e precedentes, mas é preciso ter em mente que a função do magistrado é insubstituível, não podendo deixar de cumprir seu dever de analisar todas as circunstâncias do caso concreto – inclusive para verificar se um determinado caso tem nuances que o distinguem de outro que gerou um precedente obrigatório.

Considerando que a aplicação de IA nos tribunais brasileiros ainda é muito recente, há uma lacuna de pesquisas e estudos mais aprofundados que permitam verificar a **influência da ferramenta na disposição de ânimo dos magistrados**, declarada ou evidenciada por sua conduta, entre aqueles que já estão lidando com a ferramenta.

Por exemplo, é possível que a IA tenha o efeito colateral de deixar os magistrados acomodados com resultados automatizados, ou que se deixem influenciar, ainda que inconscientemente, pela previsibilidade de reforma ou confirmação de suas decisões por instâncias recursais, acarretando julgamentos diferentes do que fariam se não houvesse a ferramenta, ou resultando em falta de oxigenação e engessamento da jurisprudência.

Sob outro prisma, cabe mencionar que os magistrados também podem ser instados a fazer um controle mais minucioso de ferramentas de IA, quanto a questões específicas, **mediante provação de usuário externo**, que igualmente tem direito ao controle da IA aplicada a seu caso, só que exercido de sua posição exógena, da forma que lhe é possível, mediante provação dirigida a quem tem acesso ao controle direto da ferramenta.

Foi o que aconteceu no exemplo norte-americano do caso *Wisconsin versus Loomis*, mencionado na seção anterior, em que a defesa do réu Eric Loomis contestou as conclusões do software COMPAS, apresentando a tese de que havia ocorrido tratamento injusto e discriminatório⁴¹⁷.

⁴¹⁶CNJ. Resolução nº 332, de 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 29 jun. 2020.

⁴¹⁷FINDLAW. *State v. Loomis*. Disponível em: <https://caselaw.findlaw.com/wi-supreme-court/1742124.html>. Acesso em: 26 jun. 2022.

Um exemplo mais recente e importante, porque também serve para demonstrar outra possibilidade relevante de controle externo da IA aplicada ao sistema de Justiça – **o controle pela sociedade civil organizada** – é o caso *José Rodriguez versus Massachusetts Parole Board*, que está em tramitação perante o Supremo Tribunal Judicial de Massachusetts, nos Estados Unidos.

Em tal caso, além da apelação do próprio réu⁴¹⁸, apresentada em novembro de 2021, há a participação, na qualidade de *amicus curiae*⁴¹⁹, requerida em fevereiro de 2022, da *Electronic Privacy Information Center* (EPIC), uma organização não governamental norte-americana com sede em Washington D.C., dedicada ao uso responsável da tecnologia, “criada em 1994 para proteger a privacidade, a liberdade de expressão e os valores democráticos na era da informação”⁴²⁰.

Na questão levada a julgamento, José Rodriguez é um condenado à prisão perpétua por um crime que cometeu quando ainda era uma criança.

O contexto da situação envolve o fato de que os Estados Unidos são o único país do mundo a permitir que crianças e adolescentes sejam condenadas à prisão perpétua sem liberdade condicional – a chamada *juvenile life without parole*. Porém, em 2012, ao julgar o caso *Miller versus Alabama*, a Suprema Corte dos Estados Unidos entendeu que essa pena é inconstitucional por ser cruel e desumana, ao vedar a chance de liberdade condicional, haja vista não levar em conta que crianças e adolescentes são pessoas ainda em desenvolvimento e que podem se reabilitar com apoio e orientação⁴²¹.

Cerca de quatro anos depois, em 2016, a Suprema Corte dos Estados Unidos, ao julgar o caso *Montgomery versus Louisiana*, tornou retroativa a decisão do caso *Miller versus Alabama*, passando a permitir a revisão da situação de todas as pessoas condenadas à prisão perpétua nos Estados Unidos por crimes praticados quando ainda eram menores de idade. O critério mais importante para o benefício é que não haja probabilidade razoável de que o réu, uma vez solto, volte a delinquir⁴²².

⁴¹⁸EPIC. **Brief of appellant.** José Rodriguez v. Massachusetts Parole Board. Disponível em: https://epic.org/wp-content/uploads/2022/02/SJC-13197_01_Appellant_Rodriguez_Brief.pdf. Acesso em: 1 jul. 2022.

⁴¹⁹EPIC. **Brief of Amicus Curiae Electronic Privacy Information Center (EPIC) in support of appellant.** SJC-13197, 14 fev. 2022. Disponível em: <https://epic.org/wp-content/uploads/2022/02/EPIC-Amicus-Rodriguez-Parole-Board.pdf>. Acesso em: 1 jul. 2022.

⁴²⁰EPIC. **About us.** Disponível em: <https://epic.org/about/>. Acesso em: 1 jul. 2022.

⁴²¹MARTINELLI, Julieta. **As nation changes stance on life sentences for minors, Tennessee's life juvenile lifers are left behind.** Latino Rebels. 6 maio 2022. Disponível em: <https://www.latinorebels.com/2022/05/06/tennesseejuvenilelifers/>. Acesso em: 1 jul. 2022.

⁴²²Ibidem.

Ocorre que José Rodriguez, estando preso há mais de 35 anos e tendo cumprido todos os demais requisitos para pleitear a liberdade condicional, teve o pedido negado por ser considerado com alto risco de reincidência por uma ferramenta preditiva de IA, *Level of Service/Case Management Inventory* (LS/CMI). A entidade EPIC, atuando como *amicus curiae*, reclama por transparência, para que seja possível compreender por que a ferramenta chegou a tal conclusão, e aponta sobretudo que o *software* não foi desenvolvido observando dados de menores de idade e especificidades dessa categoria de pessoas, razão pela qual não poderia ter sido aplicado para o caso de Rodriguez⁴²³.

No Brasil, como se sabe, a Constituição proíbe em cláusula pétreia a prisão perpétua (art. 5º, XLVII)⁴²⁴, de modo que o ponto de interesse quanto ao debate levantado pela EPIC no caso de Rodriguez é focado nos **reclamos de transparência algorítmica por uma organização não governamental**, atuando não só em benefício do réu como também de toda a sociedade.

Isso porque a transparência permite conhecer os dados, as técnicas, os códigos, possibilita testes por terceiros e por auditorias independentes para verificação de vieses, cria condições para que a parte prejudicada se defenda, permite que especialistas apresentem sugestões de melhorias e modificações para otimizar a precisão, além de permitir que se descortinem conflitos de interesse em caso de sistemas desenvolvidos por empresas privadas⁴²⁵.

Em 2020, a EPIC também desenvolveu um trabalho relevante para o estudo da aplicação de IA no sistema de Justiça criminal, publicando o relatório “Liberdade em risco”⁴²⁶, com uma visão geral das diferentes ferramentas de IA para avaliação de risco nos 50 Estados norte-americanos, a fim de possibilitar um entendimento mais amplo do contexto de utilização de vários *softwares*, como referencial que serve de ponto de partida para que profissionais e a comunidade acadêmica aprofundem pesquisas sobre o assunto.

⁴²³EPIC. **Rodriguez v. Massachusetts Parole Board.** SJC-13197 (2021). Disponível em: <https://epic.org/documents/rodriguez-v-massachusetts-parole-board/#epic-s-brief>. Acesso em: 1 jul. 2022.

⁴²⁴BRASIL. **Constituição da República Federativa do Brasil de 1988.** Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 1 jul. 2022.

⁴²⁵EPIC. **Rodriguez v. Massachusetts Parole Board.** SJC-13197 (2021). Disponível em: <https://epic.org/documents/rodriguez-v-massachusetts-parole-board/#epic-s-brief>. Acesso em: 1 jul. 2022.

⁴²⁶EPIC. **Liberty at risk: pre-trial risk assessment tools in the U.S. Set. 2020.** Disponível em: <https://epic.org/wp-content/uploads/2022/02/Liberty-At-Risk-Report-FALL-2020-UPDATE.pdf>. Acesso em: 2 jul. 2022.

No relatório, foram listadas as **principais preocupações relacionadas à utilização de ferramentas de IA para avaliação de risco**, a saber: (i) falta de transparência; (ii) falta de responsabilidade; (iii) falta de consentimento informado; (iv) falta de explicabilidade; (v) perpetuação e codificação de padrões de policiamento tendenciosos; (vi) viés no uso das ferramentas baseadas em raça, etnia e idade; (vii) validação insuficiente; (viii) treinamento insuficiente; (ix) erros de *software* e codificação⁴²⁷.

Como **desafios legais para a utilização de ferramentas de IA para avaliação de risco**, o relatório mencionou a dificuldade para garantir o efetivo direito ao devido processo legal com paridade de armas e ampla defesa, além de regras quanto à responsabilidade pela ferramenta⁴²⁸.

Constam no relatório da EPIC, ainda, **sugestões de questões importantes que um usuário deve analisar diante de uma ferramenta de IA para avaliação de risco**, sobre (i) que fatores estão sendo levados em conta pela ferramenta; (ii) qual é a matriz de decisão; (iii) se estão sendo levados em conta fatores como raça, gênero, idade, renda e outros dados sensíveis que possam levar a tratamento desigual e discriminatório; (iv) se a ferramenta é validada com frequência, e se contém dados localizados e representativos; (v) como os dados foram coletados e tratados, e se estão disponíveis para verificação.

Para o sistema de Justiça brasileiro, o tema concernente ao controle e verificação da transparência algorítmica em avaliação de risco é relevante porque, no futuro, podem surgir projetos de modelos de IA para auxiliar na avaliação de risco em áreas como violência doméstica contra a mulher, concessão de fiança e medidas alternativas à prisão, execução penal, concessão de remissão e cumprimento de medidas socioeducativas por atos infracionais praticados por adolescentes, entre outras.

Quanto à preocupação com os dados utilizados em ferramentas de IA para uso no âmbito da persecução penal no Brasil, cabe mencionar que a Lei Geral de Proteção de Dados prevê em seu art. 5º, I a III, três categorias de dados: pessoais, sensíveis e anonimizados. Contudo, o art. 4º, III, “a” e “d”, da LGPD estabelece que a

⁴²⁷EPIC. **Liberty at risk: pre-trial risk assessment tools in the U.S.** Set. 2020. Disponível em: <https://epic.org/wp-content/uploads/2022/02/Liberty-At-Risk-Report-FALL-2020-UPDATE.pdf>. Acesso em: 2 jul. 2022.

⁴²⁸Ibidem.

lei não se aplica ao tratamento de dados realizado para fins exclusivos de segurança pública e atividades de investigação e repressão de infrações penais. Para estes, o § 1º do art. 4º da LGPD prevê que seu tratamento será regido por uma lei específica, que deverá equilibrar as medidas necessárias de atendimento ao interesse público, o devido processo legal e a proteção dos direitos dos titulares dos dados⁴²⁹.

Acontece que essa lei específica ainda não existe. Há apenas um Anteprojeto de Lei de Proteção de Dados para segurança pública e persecução penal⁴³⁰, elaborado por uma comissão de juristas coordenada pelo Ministro Nefi Cordeiro, do Superior Tribunal de Justiça, cujo texto foi entregue à Presidência da Câmara dos Deputados em 05/11/2020 – inclusive com a recomendação de que o CNJ seja o órgão responsável pelo controle dos dados de persecuções penais⁴³¹.

No referido anteprojeto, não há referência à expressão “inteligência artificial”, e sim a “tratamento automatizado de dados pessoais”, estabelecendo um conjunto de regras para proteção de “direitos alinhados às tendências contemporâneas de regulação das decisões automatizadas, como o direito à proteção contra a discriminação e o direito à explicação de processos automatizados”⁴³².

Entre essas regras de proteção, além de diversas medidas de controle dos utilizadores dos dados e controle do acesso aos dados, consta a obrigação de elaborar “relatórios de impacto à proteção de dados pessoais”, quando houver utilização “de dados pessoais sensíveis, sigilosos, ou operações que apresentem elevado risco aos direitos, liberdades e garantias dos titulares de dados”⁴³³.

É provável que, no futuro próximo, quando houver a chamada “LGPD Penal”⁴³⁴, muitas interseções com as regras estabelecidas na Resolução 332/2020

⁴²⁹BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/l13709.htm. Acesso em: 2 jul. 2022.

⁴³⁰BRASIL. Anteprojeto de Lei de Proteção de Dados para segurança pública e persecução penal. Disponível em: <https://www2.camara.leg.br/atividade-legislativa/comissoes/grupos-de-trabalho/56a-legislatura/comissao-de-juristas-dados-pessoais-seguranca-publica/documentos/outros-documentos/DADOSAnteprojetocomissaoprotecaodadossegurancapersecucaoFINAL.pdf>. Acesso em: 2 jul. 2022.

⁴³¹BRASIL. Rodrigo Maia recebe anteprojeto para controle de dados de investigações criminais. Câmara dos Deputados - Agência Câmara de Notícias. 5 nov. 2020. Disponível em: <https://www.camara.leg.br/noticias/705293-rodrigo-maia-recebe-anteprojeto-para-controle-de-dados-de-investigacoes-criminais/>. Acesso em: 2 jul. 2022.

⁴³²Ibidem.

⁴³³Ibidem.

⁴³⁴COSTA, Eduarda; REIS, Carolina. Histórico da LGPD Penal: o que foi feito até aqui e quais são os próximos passos? Laboratório de Políticas Públicas e Internet. 16 abr. 2021. Disponível em: <https://lapin.org.br/2021/04/16/lgpd-penal-o-que-foi-feito-ate-aqui-e-quais-sao-os-proximos-passos/>. Acesso em: 2 jul. 2022.

sejam verificadas, quanto a outras diretrizes e métodos de controle de dados, recomendando adequação da Resolução 332/2020 pelo CNJ.

Como se percebe, o controle de ferramentas de IA pelo usuário, incluindo preocupação com os dados, com a transparência e com os resultados apresentados pela máquina, é tarefa de considerável complexidade que requer uma **visão sistêmica**, desafiando treinamento e capacitação para os usuários internos quanto a vários temas diferentes, mas complementares entre si.

Do ponto de vista dos usuários externos, pode-se afirmar que o desafio que se apresenta ao Poder Judiciário é o da **comunicação efetiva, constante, de fácil acesso e em linguagem simples quanto à utilização de IA**, não só sob o prisma da Ciência Política – com a finalidade de conquistar e manter a confiança da população em relação ao uso dessa nova tecnologia na prestação do serviço jurisdicional – como também em cumprimento ao **dever de informação**.

Com efeito, no art. 18 da Resolução 332/2020 consta que os usuários externos têm o direito de saber sobre a aplicação de IA “nos serviços que lhes forem prestados”, e que essa informação deve ser dada “em linguagem clara e precisa”, inclusive sendo destacado que a solução apresentada pela IA não tem caráter vinculante, e que “sempre é submetida à análise da autoridade competente”⁴³⁵.

Acontece que também nesse ponto cabe a afirmação de que “somente é possível controlar o que se conhece”⁴³⁶, para ressaltar que de modo geral a população brasileira não tem como exercer efetivo controle, de sua posição de usuária externa, porque não tem conhecimento suficiente sobre a aplicação de IA no Judiciário. Passados cerca de dois anos do advento da Resolução 332/2020, verifica-se que o dever de informação previsto no art. 18 não está sendo observado pelos tribunais.

A informação pública mais acessível sobre o uso da IA nos tribunais brasileiros está concentrada na ferramenta de *business intelligence* mantida pelo CNJ em sua página na Internet, que consiste em painéis de projetos com IA no Poder

⁴³⁵CNJ. **Resolução nº 332, de 21 de agosto de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 29 jun. 2020.

⁴³⁶PRADO, Eunice M.B.; MÜNCH, L. A. C.; VILLARROEL, M. A. C. “Sob controle do usuário”: formação dos juízes brasileiros para o uso ético da IA no judiciário. **Direito Público**, v. 18, n. 100, 2022. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/6021>. Acesso em: 28 jun. 2022.

Judiciário⁴³⁷, apresentado diversas informações gerais que serão abordadas no próximo capítulo deste trabalho.

Por ora, o que se quer salientar é que o fato de o CNJ manter esse repositório representa um nível de informação bem diferente da que consta no art. 18 da Resolução 332/2022, que deve ser providenciada por cada tribunal que tem projeto de IA em produção.

A título de exemplo, os quatro tribunais que mais possuem projetos de IA em julho de 2022, de acordo com o Painel de Projetos IA no Poder Judiciário 2022⁴³⁸, são o Tribunal de Justiça de Rondônia (TJRO), com 21 projetos, o Tribunal de Justiça do Rio Grande do Sul (TJRS), com 7 projetos e, empatados, o Tribunal de Justiça do Distrito Federal e Territórios (TJDFT) junto com o Tribunal de Justiça do Paraná (TJPR), com 5 projetos cada um. Porém, verificando-se em julho de 2022 os portais de tais tribunais na Internet, nenhum deles possui ícone ou *link* já na página principal sobre seus projetos de IA.

Utilizando o campo de busca, ao digitar a expressão “inteligência artificial”, surgem resultados de notícias variadas sobre os projetos, sem qualquer sistematização, à exceção do TJRS, cujo resultado da busca foi o painel “Iniciativas de IA no TJRS-2022”⁴³⁹, com a explicação de que seguiu o padrão adotado no painel mantido pelo CNJ⁴⁴⁰.

O fato é que, de modo geral, a informação sobre os projetos de IA não está acessível aos jurisdicionados. Não está fácil de encontrar nem de compreender. Não há artigos, cópias dos projetos disponíveis para consulta, *links* para vídeos, tutoriais nem maiores explicações além de notícias sobre os projetos. Nas páginas de tais tribunais na Internet, não há um campo próprio para os projetos de IA, que reúna informações importantes para os usuários externos, quer se trate dos jurisdicionados,

⁴³⁷CNJ. **Painéis e publicações. Plataforma Sinapses/Inteligência Artificial.** Disponível em: <https://www.cnj.jus.br/sistemas/plataforma-sinapses/paineis-e-publicacoes>. Acesso em: 2 jul. 2022.

⁴³⁸CNJ. **Resultados pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,currsel&select=language,BR. Acesso em: 2 jul. 2022.

⁴³⁹TJRS. **Iniciativas de IA no TJRS-2022.** Disponível em: <https://datastudio.google.com/reporting/2201f416-9182-4f8a-9638-1cf7d047aa3/page/6sHkC>. Acesso em: 3 jul. 2022.

⁴⁴⁰TJRS. **Projetos com Inteligência Artificial no TJRS.** Disponível em: <https://www.tjrs.jus.br/novo/institucional/direcoes/ditic/projetos-com-inteligencia-artificial-no-tjrs/>. Acesso em: 3 jul. 2022.

quer se trate de pesquisadores, acadêmicos e entidades da sociedade civil organizada, que tenham interesse em saber sobre os projetos.

Em se tratando a IA de uma nova tecnologia que envolve riscos, pode-se dizer que o dever de informação está compreendido na obrigação geral de segurança, mencionada no início deste capítulo, no item 3.1, que impõe uma atitude de precaução por parte dos tribunais, no sentido de adotar todas as providências que estiverem a seu alcance para buscar mitigar os riscos – inclusive facilitar e permitir aos usuários externos o controle do uso da IA.

A propósito, em janeiro de 2022 o CNJ realizou uma audiência pública, pelo Comitê de Comunicação Social do Poder Judiciário, para receber sugestões que pudessem “subsidiar um parecer sobre padronização e uniformização dos portais do Poder Judiciário”, a ser submetido ao Plenário do CNJ, com o objetivo de facilitar o acesso a informações da Justiça⁴⁴¹. Parece salutar que venha a ser incluído, como item obrigatório nos portais, um campo específico sobre IA.

Por tudo o que se viu nesta seção, com menos de 10% dos magistrados brasileiros sentindo-se preparados para exercerem controle e supervisão de ferramentas de IA, e com insuficiência de informações disponíveis para a população e a comunidade científica sobre projetos de IA, muito há que se caminhar nessa seara, em termos de capacitação e plano de comunicação, para tornar efetivas as regras sobre controle do usuário previstas na Resolução 332/2020.

3.5 A Portaria nº 271/2020 do Conselho Nacional de Justiça

Enquanto a Resolução 332/2020 “dispõe sobre a ética, a transparência e a governança na produção e no uso da Inteligência Artificial no Poder Judiciário e dá outras providências”⁴⁴², cerca de quatro meses depois de sua publicação sobreveio a Portaria nº 271, de 04/12/2020, que “regulamenta o uso de Inteligência Artificial no âmbito do Poder Judiciário”⁴⁴³, publicada já em outra gestão do CNJ, sob a presidência do Ministro Luiz Fux – marcada por diversos projetos de transformação digital da

⁴⁴¹CNJ. **Padronizar portais no judiciário é facilitar acesso a informações da justiça.** 26 jan. 2022. Disponível em: <https://www.cnj.jus.br/padronizar-portais-no-judiciario-e-facilitar-acesso-a-informacoes-da-justica/>. Acesso em: 3 jul. 2022.

⁴⁴²CNJ. **Resolução nº 332, de 21 de agosto de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 4 jul. 2020.

⁴⁴³CNJ. **Portaria nº 271, de 4 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 jul. 2022.

Justiça brasileira, acelerada por força da pandemia de Covid-19, com destaque para o projeto Justiça 4.0.

Mediante instrumento de cooperação internacional, o CNJ firmou uma parceria com o Programa das Nações Unidas para o Desenvolvimento (PNUD), com vigência de novembro/2020 a julho/2023, para executar o projeto intitulado “Justiça 4.0: inovação e efetividade na realização da Justiça para todos”, objetivando promover inovação para o aprimoramento da prestação jurisdicional no Brasil mediante o desenvolvimento de “novas metodologias, estudos e ferramentas”, no sentido de formular políticas públicas judiciais orientadas por dados e com utilização de avanços tecnológicos, incluindo IA⁴⁴⁴.

No referido projeto Justiça 4.0, foi mencionada a importância da tecnologia para fazer frente à realidade de quase 80 milhões de processos em tramitação na Justiça brasileira, sendo feita uma contextualização do grau de litigiosidade que sobrecarrega o Judiciário no país, mediante comparação com países europeus, exemplificando que no Brasil há cerca de 1 processo para cada 3 habitantes, enquanto em Portugal a proporção é de 1 para 141 e na Alemanha é de 1 para 238. Quanto à força de trabalho, na comparação entre a quantidade de magistrados para cada grupo de 100 mil habitantes, o projeto menciona que no Brasil a proporção é de 9 magistrados por 100 mil, enquanto em Portugal é de 19 por 100 mil e na Alemanha, 24 por 100 mil⁴⁴⁵.

Dado esse quadro de sobrecarga, consta no projeto Justiça 4.0 a possibilidade de “automação processual com recursos de robotização e o uso de técnicas de inteligência artificial (IA), como, por exemplo, *machine learning*”, que permitem “economizar o tempo de servidores que poderiam estar dedicados a tarefas que demandam criatividade e elaboração”⁴⁴⁶.

Assim, entre as entregas do projeto Justiça 4.0 foram previstas “a implantação de juízos inteiramente digitais e a implantação da Plataforma Digital do Poder Judiciário (PDPJ)”, com incremento da IA para “aumentar o grau de automação do processo judicial eletrônico”⁴⁴⁷, com as seguintes considerações:

⁴⁴⁴CNJ. Projeto BRA/20/015 CNJ-PNUD. Disponível em: https://www.cnj.jus.br/wp-content/uploads/2020/12/TCOT_051_2020-atualizado.pdf. Acesso em: 4 jul. 2022.

⁴⁴⁵Ibidem.

⁴⁴⁶Ibidem.

⁴⁴⁷Ibidem.

No campo da eficiência da prestação jurisdicional, serão implantadas estratégias voltadas à melhoria do processo eletrônico; ao aprimoramento dos registros processuais primários; à consolidação, implantação, higienização e publicização do DataJud (Base Nacional de Dados Processuais do Poder Judiciário); à implantação de ferramenta para o agrupamento de demandas repetitivas e precedentes com o uso de solução de inteligência artificial; à implantação de ferramenta para a extração e consolidação de bases e peças processuais para prover insumos para IA (DataJud e CODEX), entre outras propostas de redução da quantidade de procedimentos e simplificação dos processos na Justiça, além da realização de estudos e testagem de ferramentas tecnológicas em outros temas relevantes para a atuação do CNJ, como a execução fiscal⁴⁴⁸.

Foi em meio a esse intenso contexto de inovações tecnológicas previstas no projeto Justiça 4.0 que, no início de dezembro de 2020, foi publicada a Portaria 271/2020, fazendo referência à Resolução 332/2020 e estruturada em 7 capítulos, a saber: I. Das disposições gerais; II. Da plataforma de Inteligência Artificial; III. Do funcionamento da plataforma de Inteligência Artificial; IV. Dos modelos de Inteligência Artificial; V. Da utilização da Inteligência Artificial; VI. Da responsabilidade dos órgãos do Poder Judiciário; VII. Das disposições finais⁴⁴⁹.

Nas disposições gerais, a Portaria 271/2020 estabeleceu que serão considerados como de IA os projetos voltados à automação de processos judiciais e administrativos e de rotinas de trabalho, à análise da massa de dados existentes no Poder Judiciário e às soluções de apoio à decisão de magistrados e de auxílio à elaboração de minutas de atos judiciais (art. 2º, parágrafo único).

Como diretrizes, a Portaria 271/2020 em seu art. 3º reforçou a atuação colaborativa como medida de economia, para evitar projetos simultâneos com a mesma finalidade, bem como a importância da interoperabilidade de soluções, de modo que sirvam para diversos sistemas de processo eletrônico em uso no país. Também estabeleceu a adoção de códigos abertos e livres, bem como as diretrizes de acesso à informação, de transparência e de foco na celeridade processual.

Além dessas, há uma diretriz no inciso VIII do art. 3º que interessa especialmente a esta pesquisa, tratando-se do “estabelecimento de mecanismos de governança colaborativa e democrática”, que contem “com a participação do Poder Judiciário, daqueles que exercem funções essenciais à justiça, da comunidade

⁴⁴⁸Ibidem.

⁴⁴⁹CNJ. Portaria nº 271, de 4 de dezembro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 jul. 2022.

acadêmica e da sociedade civil”⁴⁵⁰. No entanto, percebe-se aqui mais uma demonstração do que foi apontado na introdução deste trabalho, citando Morley et al.: estabeleceu-se “o que”, porém sem dizer “como”⁴⁵¹.

Dada essa lacuna, no próximo capítulo será estudada possível proposta de mecanismo complementar de governança colaborativa e democrática.

Por ora, é importante destacar que **todo o sistema de regulamentação da Portaria 271/2020 concentra-se na plataforma Sinapses**, mencionada no art. 3º, III, da Resolução 332/2020, cuja origem e finalidade já foram antes descritas, no item 3.4.1, ao tratar dos principais pilares da referida Resolução.

Dessa vez, na Portaria 271/2020 consta em seu art. 5º expressa menção de que a plataforma Sinapses é mantida pelo CNJ com apoio e suporte técnico da equipe do Tribunal de Justiça de Rondônia, e que todos os códigos-fontes e atualizações dos modelos de IA utilizados nos tribunais brasileiros ficarão disponíveis em seu **repositório de códigos-fontes**, no endereço <<https://git.cnj.jus.br>>.

Desse modo, o repositório de códigos utiliza o aplicativo GitLab, descrito na página inicial do referido endereço como “uma plataforma completa de *DevOps*”, e como “um aplicativo único para todo o ciclo de vida de desenvolvimento de *software*”, incluindo “planejamento de projetos, gerenciamento de código-fonte até CI/CD, monitoramento e segurança”⁴⁵².

Para melhor compreensão da finalidade da ferramenta, que atende à diretriz colaborativa do desenvolvimento de modelos de IA na Justiça brasileira, cabe trazer um conceito do termo técnico *DevOps*, próprio da área de TI (em tradução livre):

DevOps é uma metodologia de desenvolvimento destinada a fazer a ponte entre Desenvolvedores e Operadores, enfatizando a comunicação e colaboração, integração contínua, garantia de qualidade e entrega com implantação automatizada, utilizando um conjunto de práticas de desenvolvimento⁴⁵³.

⁴⁵⁰CNJ. **Portaria nº 271, de 4 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 jul. 2022.

⁴⁵¹MORLEY, Jessica et al. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. **Science and engineering ethics**, v. 26, n. 4, p. 2141-2168, 2020. Disponível em: <https://link.springer.com/article/10.1007/s11948-019-00165-5%23Sec2>. Acesso em: 4 jul. 2022.

⁴⁵²CNJ. **Repositório de projetos do CNJ.** Disponível em: https://git.cnj.jus.br/users/sign_in. Acesso em: 4 jul. 2022.

⁴⁵³JABBARI, Ramtin et al. What is DevOps? A systematic mapping study on definitions and practices. In: **Proceedings of the Scientific Workshop Proceedings of XP2016**. 2016. p. 1-11.

Quanto aos termos técnicos *CI/CD*, que se relacionam à metodologia *DevOps*, são acrônimos de *continuous integration/continuous delivery*, possibilitando a integração e implantação contínuas de códigos⁴⁵⁴ (em tradução livre):

O método de Integração Contínua (CI) permite que os desenvolvedores usem ambientes de desenvolvimento acelerados, como Agile, sem comprometer a qualidade do código. O uso de repositórios estáveis nos quais os desenvolvedores frequentemente fazem pequenas alterações, executando testes com cada modificação, resulta em um código altamente testado e prontamente implantável.

Além do repositório de códigos-fontes, o art. 5º da Portaria 271/2020 estabelece que os **modelos de IA liberados para produção** estarão disponíveis para consulta no endereço <<https://sinapses.ia.pje.jus.br/modelos>>, contendo a descrição, acurácia, caso de uso e URL da API para consumo⁴⁵⁵.

Justamente pela natureza colaborativa da metodologia *DevOps* empregada, o art. 6º da Portaria 271/2020 permite que a equipe da plataforma Sinapses promova, de ofício ou mediante sugestão dos participantes, a criação, alteração ou extinção de indicadores para os modelos de IA, com a finalidade de ranqueá-los quanto à sua aplicação e benefícios.

A respeito dos dados utilizados para o desenvolvimento dos modelos de IA pelos tribunais, o art. 11 da Portaria 271/2020 determina que, antes de desenvolver e registrar na plataforma Sinapses um modelo, deve haver instalação do módulo extrator para assegurar que os dados constem do repositório central. Assim, os conjuntos de dados utilizados para treinamento dos modelos de IA devem ser disponibilizados juntos aos recursos dos referidos modelos. É dever do órgão ou tribunal que desenvolveu o modelo cuidar das providências de sigilo e segredo de justiça, providenciando medidas para ocultar ou anonimizar os dados sensíveis.

Em caso de modelos de IA para elaborar minutas de atos judiciais e fazer previsões, o art. 12 da Portaria 271/2020 estabelece que devem ser adotadas medidas que permitam o rastreamento e a auditoria das previsões.

⁴⁵⁴AREFEEN, Mohammed Shamsul; SCHILLER, Michael. Continuous Integration Using Gitlab. **Undergraduate Research in Natural and Clinical Science and Technology Journal**, v. 3, p. 1-6, 2019.

⁴⁵⁵CNJ. **Portaria nº 271, de 4 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 jul. 2022.

Já o art. 13 determina que, em sendo verificada eventual discordância quanto ao uso das previsões, os sistemas judiciais que utilizarem esse tipo de modelo de IA devem registrar a informação de discordância, para assegurar a auditoria e a melhoria do modelo de IA.

No capítulo que trata sobre a responsabilidade dos órgãos do Poder Judiciário, consta o dever de providenciar capacitação dos colaboradores (art. 14), bem como o dever de disponibilizar toda a documentação dos projetos de IA, indicando as necessidades que motivaram sua criação e os objetivos pretendidos (art. 15). Além disso, para os modelos de IA já em uso, determinou-se a adequação à Portaria 271/2020 e à Resolução 332/2020 (art. 16). Também foi determinada a observância às Tabelas Processuais Unificadas do Poder Judiciário, como medida necessária para alcançar a padronização e uniformização taxonômica e terminológica de classes, assuntos e movimentações processuais.

Explanadas as principais disposições da Portaria 271/2020, é interessante notar que, não obstante se trate de um ato normativo que trouxe avanços significativos em relação à Resolução 332/2020, ao cuidar de diversas questões práticas e operacionais, continuou deixando lacunas sobre temas importantes, ao silenciar, por exemplo, sobre medidas adicionais de cibersegurança, efetivação da norma que determina diversidade nas equipes desenvolvedoras de IA, ou métodos para assegurar transparência por interpretabilidade dos modelos de IA.

Também nada foi dito a respeito dos “eventos adversos” referidos no art. 27 da Resolução 332/2020 – nem mesmo sob a forma de outras nomenclaturas como erros, falhas e incidentes. O art. 13 da Portaria 271/2020 fez menção apenas à expressão “discordância”, utilizada no contexto de resultados divergentes entregues por modelos de IA preditivos.

Percebe-se ainda que não há um plano de comunicação direcionado à comunidade jurídica e à sociedade, em linguagem clara e de fácil compreensão, sobre o uso da IA nos tribunais, como também não foi estabelecido qualquer método para que usuários internos e externos sinalizem aos tribunais e/ou ao CNJ que algum erro, falha, incidente ou evento adverso aconteceu. É sobre esses pontos que o próximo capítulo irá tratar.

4 APLICAÇÃO DA IA NOS TRIBUNAIS BRASILEIROS, CONTROLE DOS RISCOS E TRATAMENTO DOS INCIDENTES

Nos anos pandêmicos de 2020 e 2021, além do ano de 2022 em curso, marcados pela Covid-19, tem havido uma aceleração sem precedentes na transformação digital do Poder Judiciário brasileiro, em que a criatividade para desenvolver estratégias e fazer a Justiça não parar de funcionar, no período de isolamento social, ensejou uma série de projetos nos tribunais de Norte a Sul do país, além de uma intensa atividade normativa do CNJ, que baixou inúmeras resoluções para buscar coordenar todas as iniciativas.

Nessa onda de inovação por absoluta necessidade do serviço público essencial que cabia ao Judiciário continuar prestando à população, mesmo em condições adversas, criaram-se instrumentos como a base de dados DataJud⁴⁵⁶, a Plataforma Sinapses⁴⁵⁷, a Plataforma Codex⁴⁵⁸, a Plataforma Digital do Poder Judiciário Brasileiro (PDPJ-Br)⁴⁵⁹, o Juízo 100% Digital⁴⁶⁰, o Balcão Virtual⁴⁶¹, os Núcleos de Justiça 4.0⁴⁶², os Núcleos de Cooperação Judiciária⁴⁶³, os Centros de Inteligência⁴⁶⁴, a Rede de Laboratórios de Inovação (RenovaJud)⁴⁶⁵, apenas para citar alguns dos mais importantes.

Além dessas inovações, multiplicaram-se diversos projetos de IA em desenvolvimento nos tribunais por todo o país, que já vinham dando seus primeiros passos desde 2018, conforme mencionado no item 3.4 do capítulo anterior, e cujo cenário atual será descrito a seguir.

⁴⁵⁶CNJ. Resolução nº 331, de 20 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3428>. Acesso em: 7 jul. 2022.

⁴⁵⁷CNJ. Resolução nº 332, de 20 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 7 jul. 2022.

⁴⁵⁸CNJ. Resolução nº 446, de 14 de março de 2022. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4417>. Acesso em: 7 jul. 2022.

⁴⁵⁹CNJ. Resolução nº 335, de 29 de setembro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3496>. Acesso em: 7 jul. 2022.

⁴⁶⁰CNJ. Resolução nº 345, de 9 de outubro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3512>. Acesso em: 7 jul. 2022.

⁴⁶¹CNJ. Resolução nº 372, de 12 de fevereiro de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3742>. Acesso em: 7 jul. 2022.

⁴⁶²CNJ. Resolução nº 385, de 6 de abril de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3843>. Acesso em: 7 jul. 2022.

⁴⁶³CNJ. Resolução nº 350, de 27 de outubro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3556>. Acesso em: 7 jul. 2022.

⁴⁶⁴CNJ. Resolução nº 349, de 23 de outubro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3547>. Acesso em: 7 jul. 2022.

⁴⁶⁵CNJ. Resolução nº 395, de 7 de junho de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 7 jul. 2022.

4.1 Panorama de aplicação da IA nos tribunais

De início cabe registrar que o Poder Judiciário brasileiro, de acordo com a Constituição Federal⁴⁶⁶, possui um total de 92 órgãos, a saber: o Supremo Tribunal Federal, o Conselho Nacional de Justiça, o Superior Tribunal de Justiça, o Tribunal Superior do Trabalho, o Tribunal Superior Eleitoral, o Superior Tribunal Militar, 5 Tribunais Regionais Federais, 27 Tribunais de Justiça Estaduais, 24 Tribunais Regionais do Trabalho, 27 Tribunais Regionais Eleitorais e 3 Tribunais de Justiça Militar⁴⁶⁷.

Não foi incluído nesse número total o mais novo órgão, o Tribunal Regional Federal da 6ª Região, recentemente criado pela Lei nº 14.226, de 20/10/2021⁴⁶⁸, pelo fato de que ainda não foi inaugurado, encontrando-se em fase de escolha dos desembargadores da primeira composição, prevista para agosto de 2022⁴⁶⁹.

Dado esse organograma da Justiça brasileira, importa mencionar que foi a partir de dezembro de 2020 que o CNJ passou a disponibilizar, em sua página na Internet dedicada à plataforma Sinapses, o painel de *business intelligence* no qual é possível consultar as aplicações de IA nos tribunais oficialmente comunicadas ao CNJ, nos termos da Resolução nº 332/2020.

Por ocasião do lançamento do serviço, havia “41 projetos em 32 órgãos – Superior Tribunal de Justiça (STJ), Tribunal Superior do Trabalho (TST), Conselho de Justiça Federal (CJF) e Tribunais eleitorais (3), estaduais (14), federais (4) e do trabalho (8)”⁴⁷⁰.

⁴⁶⁶BRASIL. **Constituição da República Federativa do Brasil de 1988**. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 8 jul. 2022.

⁴⁶⁷CNJ. **Justiça em Números 2020**. Brasília: CNJ, 2020. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2020/08/WEB-V3-Justi%C3%A7a-em-N%C3%BAmeros-2020-atualizado-em-25-08-2020.pdf>. Acesso em: 8 jul. 2022.

⁴⁶⁸BRASIL. **Lei nº 14.226, de 20 de outubro de 2021**. Disponível em: <https://www.in.gov.br/en/web/dou/-/lei-n-14.226-de-20-de-outubro-de-2021-353673576>. Acesso em: 8 jul. 2022.

⁴⁶⁹STJ. **Sessão para escolha dos desembargadores do TRF6 será presencial, em 1º de agosto, com votação secreta**. 13 jun. 2022. Disponível em: <https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/13062022-Sessao-para-escolha-de-desembargadores-do-TRF6-sera-presencial-em-1-de-agosto-com-votacao-secreta.aspx>. Acesso em: 8 jul. 2022.

⁴⁷⁰CNJ. **Painel dá transparência a projetos de inteligência artificial no Judiciário**. Agência CNJ de Notícias, 23 dez. 2020. Disponível em: <https://www.cnj.jus.br/painel-da-transparencia-a-projetos-de-inteligencia-artificial-no-judiciario/>. Acesso em: 8 jul. 2022.

A tela do painel contém o mapa do Brasil na parte superior (Figura 8), de modo que é possível a qualquer internauta clicar em cada Estado para consultar informações mais detalhadas sobre os projetos de IA por localidade/tribunal.

Figura 8 - Parte superior do 1º painel de projetos de IA nos tribunais



Fonte: CNJ⁴⁷¹.

Segundo noticiado pelo CNJ na época do lançamento do painel, a respeito das principais motivações para o desenvolvimento dos projetos, foram observadas três razões principais: a busca por inovação, a intenção de reduzir o acúmulo de trabalho e o intuito de aumentar a celeridade da tramitação dos processos⁴⁷².

Também se destacou que cerca de 60% dos modelos de IA foram desenvolvidos internamente, por equipes próprias de TI dos tribunais, e quase 90% possuíam código-fonte na linguagem *Python*. A respeito dos dados utilizados, 46% dos projetos envolviam o tratamento de dados pessoais, e 41% lidavam com dados de identificação pessoal⁴⁷³.

Consultando esse primeiro painel “Projetos com Inteligência Artificial no Poder Judiciário”⁴⁷⁴, observam-se diversos gráficos com informações relevantes sobre os seguintes tópicos:

- Breve descrição sobre de que se trata cada projeto;

⁴⁷¹CNJ. **Projetos com Inteligência Artificial no Poder Judiciário.** Disponível em: <https://paineisanalytics.cnj.jus.br/single/?appid=29d710f7-8d8f-47be-8af8-a9152545b771&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&opt=ctxmenu,cursel>. Acesso em: 8 jul. 2022.

⁴⁷²CNJ. **Painel dá transparência a projetos de inteligência artificial no Judiciário.** Agência CNJ de Notícias, 23 dez. 2020. Disponível em: <https://www.cnj.jus.br/painel-da-transparencia-a-projetos-de-inteligencia-artificial-no-judiciario/>. Acesso em: 8 jul. 2022.

⁴⁷³Ibidem.

⁴⁷⁴CNJ. **Projetos com Inteligência Artificial no Poder Judiciário.** Disponível em: <https://paineisanalytics.cnj.jus.br/single/?appid=29d710f7-8d8f-47be-8af8-a9152545b771&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&opt=ctxmenu,cursel>. Acesso em: 8 jul. 2022.

- O que motivou o uso da ferramenta de IA para a aplicação desenvolvida;
- Como se deu o desenvolvimento da ferramenta;
- Para qual plataforma de processo eletrônico a ferramenta foi criada;
- Qual foi a linguagem de programação utilizada;
- Qual é o método ou técnica de IA empregada;
- Que recursos se aplicam à ferramenta (por exemplo, análise de texto, de fala, organização de dados etc.);
- Se a ferramenta faz ou não análise de dados não-estruturados;
- Se os dados usados para treinar a ferramenta são conhecidos pela equipe que a utiliza;
- Se o código da ferramenta está disponível publicamente e pode ser revisado;
- A quem pertence o algoritmo da ferramenta e seu código;
- Se, não sendo detentora, a instituição tem acesso ao código e ao funcionamento do algoritmo;
- Se a ferramenta trata dados pessoais em conformidade com a Lei Geral de Proteção de Dados;
- Se a ferramenta está tratando informações de identificação pessoal;
- Quem gerou o dado para o treinamento da ferramenta;
- Se as equipes técnica e não-técnica da instituição conseguem explicar as entradas, saídas, e como são gerados os resultados;
- Se a ferramenta passou por monitoramento e revisão;
- Se a ferramenta possui documentação.

Acontece que, nos termos do art. 19 da Portaria 271/2020⁴⁷⁵, que regulamentou a IA no Judiciário, cabe ao CNJ fazer levantamento periódico dos projetos, com o quantitativo de processos judiciais beneficiados. Assim, de acordo com tal regra, já foi disponibilizado um novo painel com informações mais recentes, elaborado com dados de maio de 2022, que foram informados ao CNJ pelos tribunais.

⁴⁷⁵CNJ. Portaria nº 271, de 4 de dezembro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 8 jul. 2022.

Trata-se do painel intitulado “Resultados Pesquisa IA no Poder Judiciário – 2022”⁴⁷⁶, do qual se destaca que houve um notável aumento de 171% no número de projetos, de modo que até maio de 2022 o total oficial é de 111 projetos em 53 tribunais. Também é possível perceber engajamento bem maior dos órgãos ao prestarem informações dos projetos ao CNJ, haja vista o destaque quanto ao aumento de 265% de respostas em relação à consulta anterior.

De acordo com o novo painel, dos 111 projetos de IA, 63 já estão em produção, encontrando-se efetivamente em uso, e 42 projetos já estão inclusive com o código-fonte disponível na Plataforma Sinapses⁴⁷⁷. Manteve-se visualmente o mapa no Brasil na parte superior do painel (Figura 9), a partir do qual continua sendo possível consultar diversas informações dos projetos de IA por localidade/tribunal.

Figura 9 - Parte superior do painel de projetos de IA nos tribunais em 2022



Fonte: CNJ⁴⁷⁸.

No novo painel, abaixo do mapa do Brasil, os gráficos informativos passaram a ser organizados em 4 blocos, a saber: (i) detalhamento dos projetos, (ii) equipes, (iii) recursos, plataformas, métodos e linguagens e (iv) dados e códigos. Cada gráfico permite a consulta e exportação dos dados em imagem, formato PDF e em tabela no Excel.

⁴⁷⁶CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language,BR. Acesso em: 8 jul. 2022.

⁴⁷⁷Ibidem.

⁴⁷⁸Ibidem.

Nas próximas subseções serão descritos os projetos de forma resumida, com base nas informações constantes no novo painel, por ramos da Justiça, partindo do ramo que tem mais projetos e seguindo a ordem decrescente.

4.1.1 Projetos de IA na Justiça Estadual

Quanto aos projetos de IA por segmento da Justiça, o painel⁴⁷⁹ demonstra que, dos 111 projetos, 65 estão sendo desenvolvidos na Justiça Estadual, o que representa um percentual próximo de 60%.

Cabe pontuar que, de acordo com o painel de Estatísticas do Poder Judiciário⁴⁸⁰, atualizado até 30/04/2022 pela base nacional de dados judiciais, o DataJud, encontram-se em tramitação na Justiça brasileira mais de 78 milhões de processos pendentes, dos quais pouco mais de 60 milhões estão na Justiça Estadual, o que demonstra que quase 80% dos processos encontram-se nesse ramo da Justiça.

Considerando que há 27 tribunais na Justiça Estadual, pelo painel depreende-se que apenas 3 tribunais estaduais ainda não aparecem com projetos de IA: o Tribunal de Justiça do Amazonas (TJAM), o Tribunal de Justiça de Sergipe (TJSE) e o Tribunal de Justiça de Mato Grosso (TJMT).

O Tribunal de Justiça de Rondônia (TJRO), de onde se originou a plataforma Sinapses, é de longe o que tem maior número, com 21 projetos de IA, todos desenvolvidos por equipe interna e com os códigos depositados na plataforma, dos quais 20 já estão em produção.

Quanto às funcionalidades descritas para cada um dos 21 projetos do TJRO, são as seguintes:

- Acórdão Seção: para identificar seções em um acórdão, reconhecendo partes como ementa, relatório e voto;
- Assunto judicial/Multilabel/Texto: para identificar os assuntos judiciais contidos em uma petição inicial;

⁴⁷⁹CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language,BR. Acesso em: 8 jul. 2022.

⁴⁸⁰CNJ. **Estatísticas do Poder Judiciário.** Disponível em: <https://painelestatistica.stg.cloud.cnj.jus.br/estatisticas.html>. Acesso em: 9 jul. 2022.

- Assunto Juizado Especial Criminal/Multiclasse/Texto: para identificar o assunto de documento no âmbito de Juizado Especial Criminal;
- Classe judicial/Multiclasse/Texto: para identificar a classe judicial de uma petição inicial;
- Gerador de texto de magistrado: para gerar textos com sugestão de palavras, auxiliando na produção de documentos;
- Identifica pedido e decisão: para reconhecer tais partes em um documento;
- Identifica conexão/Multiclasse/Composto: para detectar a conexão entre processos;
- Identifica petição inicial/Multiclasse/Composto: para identificar as petições iniciais em processos vindos dos tribunais, consolidados no Codex;
- Justiça gratuita/Binário/Texto: para analisar a petição inicial e detectar se há pedido de assistência judiciária;
- Liminar/Binário/Texto: para analisar a petição inicial e detectar se há pedido de liminar;
- Mapeamento entidades: para identificar entidades em um documento;
- Peticionamento inteligente: para facilitar a distribuição de documentos originários de delegacias, a exemplo de Termos Circunstanciados de Ocorrência, Inquéritos Policiais, etc., substituindo o recebimento e cadastramento manual de informações no sistema PJe pela extração automatizada dos dados necessários a partir dos documentos anexados no sistema pela Delegacia de Polícia, de modo que o distribuidor apenas verifica e valida as informações que foram previamente preenchidas de forma automática pelo sistema;
- Publica Diário: para automatizar a comunicação de acordo com os documentos gerados pelo magistrado;
- Similaridade petição inicial: para calcular a taxa de similaridade entre petições iniciais;
- Sumarizador: para elaborar sumário de textos;

- TGM Área/Multiclasse/Texto: para realizar triagem de grande massa, mediante classificação entre Cível e Especial;
- TGM Cíveis/Multiclasse/Texto: para realizar triagem de grande massa na 2^a instância, nas Câmaras Cíveis;
- TGM Especiais: para realizar triagem em processos na 2^a instância, nas Câmaras Especiais, classificando os feitos de acordo com as classes treinadas;
- TGM: para realizar triagem de processos de grande massa dos processos recebidos nos gabinetes da 2^a instância, em agrupamentos (banco, fornecimento de água, companhia elétrica, DPVAT, etc.), de modo a facilitar a distribuição dos feitos por temas entre os servidores dos gabinetes;
- Tipo movimento magistrado - classificador: é um modelo de predição de movimentos das Tabelas Processuais Unificadas de acordo com o despacho proferido pelo magistrado, tratando-se de ferramenta que é usada como base para outros classificadores em vários tribunais, a exemplo do sistema JANUS (desenvolvido pelo TRE-BA e utilizado por diversos tribunais eleitorais, como será visto adiante);
- Vetoriza texto: ferramenta para criar vetores para textos.

O segundo tribunal com mais projetos, de acordo com o painel, é o Tribunal de Justiça do Rio Grande do Sul (TJRS), com 7 projetos, todos já em produção (porém nenhum com modelo depositado na plataforma Sinapses), com as seguintes funcionalidades:

- Chatbot DIGEP: *chatbot* com perguntas e repostas referentes à Direção de Gestão de Pessoas;
- Chatbot para o atendimento da CADI: assistente virtual para atendimento de usuários quanto a assuntos relativos a TI;
- Grafo: sistema para visualização de vídeos de audiências e transcrições;
- IA Classificador por conteúdo: para classificar documentos de acordo com a similaridade dos textos;

- IA Execução Fiscal: para analisar petições iniciais de execuções fiscais com classificação por IA;
- IA Gerador de resumos: para gerar resumos de textos de documentos processuais;
- IA Validação de assuntos: para validar os assuntos de processos cadastrados no sistema de processo eletrônico Eproc de acordo com o texto da petição inicial.

O terceiro tribunal em número de projetos é o Tribunal de Justiça do Distrito Federal e Territórios (TJDFT), com 5 projetos de IA, todos desenvolvidos pela equipe interna do tribunal e já em produção (com 2 códigos hospedados na plataforma Sinapses) e as seguintes funcionalidades:

- Sistema Amon, para controle de quem transita nas dependências do tribunal por meio de reconhecimento facial;
- Sistema Artiu, desenvolvido para a Central de Mandados com duas funcionalidades, correção de CEP de endereços (para evitar retrabalho e desperdício de recursos com o retorno de mandados sem cumprimento por divergência entre CEP e endereço) e classificação de mandados urgentes;
- Sistema Hórus de processamento inteligente e automação em Vara de Execução Fiscal, para (*i*) realizar a inserção automática de processos digitalizados, categorizando e classificando os documentos com base em tipos específicos existentes no sistema de tramitação, utilizando o processo de certificação digital para validar o processo, (*ii*) reconhecer o código de caracteres dos documentos via OCR para confrontar os dados das peças processuais com os dados dos sistemas, detectando eventuais divergências das peças processuais digitalizadas manualmente, (*iii*) recuperar metainformações de sistema de tramitação de processos físicos para subsidiar dados para o pré-cadastro e futuro tratamento de dados;
- SAREF (Sistema de Apresentação e Reconhecimento Facial), em funcionamento na Vara de Execuções das Penas em Regime Aberto

e Vara de Execuções das Penas e Medidas Alternativas, com a finalidade de agilizar a apresentação de apenados utilizando algoritmos de reconhecimento facial, facilitando a comprovação do cumprimento de penas de forma remota. Para tanto, o usuário pode se apresentar perante a Justiça por meio de um celular ou computador com acesso à Internet, sem precisar se deslocar até o fórum;

- Sistema Toth, integrado ao PJe, para recomendar a classe e o assunto do processo com base na petição inicial do advogado, de acordo com as Tabelas Processuais Unificadas, evitando/diminuindo eventual classificação errônea desde o início do ciclo de vida processual, o que contribui para melhorar a qualidade dos dados enviados ao DataJud, bem como relatórios baseados nesses metadados processuais.

Empatado com o TJDFT, consta no painel o Tribunal de Justiça do Paraná (TJPR) também com 5 projetos de IA, totalmente desenvolvidos pela equipe interna de TI do tribunal, com 1 deles já em produção, embora nenhum conste como depositado na plataforma Sinapses. Trata-se de uma espécie de família de projetos, intitulada “Larry”, que possuem as funcionalidades assim descritas:

- Larry - Análise de Juntada: para sugerir o próximo passo que o servidor deveria tomar, diante de uma análise de juntada de petição ou documento (o projeto é direcionado, a princípio, a analistas judiciais de 1º grau);
- Larry - Aplicação de IA para Identificação de Demanda Repetitiva: para auxiliar os diferentes órgãos do tribunal a identificarem demandas repetitivas ou de massa, acelerando a resolução de processos (o projeto é direcionado, a princípio, à 1ª Vice-Presidência, que irá validá-lo);
- Larry - Prescrição Intercorrente: para estabelecer um modelo em que a unidade judiciária possa se apoiar a fim de identificar um processo que esteja prescrito ou em vias de prescrição;
- Larry - Semelhança de Petições Iniciais: para encontrar processos semelhantes, com suspeita de advocacia predatória, bem como para

localizar oportunidades de agrupamento de processos para formação de incidentes de demandas repetitivas (o projeto tem como público-alvo as secretarias do 1º grau de jurisdição);

- Larry - 2º grau: para identificar matérias individualizáveis entre processos distribuídos como alheios às áreas de especialização, de modo a melhorar a distribuição e reduzir o tempo de tramitação (o projeto tem como público-alvo o 2º grau de jurisdição).

Com 3 projetos de IA registrados no painel, desenvolvidos pela equipe interna de TI do tribunal, todos já em produção e com os modelos depositados na plataforma Sinapses, está o Tribunal de Justiça da Bahia (TJBA), cujos projetos estão descritos da seguinte forma resumida:

- IAJUS - Inteligência Artificial e Automação Inteligente: nesse projeto, uma das ferramentas desenvolvidas está atuando diariamente em 86 unidades judiciais com competência para Defesa do Consumidor, fazendo processamento de linguagem natural para ler aspetos iniciais dos processos, identificar seu tema/assunto e realizar etiquetagem para possibilitar o julgamento por temas;
- Mapeamento de demandas repetitivas ou com potencial de repetitividade: para auxiliar a 2ª Vice-Presidência do tribunal ao realizar etiquetagem de processos;
- Sofia - Assistente Virtual dos Juizados Especiais da Bahia: para atendimento às principais demandas administrativas e judiciais da Coordenação dos Juizados Especiais, diminuindo a demanda de atendimentos por e-mail e telefone que antes ocupavam muito tempo dos servidores da Coordenação. A assistente virtual já está sendo utilizada para atendimento ao público, e uma nova funcionalidade está em fase de homologação, a de sugestão de julgados relacionados, visto que a assistente consegue mostrar julgados semelhantes aos processos consultados.

O Tribunal de Justiça do Ceará (TJCE) aparece no painel com 2 projetos de IA, porém as descrições dos projetos são muito vagas, genéricas, sem apresentar de forma específica suas funcionalidades, havendo apenas menção de uso de

processamento de linguagem natural (PLN) para classificar assuntos de forma automática. Depreende-se que os projetos ainda estão em fase inicial de concepção.

Também com 2 projetos de IA no painel está o Tribunal de Justiça do Maranhão (TJMA), que ainda não os colocou em produção mas já constam como modelos depositados na plataforma Sinapses (havendo menção de que foram desenvolvidos em colaboração técnica com instituição de ensino superior), com as seguintes descrições:

- Apolo - Analisador de Precedentes: para analisar a petição inicial e classificá-la entre os temas mapeados pelo Núcleo de Precedentes do TJMA, etiquetando o processo desde a distribuição, com o objetivo de facilitar e agilizar a análise dos processos e as decisões dos magistrados;
- Apolo - Analisador de Prevenção: para analisar a prevenção de ações distribuídas utilizando redes neurais artificiais.

Igualmente com 2 projetos de IA no painel está o Tribunal de Justiça da Paraíba (TJPB), 1 deles já em produção, desenvolvido por equipe de TI do tribunal e com modelo na plataforma Sinapses, com as seguintes descrições:

- Automação de Provimentos em Correções Judiciais no PJe: tem como público-alvo a Corregedoria-Geral do TJPB, para analisar a situação de processos judiciais em unidades judiciais (paralisados há mais de 100 dias, quanto ao cumprimento das metas 2, 4, 6 e 8 do CNJ, os 20 processos mais antigos e 50 processos arquivados aleatoriamente nos últimos 5 anos), de modo a lançar provimentos com despachos autoassinados em tarefa agendada;
- Primeiro Modelo de IA no PJe: cujo projeto é descrito como a intenção do TJPB em implantar modelos de IA do TJRO disponíveis na plataforma Sinapses para uso de automação no PJe, objetivando acelerar a tramitação de processos, com destaque para as aplicações “Movimento Inteligente”, “Classe Judicial” ou “Assunto Judicial”.

Chama a atenção que o maior tribunal do país⁴⁸¹, o Tribunal de Justiça de São Paulo (TJSP), figure no painel com apenas 2 projetos de IA, mesmo assim ainda em fase de desenvolvimento, em que os modelos sequer estão depositados na plataforma Sinapses.

Além disso, segundo as outras informações disponíveis no painel, a equipe de desenvolvimento envolve apenas 2 pessoas e, ao contrário dos outros tribunais até aqui analisados, em que majoritariamente os projetos são desenvolvidos pelas equipes internas de TI dos próprios tribunais, no caso do TJSP os projetos estão sendo realizados mediante parceria com entidade externa, a Universidade de São Paulo, para as seguintes funcionalidades:

- Análise de Guias de Custas Duplicadas: para extração do conteúdo de guias de custas judiciais, metadados das guias e processos com o intuito de identificar os valores recolhidos e verificar eventuais duplicidades;
- Identificação de Processos com Precedentes Vinculados: para análise e localização de processos repetitivos que devem ficar suspensos na 2^a instância até decisão final do STJ.

É interessante notar que, dos mais de 78 milhões de processos pendentes de julgamento no Brasil, quase 22 milhões estão em tramitação somente no TJSP, segundo o painel de Estatísticas do Poder Judiciário⁴⁸².

Nesse contexto, é curioso que não haja projetos totalmente próprios, e que uma das iniciativas de uso de IA, em tal tribunal, seja para verificar se alguém pagou custas judiciais a mais, enquanto a outra seja para buscar processos para suspensão – nenhuma para conferir diretamente mais celeridade na tramitação de feitos.

Empatado com o TJCE, TJMA, TJPB e TJSP – todos tribunais estaduais que também possuem 2 projetos de IA registrados no painel –, está o Tribunal de Justiça do Tocantins (TJTO), cujos projetos ainda não estão em produção nem com modelos depositados na plataforma Sinapses, havendo o registro de desenvolvimento em colaboração com o Programa de Pós Graduação em Modelagem Computacional

⁴⁸¹CNJ. **Justiça em Números 2021**. Brasília: CNJ, 2021. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2021/09/relatorio-justica-em-numeros2021-12.pdf>. Acesso em: 9 jul. 2022.

⁴⁸²CNJ. **Estatísticas do Poder Judiciário**. Disponível em: <https://painelestatistica.stg.cloud.cnj.jus.br/estatisticas.html>. Acesso em: 9 jul. 2022.

de Sistemas da Universidade Federal do Tocantins. Os projetos do TJTO encontram-se assim descritos no painel:

- Agrupamento de Sentenças: para agrupar recursos que chegam na 2^a instância de acordo com as sentenças proferidas em 1º grau de jurisdição, de modo que os Desembargadores possam analisar as apelações de acordo com a temática recorrida, permitindo a análise e a movimentação processual em bloco;
- MinerJus: para conferir maior confiabilidade à classificação dos assuntos dos processos no momento do cadastro e distribuição de petições iniciais, usando processamento de linguagem natural para extrair informações contidas em documentos no formato PDF.

Os tribunais estaduais abaixo listados em ordem alfabética, conforme aparecem no painel, possuem apenas 1 projeto de IA cada, com as descrições e funcionalidades detalhadas a seguir.

- Tribunal de Justiça do Acre (TJAC): projeto para atuar no portal de peticionamento do tribunal, objetivando melhorar a classificação de assunto de petições intermediárias e de classe/assunto de petições iniciais, mediante leitura de documentos em PDF. O projeto já está em produção, mas o modelo ainda não está depositado na plataforma Sinapses;
- Tribunal de Justiça de Alagoas (TJAL): projeto Hércules, buscando acelerar o trâmite de processos de execução fiscal, iniciado na 15^a Vara de Execução Fiscal para (*i*) classificação automática de petições intermediárias que ficavam acumuladas numa fila do sistema de processo eletrônico, conseguindo classificar corretamente 95 de cada 100, (*ii*) analisar e detectar erros de preenchimento de certidões de dívida ativa, (*iii*) cruzamento e identificação de processos oriundos da Prefeitura de Maceió que já haviam sido baixados na base de dados da Prefeitura; (*iv*) estudar soluções para o envio de correspondências com endereços inconsistentes, (*v*) classificação de apelações em processos de execução fiscal em 3 classes, com 97% de acurácia. No painel consta que o modelo de IA ainda não

está depositado na plataforma Sinapses. Embora a descrição do projeto demonstre que o modelo já está em uso, aparece no painel como ainda em desenvolvimento;

- Tribunal de Justiça do Amapá (TJAP): projeto Tucujuris Inteligência Artificial (TIA), para identificação de demandas repetitivas no momento de ingresso de uma nova petição inicial;
- Tribunal de Justiça do Espírito Santo (TJES): projeto Argos Inteligência Artificial em Gratuidade de Justiça, para cruzar dados de diversas bases em busca de parâmetros que justifiquem a concessão de pedidos de Justiça Gratuita, facilitando o acesso do cidadão à Justiça. Consta no painel que o projeto ainda está em fase desenvolvimento, em parceria com uma *startup*, e o modelo de IA ainda não está depositado na plataforma Sinapses;
- Tribunal de Justiça de Goiás (TJGO): projeto Inteligência Artificial Berna (Busca Eletrônica utilizando Recursos de Linguagem Natural), para identificar petições iniciais com mesmo fato e tese jurídica distribuídas para unidades distintas, em ofensa o princípio do juiz natural, que os mecanismos atuais de conexão não conseguem detectar. Pretende-se (*i*) realizar as devidas conexões de processos; (*ii*) verificar se os agrupamentos criados guardam relação com precedentes, em cumprimento ao art. 332 do Código de Processo Civil⁴⁸³; (*iii*) verificar possível aplicação do Enunciado 73 do FONAJE⁴⁸⁴ em processos dos Juizados Especiais; (*iv*) facilitar a identificação de casos para que as Turmas de Uniformização criem súmulas; (*v*) propiciar novos métodos de gestão de processos em unidades judiciais. O projeto foi desenvolvido a partir de projeto de tese de doutorado de Analista de TI do tribunal, e já se encontra em produção, porém o modelo ainda não está depositado na plataforma Sinapses;
- Tribunal de Justiça de Minas Gerais (TJMG): projeto para (*i*) identificar a petição inicial entre as peças e documentos eletrônicos do PJe,

⁴⁸³BRASIL. Lei nº 13.105, de 16 de março de 2015. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2015-2018/2015/lei/l13105.htm. Acesso em: 9 jul. 2022.

⁴⁸⁴FONAJE. Enunciados - Fórum Nacional de Juizados Especiais. Disponível em: <https://fonaje.amb.com.br/enunciados/>. Acesso em: 9 jul. 2022.

- (ii) identificar e classificar assuntos das ações referentes à judicialização da saúde por meio de metodologia *multi-label*. O projeto foi totalmente desenvolvido pela equipe técnica do tribunal e o modelo já está na plataforma Sinapses, embora ainda não esteja em produção;
- Tribunal de Justiça do Mato Grosso do Sul (TJMS): o projeto consta no painel como em fase inicial de conhecimento e uso das plataformas Sinapses e Codex do CNJ;
 - Tribunal de Justiça do Pará (TJPA): projeto INDIA (Indexador de Documentos Judiciais com Inteligência Artificial), com a criação de ferramenta para ser usada na tarefa de migração de processos do meio físico para o digital (PJe), facilitando a indexação e identificação dos diversos tipos de documentos existentes dentro dos processos, de modo que possam ser consultados de modo mais fácil, rápido e organizado. Criado pela equipe interna do tribunal, o modelo ainda não está em produção nem depositado na plataforma Sinapses;
 - Tribunal de Justiça de Pernambuco (TJPE): projeto ELIS de automação de atividades identificadas como gargalos em processos de execução fiscal, utilizando-se a ferramenta na triagem inicial dos feitos para analisar e classificar os aspectos de (i) divergências entre os dados dos documentos contidos na petição inicial, na Certidão de Dívida Ativa e no sistema PJe, (ii) competência diversa e (iii) prescrição. O modelo foi desenvolvido pela equipe interna do tribunal, está em uso e se encontra depositado na plataforma Sinapses;
 - Tribunal de Justiça do Piauí (TJPI): projeto para identificação e tratamento de processos de contratos bancários, objetivando a identificação e sugestão de decisões-modelos para as diversas situações concretas apresentadas nos processos. O modelo ainda está em fase de desenvolvimento pela equipe interna do tribunal e não está depositado na plataforma Sinapses;
 - Tribunal de Justiça do Rio de Janeiro (TJRJ): projeto ODR TJRJ, tratando-se de uma plataforma de acordo pré-processual com empresas externas, para ser usada por consumidores em litígios com

empresas, objetivando a realização de acordos rápidos, contribuindo para reduzir a taxa de judicialização. O projeto está em fase de desenvolvimento em colaboração com outras instituições, e o modelo ainda não está depositado na plataforma Sinapses;

- Tribunal de Justiça do Rio Grande do Norte (TJRN): projeto GPSMed, para analisar ações de saúde pública em que o Estado figura no polo passivo, a partir do conteúdo das petições iniciais, decisões e sentenças proferidas. O projeto ainda está em desenvolvimento pela equipe interna do tribunal e o modelo ainda não está depositado na plataforma Sinapses;
- Tribunal de Justiça de Roraima (TJRR): projeto Mandamus, para distribuição automática de mandados aos oficiais de justiça. O projeto ainda está em desenvolvimento, em parceria com outras instituições, e pelo que consta no painel o modelo ainda não está depositado na plataforma Sinapses;
- Tribunal de Justiça de Santa Catarina (TJSC): projeto de classificação por conteúdo, para realizar a triagem de processos a partir do conteúdo dos documentos protocolados. A ferramenta foi desenvolvida em parceria com outra instituição e já está em produção, implantada em 20 unidades do 1º grau, com potencial de expansão para outras unidades tanto do 1º como do 2º graus de jurisdição. O modelo ainda não está depositado na plataforma Sinapses.

4.1.2 Projetos de IA na Justiça Federal

De acordo com o painel de projetos de IA do CNJ em 2022⁴⁸⁵, existem 14 projetos na Justiça Federal, na qual tramitam quase 12 milhões dos cerca de 78 milhões de processos pendentes no Brasil, conforme o painel do CNJ de estatísticas

⁴⁸⁵CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language.BR. Acesso em: 8 jul. 2022.

extraídas do DataJud⁴⁸⁶. Todos os Tribunais Regionais Federais possuem projetos de IA.

No Tribunal Regional Federal da 2^a Região (TRF2), com jurisdição abrangendo os Estados do Rio de Janeiro e Espírito Santo, são 4 projetos, dos quais 3 já estão em produção, como funcionalidades incorporadas ao sistema de processo eletrônico Eproc. São desenvolvidos em parceria com outras instituições e nenhum dos modelos de IA está depositado na plataforma Sinapses. As descrições e funcionalidades são as seguintes:

- Classificação por conteúdo na automatização dos localizadores: para reduzir o trabalho manual em documentos similares por meio de etiquetas de classificação;
- Projeto Intelligentia: para identificar processos similares usando a busca por número, aprimorando a gestão processual da Vice-Presidência;
- Sugestão de temas repetitivos e com repercussão geral: para reduzir a dificuldade na classificação de temas do STF, STJ e da Turma Nacional de Uniformização (TNU);
- Validação de assuntos: para reduzir o número de retificações de ação por classificação errônea do assunto principal do processo, que muitas vezes só vem a ser percebida no 2º Grau.

No Tribunal Regional Federal da 5^a Região (TRF5), com jurisdição nos Estados do Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e Sergipe, também há 4 projetos de IA, 2 desenvolvidos por equipe interna e 2 em parceria com outras instituições. Nenhum deles entrou em produção, mas 1 já está com o modelo depositado na plataforma Sinapses, de acordo com o painel, onde se encontram assim descritos:

- Busca Inteligente de Jurisprudência no sistema CREA: para busca semântica na base de sentenças e acórdãos do sistema, facilitando que magistrados e servidores localizem mais rapidamente temas e

⁴⁸⁶CNJ. **Estatísticas do Poder Judiciário.** Disponível em: <https://painel-estatistica.stg.cloud.cnj.jus.br/estatisticas.html>. Acesso em: 9 jul. 2022.

decisões semelhantes ao de um caso determinado que estiverem analisando, na Justiça Federal do Rio Grande do Norte;

- *Chatbot NISIA*: para atendimento automatizado ao cidadão, prestando informações e orientando sobre os serviços oferecidos pelo tribunal, reduzindo a necessidade de atendimento por servidores em questões simples e comuns;
- Classificação de petições iniciais para perícias: a partir do texto da petição inicial e dos metadados do processo, de modo a indicar os processos que necessitarão de perícia e em que áreas. O objetivo é auxiliar servidores das Varas Federais envolvidos na triagem e agendamento de perícias, com a meta de integrar a solução ao sistema de processo eletrônico PJe;
- Triagem e análise textual de recursos interpostos ao TRF5: com a finalidade de categorizar os processos judiciais pela matéria, para identificar os recursos cujos temas envolvam matéria já decidida por Tribunal Superior, além de identificar recursos que tratam de matéria repetitiva com potencial de ser representativo de controvérsia.

No Tribunal Regional Federal da 4^a Região (TRF4), com jurisdição nos Estados do Paraná, Santa Catarina, e Rio Grande do Sul, existem 3 projetos de IA, todos desenvolvidos por equipe interna e já em produção, embora os modelos não estejam depositados na plataforma Sinapses, segundo o painel. As funcionalidades são as seguintes:

- Agrupamento de apelações por similaridade de sentença: gera um mapa do acervo de gabinetes de Desembargadores, em que cada processo é representado por um ponto, permitindo a visualização facilitada de demandas repetitivas, bem como o filtro por termos, situação do processo e localizadores;
- Classificador de petições: para sugerir ao peticionante a classificação do tipo de petição de acordo com o texto;
- Resumo de petições: para exibir somente os trechos mais importantes de petições com até uma página, sem precisar abrir o documento todo.

No Tribunal Regional Federal da 1^a Região (TRF1), com jurisdição no Acre, Amapá, Amazonas, Bahia, Distrito Federal, Goiás, Maranhão, Mato Grosso, Minas Gerais, Pará, Piauí, Rondônia, Roraima e Tocantins (o TRF6, quando implantado, terá jurisdição em Minas Gerais, como já dito), existem 2 projetos de IA, ainda em desenvolvimento, com a colaboração de outras instituições. Os modelos não estão depositados na plataforma Sinapses e possuem a seguinte descrição de funcionalidades:

- ALEI - Análise Legal Inteligente: para auxiliar os servidores do TRF1 na análise dos recursos recebidos;
- Projeto para iniciar aplicação de IA na Justiça Federal de 1^a instância do Distrito Federal, ainda sem ferramentas concretas e funcionalidades específicas.

No Tribunal Regional Federal da 3^a Região (TRF3), com jurisdição nos Estados de São Paulo e Mato Grosso do Sul, há 1 projeto de IA em produção, totalmente desenvolvido pela equipe interna do tribunal, encontrando-se disponível para toda a Justiça Federal da 3^a Região e já com o modelo depositado na plataforma Sinapses. Trata-se do projeto SIGMA/SINARA, cuja funcionalidade é gerenciar modelos de despachos e decisões judiciais, auxiliando na produção de minutas.

De acordo com a descrição do projeto constante no painel do CNJ, o sistema SIGMA utiliza o algoritmo SINARA, que extrai o fundamento legal de um texto jurídico, classificando e ranqueando os modelos de atos judiciais que possuem a maior probabilidade de serem selecionados pelo assessor ou magistrado, com base nos modelos que o órgão julgador costuma utilizar.

Consta na descrição que o objetivo da ferramenta é aumentar a rapidez e a qualidade da produção de texto, e que a unidade na qual está funcionando há mais tempo é a Vice-Presidência do TRF3, onde auxiliou a zerar o acervo processual relativo a matéria tributária.

4.1.3 Projetos de IA na Justiça Eleitoral

O terceiro ramo da Justiça brasileira em que há mais projetos de IA é a Justiça Eleitoral. O painel registra 11 projetos, em 9 dos 27 Tribunais Regionais Eleitorais.

O Tribunal Regional Eleitoral de Pernambuco (TRE-PE) possui 2 projetos, que ainda não estão em produção mas cujos códigos já estão hospedados na plataforma Sinapses. As funcionalidades são:

- Combate à desinformação: projeto desenvolvido em parceria com o Centro de Informática da Universidade Federal de Pernambuco (CIn-UFPE), para monitorar redes sociais, analisar textos e identificar depoimentos que veiculem conteúdo falso relacionado à Justiça Eleitoral, ao sistema de voto eletrônico e ao processo eleitoral. Ao ser identificada uma desinformação, o sistema deve emitir material para corrigir a informação falsa divulgada;
- Sistema JANUS: desenvolvido pelo TRE-BA e com cessão formalizada ao TRE-PE, trata-se de uma ferramenta de automação para processos de prestação de contas de candidatos em eleições.

O Tribunal Regional Eleitoral de São Paulo (TRE-SP) também possui 2 projetos de IA registrados no painel, ambos desenvolvidos por equipe interna. Um dos projetos já está em produção, mas nenhum deles tem o código depositado na plataforma Sinapses. São descritos da seguinte forma:

- *Chatbot* para atendimento automatizado dos usuários internos de serviços de TI do tribunal, reduzindo a sobrecarga de demandas de suporte técnico;
- Inteligência Artificial Sophia: para elaborar planos individuais de capacitação de servidores com base em rol de cursos pré-definidos e em avaliação gerencial por competências.

Os tribunais eleitorais abaixo listados possuem apenas 1 projeto de IA, destacando-se que vários informaram ao CNJ que estavam utilizando o sistema JANUS, desenvolvido pelo TRE-BA, porém cada um detalhou a funcionalidade de forma diferente, como se percebe abaixo nos resumos das descrições:

- Tribunal Regional Eleitoral da Bahia (TRE-BA): sistema JANUS, desenvolvido por equipe interna, já em uso e com modelo depositado na plataforma Sinapses, para automação processual de tarefas repetitivas e para classificação de peças processuais;
- Tribunal Regional Eleitoral do Espírito Santo (TRE-ES): Bel (Bot Eleitoral), *chatbot* desenvolvido pela equipe interna do tribunal, cuja funcionalidade é o atendimento do eleitor;
- Tribunal Regional Eleitoral do Maranhão (TRE-MA): sistema JANUS, em colaboração com o TRE-BA, cujo objetivo pretendido é minutar sentenças em processos de registro de candidatura e prestação de contas, baseando-se nos pareceres técnicos e do Ministério Público;
- Tribunal Regional Eleitoral do Piauí (TRE-PI): sistema JANUS, em colaboração com o TRE-BA, para automatizar a elaboração de minutas de sentenças em processos de prestação de contas eleitorais;
- Tribunal Regional Eleitoral do Rio de Janeiro (TRE-RJ): sistema JANUS, em colaboração com o TRE-BA, para movimentação automática de processos de prestação de contas;
- Tribunal Regional Eleitoral do Rio Grande do Norte (TRE-RN): projeto Celina, desenvolvido pela equipe interna do tribunal, que já está em uso e tem a funcionalidade de servir como atendente virtual ao eleitor, podendo ser acessada na página do tribunal no Facebook, no site oficial do tribunal e no aplicativo de mensagens Telegram;
- Tribunal Regional Eleitoral de Sergipe (TRE-SE): projeto CandLe, ainda em estágio inicial de concepção, mas que pretende automatizar atividades em processos de registro de candidaturas, sobretudo quando à análise da comprovação de condições de elegibilidade.

4.1.4 Projetos de IA na Justiça do Trabalho

Dos 24 Tribunais Regionais do Trabalho, 9 possuem ao menos 1 projeto de IA, com destaque para o fato de que alguns projetos são desenvolvidos em colaboração entre tribunais:

- Tribunal Regional do Trabalho da 3^a Região (TRT3/Minas Gerais): Nacionalização Projeto Meta 9, para implantar até dezembro de 2022 solução que permita identificar, organizar e disponibilizar informações de processos a respeito de trabalho infantil, assédio sexual, aprendizagem e trabalho análogo ao escravo;
- Tribunal Regional do Trabalho da 4^a Região (TRT4/Rio Grande do Sul): Índice de Conciliabilidade por Inteligência Artificial (ICia), projeto já em produção, para estimar a possibilidade de haver conciliação em processos na Justiça do Trabalho, com base em processos julgados ou em que houve acordo, objetivando ganho de agilidade e eficiência na seleção de processos para inclusão em pauta de conciliação, com potencial para aumentar a taxa de conciliação a partir da escolha mais assertiva de processos para a pauta;
- Tribunal Regional do Trabalho da 5^a Região (TRT5/Bahia), Tribunal Regional do Trabalho da 7^a Região (TRT7/Ceará) e Tribunal Regional do Trabalho da 15^a Região (TRT15/Campinas-SP): projeto GEMINI, já em produção, tratando-se de um módulo no PJe compatível com todas as instâncias, que permite o agrupamento de processos por similaridade de temas, agilizando a elaboração de minutas de decisões;
- Tribunal Regional do Trabalho da 9^a Região (TRT9/Paraná): Magus, projeto em desenvolvimento para funcionar como assistente de gabinete, mediante extração e classificação de informações e pertinência de resultados de busca por jurisprudência, de modo a auxiliar a elaboração de votos;
- Tribunal Regional do Trabalho da 12^a Região (TRT12/Santa Catarina): projeto Concilia JT, já em uso, para identificar potencial de sucesso na conciliação com base no acervo de processos do tribunal;
- Tribunal Regional do Trabalho da 22^a Região (TRT22/Piauí): projeto ainda em desenvolvimento, mediante parceria com empresa privada, objetivando predizer a possibilidade de acordo em processos trabalhistas;

- Tribunal Regional do Trabalho da 23^a Região (TRT23/Mato Grosso): projeto Sig-JT, ainda em desenvolvimento, mediante colaboração com o TRT15 e o TRT24 para compilar informações visualmente em gráficos e/ou listas de processos, relativas a indicadores e metas, para auxiliar a autogestão e auto-avaliação das unidades judiciárias.

4.1.5 Projetos de IA nos Tribunais Superiores

Dos 5 Tribunais Superiores, somente o Superior Tribunal Militar (STM) não possui projeto de IA.

O Tribunal Superior Eleitoral (TSE) é o que tem mais projetos, em número de 3, que ainda estão em fase de desenvolvimento. Consta no painel que 2 soluções foram adquiridas e 1 desenvolvida por equipe interna. As funcionalidades são as seguintes:

- *Chatbot* para atendimento na Justiça Eleitoral;
- Nova ferramenta de pesquisa de jurisprudência, mediante a solução *ElasticSearch*;
- Projeto para detectar comportamentos inautênticos no e-Título, de modo a evitar que se comprometa o funcionamento correto do sistema.

O Supremo Tribunal Federal (STF) possui 2 projetos de IA, constando no painel que 1 foi desenvolvido em colaboração com outra instituição e 1 por equipe interna:

- Projeto Victor: em produção, auxilia a Secretaria de Precedentes na tarefa de classificar os recursos quanto a temas de repercussão geral;
- Projeto RAFA 2030: ainda em desenvolvimento, pretende auxiliar os servidores da área-fim do STF na classificação e autuação de processos, ao analisar textos jurídicos e classificá-los quanto aos Objetivos de Desenvolvimento Sustentável da Agenda 2030 da ONU.

O Superior Tribunal de Justiça (STJ) possui 1 projeto de IA cadastrado no painel do CNJ, tratando-se do ATHOS, totalmente desenvolvido por equipe interna do tribunal. Já em produção, a ferramenta é usada para agrupamento semântico, pesquisa vetorial e monitoramento de peças processuais.

O Tribunal Superior do Trabalho (TST) também possui 1 projeto de IA, o Bem-te-vi, desenvolvido pela equipe do próprio tribunal. Já em produção, tem a funcionalidade de extrair informações dos sistemas do TST e dos tribunais regionais, auxiliando a gestão e triagem de processos nos gabinetes dos Ministros, sendo capaz inclusive de indicar, com base em processos já julgados pelo Ministro, a possibilidade de denegação por transcendência.

4.1.6 Projetos de IA nos Conselhos

Com 4 projetos de IA registrados no painel, consta o próprio Conselho Nacional de Justiça (CNJ), com as seguintes descrições de funcionalidades:

- Automação de classificações processuais usando IA, no âmbito do Projeto Justiça 4.0: para aplicação de técnicas de processamento de linguagem natural (PLN) com uso de algoritmo de aprendizagem em petições iniciais e documentos comprobatórios, com algoritmos de reconhecimento de entidades e de uma base de conhecimento de legislações, para melhorar a assertividade da distribuição, a classificação das medidas de urgência e evidência, a identificação de prescrição e a prevenção, levando em conta ainda o enquadramento nas Tabelas Processuais Unificadas;
- Busca e análise de precedentes qualificados do STF e STJ, no âmbito do projeto Justiça 4.0: para que o algoritmo seja capaz de identificar precedentes qualificados adequados para o apoio à decisão, de modo a possibilitar o agrupamento por similaridade. O objetivo da solução é disponibilizar os precedentes de forma ordenada, tornando mais fácil para o usuário localizar o precedente de interesse para solução de caso específico que estiver analisando;
- Sinapses - Agrupamento por Similaridade: para identificar e agrupar processos similares, idênticos ou que tenham relação entre si (quanto à causa de pedir e aos pedidos, por exemplo), por citação

- de jurisprudência, legislação ou entendimento doutrinário, facilitando o escalonamento do entendimento das unidades judiciárias de 1^a e 2^o instâncias;
- Plataforma Sinapses: é a própria plataforma nacional de armazenamento, treinamento supervisionado, controle de versionamento, distribuição e auditoria dos modelos de IA, além de estabelecer os parâmetros de sua implementação e funcionamento.

Finalizando a descrição dos 111 projetos de IA existentes oficialmente no Poder Judiciário brasileiro, conforme registro no painel do CNJ até maio de 2022⁴⁸⁷, resta o projeto GEMINI, do Conselho Superior da Justiça do Trabalho (CSJT), em produção, com a funcionalidade de agrupamento de processos por similaridade de temas verificados em peças processuais no sistema de processo eletrônico PJe, possibilitando mais agilidade na tramitação dos feitos e maior padronização de entendimentos.

4.2 Controle dos riscos da IA nos tribunais

Analizando os projetos de IA nos tribunais brasileiros a partir das informações oficiais constantes no painel do CNJ atualizado até maio de 2022⁴⁸⁸, é possível depreender que a esmagadora maioria é voltada para automação da atividade-meio, de administração judiciária, entendida como os atos de gestão que são realizados em setores como cartórios, secretarias e órgãos administrativos das varas, juizados e tribunais, e que viabilizam o exercício da atividade-fim⁴⁸⁹. Esta, por sua vez, é a prestação da jurisdição propriamente dita pelos magistrados, por meio

⁴⁸⁷CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language,BR. Acesso em: 9 jul. 2022.

⁴⁸⁸CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language,BR. Acesso em: 9 jul. 2022.

⁴⁸⁹VIEIRA, Luciano J. M.; PINHEIRO, Ivan A. Contribuições do Conselho Nacional de Justiça para a gestão do Poder Judiciário. In: **XXXII ENCONTRO DA ANPAD**, Rio de Janeiro, 2008. Disponível em: http://www.anpad.org.br/abrir_pdf.php?e=ODcxNQ==. Acesso em: 11 jul. 2022.

dos despachos, decisões e sentenças⁴⁹⁰, aplicando a lei a casos concretos e dizendo quem tem razão ao julgar litígios, ministrando Justiça.

Quando se analisa o gráfico do painel que demonstra as motivações para a criação das ferramentas de IA, vê-se que 94 dos 111 projetos tiveram como principal motivação o aumento de produtividade, para realizar maior volume de tarefas em menor tempo.

Percebe-se que os tribunais estão buscando prioritariamente aproveitar a tecnologia para fazer mais com menos – o que é compreensível quando se leva em conta o contexto da realidade de sobrecarga da Justiça brasileira, já mencionada no capítulo anterior, no item 3.5, quando se fez uma comparação com o cenário europeu, acrescentando agora o déficit da força de trabalho: por um lado, há cerca de 80 milhões de processos, por outro, há 20% de cargos vagos de magistrados e 18% de cargos efetivos de servidores não providos, de acordo com relatório Justiça em Números 2021⁴⁹¹.

Combinando o gráfico sobre as motivações para os projetos de IA com a análise das descrições dos projetos, constata-se que, de fato, a ampla maioria é direcionada a serviços cartorários e administrativos, como triagem, classificação e agrupamento de processos, depreendendo-se que os usuários internos que terão mais contato direto com as ferramentas de IA são os servidores, não os magistrados, e que muitas ferramentas sequer são utilizadas diretamente por pessoas, por serem concebidas como soluções computacionais para se integrarem a sistemas informatizados.

A percepção de uso majoritário da IA na atividade-meio dos tribunais brasileiros é ainda reforçada por outro gráfico constante no painel, demonstrando que, dos 3 recursos mais utilizados nas ferramentas, a análise de texto se sobressai, utilizada em 98 projetos, seguida da organização de dados (em 48 projetos) e da otimização de processos e automação de fluxos de trabalho (em 36 projetos).

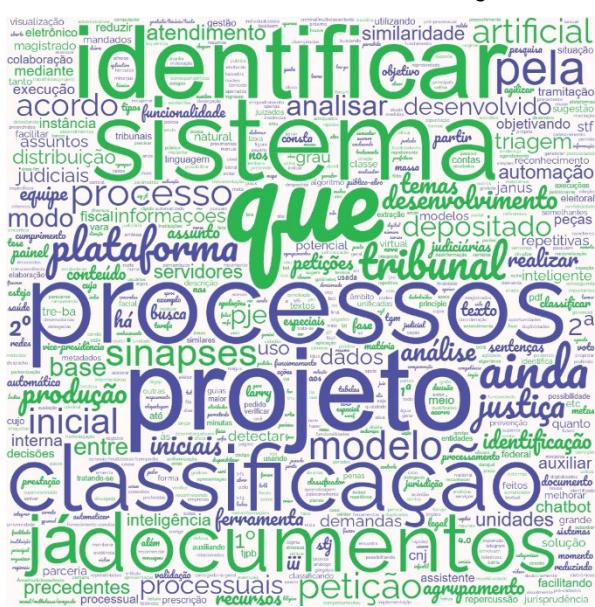
Corroborando tais constatações sobre o uso principal da IA para a área-meio do Judiciário, ao fazer análise utilizando ferramenta geradora de nuvem de palavras, alimentada com todo o texto das descrições dos projetos mencionadas na

⁴⁹⁰BRASIL. Lei nº 13.105, de 16 de março de 2015. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2015-2018/2015/lei/l13105.htm. Acesso em: 9 jul. 2022.

⁴⁹¹CNJ. **Justiça em Números 2021**. Brasília: CNJ, 2021. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2021/09/relatorio-justica-em-numeros2021-12.pdf>. Acesso em: 11 jul. 2022.

seção anterior, o resultado (Figura 10) demonstra que entre as palavras mais recorrentes estão “identificar”, “processos”, classificação”, “documentos”.

Figura 10 - Palavras mais recorrentes nas descrições dos projetos de IA



Fonte: Gerada pela autora via WordClouds.com⁴⁹².

Algumas disposições da Resolução 332/2020⁴⁹³ parecem ter encontrado eco nos tribunais, a exemplo da recomendação de cautela prevista no art. 23, que desestimulou a utilização de modelos de IA em matéria penal, principalmente relativos à sugestão de modelos de decisões preditivas, bem como à verificação de reincidência penal. Como visto nos capítulos anteriores, toda a polêmica em torno do algoritmo COMPAS influenciou a Carta Ética Europeia e, por sua vez, a Resolução brasileira de ética aplicada à IA, o que explica a inserção do referido art. 23 no ato normativo.

Com efeito, dos 111 projetos de IA, nenhum é voltado para avaliação de risco no âmbito da persecução criminal ou da execução penal. É um cenário bem diferente da larga utilização de ferramentas desse tipo no sistema de Justiça criminal em diversos Estados norte-americanos, consideradas de alto risco de lesão a direitos humanos fundamentais, e que são tão criticadas pela comunidade acadêmica e pela

⁴⁹²WORDCLOUDS.COM. Free online word cloud generator and tag cloud creator. Disponível em: <https://www.wordclouds.com/>. Acesso em: 12 jul. 2022.

493 CNJ. Resolução nº 332, de 21 de agosto de 2020. Disponível em:
<https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 12 jul. 2022.

sociedade civil, por questões de falta de transparência e de viés discriminatório, conforme o trabalho da professora Cynthia Rudin⁴⁹⁴ e o relatório *Liberty at risk*⁴⁹⁵, referidos nos itens 3.4.4 e 3.4.5 do capítulo anterior.

Como a IA é uma área extremamente dinâmica, é possível que, no futuro próximo, com os avanços das pesquisas sobre o tema, a exemplo das novas técnicas de IA interpretável em vez de explicável propostas por Cynthia Rudin, aliadas a precauções especiais contra discriminação na coleta, curadoria e tratamento dos dados para treinamento dos modelos, venham a surgir em tribunais brasileiros projetos na área de avaliação de risco na seara criminal, que sejam comprovadamente úteis e ao mesmo tempo tenham riscos mais controlados.

Outro dispositivo de cautela presente na Resolução 332/2020, que em certa medida também pode ter sua inserção no ato normativo explicada por guardar relação com os aprendizados em torno da discussão sobre o COMPAS, foi o art. 24, *caput*, que estabeleceu o uso preferencial de *softwares* de código aberto.

Como analisado no item 3.4.4 do capítulo anterior, a empresa criadora do sistema COMPAS, em se tratando de empresa privada que visa ao lucro, alegou o sigilo empresarial e a proteção à propriedade intelectual de seu código-fonte para não ser obrigada a informá-lo à Justiça, tese que foi acolhida pela Suprema Corte de Wisconsin, sob críticas da comunidade acadêmica e da sociedade civil, por impossibilitar o conhecimento dos mecanismos pelos quais o *software* chega aos resultados que são utilizados pelo sistema de Justiça criminal norte-americano para conceder ou negar liberdade a réus.

Já no Brasil, pelo que consta no painel de projetos de IA do CNJ, é possível perceber que o art. 24, *caput*, da Resolução 332/2020, assim como o art. 23, também encontrou ressonância nos tribunais, pois os órgãos do Poder Judiciário são detentores dos códigos de 100 dos 111 projetos.

Entretanto, cabe ponderar que o gráfico relativo à pergunta “A quem pertence o algoritmo da ferramenta e seu código?”⁴⁹⁶ indica conter possível

⁴⁹⁴RUDIN, Cynthia. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, v. 1, n. 5, p. 206-215, 2019.

⁴⁹⁵EPIC. **Liberty at risk: pre-trial risk assessment tools in the U.S.** Set. 2020. Disponível em: <https://epic.org/wp-content/uploads/2022/02/Liberty-At-Risk-Report-FALL-2020-UPDATE.pdf>. Acesso em: 2 jul. 2022.

⁴⁹⁶CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language.BR. Acesso em: 9 jul. 2022.

sobreposição de respostas, porque a soma total de respostas resulta em 127, quando os projetos são apenas 111: (i) Tribunal/Conselho (100); (ii) Terceiros (10); (iii) Código aberto (7); (iv) UFRN; (v) Ainda não iniciado/em fase de pesquisa (1); (vi) CSJT (1); (vii) Parte do código e pesquisas estão no GitHub (1); (viii) Pertencem à empresa privada que desenvolve o código (1); (ix) TRE-BA (1); (x) UnB (1).

Outro gráfico vizinho no painel traz, à pergunta correlata “Sua instituição tem acesso ao código e funcionamento do algoritmo?”⁴⁹⁷, as seguintes respostas: Sim (93), Não (6), Não aplicável (7), Não sei (5).

Há ainda, nas respostas sobre as equipes desenvolvedoras, que 3 soluções foram adquiridas e 1 é desenvolvida em parceria com *startup*.

Analizando essas respostas conjugadas, em cotejo com as descrições dos projetos, apesar das disparidades mencionadas é possível inferir que o Poder Judiciário brasileiro possui e/ou tem acesso à franca maioria dos códigos dos modelos de IA utilizados por seus órgãos e nem de longe está refém de empresas privadas ou sigilos empresariais/industriais em área tão estratégica para a instituição.

Quanto a uma outra utilização de alto risco – reconhecimento facial – que inclusive foi tratada com precaução máxima pela Resolução 332/2020⁴⁹⁸ (art. 22, § 2º), sendo a única aplicação de IA que exige prévia autorização do CNJ para implementação (quando as outras só precisam ser comunicadas), o painel de projetos de IA do CNJ demonstra que há um tribunal, o TJDFT, que está utilizando a técnica para dois projetos, Amon e SAREF.

Como já referido na seção anterior, o projeto Amon se destina ao controle de quem transita nas dependências do tribunal por meio de reconhecimento facial, constando na descrição do projeto que busca “trazer mais segurança aos jurisdicionados”⁴⁹⁹.

Segundo artigo⁵⁰⁰ de Melo, Neves e Neto, publicado em 2021 na Revista Eletrônica do CNJ sobre o Amon, o TJDFT possui 19 circunscrições no território do

⁴⁹⁷Ibidem.

⁴⁹⁸CNJ. **Resolução nº 332, de 21 de agosto de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 14 jul. 2022.

⁴⁹⁹CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,currsel&select=language,BR. Acesso em: 14 jul. 2022.

⁵⁰⁰MELO, Jairo Simão Santana; NEVES, Thiago Arruda; NETO, Celso Oliveira. AMON: controle de acesso do jurisdicionado no TJDFT a partir de técnicas de reconhecimento facial. **Sistema e-Revista CNJ**, v. 5, n. 1, p. 129-140, 2021. Disponível em: <https://www.cnj.jus.br/ojs/index.php/revista-cnj/article/view/157/110>. Acesso em: 14 jul. 2022.

Distrito Federal, por onde circulam diariamente milhares de pessoas, existindo controle de acesso nas portarias dos prédios, em que é preciso apresentar documento de identificação para obter autorização a adentrar, havendo ainda outros controles como *scanners* e portas giratórias. Os dados dos visitantes, incluindo a unidade judiciária em que desejam ir, são registrados no sistema informatizado SidenWeb, que a partir de 2019 também armazena suas fotografias.

Esse banco de dados formado a partir das fotografias permitiu a criação do projeto Amon, que conforme o artigo entrou em operação em junho de 2020, com 2 parâmetros: (i) uso do modelo CNN (*Convolutional Neural Network*/Rede Neural Convolucional), escolhido por ser mais preciso e robusto, embora mais lento que a outra opção considerada pela equipe técnica; (ii) configuração da acurácia desejada⁵⁰¹.

No tocante à precisão do sistema, verificou-se no recorte temporal examinado pelos autores do artigo uma taxa de erro em torno de 25,16%, distribuída entre erros no reconhecimento facial e reconhecimento facial sem cadastro⁵⁰².

Quanto à efetiva utilidade do sistema, também consta no artigo que foram detectadas 92 ocorrências de possíveis fraudes, indicando situações de pessoas que poderiam já ter visitado os prédios do tribunal antes, porém portando documentos diferentes⁵⁰³.

O artigo contém ainda uma seção sobre perspectivas futuras, como: (i) testes de cruzamento dos metadados retornados do sistema de reconhecimento facial, usando o CPF cadastrado para verificar incidências de uma pessoa no SEEU – Sistema Eletrônico de Execução Unificado; (ii) aplicar o reconhecimento facial em vídeos ao vivo nas câmeras dos circuitos internos dos prédios do tribunal, para verificar, por exemplo, se alguém está transitando sem permissão em determinada área; (iii) usar a técnica para controle e reconhecimento das visitas de presos, evitando que pessoas se façam passar por outras usando documentos diferentes, incluindo identificação de advogados; (iv) aplicar o reconhecimento facial para as situações de apresentação de apenados beneficiados pela suspensão da pena, liberdade

⁵⁰¹MELO, Jairo Simão Santana; NEVES, Thiago Arruda; NETO, Celso Oliveira. AMON: controle de acesso do jurisdicionado no TJDFT a partir de técnicas de reconhecimento facial. **Sistema e-Revista CNJ**, v. 5, n. 1, p. 129-140, 2021. Disponível em: <https://www.cnj.jus.br/ojs/index.php/revista-cnj/article/view/157/110>. Acesso em: 14 jul. 2022.

⁵⁰²Ibidem.

⁵⁰³Ibidem.

provisória ou livramento condicional, submetidos à obrigação de comparecimento periódico; bem como à fiscalização de recolhimento domiciliar⁵⁰⁴.

Para essa última aplicação, foi explicado o contexto de que o TJDFT possuía um contingente de mais de 20 mil jurisdicionados sujeitos à obrigação de se apresentar em juízo, e que foi preciso encontrar alternativas para atender a essa população durante o período de isolamento social na pandemia de Covid-19, quando houve suspensão de atividades presenciais⁵⁰⁵.

Dessa necessidade resultou o SAREF (Sistema de Apresentação e Reconhecimento Facial), descrito no painel de projetos de IA do CNJ como uma aplicação que tem a finalidade de “agilizar a apresentação de uma população de apenados da unidade”, funcionando na Vara de Execuções das Penas em Regime Aberto e Vara de Execuções das Penas e Medidas Alternativas. Consta que torna a apresentação remota do apenado “mais rápida e segura, além de contribuir também para a eliminação de filas e aglomerações”, visto que seu objetivo é permitir que o usuário se apresente perante a Justiça por meio de um celular ou computador com acesso à Internet, sem precisar se deslocar até o fórum⁵⁰⁶.

Dito o que foi possível obter em termos de informações sobre os sistemas de reconhecimento facial usados pelo TJDFT, cabe tecer algumas considerações.

Para começar, não há qualquer indicativo, no painel do CNJ sobre projetos de IA nos tribunais, a respeito da autorização prévia necessária ao reconhecimento facial, prevista no art. 22, § 2º, da Resolução 332/2020, que tenha sido formalmente concedida ao TJDFT.

Pode-se até presumir que a autorização existe, afinal o artigo mencionado foi publicado na Revista Eletrônica do próprio CNJ, mas o fato é que, em se tratando de uma aplicação de IA que mereceu a maior das precauções no ato normativo, estando subjacente seu grau de risco, convém que a autorização esteja no mínimo sinalizada no painel, ou, de modo ideal, destacada e divulgada publicamente de modo claro, inequívoco e transparente.

⁵⁰⁴Ibidem.

⁵⁰⁵Ibidem.

⁵⁰⁶CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language.BR. Acesso em: 14 jul. 2022.

Quanto à utilidade da ferramenta Amon, os primeiros resultados divulgados no artigo⁵⁰⁷ de Melo, Neves e Neto demonstram que o sistema de reconhecimento facial integrado ao sistema de cadastro dos visitantes realmente contribui, a princípio, para o incremento da segurança dentro dos prédios do TJDFT, tanto que detectou 92 possíveis casos de fraudes, quanto a pessoas que podiam estar portando documentos de outras, o que representa, se confirmadas as fraudes, um perigo para a segurança de jurisdicionados, servidores e magistrados.

Mas a questão que se coloca é que, atualmente, o debate sobre a tecnologia de reconhecimento facial não pode deixar de ser concomitante ao debate ético quanto a prejuízos para negros e mulheres, desde a pesquisa pioneira de Joy Buolamwini e Timnit Gebru⁵⁰⁸ (referida no capítulo anterior no item 3.4.4), com todos os impactos que teve, gerando um debate ético que culminou até na decisão da gigante de tecnologia IBM, em junho de 2020, de suspender a comercialização de seu *software* de reconhecimento facial⁵⁰⁹.

Cabe a seguinte reflexão: se até a IBM, mesmo após os esforços de treinamento com massivos dados fenotípicos diversos, constatou que os resultados continuavam sendo prejudiciais a negros e mulheres em níveis inaceitáveis, será que a equipe técnica do TJDFT, que não chega a 5 pessoas de acordo com o painel do CNJ, conseguiu ou conseguirá atingir um nível aceitável que a gigante de tecnologia ainda não alcançou?

Percebe-se que, no artigo⁵¹⁰ de Melo, Neves e Neto a questão sequer é tangenciada, não havendo qualquer menção ao assunto relativo a risco de viés discriminatório, apesar de ser um tema candente na área de reconhecimento facial.

⁵⁰⁷MELO, Jairo Simão Santana; NEVES, Thiago Arruda; NETO, Celso Oliveira. AMON: controle de acesso do jurisdicionado no TJDFT a partir de técnicas de reconhecimento facial. **Sistema e-Revista CNJ**, v. 5, n. 1, p. 129-140, 2021. Disponível em: <https://www.cnj.jus.br/ojs/index.php/revista-cnj/article/view/157/110>. Acesso em: 14 jul. 2022.

⁵⁰⁸BUOLAMWINI, Joy; GEBRU, Timnit Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In: **Conference on fairness, accountability and transparency**. PMLR, 2018. p. 77-91. Disponível em: http://proceedings.mlr.press/v81/buolamwini18a.html?mod=article_inline. Acesso em: 12 jul. 2022.

⁵⁰⁹MEYER, David. **IBM pulls out of facial recognition, fearing racial profiling and mass surveillance**. Fortune. 9 jun. 2020. Disponível em: <https://fortune.com/2020/06/09/george-floyd-ibm-exits-facial-recognition-bias-human-rights/>. Acesso em: 13 jul. 2022.

⁵¹⁰MELO, Jairo Simão Santana; NEVES, Thiago Arruda; NETO, Celso Oliveira. AMON: controle de acesso do jurisdicionado no TJDFT a partir de técnicas de reconhecimento facial. **Sistema e-Revista CNJ**, v. 5, n. 1, p. 129-140, 2021. Disponível em: <https://www.cnj.jus.br/ojs/index.php/revista-cnj/article/view/157/110>. Acesso em: 14 jul. 2022.

Quando o artigo menciona os erros detectados, como não tocou no tema da possibilidade de viés algorítmico discriminatório, não testou hipóteses nesse sentido sobre causas dos erros, como também não fez correlações a maior ou menor acurácia, a depender da cor da pele das pessoas cuja imagem é capturada e classificada, a fim de verificar se há ou não incidência de erros em percentuais díspares demais.

Acontece que, para um sistema de reconhecimento facial em uso justamente pelo Poder Judiciário, não se pode deixar de lado o debate sobre o que é ético e justo, pois o risco de tratamento discriminatório de jurisdicionados é real com essa tecnologia específica, segundo apontam cientistas que são consideradas algumas das maiores especialistas do mundo nesse tema.

A questão ética torna-se ainda mais relevante quanto às perspectivas futuras narradas no artigo⁵¹¹, de reconhecimento facial em tempo real por câmeras ao vivo, que também levanta debates sobre a hipervigilância estatal *versus* o direito constitucional à privacidade e à proteção dos dados pessoais (que agora no Brasil também tem *status* de direito fundamental, desde a recente Emenda Constitucional nº 115/2022)⁵¹².

O artigo ainda fala sobre possível uso futuro em carceragens para o reconhecimento de visitas de presos, inclusive advogados, o que faz acender um alerta para a necessária discussão ética e legal, tendo em vista os direitos e prerrogativas dos advogados, previstos na Lei nº 8.906/1994⁵¹³.

Outro possível uso futuro muito preocupante é o de fazer correlações com informações de sistemas de execução penal, caso venham, na realidade, a se destinar à captura de pessoas com mandado de prisão expedido, pois a pesquisa de Buolamwini e Gebru, como se viu no item 3.4.4, demonstrou a alta taxa de erro para negros em sistemas utilizados em grandes empresas de tecnologia, como Microsoft e IBM. Não se pode esquecer que pretos e pardos representam 67,34% da população

⁵¹¹MELO, Jairo Simão Santana; NEVES, Thiago Arruda; NETO, Celso Oliveira. AMON: controle de acesso do jurisdicionado no TJDFT a partir de técnicas de reconhecimento facial. **Sistema e-Revista CNJ**, v. 5, n. 1, p. 129-140, 2021. Disponível em: <https://www.cnj.jus.br/ojs/index.php/revista-cnj/article/view/157/110>. Acesso em: 14 jul. 2022.

⁵¹²BRASIL. **Emenda Constitucional nº 115, de 10 de fevereiro de 2022**. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc115.htm. Acesso em: 15 jul. 2022).

⁵¹³BRASIL. **Lei nº 8.906, de 4 de julho de 1994**. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8906.htm. Acesso em: 15 jul. 2022.

carcerária nos Estados brasileiros, segundo dados do Departamento Penitenciário Nacional apurados de julho a dezembro de 2021⁵¹⁴.

Para se ter uma dimensão dos riscos de erros para tal aplicação cogitada, vale mencionar o teste feito pela *American Civil Liberties Union* (ACLU), importante organização não governamental norte-americana, já centenária, que existe desde o fim da Primeira Guerra Mundial para combater abuso do governo e defender as liberdades individuais, e que ao longo de mais de um século se envolveu ativamente em inúmeros debates jurídicos relevantes, sendo a instituição que mais participou de casos julgados na Suprema Corte dos EUA⁵¹⁵.

Foi justamente essa instituição que foi ouvida em maio de 2019 no Congresso dos EUA, na mesma audiência pública⁵¹⁶ sobre a tecnologia de reconhecimento facial de que também participou Joy Buolamwini, como mencionado no capítulo anterior deste trabalho, no item 3.4.4, e foi o representante da ACLU que chegou a pedir moratória do uso da tecnologia enquanto não houvesse regulamentação. Aliás, é essa sequência de fatos, entre outros, que contextualiza a decisão particular de moratória da IBM, de parar de comercializar seu *software*, tomada sob pressão em junho de 2020.

Entre os argumentos para justificar o pedido de moratória dessa tecnologia ao Congresso dos EUA, a ACLU valeu-se de um teste⁵¹⁷ que realizou em 2018, utilizando o *Rekognition*, o *software* de reconhecimento facial da Amazon, que está disponível a qualquer pessoa que queira fazer correspondências entre imagens do rosto de pessoas.

A ACLU divulgou que “executar todo o teste nos custou US\$ 12,33 – menos do que uma pizza grande”, e que usou 25 mil fotos de pessoas presas, disponíveis publicamente, encontradas com ferramentas simples de busca. Formado esse banco de dados, foram então comparados os rostos de todos os membros do Congresso dos

⁵¹⁴BRASIL. **Levantamento nacional de informações penitenciárias – período de julho a dezembro de 2021.** Disponível em: <https://www.gov.br/depes/pt-br/servicos/sisdepen/sisdepen>. Acesso em: 15 jul. 2022.

⁵¹⁵ACLU. **ACLU History.** Disponível em: <https://www.aclu.org/about/aclu-history>. Acesso em: 15 jul. 2022.

⁵¹⁶QUACH, Katyanna. **We listened to more than 3 hours of US Congress testimony on facial recognition so you didn't have to go through it.** The Register. 22 maio 2019. https://www.theregister.com/2019/05/22/congress_facial_recognition/. Acesso em: 13 jul. 2022.

⁵¹⁷JACOB, Snow. **Amazon's face recognition falsely matched 28 members of Congress with mugshots.** ACLU. 26 jul. 2018. Disponível em: <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>. Acesso em: 15 jul. 2022.

EUA, representantes do povo norte-americano, de modo que havia republicanos e democratas, homens e mulheres, de idades e características físicas variadas⁵¹⁸.

O resultado do teste é que foram encontradas 28 falsas correspondências, identificando esses 28 congressistas norte-americanos como pessoas presas que cometem crimes, atingindo desproporcionalmente parlamentares de pele escura, no percentual de 40%, embora os congressistas negros fossem 20% dos membros⁵¹⁹.

As falsas correspondências atingiram 3 senadores e 25 deputados (Figura 11), “entre eles a lenda dos direitos civis Rep. John Lewis (D-Ga)”⁵²⁰. Falecido em 2020, primeiro parlamentar negro velado no Capitólio⁵²¹, John Lewis foi o líder negro mais jovem na Marcha para Washington em 1963, ocasião em que Martin Luther King proferiu o discurso histórico “*I have a dream*”⁵²². Todavia, ao tempo do teste realizado em 2018, se fosse abordado por um policial usando o *software* de reconhecimento facial da Amazon, John Lewis seria dado como um criminoso.

Figura 11 – Congressistas dos EUA incorretamente reconhecidos por IA



Fonte: Jacob Snow/ACLU⁵²³.

⁵¹⁸Ibidem.

⁵¹⁹JACOB, Snow. **Amazon's face recognition falsely matched 28 members of Congress with mugshots.** ACLU. 26 jul. 2018. Disponível em: <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>. Acesso em: 15 jul. 2022.

⁵²⁰Ibidem.

⁵²¹ASSOCIATED PRESS. **John Lewis é o primeiro parlamentar negro a ser velado na Rotunda do Capitólio, em Washington.** 27 jul. 2020. Disponível em: <https://g1.globo.com/mundo/noticia/2020/07/27/john-lewis-e-o-primeiro-parlamentar-negro-a-ser-velado-na-rotunda-do-capitolio-em-washington.ghtml>. Acesso em: 15 jul. 2022.

⁵²²G1. **Pioneiro na luta pelos direitos civis, John Lewis morre aos 80 anos nos EUA.** 18 jul. 2020. <https://g1.globo.com/mundo/noticia/2020/07/18/pioneiro-na-luta-pelos-direitos-civis-deputado-john-lewis-morre-nos-eua.ghtml>. Acesso em: 15 jul. 2022.

⁵²³JACOB, Snow. Op. cit.

Consta que os perigos envolvendo o fornecimento do *software Rekognition* para forças de segurança (que ao tempo do teste feito pela ACLU já estava sendo usado no Estado do Oregon), ensejou manifestações públicas da parte de funcionários da própria Amazon, acionistas, 70 organizações não governamentais de defesa dos direitos civis, mais de 400 membros da academia e mais de 150.000 cidadãos, em soma de esforços para pressionar pela interrupção do fornecimento de vigilância facial a governos⁵²⁴.

Reforçando os riscos do *Rekognition*, foi divulgada em janeiro de 2019 uma nova pesquisa científica⁵²⁵ de Inioluwa Deborah Raji e Joy Buolamwini sobre o tema, que recebeu ataques⁵²⁶ e questionamentos⁵²⁷ da Amazon, refutados por Buolamwini⁵²⁸, com o reforço do apoio formal que recebeu de diversos cientistas de IA importantes no cenário mundial, incluindo Yoshua Bengio, um dos vencedores do Prêmio Turing de 2018⁵²⁹, em manifestação técnica coletiva divulgada em maio de 2019⁵³⁰.

O orientador de Buolamwini no doutorado, Ethan Zuckerman, diretor do *Center for Civic Media* do MIT, também se manifestou⁵³¹, dizendo que se sentia compelido a falar sobre o assunto para chamar a atenção quanto ao comportamento

⁵²⁴JACOB, Snow. **Amazon's face recognition falsely matched 28 members of Congress with mugshots.** ACLU. 26 jul. 2018. Disponível em: <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>. Acesso em: 15 jul. 2022.

⁵²⁵RAJI, Inioluwa Deborah; BUOLAMWINI, Joy. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In: **Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society**. 2019. p. 429-435. Disponível em: <https://doi.org/10.1145/3306618.3314244>. Acesso em: 16 jul. 2022.

⁵²⁶WOOD, Matt. **Thoughts on recent research paper and associated article on Amazon Rekognition.** AWS Machine Learning Blog. 26 jan. 2019. Disponível em: <https://aws.amazon.com/pt/blogs/machine-learning/thoughts-on-recent-research-paper-and-associated-article-on-amazon-rekognition/>. Acesso em: 16 jul. 2022.

⁵²⁷PUNKE, Michael. **Some thoughts on facial recognition legislation.** AWS Machine Learning Blog. 7 fev. 2019. Disponível em: <https://aws.amazon.com/pt/blogs/machine-learning/some-thoughts-on-facial-recognition-legislation/>. Acesso em: 16 jul. 2022.

⁵²⁸BUOLAMWINI, Joy. **Response: racial and gender bias in Amazon Rekognition - commercial AI system for analyzing faces.** Medium. 25 jan. 2019. Disponível em: <https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a28922eeecd>. Acesso em: 16 jul. 2022.

⁵²⁹BASS, Dina. **Amazon schooled on AI facial technology by Turing Award winner.** Bloomberg. 3 abr. 2019. Disponível em: <https://www.bloomberg.com/news/articles/2019-04-03/amazon-schooled-on-ai-facial-technology-by-turing-award-winner#xj4y7vzkg>. Acesso em: 15 jul. 2022.

⁵³⁰CONCERNED RESEARCHERS. **On recent research auditing commercial facial analysis technology.** Medium. 26 mar. 2019. Disponível em: <https://medium.com/@bu64dcjrytwitb8/on-recent-research-auditing-commercial-facial-analysis-technology-19148bda1832>. Acesso em: 16 jul. 2022.

⁵³¹ZUCKERMAN, Ethan. **The narrative of public criticism.** MIT Media Lab. 7 fev. 2019. Disponível em: <https://civic.mit.edu/index.html%3Fp=2440.html>. Acesso em: 16 jul. 2022.

das poderosas grandes empresas de tecnologia diante de pesquisas científicas sérias, em trabalhos revisados por pares, que apontavam resultados imperfeitos de produtos e serviços, causando um desconforto para as organizações.

No caso do *Rekognition*, a Amazon reagiu desqualificando a pesquisa, chamada de falsa e enganosa, atacando as pesquisadoras, mantendo uma atitude de negação, além de se recusar a fornecer conjuntos de dados para testes e treinamentos. Foi um comportamento contrário à reação da IBM e Microsoft, outras gigantes de tecnologia também pesquisadas por Joy Buolamwini, que analisaram os problemas por ela apontados, passaram a trabalhar para tentar resolvê-los, divulgar os resultados dos avanços e colaborar com a academia, fornecendo material para testes⁵³².

Essa reflexão de Ethan Zuckerman é útil ao Poder Judiciário brasileiro na medida em que lança luzes sobre quais as melhores atitudes que os gestores do CNJ e dos tribunais poderão adotar quando começarem a surgir os primeiros desconfortos causados por críticas a modelos de IA, baseadas em resultados negativos constatados em trabalhos científicos sérios advindos da comunidade acadêmica.

Dependendo da dimensão das falhas encontradas, pode eclodir uma séria crise de imagem da instituição, a necessitar de gerenciamento que, para ser eficiente, convém que seja pensado e planejado com antecedência, pois o que estará em jogo é a confiança da sociedade no Poder Judiciário.

Como se pode perceber, há muitas discussões éticas e seus desdobramentos acontecendo nos últimos anos, envolvendo modelos de IA em geral, e em particular aplicações para reconhecimento facial.

Não se está proondo, aqui, tolher a inovação representada pelo Amon, cujo acompanhamento criterioso dos resultados poderá vir a demonstrar, de modo irrefutável, ter a utilidade benéfica de incrementar a segurança nos prédios do TJDFT, para o qual foi inicialmente proposto.

O que se está fazendo neste momento é levantar o debate ético quanto à utilização atual e a usos futuros para outras finalidades. Aliás, cabe questionar: no caso específico da tecnologia de reconhecimento facial, em que incide a regra de precaução máxima prevista na Resolução 332/2020, quanto à necessidade de prévia

⁵³²ZUCKERMAN, Ethan. **The narrative of public criticism**. MIT Media Lab. 7 fev. 2019. Disponível em: <https://civic.mit.edu/index.html%3Fp=2440.html>. Acesso em: 16 jul. 2022.

autorização, para cada nova finalidade da mesma ferramenta há que se obter nova autorização?

Parece que a resposta é afirmativa, dada a necessidade de adotar cautelas especiais no uso dessa tecnologia, aplicando na prática o princípio da precaução. Adicionalmente, é prudente fazer mais pesquisas sobre o tema, além de verificar e auditar a ferramenta com frequência.

Enquanto isso, afigura-se importante capacitar os servidores que atuam nas portarias dos prédios e são usuários internos do sistema Amon, para que fiquem cientes das taxas de erros e sejam treinados para agir corretamente em casos de falsas correspondências, tratando o visitante com urbanidade ao tempo em que verificam sua identidade de outras formas, sem presumir de antemão que se trata de uma pessoa fraudadora/criminosa.

Ainda assim, é possível que alguém se sinta prejudicado com eventual erro da máquina, afigurando-se necessário existir um canal para registrar o fato de modo que seja útil não só para a adoção de providências, mas para catalogação de erros, verificação e auditoria, com vistas ao aprimoramento do sistema – assunto que será tratado na seção seguinte.

Adicionalmente, quanto aos casos em que há fortes indícios de fraude, demonstra-se importante haver um protocolo para essa contingência, incluindo a posterior verificação do resultado da persecução criminal, tratando-se de um dado que, somado às outras ocorrências semelhantes e observados em recortes temporais, podem ter o condão de demonstrar continuamente a utilidade, razão de ser e propósito dessa ferramenta de IA.

Acima de tudo, cabe questionar: caso se verifique que o sistema Amon retorna resultados desproporcionalmente prejudiciais a negros e mulheres, seria tolerável submeter, por exemplo, 20%, 30%, 40% ou qualquer outro percentual dos visitantes do tribunal a procedimentos adicionais de verificação de identidade, somente porque são negros e mulheres?

Estariam tendo essas pessoas a garantia constitucional de seu direito à igualdade e acesso à Justiça? A sensação de humilhação e discriminação que poderiam ser experimentadas pelas pessoas verificadas iria desaparecer por mais urbanidade que houvesse no tratamento a elas dispensado?

E mais: considerando que o TJDFT possui uma Comissão Multidisciplinar de Inclusão, presidida por uma importante liderança negra no cenário nacional, que

faz um trabalho pioneiro de inclusão e enfrentamento à discriminação (conforme mencionado no item 3.4.3), qual seria o parecer de tal órgão sobre essas questões levantadas quanto ao sistema Amon, se a Comissão fosse instada a opinar?

No tocante à utilidade do SAREF⁵³³, de proporcionar a opção ao apenado de se apresentar à Justiça de forma remota, a princípio demonstra ser um uso mais controlado da tecnologia de reconhecimento facial, dentro de uma população específica e não aleatória, previamente cadastrada e confirmada, tratando-se de uma ferramenta pensada para ampliar possibilidades e não restringir direitos.

Isso porque, mesmo que o sistema eventualmente não reconheça o apenado, não significa que ele automaticamente estará sujeito à perda do benefício de permanecer em liberdade, apenas terá de comparecer presencialmente na unidade judiciária, da forma tradicional como sempre se fez.

Mesmo assim, mais uma vez é preciso ter em mente o fenótipo da maioria da população carcerária brasileira, sendo esperado que erros de reconhecimento facial ocorram em razão das limitações dessa tecnologia. Nesses casos, continua sendo de grande utilidade catalogar e pesquisar os erros de reconhecimento verificados no SAREF, cujo exame contribui para o aprimoramento do sistema.

Sob outro prisma, coloca-se uma questão no campo ético-filosófico, e também jurídico-legal quanto ao SAREF, que é fruto da inovação por necessidade, em meio à pandemia: esse comparecimento remoto do apenado atende à razão de ser da obrigação de comparecimento periódico em juízo?

Na verdade, para todas as inovações surgidas nesse anos pandêmicos, inclusive a IA, cabe sempre questionar: essa inovação é realmente benéfica para o sistema de justiça, para a sociedade, para o interesse público?

Diante de tantas mudanças, S. Tavares Pereira, em sua obra *Machine Learning nas Decisões*⁵³⁴, afirma (destacou-se):

[...] muitas das transformações, feitas sob a pressão da doença, (i) ou se basearam em tecnologias antigas, já disponíveis antes da pandemia, e que não era adotadas pela resistência da inércia, finalmente vencida, ou (ii) precisarão ser reavaliadas sob luzes constitucionais, legais, filosóficas e sociais, pois feitas acriticamente para atendimento das circunstâncias.

⁵³³TJDFT. **Sistema de apresentação e reconhecimento facial - SAREF.** Disponível em: <https://saref.tjdft.jus.br/>. Acesso em: 15 jul. 2022.

⁵³⁴PEREIRA, S. Tavares. **Machine learning nas decisões: o uso jurídico dos algoritmos aprendizes.** 1^a edição. Santa Catarina/Florianópolis, 2021, p. 786-787.

Nesse passo, há uma maneira para reavaliar os projetos de IA dos tribunais brasileiros sob luzes constitucionais, legais e éticas, que pode ser determinada pelo CNJ a todos os tribunais, tanto no que se refere aos projetos que já existem como também aos que vierem a ser criados, e que também pode constar como requisito para o Prêmio CNJ de Qualidade: instar os tribunais a realizar procedimentos que levem à classificação de risco dos modelos de IA.

Abre-se ensejo, aqui, para destacar um ponto importante atualmente, com o avanço dos estudos e pesquisas sobre o tema: a necessidade de aprimoramento, pelo CNJ, da regulamentação do uso da IA nos tribunais, avançando para incluir a classificação dos projetos quanto ao risco, que é o caminho que: (i) o Parlamento Europeu está trilhando nas discussões sobre o *AI Act*⁵³⁵, (ii) que também consta no Projeto de Lei nº 21-A/2020⁵³⁶ enviado pela Câmara dos Deputados ao Senado brasileiro, conforme já tratado no capítulo anterior, no item 3.2., e (iii) que a ciência já trilhou no passado quanto à tecnologia de DNA recombinante, que era nova e arriscada em 1975, como dito no primeiro capítulo deste trabalho, no item 2.1.

Note-se que, apenas consultando o painel do CNJ sobre os projetos de IA nos tribunais, não é possível saber de pronto quais ferramentas são de risco baixo, moderado ou alto. Essa é uma informação importante não só para a gestão dos riscos, de modo a indicar quais aplicações estão sujeitas a maior ou menor regulação e controle, como também para prestar informação clara à comunidade acadêmica e à sociedade civil – o que contribui inclusive para conquistar e manter a confiança nas novas tecnologias usadas pelo Poder Judiciário.

Uma possibilidade para implementar a classificação de risco dos projetos de IA nos tribunais é tomar por parâmetro o modelo proposto em fevereiro de 2022 pela OCDE, elaborado por mais de 60 especialistas, no trabalho *OECD Framework*

⁵³⁵COMISSÃO EUROPEIA. **Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União.** Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>. Acesso em: 15 jul 2022.

⁵³⁶BRASIL. **Projeto de Lei nº 21-A de 2020.** Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarIntegra?codteor=2129459. Acesso em: 15 jul. 2022.

*for the classification of AI systems*⁵³⁷, com as necessárias adaptações para a realidade do Poder Judiciário brasileiro.

Outro instrumento interessante que pode ser adaptado para a Justiça brasileira no controle dos projetos de IA é o formulário utilizado no Reino Unido para avaliar o potencial de risco de projetos, o *Risk Potential Assessment (RPA)*⁵³⁸.

A pesquisa documental conduzida neste trabalho chegou a tal formulário quando se buscou saber o que ocorreu depois de 2018 com a ferramenta britânica HART, que havia sido mencionada como arriscada na Carta Ética Europeia pelas críticas de viés discriminatório e falta de transparência, assim como o norte-americano COMPAS, conforme já tratado no primeiro e segundo capítulos desta dissertação, nos itens 2.3 e 3.4.4.

Apurou-se que em janeiro de 2022, sem fazer menção às críticas, a Polícia de Durham anunciou⁵³⁹ que a ferramenta HART deixou de ser usada em setembro de 2020, com a conclusão da última etapa do projeto de pesquisa da Universidade de Cambridge, que acompanhou as pessoas que foram presas entre setembro de 2016 e outubro de 2017 no condado de Durham e Darlington, observando nos dois anos seguintes se elas reincidiram ou não, a fim de comparar com as indicações da IA e as decisões adotadas pelos policiais humanos.

O comunicado da Polícia de Durham foi sucinto, e de modo geral afirmou que os achados preliminares demonstravam que a IA se mostrou ligeiramente mais precisa que policiais experientes para avaliar se um infrator poderia reincidir⁵⁴⁰.

Porém, deixou claro que o relatório final da pesquisa, com o estudo acadêmico completo, ainda estava em fase de elaboração pela Universidade de Cambridge, abordando o refinamento do modelo dentro de uma estrutura ética, que foi desenvolvida em parceria com a Universidade de Winchester e a Universidade Sheffield Hallam, que foi remetida a um comitê da Câmara dos Comuns e da Câmara dos Lordes⁵⁴¹.

⁵³⁷OCDE. OECD Framework for the classification of AI systems. **OECD Digital Economy Papers**, n. 323, OECD Publishing: Paris, 2022. Disponível em: <https://doi.org/10.1787/cb6d9eca-en>. Acesso em: 15 jul. 2022.

⁵³⁸GOV.UK. **Risk potential assessment form.** Disponível em: <https://www.gov.uk/government/publications/risk-potential-assessment-form>. Acesso em: 11 jun. 2022.

⁵³⁹DURHAM CONSTABULARY. **AI can predict reoffending, university study finds.** 24 jan. 2022. Disponível em: <https://durham.police.uk/News/News-Articles/2022/January/AI-can-predict-reoffending-university-study-finds.aspx>. Acesso em: 11 jun. 2022.

⁵⁴⁰Ibidem.

⁵⁴¹Ibidem.

Diante dessa referência a um comitê, pesquisando-se do que se trata, foi possível constatar que o Parlamento do Reino Unido possui um Comitê de Avaliação e Planejamento de Risco, para tratar de “perigos nacionais disruptivos”⁵⁴², com o objetivo de preparar a nação para riscos extremos, tornando sua sociedade mais resiliente.

Trata-se de um órgão novo, criado por determinação da Câmara dos Lordes em 15/10/2020. Desde então, o Comitê realizou consultas e audiências públicas entre novembro de 2020 e junho de 2021, debruçou-se sobre estudos de casos específicos de risco e reuniu evidências sobre IA, biossegurança, infraestrutura e indústria nuclear, petróleo e gás. Também colheu indicações junto a empresas e seguradoras a respeito de possíveis estratégias para controle e mitigação de risco. Além disso, o Comitê contou, na qualidade de consultor especial, com o Professor David Alexander, do Instituto de Redução de Riscos e Desastres da UCL (*University College London*)⁵⁴³.

Contudo, verifica-se que o trabalho do Comitê de Avaliação e Planejamento de Risco ainda se encontra em estágio inicial e que, especificamente a respeito de gerenciamento de riscos relacionados a IA, que pudesse ser útil a esta pesquisa, constata-se que ainda nenhuma diretriz chegou a ser dada pelo parlamento britânico sobre a ferramenta HART ou outras aplicações para a polícia ou o sistema de Justiça.

O que interessa a este trabalho foi a menção, no primeiro relatório do Comitê, quanto à criação, pelo Primeiro-Ministro do Reino Unido, em junho de 2021, do Conselho Nacional de Ciência e Tecnologia, bem como do Escritório de Estratégia de Ciência e Tecnologia, passando a ser os órgãos responsáveis por definir as diretrizes britânicas na área de tecnologias emergentes⁵⁴⁴.

Em novembro de 2021, o governo britânico disponibilizou a versão mais recente de um formulário de avaliação de risco potencial⁵⁴⁵, para guiar análise inicial de risco ou fazer revisão do risco estratégico de programas, projetos, políticas públicas ou iniciativas emergentes em geral.

⁵⁴²UK PARLIAMENT. **Risk assessment and risk planning committee.** Disponível em: <https://committees.parliament.uk/committee/483/risk-assessment-and-risk-planning-committee/>. Acesso em: 11 jun. 2022.

⁵⁴³UK PARLIAMENT. **Preparing for extreme risks: building a resilient society.** Select Committee on Risk Assessment and Risk Planning. **Report of Session 2021–22.** HL Paper 110. 3 dez. 2021. Disponível em: <https://committees.parliament.uk/publications/8082/documents/83124/default/>. Acesso em: 11 jun. 2022.

⁵⁴⁴Ibidem.

⁵⁴⁵GOV.UK. **Risk potential assessment form.** Disponível em: <https://www.gov.uk/government/publications/risk-potential-assessment-form>. Acesso em: 11 jun. 2022.

Ainda que não exclusivo para projetos de IA, o formulário britânico é interessante por conter um conjunto padronizado de critérios para avaliar os impactos e a complexidade de iniciativas, servindo para avaliar o risco de projetos em estágio inicial e também para reavaliar projetos já em execução.

Traz um aviso de que é um indicador potencial de risco e não um modelo exaustivo de análise de risco, porém tem o valor de servir de ponto de partida para permitir avaliações mais amplas, porque ajuda a equipe responsável pelo projeto a debater e identificar problemas.

Trata-se de um modelo de documento em *Word*, editável, com diversos itens de verificação acompanhados das instruções detalhadas de preenchimento, tudo distribuído em 15 páginas, até chegar na representação visual do nível de risco do projeto classificado pela equipe.

Outro parâmetro de controle recentemente publicado, e específico para IA tal qual o já mencionado modelo de classificação da OCDE (*OECD Framework for the classification of AI systems*), é o *capAI*⁵⁴⁶, procedimento proposto em março de 2022 pela equipe de especialistas liderada por Luciano Floridi na Universidade de Oxford, para realizar avaliação de conformidade de sistemas de IA alinhados com o *AI Act*.

A tais instrumentos e procedimentos de controle se juntam as ferramentas de documentação para transparência, referidas no capítulo anterior, no item 3.4.4, desenvolvidas por Margaret Mitchell et al. e Timnit Gebru et al., respectivamente para cartões de modelo⁵⁴⁷ e para fichas de registro dos conjuntos de dados⁵⁴⁸.

Esses modelos de documentação idealizados por Mitchell e Gebru, que foram líderes da equipe de ética em IA do Google, são importantes para o contexto da IA aplicada nos tribunais brasileiros, tendo em vista que no painel do CNJ consta que 33,3% dos projetos não possuem documentação⁵⁴⁹.

⁵⁴⁶FLORIDI, Luciano et al. *capAI: a procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act*. 23 mar. 2022. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4064091. Acesso em: 16 jul. 2022.

⁵⁴⁷MICHELL, Margaret et al. Model cards for model reporting. In: **Proceedings of the conference on fairness, accountability, and transparency**. 2019. p. 220-229. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3287560.3287596>. Acesso em: 13 jul. 2022.

⁵⁴⁸GEBRU, Timnit et al. Datasheets for datasets. **Communications of the ACM**, v. 64, n. 12, p. 86-92, 2021. Disponível em: <https://dl.acm.org/doi/10.1145/3458723>. Acesso em: 13 jul. 2022.

⁵⁴⁹CNJ. **Resultados Pesquisa IA no Poder Judiciário – 2022**. Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language.BR. Acesso em: 14 jul. 2022.

Descritos alguns dos instrumentos de controle mais relevantes no cenário mundial atualmente, desenvolvidos com o avanço da pesquisa em IA, é preciso dizer que foram elaborados para uso geral de aplicações da tecnologia em setores dos mais variados, como indústria, comércio, trabalho, sistema financeiro, educação, área da saúde, governos etc. Isso significa que, para uso por sistemas de Justiça, é preciso fazer adequações às especificidades dessa aplicação.

A propósito, cabe destacar que **aqui mesmo, no Brasil, já existe um parâmetro de notável qualidade, desenvolvido para a realidade da Justiça brasileira** de acordo com as regras da Resolução CNJ 332/2020⁵⁵⁰, a Portaria CNJ nº 271/2020⁵⁵¹, a Carta Ética Europeia⁵⁵² de 2018 e as Orientações Éticas para uma IA de Confiança⁵⁵³ de 2019 (elaboradas pelo GPAN IA - Grupo Independente de Peritos de Alto Nível sobre a Inteligência Artificial, criado pela Comissão Europeia em junho de 2018).

Trata-se do **Manual de Diretrizes de Auditabilidade e Conformidade no Desenvolvimento e Testes de Soluções de IA no Âmbito do LIAA-3R**⁵⁵⁴, elaborado em São Paulo pelo Tribunal Regional Federal da 3^a Região (TRF3), por meio do Grupo de Validação Ético-Jurídica (GVEJ) do Laboratório de Inteligência Artificial Aplicada da Justiça Federal da 3^a Região (LIAA-3R), do Laboratório de Inovação do TRF3 (iLabTRF3) e do Laboratório de Inovação da Justiça Federal de São Paulo (iJuspLab).

O manual já está em sua 2^a edição, publicada em 2022 (a 1^a edição é de abril de 2021), e contém: (i) Glossário; (ii) Diretrizes Gerais de Conformidade, de acordo com os atos normativos do CNJ e à luz de sua aplicação prática no âmbito do laboratório; (iii) Diretrizes Específicas de Conformidade, que tratam dos

⁵⁵⁰CNJ. Resolução nº 332, de 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 16 jul. 2022.

⁵⁵¹CNJ. Portaria nº 271, de 4 de dezembro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 jul. 2022.

⁵⁵²CEPEJ. European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment. Disponível em: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 16 jul. 2022.

⁵⁵³EUROPEAN COMISSION. Orientações éticas para uma IA de confiança. Directorate-General for Communications Networks, Content and Technology. **Publications Office**, 2019. Disponível em: <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-pt/format-PDF>. Acesso em: 16 jul. 2022.

⁵⁵⁴LIAA-3R. **Manual de Diretrizes de Auditabilidade e Conformidade no Desenvolvimento e Testes de Soluções de IA no Âmbito do LIAA-3R/ Grupo de Validação Ético-Jurídica (GVEJ) do LIAA-3R, iLabTRF3, iJuspLab.** 2. ed., rev e atual. São Paulo: LIAA-3R, 2022. Disponível em: https://www.trf3.jus.br/documentos/adeg/lnova/LIAA-3R/MANUAL/DIRETRIZES_DE_AUDITABILIDADE_2_ED_FICHA_ATUALIZADA.pdf. Acesso em: 16 jul. 2022.

procedimentos de aprovação e registro dos projetos, sua documentação, da segurança da informação e de verificação de eventuais conflitos de interesses; (iv) Diretrizes Referentes à LGPD, quando aos procedimentos para tratamento dos dados utilizados⁵⁵⁵.

Além disso, o Manual traz 9 anexos⁵⁵⁶ com modelos dos seguintes documentos:

- Termo de ciência e confidencialidade;
- Termo de justificativa de uso de dados pessoais;
- Termo de encerramento do tratamento e de justificativa da conservação de dados pessoais;
- Modelo de relatório parcial das atividades de anotação;
- Modelo de relatório parcial das atividades de desenvolvimento;
- Modelo de relatório final sobre a formatação dos *datasets*;
- Lista de questões à equipe de desenvolvedores;
- Lista de verificação para o GVEJ;
- Fluxograma de aprovação de projetos.

De acordo com parecer⁵⁵⁷ do GVEJ, o TRF3 elaborou e utilizou o Manual para as atividades de validação ético-jurídica de seus projetos SINARA e SIGMA, que haviam sido iniciados de acordo com as diretrizes então vigentes para a inscrição de projetos no CNJ, que era a Portaria nº 25/2019⁵⁵⁸, a qual exigia que os projetos de IA nos tribunais fossem inscritos no CNJ sob a forma de projetos de pesquisa acadêmica, que as equipes contassem com pessoas desenvolvendo funções especificadas na Portaria, e que deveria haver validação ético-jurídica dos modelos.

Essas duas últimas diretrizes podem ser visualmente representadas conforme a Figura 12, que descrevem uma estrutura mínima bastante enxuta, quando

⁵⁵⁵LIAA-3R. **Manual de Diretrizes de Auditabilidade e Conformidade no Desenvolvimento e Testes de Soluções de IA no Âmbito do LIAA-3R/ Grupo de Validação Ético-Jurídica (GVEJ) do LIAA-3R, iLabTRF3, iJuspLab.** 2. ed., rev. e atual. São Paulo: LIAA-3R, 2022. Disponível em: <https://www.trf3.jus.br/documentos/adeg/Inova/LIAA-3R/MANUAL/DIRETRIZES DE AUDITABILIDADE 2 ED FICHA ATUALIZADA.pdf>. Acesso em: 16 jul. 2022.

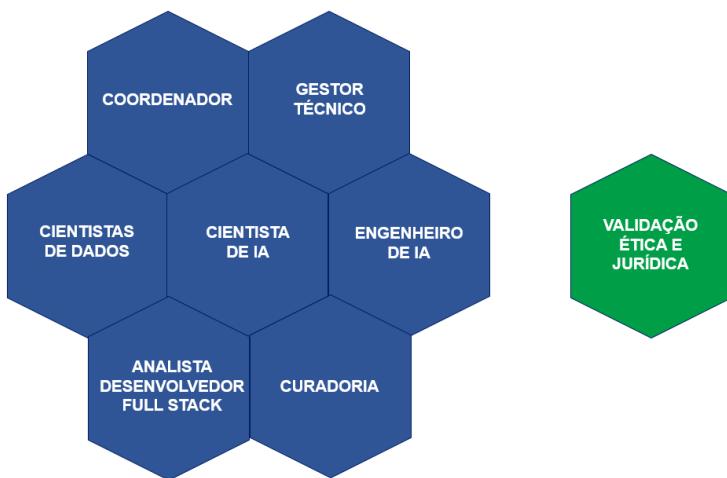
⁵⁵⁶Ibidem.

⁵⁵⁷MARIANO JÚNIOR, Raul; MARTINS, Claudio R. N. **Parecer nº GVEJ.** Tribunal Regional Federal da 3ª Região. 17 dez. 2021. Disponível em: https://www.trf3.jus.br/documentos/adeg/Inova/LIAA-3R/PARECER_TECNICO/SEI_8354929_Parecer_N.I._GVEJ_1_.pdf. Acesso em: 16 jul. 2022.

⁵⁵⁸CNJ. **Portaria nº 25, de 19 de fevereiro de 2019.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2829>. Acesso em: 16 jul. 2022.

comparada à Figura 6 constante no capítulo anterior, no item 3.4.3, que reproduziu as possíveis funções para equipes de IA de alto desempenho, conforme relatório da Deloitte⁵⁵⁹.

Figura 12 – Funções mínimas de equipes de IA pela Portaria 25/2019



Fonte: Elaborada pela autora com base na Portaria CNJ nº 25/2019⁵⁶⁰.

Consta no parecer⁵⁶¹ do GVEJ que as regras de elaboração de projeto de pesquisa sob a forma de documento acadêmico, de funções mínimas de equipes e de validação ético-jurídica deixaram de ser exigidas pelo CNJ porque a Portaria nº 25/2019⁵⁶² (que havia instituído o Laboratório de Inovação para o Processo Judicial em meio Eletrônico - Inova PJe e o Centro de Inteligência Artificial aplicada ao PJe) foi expressamente revogada pela Resolução nº 395/2021⁵⁶³ (que instituiu a Política de Gestão da Inovação no Poder Judiciário).

Apesar da revogação da Portaria 25/2019, por outro lado sobrevieram a Resolução 332/2020, a Portaria 271/2021 e a entrada total em vigor da LGPD, razões pelas quais o TRF3, de acordo com o parecer⁵⁶⁴ referido, entendeu que continuava

⁵⁵⁹DELOITTE. Building successful AI teams. 2021. <https://www2.deloitte.com/us/en/pages/consulting/articles/building-diverse-teams-in-tech-organizations.html>. Acesso em: 16 jun. 2022.

⁵⁶⁰CNJ. Portaria nº 25, de 19 de fevereiro de 2019. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2829>. Acesso em: 16 jul. 2022.

⁵⁶¹MARIANO JÚNIOR, Raul; MARTINS, Claudio R. N. Op. cit.

⁵⁶²CNJ. Portaria nº 25, de 19 de fevereiro de 2019. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2829>. Acesso em: 16 jul. 2022.

⁵⁶³CNJ. Resolução nº 395, de 7 de junho de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 16 jul. 2022.

⁵⁶⁴MARIANO JÚNIOR, Raul; MARTINS, Claudio R. N. Op. cit.

sendo relevante a validação ético-jurídica dos modelos de IA desenvolvidos pelo tribunal, mantendo a estruturação de seu órgão criado para tal finalidade, o GVEJ, ainda que a superveniente alteração normativa do CNJ tenha ocasionado uma lacuna sobre o tema.

Com efeito, na estrutura atualmente em vigor, a Rede de Inovação do Poder Judiciário (RenovaJud)⁵⁶⁵ é integrada por 4 órgãos: (i) Comitê Gestor Nacional da Inovação do Poder Judiciário; (ii) Laboratório de Inovação e dos Objetivos de Desenvolvimento Sustentável (LIODS/CNJ); (iii) Laboratórios de Inovação instituídos no âmbito dos órgãos do Poder Judiciário; e (iv) Conselho Consultivo Nacional da Inovação do Poder Judiciário.

Percebe-se que realmente não há, em qualquer instrumento normativo do CNJ em vigor, a previsão de criação, no CNJ e nos tribunais, de um órgão como um **Comitê de Ética em IA**, específico para analisar o cumprimento das diretrizes éticas e legais no desenvolvimento de modelos de IA, atuando desde o início do ciclo de desenvolvimento, ao longo do processo de *design*, acompanhando a produção e reavaliando periodicamente os resultados.

Acontece que todos os riscos envolvendo a IA e as inúmeras questões éticas tratadas ao longo deste trabalho apontam na direção de que esse aperfeiçoamento da regulamentação do CNJ necessita urgentemente ocorrer.

Outra lição que pode ser extraída da experiência pioneira do Grupo de Validação Ético-Jurídica do TRF3, que atua como um comitê de ética em IA, é a **organização dos documentos via SEI** (Sistema Eletrônico de Informações), que foi mencionada no parecer⁵⁶⁶, fazendo referência a outros 8 expedientes, com as respectivas numerações de registro no SEI, documentando de forma organizada todas as etapas do procedimento de validação constante no Manual GVEJ.

Também é interessante destacar, do parecer, que se cuidou de evitar conflitos de atividades, tendo a “clareza sobre o papel dos laboratórios de inovação na área tecnológica, que se restringe ao desenvolvimento de protótipo e MVPs”⁵⁶⁷ (produtos mínimos viáveis), como também se recomendou encaminhar o parecer à

⁵⁶⁵CNJ. Resolução nº 395, de 7 de junho de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 16 jul. 2022.

⁵⁶⁶MARIANO JÚNIOR, Raul; MARTINS, Claudio R. N. Op. cit.

⁵⁶⁷Ibidem.

análise e considerações do Comitê Gestor de Proteção de Dados Pessoais da Justiça Federal da 3^ª Região.

Além disso, as informações relevantes sobre os projetos do TRF3 estão disponíveis para consulta pública na página do tribunal na Internet, em que é possível consultar a portaria que instituiu o laboratório de IA, as duas edições do Manual, o fluxograma de aprovação de projetos do LIAA-3R, o parecer, os projetos SIGMA e SINARA, um tutorial e um vídeo sobre o SIGMA, e o e-mail para contato com o laboratório⁵⁶⁸.

Essa **disponibilidade da informação** é importante porque, além de atender aos deveres de publicidade, transparência e possibilitar controle por usuários externos, **ajuda a desmistificar a errônea ideia de que está havendo ou em vias de haver juiz-robô no Brasil**.

A título de exemplo, em setembro de 2020, no portal Conjur, houve um interessante embate de ideias entre o jurista Lenio Streck, no artigo de opinião “Um robô pode julgar? Quem programa o robô?”⁵⁶⁹, e o juiz Alexandre Morais da Rosa, que havia lançado um livro criticado por Streck, e respondeu no dia seguinte à crítica com outro artigo de opinião, “Inteligência artificial e Direito: ensinando um robô a julgar”, no qual afirma que as objeções e reflexões lançadas por Streck representam “o contrário da proposta do livro, que é de suporte ao trabalho judicial, não de substituição do juiz”, em não mais que “automatização das atividades repetitivas e burocráticas que podem ser automatizadas”⁵⁷⁰.

A questão é relevante porque, como já explicado no capítulo anterior, a Resolução 332/2020 é clara em permitir a IA apenas como ferramenta auxiliar ao magistrado, sem qualquer possibilidade de decisão autônoma pela máquina, visto que qualquer modelo de IA criado para apoio ao gabinete exige supervisão e controle por parte do magistrado.

⁵⁶⁸TRF3. **Laboratório de Inteligência Artificial Aplicada LIAA-3R.** Disponível em: https://www.trf3.jus.br/adeg/governanca-e-estrategia/laboratorio-de-inovacao-do-trf3a-regiao/laboratorio-de-inteligencia-artificial-aplicada-liaa-3r?sword_list%5B0%5D=sigma&no_cache=1. Acesso em: 16 jul. 2022.

⁵⁶⁹STRECK, Lenio Luiz. **Um robô pode julgar? Quem programa o robô?** Consultor Jurídico. 3 set. 2020. Disponível em: <https://www.conjur.com.br/2020-set-03/senso-incomum-robo-julgar-quem-programa-robo>. Acesso em: 16 jul. 2022.

⁵⁷⁰ROSA, Alexandre Morais da. **Inteligência artificial e Direito: ensinando um robô a julgar.** Consultor Jurídico. 4 set. 2020. Disponível em: <https://www.conjur.com.br/2020-set-04/limite-penal-inteligencia-artificial-direito-ensinando-robo-julgar>. Acesso em: 16 jul. 2022.

De todos os 111 projetos de IA constantes no Painel do CNJ, já foi visto no início desta seção que a maioria esmagadora – na realidade, a quase totalidade – é de projetos para a atividade-meio.

Talvez a aplicação mais próxima de ser confundida com um juiz-robô seja o SIGMA, do TRF3, ora analisado. Entretanto, qualquer impressão equivocada cai por terra pela leitura do parecer⁵⁷¹ do GEVJ, no qual se vê que a SINARA é um sistema que identifica artigos de lei, precedentes e súmulas em textos jurídicos, usada como parte de uma solução computacional que permite que o SIGMA faça o ranqueamento dos modelos de decisões judiciais constantes dos arquivos do magistrado que são mais prováveis de corresponder às informações verificadas pela SINARA.

Todo magistrado tem seu arquivo de modelos. Assim, o SIGMA, que está incorporado ao sistema de processo eletrônico PJe, apenas apresenta uma lista de opções entre vários modelos, do mais provável de ser escolhido ao menos provável de acordo com o fundamento jurídico do texto verificado pela SINARA, permitindo que o magistrado escolha qual dos modelos sugeridos vai querer usar e adaptar de acordo com as circunstâncias do caso concreto que está examinando.

A ferramenta foi verificada pelo CNJ e comprovadamente economiza tempo, reduz o risco de decisões conflitantes, “facilita a produção de minutas, elimina tarefas repetitivas e fornece insumos qualificados para redação, aumentando a produtividade e qualidade da prestação jurisdicional”⁵⁷², tanto que em 2021 o Projeto SIGMA foi o vencedor do 18º Prêmio Innovare na categoria CNJ/Tecnologia, além de já ter sido escolhido para figurar no Portal CNJ de Boas Práticas do Poder Judiciário.

Por tudo o que foi exposto a respeito de ferramentas de controle da IA, depreende-se que o Manual de Diretrizes de Auditabilidade e Conformidade no Desenvolvimento e Testes de Soluções de IA no Âmbito do LIAA-3R, com todo o seu roteiro e modelos de documentos, afigura-se como o material atualmente mais indicado, e que já está pronto para ser usado de imediato como parâmetro por outros tribunais, em vez de começar do zero fazendo adaptações dos novos instrumentos internacionais.

⁵⁷¹MARIANO JÚNIOR, Raul; MARTINS, Claudio R. N. Op. cit.

⁵⁷²INSTITUTO INNOVARE. **Premiada CNJ/Tecnologia: Iniciativa do TRF 3ª Região que facilita produção de minutas é premiada pelo Innovare e CNJ.** 21 out. 2021. Disponível em: <https://www.premioinnovare.com.br/noticias/premiada-cnjtecnologia:-iniciativa-do-trf-3a.-regiao-que-facilita-producao-de-minutas-e-premiada-pelo-innovare-e-cnj/108>. Acesso em: 16 jul. 2022.

Assim, o que se evidencia mais produtivo é usar o Manual do TRF3 como base, aprimorando e atualizando suas diretrizes de acordo com as novidades e avanços da pesquisa internacional, dado o dinamismo na área da IA.

Feita a análise do painel do CNJ, dos instrumentos internacionais de controle e do Manual do TRF3, percebe-se que remanesce a lacuna do tratamento de erros e incidentes, assunto a ser tratado na próxima e última seção.

4.3 Tratamento dos erros e incidentes de IA nos tribunais

Para a abordagem desse tema, entende-se que primeiro é preciso tecer algumas considerações interdisciplinares, resgatando conceitos referenciados no segundo capítulo deste trabalho e acrescentando outros, dentro de um contexto histórico-cronológico.

Desde 1979 o filósofo Hans Jonas, com sua noção de “heurística do medo”⁵⁷³, alertava que o medo podia ser um bom conselheiro no agir do gestor público perante novas tecnologias, pois fazer um exercício de imaginar tudo o que poderia dar errado e buscar manejar essas hipóteses, tentando adotar providências de antemão para evitar que ocorram, é uma atitude de cautela diante do desconhecido impacto futuro das criações tecnológicas. Nascia, assim, a base filosófica do princípio da precaução e do princípio da responsabilidade, conforme mencionado no item 3.1 deste trabalho.

Porém, como também se pontuou no item 3.1, desde 1986 o sociólogo Ulrich Beck já dizia que, em se tratando de novas tecnologias em plena sociedade de risco, não raro reina a negação, ausência ou minimização dos riscos, de modo que o detentor do poder político ou econômico, por vaidade, lucro ou puro e simples descaso, acaba se comportando por ação ou omissão como se as pessoas que desenvolvem novas tecnologias fossem sujeitas à “desumana lei da infalibilidade”⁵⁷⁴.

Assim, os gestores acabam focando de modo imediatista apenas no lado bom, no sucesso proveitoso das inovações criadas em sua gestão, mandato ou empresa, em que muitas vezes a possibilidade de erro sequer é cogitada, ou relegada a segundo plano.

⁵⁷³JONAS, Hans. Op. cit.

⁵⁷⁴BECK, Ulrich. Op. cit.

Beck também observou, como dito no item 3.1, que a busca desenfreada pelo aumento da produtividade ou pelo lucro, mediante inovações tecnológicas, vem acompanhada de uma miopia ou cegueira convenientes quanto à possibilidade de erro, deixando de lado os riscos produzidos, fazendo com que seja procrastinada ou deixe de ser tomada a decisão de gerenciá-los e de prospectar “alcance, tipo e teor de ameaça, círculo de pessoas afetadas, efeitos retardados, medidas a tomar, responsáveis, demandas de reparação”⁵⁷⁵.

Então, quando acontece um resultado negativo impactante, que não deveria ser surpresa e para o qual já deveria haver um plano de contingência, procura-se negar ou abafar o caso, ao mesmo tempo em que se percebe que será preciso estancar depressa os efeitos, lidar com a imprensa e a opinião pública, e depois responder pelos prejuízos, ao custo não só de recursos financeiros como também do abalo à imagem e credibilidade da instituição.

Some-se agora, a tais considerações filosófico-sociológicas, um conceito que historicamente é até anterior ao do princípio da responsabilidade de Hans Jonas e da sociedade de risco de Ulrich Beck.

Trata-se da noção de *wicked problems*⁵⁷⁶, termo cunhado por Horst Rittel que surgiu pela primeira vez em artigo escrito com Melvin Webber em 1973, no âmbito da Ciência Política, para designar os problemas complexos da sociedade contemporânea, que possuem 10 características (Figura 13) que tornam muito difícil o planejamento e execução de políticas públicas para abordá-los, porque têm múltiplas causas, sofrem influência de vários fatores interconectados que se modificam o tempo todo e causam significativos impactos na sociedade.

Em se tratando de aplicações de IA, os desafios em torno da regulação e controle dessa nova tecnologia, que vêm sendo enfrentados por governos, parlamentares e pela comunidade acadêmica do mundo inteiro, como visto nos capítulos anteriores, bem demonstram que esse assunto também assume as características de um *wicked problem*.

⁵⁷⁵Ibidem.

⁵⁷⁶RITTEL, H.; WEBBER, M. Dilemmas in a general theory of planning. **Policy Sciences** 4, p. 155-169, 1973. Disponível em: <https://www.semanticscholar.org/paper/Dilemmas-in-a-general-theory-of-planning-Rittel-Webber/3667a56bd911445a6c0fc9447771d964e6831146>. Acesso em: 17 jul. 2022.

Figura 13 – As 10 características dos *wicked problems*

Fonte: Elaborada pela autora conforme Rittel e Webber⁵⁷⁷.

Rittel é considerado um dos precursores do *Design Thinking*, hoje uma das metodologias mais utilizadas para a inovação, cujas origens na Europa remontam a 1962, quando foi realizado em Londres o evento científico “Conferência sobre Métodos de Design”, que ensejou a criação da mais antiga sociedade de pesquisa multidisciplinar sobre design existente no mundo, a *Design Research Society* (Sociedade de Pesquisa em Design), fundada em 1966⁵⁷⁸.

Envolvido nesse movimento europeu, o alemão Rittel era matemático e, antes de imigrar da Alemanha para os Estados Unidos, onde passou a ser professor de Ciência do Design na Universidade da Califórnia em Berkeley, havia criado a disciplina de Metodologia do Design na Escola de Design de Ulm (HfG Ulm), fundada na cidade alemã de Ulm após a 2^a Guerra Mundial, por intelectuais que foram da resistência ao nazismo, como um “projeto moral” para “sintetizar a ciência e o design num novo humanismo científico”, diante de uma Alemanha arrasada pela guerra, que precisava ser reconstruída. A escola “funcionou de 1953 até 1968 e durou a transição da sociedade industrial para a sociedade pós-industrial”, segundo apurado pela

⁵⁷⁷RITTEL, H.; WEBBER, M. Dilemmas in a general theory of planning. *Policy Sciences* 4, p. 155-169, 1973. Disponível em: <https://www.semanticscholar.org/paper/Dilemmas-in-a-general-theory-of-planning-Rittel-Webber/3667a56bd911445a6c0fc9447771d964e6831146>. Acesso em: 17 jul. 2022.

⁵⁷⁸DESIGN RESEARCH SOCIETY. About the DRS. Disponível em: <https://www.designresearchsociety.org/cpages/about>. Acesso em: 17 jul. 2022.

pesquisadora portuguesa Isabel Neves, que examinou todos os arquivos e materiais de aula originais de Rittel na antiga escola, hoje um museu⁵⁷⁹.

Neves também apurou que o artigo sobre *wicked problems*, elaborado quando Rittel já estava nos Estados Unidos, era resultado do aprimoramento de seus métodos inovadores desde Ulm, aplicando conhecimentos da matemática e engenharia para o design, que podiam ser utilizados em inúmeras áreas, não só no campo das ciências exatas, como também das ciências sociais⁵⁸⁰.

Para problemas sociais cada vez mais complexos, os métodos de Rittel envolviam estratégias de tomada de decisão baseadas em dados, investigação empírica e novos métodos para solução de problemas, como segmentação, trabalho colaborativo em equipe, pensamento crítico, capacidade de abstração para imaginar possíveis resultados e impactos, entre outros⁵⁸¹.

Com o passar do tempo, a noção de *wicked problems* tornou-se tão importante para o desenho de qualquer política pública hoje em dia que, a título de exemplo, mais de 40 anos depois da publicação do artigo seminal de Rittel e Webber, em 2017 os pesquisadores australianos Kate Crowley e Brian Head⁵⁸² apuraram que era o artigo sobre Ciência Política mais citado nas mais variadas bases de dados acadêmicas, utilizado como referencial para trabalhos de teoria, planejamento, prática e reformas de políticas públicas, gestão, governança e gerenciamento de crises, em inúmeras áreas de aplicação.

Ao revisitarem o clássico de Rittel e Webber, em artigo publicado no mesmo periódico, *Policy Sciences*, sob o prisma das críticas, reflexões e novas contribuições de muitas gerações de pesquisadores ao redor do mundo ao longo de mais de 4 décadas, Crowley e Head compilaram algumas sugestões para modernizar as ideias iniciais de Rittel e Webber, tornando-as mais adequadas para guiar o enfrentamento de problemas complexos atuais da sociedade⁵⁸³.

⁵⁷⁹NEVES, Isabel Clara. Contribuição de Horst Rittel para a abordagem científica ao projecto no início da era computacional. **PARC Pesquisa em Arquitetura e Construção**, SP, v. 6, n. 1, p. 39–55, 2015. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/parc/article/view/8635025/4273>. Acesso em: 17 jul. 2022.

⁵⁸⁰Ibidem.

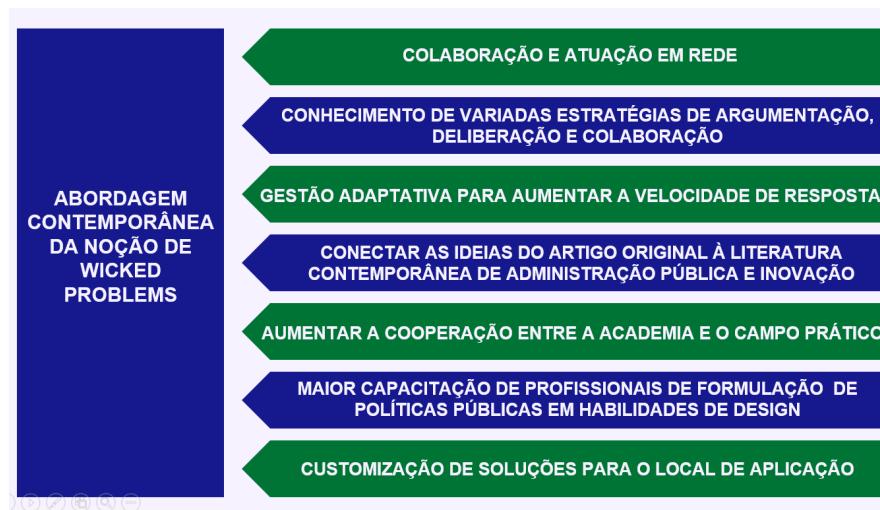
⁵⁸¹Ibidem.

⁵⁸²CROWLEY, Kate; HEAD, Brian W. The enduring challenge of 'wicked problems': revisiting Rittel and Webber. **Policy Sciences**, v. 50, n. 4, p. 539-547, 2017. Disponível em: https://www.researchgate.net/publication/320883976_The_enduring_challenge_of'_wicked_problems'_revisiting_Rittel_and_Webber/link/5a6174154585158bca4a0790/download. Acesso em: 17 jul. 2022.

⁵⁸³CROWLEY, Kate; HEAD, Brian W. Ibidem.

As principais conclusões estão resumidas na Figura 14 e, como é possível depreender, também são úteis para o campo de aplicação ora estudado – regulação e controle da IA.

Figura 14 – Abordagem contemporânea da noção de *wicked problems*



Fonte: Elaborada pela autora com base em Crowley e Head⁵⁸⁴.

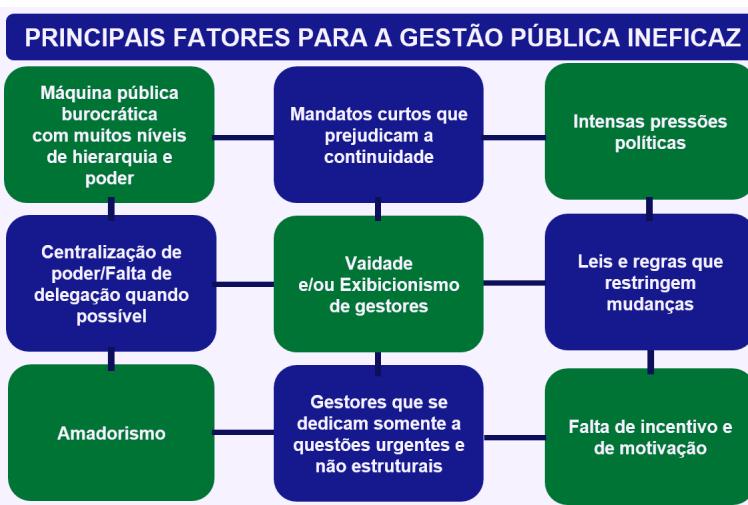
Como as sugestões encontradas no artigo de Crowley e Head dependem que sejam tomadas decisões em tais sentidos pelos gestores, ao formularem políticas públicas para problemas complexos, cabe aqui dar ênfase à sugestão que indicou conectar as ideias originais de Rittel e Webber à literatura contemporânea de Administração Pública.

Para tanto, buscando saber quais são os principais fatores que tornam uma gestão pública ineficaz, de modo geral ou em relação a algum problema, foi possível encontrar algumas indicações na obra de Hal G. Rainey, que fala sobre compreender e gerenciar organizações públicas⁵⁸⁵. Tais fatores, que se interligam, estão sintetizados na Figura 15.

⁵⁸⁴ Ibidem.

⁵⁸⁵ RAINY, Hal G. **Understanding and managing public organizations** - 5th Edition. San Francisco: Jossey-Bass, 2014, p. 9.

Figura 15 – Principais fatores para a gestão pública ineficaz



Fonte: Elaborada pela autora segundo Hal G. Rainey⁵⁸⁶.

Sob outro prisma, quando pesquisado qual seria o padrão contemporâneo mais eficaz para uma organização gerenciar os riscos de suas atividades, constatou-se que o modelo mais referenciado é a norma internacional ISO 31000:2018, adotada em 2018 no Brasil como **Gestão de Risco – Diretrizes ABNT NBR ISO 31000**⁵⁸⁷, que prevê como **etapa final o registro e relato de resultados**, com a finalidade de: (i) comunicar à organização, (ii) fornecer dados para dar suporte a decisões; (iii) aprimorar a gestão do risco; (iv) colaborar para a interação entre as partes interessadas, incluindo as que têm responsabilidade de gerir o risco.

O fato que se constatou é que, após toda a pesquisa documental empreendida neste trabalho, o Poder Judiciário brasileiro, na sua política de inovação, embora já esteja com 111 projetos de IA em desenvolvimento ou produção, e já tenha providenciado vários atos normativos e ações concretas buscando gerenciar diversas etapas dos riscos dessa atividade – inclusive de forma pioneira em relação aos outros dois Poderes da República –, ainda não cuidou de designar um órgão específico para cuidar do gerenciamento de riscos de maneira permanente, assim como ainda não cuidou de dar um endereçamento prático adequado aos erros e incidentes de IA que podem acontecer.

⁵⁸⁶Ibidem.

⁵⁸⁷ABNT. **Gestão de Risco - Diretrizes ABNT NBR ISO 31000**. 2^a ed. 28 mar. 2018. Rio de Janeiro, 2018.

A regulação e controle da IA é uma questão complexa, multifacetada, difícil, dinâmica. As milhares de emendas de eurodeputados à proposta do *AI Act*, abordando uma miríade de aspectos diferentes, referidas no item 3.2, dão bem uma demonstração de que se trata de um *wicked problem*, que exige constante acompanhamento e reavaliação por parte do poder público, pois este, ao usufruir os benefícios da tecnologia para o incremento de sua produtividade, não pode negligenciar quaisquer das etapas do gerenciamento de riscos.

Não é prudente esperar que ocorra um erro ou incidente de IA impactante para então tomar providências de maneira açodada, como ocorreu com a falha de cibersegurança que permitiu o ataque *hackerao* STJ narrado no item 3.4.2, pois assim agindo a gestão do Poder Judiciário terá sido ineficaz frente ao risco que produziu.

Fazendo um exercício de “heurística do medo” segundo Hans Jonas, pessoas podem ser prejudicadas, talvez muitas delas num só evento. Os danos que eventualmente experimentarem poderão gerar responsabilização civil do Estado, talvez em massa. A confiança pública na inovação da instituição poderá ser seriamente abalada. Além disso, detectar muito tarde problemas em sistemas, depois de incidentes sérios, pode significar também desperdício de recursos orçamentários e capital humano (já escasso na área de TI, como demonstrado nos itens 3.4.1 e 3.4.2).

A própria questão da cibersegurança, no que se relaciona aos dados usados para treinamento e produção de modelos de IA, é das áreas mais críticas, dado o sério prognóstico de pandemia cibernética, com o Brasil como alvo preferencial de *hackers* na América do Sul (conforme descrito no item 3.4.2).

Além da segurança dos dados usados nos modelos de IA, os códigos-fontes depositados na plataforma definida pelo CNJ, a GitLab (que usa o método de desenvolvimento *DevOps*, com integração ágil e entrega contínua *CI/CD*, como explicado no item 3.5), ao que parece também não estão totalmente protegidos de serem maliciosamente modificados.

Tanto é que, em recente pesquisa científica financiada pela Comissão Europeia, cujos resultados foram publicados em agosto de 2021, Munoz et al. afirmaram que “nenhum método pode verificar com precisão se a integridade do código-fonte do projeto não foi violada”, e que até então “não há trabalhos na literatura que proponham uma forma de garantir a integridade dos projetos de *software* como parte de *Pipelines CI/CD*”. Essa foi a lacuna de pesquisa que o grupo de quatro

pesquisadores espanhóis (da Universidade de Málaga) e gregos (da Universidade de Piraeus) buscou resolver, dizendo que “até onde sabemos, este é o primeiro artigo que propõe uma ferramenta para preencher a lacuna de segurança identificada”⁵⁸⁸.

Tudo é muito novo na área da IA, e mesmo que houvesse um plano de gerenciamento de risco completo, perfeitamente executado em todas as etapas por todos os tribunais brasileiros, sob o crivo exigente de Comitês de Ética em IA, de Proteção de Dados e de Cibersegurança, ainda assim é impossível evitar que erros e falhas aconteçam, simplesmente porque se trata de uma criação humana e porque o erro faz parte do processo de inovação.

Essa linha de pensamento segue a do pesquisador libanês Nassim Nicholas Taleb que, entre vasta produção científica e literária (na área de matemática, filosofia, estatística, probabilidade, risco, incerteza, erro humano, tomada de decisão, entre outros assuntos correlatos)⁵⁸⁹, publicou em 2007 a obra “A Lógica do Cisne Negro: o impacto do altamente improvável”⁵⁹⁰. No livro, Taleb chama de cisne negro um evento completamente imprevisível, que causa impacto extremo, e que depois de ocorrido, as pessoas em sua natureza humana tentam compreender, explicar, buscar causas e culpados.

Segundo o citado autor, que é um dos maiores especialistas contemporâneos em risco, como é impossível evitar completamente que erros e fatos negativos ocorram, a providência mais inteligente a ser adotada é procurar estar preparado da melhor maneira possível, o que inclui buscar aprender com eles. É o que Ulrich Beck chamou de “precaução pela prevenção”: partindo de um risco concreto, é possível se precaver de outros hipotéticos⁵⁹¹ como dito no item 3.1.

Assim, com base nas ideias de Taleb é possível afirmar que, em vez de negar, ignorar, deixar de cogitar, menosprezar, minimizar, esconder, os erros da IA que vierem a ocorrer, o mais produtivo a fazer é catalogá-los para análise e estudo, formando um **banco de dados que permita análise longitudinal** e possa fornecer

⁵⁸⁸MUÑOZ, Antonio et al. P2ISE: preserving project integrity in CI/CD based on secure elements. **Information**, v. 12, n. 9, p. 357, 2021. Disponível em: <https://www.mdpi.com/2078-2489/12/9/357>. Acesso em: 18 jul. 2022.

⁵⁸⁹TALEB, Nassim Nicholas. **Official research biography of Nassim Nicholas Taleb**. Disponível em: <https://www.fooledbyrandomness.com/CV.htm>. Acesso em: 18 jul. 2022.

⁵⁹⁰TALEB, Nassim Nicholas. **A lógica do cisne negro: o impacto do altamente improvável**. Editora Best Business, 2018.

⁵⁹¹BECK, Ulrich. Op. cit.

elementos úteis a tomadas de decisão por parte dos gestores, quanto a aplicações que precisem de ajustes, que devam ser suspensas ou até proibidas.

Ao mesmo tempo, um banco de dados centralizado que possa receber registros de incidentes ocorridos em diversas partes do país pode colaborar para **aumentar a velocidade de resposta** (conforme a abordagem contemporânea de *wicked problems* vista na Figura 14).

O Poder Judiciário brasileiro, que foi o primeiro dos 3 Poderes a regular os riscos da IA (como visto no item 3.4), não está sozinho como instituição que ainda não cuidou da etapa final do ciclo de gerenciamento de riscos da IA, quanto ao tratamento dos erros – no que poderá ser pioneiro mais uma vez no cenário nacional, talvez até entre os sistemas de Justiça mundiais, caso venha a implantar um banco de dados como o ora tratado.

Na verdade, ao que tudo indica, ainda não se falava sobre o tema no panorama global, em termos de instrumentos de políticas públicas sobre a questão, até recentemente, quando em fevereiro de 2022 a OCDE, ao publicar o modelo de classificação de risco de sistemas de IA, deixou registrado na parte final que, entre os próximos passos, está o rastreamento de incidentes de IA (em tradução livre):

Rastreamento de incidentes de IA

Um próximo passo relacionado é desenvolver uma estrutura comum para relatar incidentes de IA, especialmente aqueles que são negativos ou prejudiciais, ou controvérsias. A estrutura de incidentes alavancaria a estrutura de classificação e ajudaria a garantir consistência global e interoperabilidade na comunicação de incidentes. Seria parte de um Rastreamento Global de Incidentes de IA na OCDE, com a contribuição de instituições parceiras, para construir a base de evidências sobre riscos que se materializaram em incidentes ou quase incidentes⁵⁹².

Demonstra-se assim que, no uso da IA pela Justiça brasileira, não é suficiente apenas comunicar a ocorrência de um incidente, como previsto no art. 27 da Resolução 332/2020, segundo o qual “os órgãos do Poder Judiciário informarão ao Conselho Nacional de Justiça todos os registros de eventos adversos no uso da Inteligência Artificial”⁵⁹³. Aplicando-se o princípio da precaução, é preciso organizar

⁵⁹²OCDE. OECD Framework for the classification of AI systems. **OECD Digital Economy Papers**, n. 323, OECD Publishing: Paris, 2022. Disponível em: <https://doi.org/10.1787/cb6d9eca-en>. Acesso em: 15 jul. 2022.

⁵⁹³CNJ. **Resolução nº 332, de 21 de agosto de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 10 jul. 2021.

essas informações de maneira a extrair delas o máximo proveito, para evitar novos episódios.

Observa-se que o referido documento da OCDE também menciona a importância de rastrear incidentes de IA para outra finalidade relevante: ajudar a classificar o grau de risco das aplicações, com base nos usos da vida real⁵⁹⁴.

A propósito, falando sobre esses usos da vida real, entende-se importante deixar registrado que, no desenho inicial da metodologia de pesquisa deste trabalho, pretendia-se empreender pesquisa empírica, indo a campo para coletar, em entrevistas semiestruturadas de membros de equipes desenvolvedoras de IA nos tribunais, evidências dos incidentes que eventualmente já tivessem acontecido nas aplicações em desenvolvimento e produção.

Para tanto, seria usada como referência a única pesquisa que existia sobre o panorama de uso da IA nos tribunais brasileiros, elaborada pela Fundação Getúlio Vargas (FGV)⁵⁹⁵, cujo relatório da primeira fase foi divulgado em dezembro de 2020, e que era mais completo que o primeiro painel de projetos de IA⁵⁹⁶ divulgado pelo CNJ no mesmo mês.

O objetivo da pesquisa empírica, tal qual anunciado no documento da OCDE acima referido, era coletar evidências extraídas de aplicações reais de IA no Poder Judiciário brasileiro, descritas no relatório da FGV.

O escopo da pesquisa de campo seria investigar, junto aos desenvolvedores e usuários internos nos tribunais, quais foram os erros e eventos adversos ao usar IA que foram detectados tanto na fase de desenvolvimento como durante a operação; quem os detectou; se foram solucionados; como foram solucionados; como, por quem e para quem foram reportados; quanto tempo demorou para que os erros fossem detectados; qual o método utilizado para desenvolver as ferramentas de IA, auditá-las e corrigi-las; se houve algum processo de certificação;

⁵⁹⁴OCDE. *OECD Framework for the classification of AI systems. OECD Digital Economy Papers*, n. 323, OECD Publishing: Paris, 2022. Disponível em: <https://doi.org/10.1787/cb6d9eca-en>. Acesso em: 15 jul. 2022.

⁵⁹⁵FUNDAÇÃO GETÚLIO VARGAS (FGV). ***Inteligência Artificial: Tecnologia aplicada à gestão de conflitos no âmbito do Poder Judiciário.*** Coordenação Luis Felipe Salomão. 2020. Disponível em: https://ciapi.fgv.br/sites/ciapi.fgv.br/files/estudos_e_pesquisas_ia_1afase.pdf. Acesso em: 18 jul. 2022.

⁵⁹⁶CNJ. ***Projetos com Inteligência Artificial no Poder Judiciário.*** Disponível em: <https://paineisanalytics.cnj.jus.br/single/?appid=29d710f7-8d8f-47be-8af8-a9152545b771&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&opt=ctxmenu.curssel>. Acesso em: 8 jul. 2022.

se foram detectados vieses que denotassem discriminação algorítmica; quem compôs as equipes que desenvolveram as ferramentas e se havia diversidade nelas.

A finalidade de analisar empiricamente todas essas variáveis seria identificar quais dados seriam relevantes para catalogar e parametrizar, inclusive levando em conta a opinião dos entrevistados a esse respeito, tudo para compor, se fosse o caso, o hipotético banco de dados sobre eventos adversos em sistemas de IA do Poder Judiciário.

Acontece que, ao iniciar os contatos para explicar o escopo da pesquisa e agendar entrevistas, percebeu-se que o tema – pesquisar erros – é um assunto inconveniente, recebido com incômodo, demonstra causar desconforto.

As pessoas falavam com entusiasmo sobre os projetos, mas se mostravam reticentes quanto aos erros, algumas desconvolvendo e outras externando preocupação com a garantia de anonimato e com o que seria publicado, para não desagravar superiores, que estavam orgulhosos de estarem na vanguarda.

Vive-se uma época de “corrida maluca”, em que “todos querem um robô para chamar de seu”, como bem pontuou Fábio Porto no artigo em que critica a pulverização de iniciativas para tratar do mesmo assunto, a exemplo de precedentes e execução fiscal, apontando que “vários Tribunais sem orquestração e sem alinhamento estão buscando o mesmo objetivo, duplicando esforços e projetos, com desperdício de força de trabalho e dinheiro público”⁵⁹⁷.

Aliás, essa multiplicidade de projetos semelhantes apontada por Fábio Porto, que destoa completamente do modelo colaborativo preconizado pela Resolução 332/2020, realmente é bastante visível pelo exame do painel mais recente do CNJ de modelos de IA.

A respeito da dificuldade de fazer as pessoas falarem sobre erros, aparentemente motivada sobretudo pelo receio de desagravar superiores, percebeu-se que não foi à toa que tanto Ulrich Beck como Hal G. Rainey falaram em vaidade de gestores como um fator tanto para o desprezo ao gerenciamento de risco como para a gestão pública ineficaz.

⁵⁹⁷PORTO, Fábio R. A “corrida maluca” da Inteligência Artificial no Poder Judiciário. In: ARAÚJO, Valter Shuenquener de; GOMES, Marcus Lívio (Coord.); CANEN, Doris (Org.) **Inteligência Artificial e aplicabilidade prática no Direito**. CNJ. Brasília, 2022. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/04/inteligencia-artificial-e-a-aplicabilidade-pratica-web-2022-03-11.pdf>. Acesso em: Acesso em: 18 jul. 2022.

Assim, tendo em vista a reação cética, preocupada ou temerosa de alguns membros de equipes desenvolvedoras de IA contatados para entrevista, quanto ao relato de erros, constatou-se que os dados coletados poderiam vir bastante incompletos ou enviesados, sendo imprestáveis para o fim pretendido.

Além disso, as situações descritas no item 3.4.4 e no item 4.2, envolvendo a demissão de Timnit Gebru pelo Google e os ataques da Amazon a Joy Buolamwini, atribuídos justamente à publicação de pesquisas críticas que desagradaram superiores, dão uma demonstração de que o receio das pessoas em conceder entrevista sobre erros fazia sentido.

É preciso levar em conta que algumas aplicações de IA são tão únicas e específicas nos tribunais, e as equipes são tão reduzidas, que não era possível garantir o anonimato da fonte: ao discorrer sobre tal ou qual ferramenta, fatalmente seria possível identificar quem prestou a informação, ainda que todos os cuidados éticos em pesquisa acadêmica fossem adotados.

Diante dessas vicissitudes, optou-se por dar um passo atrás, entendendo-se que uma lacuna de pesquisa ainda mais premente que a etapa final do ciclo de gerenciamento de risco da IA (tratamento dos erros) era demonstrar a situação atual da regulação, a dimensão dos riscos, a importância dos instrumentos de controle e do gerenciamento em si, de acordo com o estado da arte e a literatura sobre o tema, com o reforço de exemplos de situações concretas.

Assim, inviabilizada a pesquisa empírica inicialmente pretendida, foi redesenhado o escopo e a metodologia do trabalho para esmiuçar tais assuntos pela técnica de pesquisa documental, com a preocupação de adicionar contexto histórico-cronológico dos fatos e teorias.

Ao longo do trabalho, a utilidade, necessidade e importância do banco de dados sobre erros e incidentes acabaram sendo confirmadas, desde o item 3.4.2, quando se apurou que especialistas brasileiros, ao analisarem o ataque *hacker* ao STJ, afirmaram que estudar episódios passados de incidentes cibernéticos ajuda na prevenção e precaução quanto a eventos futuros. Tanto é que, mesmo com base nas poucas informações até então disponíveis, puderam fazer a engenharia reversa do incidente, encontrar 99% de similaridade com um ataque ocorrido no Departamento

de Transporte do Texas (TxDOT), nos Estados Unidos, e sugerir possíveis soluções adicionais de segurança que o Poder Judiciário brasileiro poderia adotar⁵⁹⁸.

Aliás, cabe acrescentar que a solução indicada por tais especialistas brasileiros (regras YARA) foi apontada em 2020 como a ferramenta ideal para “prever cisnes negros”⁵⁹⁹ no âmbito da segurança cibernética, de acordo com Eugene Kaspersky, russo que em 1997 fundou em Moscou a empresa de alcance global que leva seu sobrenome, considerada uma das maiores empresas privadas de cibersegurança do mundo⁶⁰⁰.

Segundo noticiado em 2017, a Kaspersky figurava, 20 anos após sua fundação, como “o quarto maior desenvolvedor de antivírus do mundo, o segundo maior da Europa e o primeiro em mercados como Espanha, França e Alemanha. Entre seus clientes está, por exemplo, o Exército do Brasil”⁶⁰¹.

Atualmente, no âmbito da Política Internacional, sobretudo após os eventos ocorridos a partir de fevereiro de 2022 com o ataque da Rússia contra a Ucrânia que resultaram em guerra, as soluções da Kaspersky estão sob suspeita, dada a preocupação de que haja interferência do governo russo em tal empresa privada, de maneira forçada ou não⁶⁰².

Os Estados Unidos, que já haviam banido o *software* da Kaspersky para uso em agências governamentais desde 2017, passaram a ser seguidos nesse tipo de cautela, a partir de março de 2022, por diversos países como Reino Unido, Alemanha, Itália, Lituânia e Holanda⁶⁰³.

Embora Eugene Kaspersky (*i*) sempre tenha negado envolvimento com o governo russo, (*ii*) seja um aficionado no que faz e sua expertise o tenha levado a ser inclusive colaborador da ONU, além de parceiro de longa data da Europol (a agência europeia de cooperação policial), e (*iii*) a Comissão Europeia tenha anunciado

⁵⁹⁸PALAZOLO, Gustavo; TEIXEIRA, Ialle; DUARTE, Felipe. Op. cit.

⁵⁹⁹KASPERSKY, Eugene. **Aprenda a usar as regras YARA – como prever cisnes negros.** Kaspersky Daily. 3 set. 2020. Disponível em: <https://www.kaspersky.com.br/blog/cybersecurity-expert-training/15992>. Acesso em: 18 jul. 2022.

⁶⁰⁰ALTARES, Guillermo. **O russo que criou o antivírus do seu computador.** El País. 1 out. 2017. Disponível em: https://brasil.elpais.com/brasil/2017/09/28/internacional/1506610335_514029.html. Acesso em: 18 jul. 2022.

⁶⁰¹ALTARES, Guillermo. **O russo que criou o antivírus do seu computador.** El País. 1 out. 2017. Disponível em: https://brasil.elpais.com/brasil/2017/09/28/internacional/1506610335_514029.html. Acesso em: 18 jul. 2022.

⁶⁰²FARRIER, Ellie. **É seguro usar o Kaspersky em 2022?** Avast. 8 abr. 2022. Disponível em: <https://www.avast.com/pt-br/c-is-kaspersky-safe>. Acesso em: 18 jul. 2022.

⁶⁰³Ibidem.

formalmente em 12/04/2019 que não haviam sido encontradas evidências comprovadas de espionagem ou interferência estatal russa na empresa (movimento que foi considerado pelo mercado como reparação quanto a uma acusação injusta)⁶⁰⁴, a posição atual de vários países em relação à Kaspersky também se explica, além do receio da segurança de suas infraestruturas de TI dada a mudança de cenário, pelo esforço de governos ocidentais em impor sanções e restrições a empresas russas, no conjunto de medidas adotadas em reação à guerra⁶⁰⁵.

Não obstante tais fatos no panorama geopolítico atual devido ao conflito entre a Rússia e a Ucrânia, o que importa destacar para os fins deste trabalho é que em setembro de 2020, quando nada disso havia acontecido ainda, Eugene Kaspersky publicou um artigo em seu *blog* no qual citou a teoria de Nassim Nicholas Taleb sobre os cisnes negros ao fazer um paralelo com ameaças cibernéticas:

Atualmente, uma das ciberameaças mais perigosas que existem são as de *0-day* – vulnerabilidades raras e desconhecidas (para o pessoal da cibersegurança e outros) em softwares que podem causar danos em larga escala, mas que tendem a permanecer desconhecidas até o momento em que são explorados (ou às vezes até depois)⁶⁰⁶.

Em tal artigo, intitulado “Aprenda a usar as regras YARA – como prever cisnes negros”, Eugene Kaspersky afirmou que, para fazer frente a tais perigosas ameaças de Dia Zero, as regras YARA são uma das ferramentas mais indicadas, por serem úteis para “pesquisar *malwares* semelhantes, identificando padrões”, tendo por escopo “poder dizer que certos programas maliciosos parecem ter sido feitos pelas mesmas pessoas, com objetivos semelhantes”⁶⁰⁷.

Como se percebe, para usar essa ferramenta é preciso observar dados de incidentes passados, o que corrobora a hipótese trabalhada nesta pesquisa, sobre a importância de documentar incidentes para gerir riscos futuros incertos, segundo um profissional que indiscutivelmente é um dos pioneiros e maiores especialistas

⁶⁰⁴CIMPANU, Catalin. **EU: No evidence of Kaspersky spying despite ‘confirmed malicious’ classification.** ZDNet, 16 abr. 2019. Disponível em: <https://www.zdnet.com/article/eu-no-evidence-of-kaspersky-spying-despite-confirmed-malicious-classification/>. Acesso em: 18 jul. 2022.

⁶⁰⁵Ibidem.

⁶⁰⁶KASPERSKY, Eugene. **Aprenda a usar as regras YARA – como prever cisnes negros.** Kaspersky Daily. 3 set. 2020. Disponível em: <https://www.kaspersky.com.br/blog/cybersecurity-expert-training/15992/>. Acesso em: 18 jul. 2022.

⁶⁰⁷Ibidem.

mundiais em cibersegurança, independentemente das questões políticas envolvendo seu nome.

No mesmo sentido, também conforme narrado no item 3.4.2, um banco de dados para documentar eventos adversos tem valor inestimável para o gerenciamento de risco, segundo estudo realizado em 2021 pela ISACA⁶⁰⁸ (instituição global dedicada à certificação, auditoria e controle de sistemas de TI).

Ainda confirmando a hipótese trabalhada nesta pesquisa, no que se refere aos erros e eventos adversos em geral de qualquer atividade, foi demonstrado nesta seção que o registro e relato de incidentes constam como etapa final do ciclo de gerenciamento de risco proposto pela norma ABNT NBR ISO 31000.

Adicionalmente, de modo específico no que se refere à IA, o rastreamento e catalogação de incidentes foi recentemente indicado pela OCDE como o próximo passo importante para a classificação e controle de riscos⁶⁰⁹.

Com tais confirmações da hipótese, depreendidas com base em referências variadas, permitindo constatar que (i) a comunicação de eventos adversos ao CNJ não basta para que o gerenciamento de riscos seja eficaz, que (ii) é preciso documentar os registros de incidentes de forma racional, útil e organizada, e que (iii) é preciso criar um método e/ou estrutura para veicular a comunicação de incidentes e viabilizar o armazenamento inteligente do registro, passa-se à reflexão sobre como fazer isso concretamente, ou seja, como criar um banco de dados para tal finalidade.

4.3.1 Critérios de solução pela metodologia de design thinking

A título de nomenclatura hipotética para o banco de dados, para facilitar o processo de reflexão a seguir, doravante passará a ser chamado de BIIAJud (Banco de Incidentes de IA no Judiciário), cujo local que se afigura como mais adequado para abrigá-lo virtualmente, na página do CNJ na Internet, é na área dedicada à Plataforma Sinapses, com disponibilização do *link* para acesso nas páginas dos tribunais.

Do ponto de vista da governança, ao que tudo indica o órgão mais indicado para executar a tarefa de gerenciamento de riscos da IA, inclusive quanto ao

⁶⁰⁸ISACA. **Cyberrisk quantification.** 2021. Disponível em: <https://store.isaca.org/s/store#/store/browse/detail/a2S4w000004Ko4jEAC>. Acesso em: 24 jun. 2022.

⁶⁰⁹OCDE. OECD Framework for the classification of AI systems. **OECD Digital Economy Papers**, n. 323, OECD Publishing: Paris, 2022. Disponível em: <https://doi.org/10.1787/cb6d9eca-en>. Acesso em: 15 jul. 2022.

tratamento de incidentes, seria um Comitê de Ética em IA, de composição multidisciplinar, inclusiva e com diversidade, a ser criado tanto no âmbito do CNJ como nos tribunais, mediante acréscimo na Portaria nº 271/2020⁶¹⁰ (que regulamenta o uso da IA nos tribunais) e/ou na Resolução nº 395/2021 (que trata da Política de Inovação no Judiciário)⁶¹¹, ou ainda na própria Resolução 332/2020⁶¹² (que dispõe sobre ética, transparência e governança no uso da IA no Judiciário).

Para a criação do BIIAJud, centralizado no CNJ, seu Comitê de Ética em IA, se criado, precisará de um método, assumindo-se aqui o *design thinking* como ideal, pois “agrega o pensamento criativo ao analítico, valorizando a busca da construção coletiva de soluções múltiplas (cocriação) e propõe a geração de aprendizado a partir de materialização de ideias e teste (experimentação)”,⁶¹³ para inovação em políticas públicas, conforme Cavalcante, Mendonça e Brandalise.

Afinal, em se tratando de uma formulação de política pública para um *wicked problem*, a criação da solução via *design thinking*, para ser eficaz, necessita da participação colaborativa de vários atores com expertises diferentes e complementares.

No caso da criação do hipotético BIIAJud, afigura-se relevante a participação, além de eticistas integrantes do Comitê de Ética em IA (se vier a ser criado), dos especialistas em TI e IA do CNJ e de seu parceiro na Plataforma Sinapses, que é o Tribunal de Justiça de Rondônia, além de especialistas de áreas específicas como proteção de dados e cibersegurança, integrantes de outros Comitês e órgãos presentes na estrutura de governança do CNJ.

Assim, o processo de criação do BIIAJud seguiria as etapas de empatia, (re)definição, ideação, prototipação e teste⁶¹⁴ (Figura 16), a serem executadas no âmbito do Laboratório de Inovação e dos Objetivos de Desenvolvimento Sustentável

⁶¹⁰CNJ. Portaria nº 271, de 4 de dezembro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 jul. 2022.

⁶¹¹CNJ. Resolução nº 395, de 7 de junho de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 16 jul. 2022.

⁶¹²CNJ. Resolução nº 332, de 21 de agosto de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 10 jul. 2021.

⁶¹³CAVALCANTE, Pedro; MENDONÇA, Letícia; BRANDALISE, Isabella. Políticas Públicas e Design Thinking: interações para enfrentar desafios contemporâneos. In: CAVALCANTE, Pedro (Org.). *Inovação e Políticas Públicas: superando o mito da ideia*. Brasília: Ipea, 2019, p. 29-52. Disponível em: https://www.ipea.gov.br/portal/images/stories/PDFs/livros/livros/190910_livro_inovacao_e_politic_as_publicas_cap01.pdf. Acesso em: 18 jul. 2022.

⁶¹⁴Ibidem.

(LIODS/CNJ), com base no art. 7º da Resolução 395/2021⁶¹⁵, mediante apoio técnico do Departamento de Pesquisas Judiciais (DPJ), para municiar o projeto com dados úteis e necessários.

Figura 16 – Etapas da metodologia de *design thinking*



Fonte: Elaborada pela autora segundo Cavalcante, Mendonça e Brandalise.

A propósito, o estado da arte da aplicação do *design* especificamente na inovação da Justiça pode ser observado no artigo que Margaret Hagan (diretora do *Stanford's Legal Design Lab* e pioneira nessa área) escreveu em 2020 com Daniel Bernal, propondo um padrão metodológico para redesenhar a inovação na seara jurídica, com base na experiência de anos acumulada no laboratório dedicado ao assunto na Universidade de Stanford⁶¹⁶.

Ao estudarem casos de insucesso que levaram ao desperdício de capital humano, tempo e recursos, Hagan e Bernal tomaram por base as seguintes premissas principais para redesenhar inovações jurídicas realmente úteis e bem sucedidas (sintetizadas na Figura 17):

- Que os laboratoristas, ao desenvolverem uma solução da criação ao teste, trabalhem em todas as etapas mediante *co-design* com usuários que estejam diretamente envolvidos com a intervenção ou possam ser por ela afetados, porque serão justamente essas pessoas

⁶¹⁵CNJ. Resolução nº 395, de 7 de junho de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 16 jul. 2022.

⁶¹⁶BERNAL, Daniel; HAGAN, Margaret. Redesigning Justice Innovation: a standardized methodology. *Stanford Journal of Civil Rights & Civil Liberties*, vol. 16, n. 2, p. 335-384, 2020. Hein Online.

que poderão fornecer um ponto de vista realista a respeito da solução que estará sendo criada, haja vista terem conhecimento de causa;

- Que o *co-design* com usuários precisa ser combinado com embasamento teórico consistente, pesquisa empírica para coleta e análise de dados, estudo de exemplos de casos parecidos e que tenham obtido sucesso, avaliação de resultados com reavaliação periódica e construção de soluções a partir de todos esses fatores em seu conjunto.

Figura 17 – Estado da arte do *legal design*



Fonte: Elaborada pela autora segundo Bernal e Hagan⁶¹⁷.

Dada a importância de fazer inovação centrada nas pessoas, com foco no usuário e participação de magistrados, servidores e atores externos – de acordo com os princípios da gestão de inovação no Poder Judiciário previstos no art. 3º da Resolução 395/2021⁶¹⁸ – caberia aproveitar as premissas que norteavam o desenho inicial da pesquisa empírica que seria feita neste trabalho, e que se demonstrou inviável para uma pesquisadora acadêmica individual.

⁶¹⁷BERNAL, Daniel; HAGAN, Margaret. Redesigning Justice Innovation: a standardized methodology. **Stanford Journal of Civil Rights & Civil Liberties**, vol. 16, n. 2, p. 335-384, 2020. Hein Online.

⁶¹⁸CNJ. Resolução nº 395, de 7 de junho de 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 16 jul. 2022.

Nesse passo, para a obtenção de dados de erros e incidentes de aplicações reais nos tribunais brasileiros, que eventualmente já tenham sido verificados na prática por usuários internos, como elementos úteis ao processo de criação de eventual banco de incidentes de IA no CNJ, a solução que se afigura mais efetiva é que os dados sejam coletados junto aos tribunais pelo Departamento de Pesquisas Judiciárias do CNJ.

Além de ser um órgão oficial, evita-se o viés nas respostas por meio do art. 20 da Portaria 271/2020⁶¹⁹ (que regulamentou o uso da IA no Judiciário), segundo o qual a omissão ou manipulação intencional de informações podem ensejar a instauração de procedimento disciplinar.

Dessa forma, prestar informações completas sobre os erros e incidentes de IA ao CNJ será uma obrigação dos usuários internos, como parte integrante de uma política pública judiciária de gerenciamento dos riscos da IA, superando assim as vicissitudes acima narradas (de receio de falar sobre erros e desagrardar superiores), quando se tentou coletar tais dados nesta pesquisa acadêmica individual.

Outro instrumento útil ao processo de criação do hipotético BIIAJud, além do Painel⁶²⁰ sobre projetos de IA do CNJ, é o recente relatório⁶²¹ da 2^a fase da pesquisa da Fundação Getúlio Vargas sobre IA no Judiciário, publicado em abril de 2022, quanto a dados coletados em 2021 e mais variáveis investigadas⁶²², na comparação com o relatório⁶²³ anterior, publicado em 2020.

Esse segundo relatório da FGV fornece em 266 páginas informações detalhadas sobre modelos de IA em ideação, desenvolvimento e produção nos

⁶¹⁹CNJ. Portaria nº 271, de 4 de dezembro de 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 jul. 2022.

⁶²⁰CNJ. Projetos com Inteligência Artificial no Poder Judiciário. Disponível em: <https://paineisanalytics.cnj.jus.br/single/?appid=29d710f7-8d8f-47be-8af8-a9152545b771&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&opt=ctxmenu,cursel>. Acesso em: 8 jul. 2022.

⁶²¹FUNDAÇÃO GETÚLIO VARGAS (FGV). **Inteligência Artificial: Tecnologia aplicada à gestão de conflitos no âmbito do Poder Judiciário**. Coordenação Luis Felipe Salomão. 2^a edição. 2022. Disponível em: https://ciapi.fgv.br/sites/ciapi.fgv.br/files/relatorio_ia_2fase.pdf. Acesso em: 18 jul. 2022.

⁶²²FUNDAÇÃO GETÚLIO VARGAS (FGV). **Publicações. Centro de Inovação, Administração e Pesquisa do Judiciário**. 29 abr. 2022. Disponível em: <https://ciapi.fgv.br/publicacoes>. Acesso em: 18 jul. 2022.

⁶²³FUNDAÇÃO GETÚLIO VARGAS (FGV). **Inteligência Artificial: Tecnologia aplicada à gestão de conflitos no âmbito do Poder Judiciário**. Coordenação Luis Felipe Salomão. 2020. Disponível em: https://ciapi.fgv.br/sites/ciapi.fgv.br/files/estudos_e_pesquisas_ia_1afase.pdf. Acesso em: 18 jul. 2022.

tribunais brasileiros, contendo minúcias diferentes das que estão presentes no Painel de IA do CNJ.

Embora não faça referência a expressões como “eventos adversos”, “erros”, “falhas” ou “incidentes de IA” nas perguntas, a segunda pesquisa da FGV traz importantes contribuições para o gerenciamento de risco no tocante à observação de fontes de erros e incidentes, ao relatar os achados em relação aos seguintes itens⁶²⁴:

- “A base utilizada se apresentou suficiente para o treinamento dos modelos do projeto? Foram identificados desafios relacionados à base? ”;
- “Atuais limitações dos resultados esperados”;
- “Riscos dos modelos de Inteligência Computacional aplicados nesta iniciativa”;
- “Descobertas realizadas após o processamento dos dados com uso de técnicas de Inteligência Computacional”.

Buscando a palavra “erro” nas respostas constantes no relatório, encontram-se alguns registros sobre erros no conjunto de dados usados como base para o treinamento do modelos de IA, levando a falhas no aprendizado⁶²⁵.

Sob outro prisma, quanto aos usuários a serem contatados para participarem do processo de *design* do hipotético BIIAJud, auxiliando os laboratoristas ou colaborando na pesquisa empírica para detalhamento dos erros reais eventualmente verificados, é possível perceber, pelo panorama atual do uso da IA nos tribunais, que a esmagadora maioria de usuários dos sistemas inteligentes está enquadrada na categoria de usuários internos.

Entre os usuários internos, como descrito no item 3.4.1 do capítulo anterior, há duas categorias: de um lado estão incluídos aqueles da área técnica (servidores do setor de TI) e de outro lado estão aqueles da área de negócio (servidores das unidades judiciais e magistrados).

Em menor número, há modelos de IA em uso nos tribunais que se relacionam com usuários externos, a exemplo daqueles que são utilizados para ajudar

⁶²⁴FUNDAÇÃO GETÚLIO VARGAS (FGV). **Ieligência Artificial: Tecnologia aplicada à gestão de conflitos no âmbito do Poder Judiciário.** Coordenação Luis Felipe Salomão. 2^a edição. 2022. Disponível em: https://ciapi.fgv.br/sites/ciapi.fgv.br/files/relatorio_ia_2fase.pdf. Acesso em: 18 jul. 2022.

⁶²⁵Ibidem.

a corrigir classificações errôneas, no momento em que policiais, advogados, defensores públicos e membros do Ministério Público estão registrando a classe e o assunto de processos, ao ajuizarem uma nova ação ou protocolarem um novo documento em sistemas de processo eletrônico. Há também modelo de IA aplicado ao público em geral, a exemplo do sistema de reconhecimento facial de visitantes do TJDFT.

Assim, prospectar casos de incidentes experimentados por usuários internos e externos, analisando as circunstâncias, será de grande valia aos laboratoristas no processo de *design* do hipotético BIIAJud.

Há uma questão importante a ser pontuada aqui, pois os usuários internos da área técnica são profissionais em TI, ao passo que os demais – tanto os usuários internos da área de negócio como os usuários externos – são operadores jurídicos ou jurisdicionados em geral, de todas as classes sociais, variadas idades e graus de instrução, que têm pouco ou nenhum conhecimento em assuntos de informática e tecnologia, muito menos de IA.

Para os usuários internos da área técnica, é preciso ter em mente que, como descrito no item 3.5, o CNJ utiliza, para o repositório de códigos-fontes dos modelos de IA registrados na plataforma Sinapses, a metodologia *DevOps*, utilizando a plataforma GitLab para integrar desenvolvedores e operadores, o que permite entrega ágil e contínua.

Acrescente-se agora a informação de que o próprio GitLab possui um programa público de recompensa por descoberta de *bugs*, lançado em dezembro de 2018⁶²⁶, tratando-se de uma espécie de *nudge* para estimular a comunidade de TI a colaborar para a segurança da plataforma.

Tal informação conduz à percepção de que, para os usuários internos da área de TI, é possível que se verifique que há outras maneiras de informar, registrar e tratar erros e incidentes, que sejam adequadas à área técnica, podendo não ser o mesmo método colocado à disposição para leigos.

Os leigos em TI, que são os operadores jurídicos tanto internos como externos e sobretudo os jurisdicionados em geral, apenas podem, quando muito, perceber quando algo saiu errado ao usarem serviços prestados pelo Poder Judiciário,

⁶²⁶WANG, Kathy. **Inside the GitLab public bug bounty program.** GitLab, 29 abr. 2019. Disponível em: <https://about.gitlab.com/blog/2019/04/29/inside-the-gitlab-public-bug-bounty-program/>. Acesso em: 18 jul. 2022.

estando totalmente fora do seu controle real a capacidade de tomar alguma providência diante de um erro ou incidente de IA que atinja sua esfera de direitos. Podem, entretanto, comunicar sua ocorrência, e para isso servirá o BIIAJud.

A título de exercício de abstração e “heurística do medo”⁶²⁷, imagine-se a possibilidade de um modelo de IA, que foi programado para fazer bloqueio de ativos financeiros, apresentar uma falha que atinja o CPF e as contas bancárias de outras pessoas que não somente da parte executada, ou em valores superiores ao comando dado.

Ou um advogado que se veja impedido de protocolar uma peça processual urgente por falha de um modelo de IA, usado em módulo de protocolo de sistemas de processo eletrônico, originalmente projetado para melhorar a classificação de assuntos.

Ou centenas de pessoas atingidas por vazamento de dados de processos existentes em *datasets* de treinamento de modelos de IA, ou pior, atingidas pela perda de processos eletrônicos irreversivelmente criptografados por *hackers* até mesmo nos *backups* – que é a tendência atual de ameaça cibernética, conforme narrado no item 3.4.2 do capítulo anterior.

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

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

Significa concretizar o comando normativo expresso no art. 7º, V, da Resolução nº 395/2021, pelo qual cabe ao LIODS/CNJ “abrir espaço para a participação cidadã na concepção de projetos inovadores no Poder Judiciário ou que contribuam para a efetividade da Agenda 2030”⁶²⁸.

Significa, em última análise, a partir da experiência vivenciada por usuários com sistemas de IA, reunir informações que serão muito úteis para as seguintes

⁶²⁷JONAS, Hans. Op. cit.

⁶²⁸CNJ. **Resolução nº 395, de 7 de junho de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 16 jul. 2022.

reflexões filosóficas e práticas: Inteligência Artificial no Judiciário para quê? Para quem? Em quais aplicações os riscos não compensam os benefícios?

Feitas todas essas considerações a respeito de fatores a serem observados no processo de *design* para eventual criação de um banco de incidentes de IA no Judiciário, de acordo com o estado da arte da metodologia de *design thinking* aplicado a políticas públicas e de *legal design* aplicado a inovações na área jurídica, com foco no usuário, resta acrescentar contribuições da pesquisa exploratória empreendida, quando se verificou que há dois bancos de dados que podem servir de inspiração.

4.3.2 “MAUDE database” e sistema “MedWatch” como parâmetros

O banco de dados MAUDE (*Manufacturer and User Facility Device Experience*) foi criado em 1991 para gerenciar riscos de dispositivos médicos, sendo mantido desde então pela FDA (U.S. *Food & Drug Administration*), agência federal dos Estados Unidos para vigilância de alimentos e medicamentos, integrante do *U.S. Department of Health & Human Services*⁶²⁹.

Nesse aspecto, fazendo uma comparação em termos de atribuição para regulação e controle, pode-se dizer que o CNJ é o órgão responsável para a vigilância das aplicações de IA usadas no sistema judiciário no Brasil, assim como a FDA é o órgão responsável para a vigilância de dispositivos médicos nos Estados Unidos.

As informações armazenadas no MAUDE são periodicamente verificadas e utilizadas pela FDA como um dos mecanismos de vigilância pós-comercialização de dispositivos médicos, os quais, para serem seguros, precisam ter seus riscos gerenciados em todo o ciclo de existência – da mesma forma que a OCDE recomenda que seja feito em relação ao acompanhamento de todo o ciclo de existência das aplicações de IA⁶³⁰.

O aspecto mais importante do MAUDE é que foi projetado para armazenar o registro de intercorrências a partir da experiência dos usuários – o que também guarda consonância com o que se preconiza em relação à IA, a qual deve ser

⁶²⁹U.S. FOOD & DRUG ADMINISTRATION (FDA). **About manufacturer and user facility device experience (MAUDE)**. Disponível em: <https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities/about-manufacturer-and-user-facility-device-experience-maude>. Acesso em: 18 jul. 2022.

⁶³⁰OCDE. Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. **OECD Legal Instruments**. 21 maio 2019. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#adherents>. Acesso em: 18 jul. 2022.

projetada com valores centrados no ser humano, bem como estar sujeita ao controle do usuário.

É importante frisar que no MAUDE o registro de eventos adversos ao utilizar dispositivos médicos é obrigatório para fabricantes, importadores e estabelecimentos hospitalares, enquanto é facultativo para pessoas físicas, ou seja, profissionais de saúde como médicos e enfermeiros, pacientes e seus familiares, consumidores em geral⁶³¹.

Fazendo um paralelo com a utilização de sistemas de IA nos tribunais brasileiros, o registro de eventos adversos seria obrigatório para usuários internos (servidores e magistrados), enquanto ficaria à disposição para registro voluntário por usuários externos (advogados, policiais, defensores públicos, procuradores, membros do Ministério Público, peritos e jurisdicionados em geral).

Existem 4 categorias de eventos adversos que podem ser registradas no MAUDE: morte, lesão, mau funcionamento ou outro. Todos os registros, tanto os obrigatórios como os voluntários, ficam armazenados no MAUDE e estão disponíveis para consulta pública, que pode ser feita buscando por problema do produto, classe do produto, tipo de evento adverso, fabricante, número do modelo, número do relatório, marca, código do produto e data, sendo então possível emitir relatórios (Figura 18).

⁶³¹U.S. FOOD & DRUG ADMINISTRATION (FDA). **MAUDE - Manufacturer and User Facility Device Experience.** Disponível em: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Acesso em: 18 jul. 2022.

Figura 18 – Tela de busca do banco de dados MAUDE

Fonte: U.S Food & Drug Administration⁶³².

Se fossem usados parâmetros semelhantes para busca de incidentes de modelos de IA usados nos tribunais brasileiros, seria importante que cada modelo tivesse nome e número identificador. Os outros campos possíveis de busca poderiam ser por descrição do modelo (conforme o Painel de IA do CNJ), por funcionalidade, por ramo de Justiça, por tribunal, por Estado e por data. Talvez fosse mais recomendável, no início, deixar em aberto, para busca por palavras, as categorias de eventos, ou então defini-las no processo de *design* do hipotético BIIAJud.

Com base nos relatórios recebidos, a FDA analisa o desempenho dos dispositivos médicos, emite alertas caso detecte suspeita de problemas de segurança, além de utilizar os dados coletados a partir dos relatórios para avaliar o risco-benefício dos dispositivos⁶³³.

Todavia, consta uma advertência quanto às limitações do MAUDE, relacionadas à subnotificação de casos, principalmente voluntários, e à possibilidade de inconsistência, equívoco, falta de verificação da ocorrência ou falsidade nos registros. Diante de tais limitações, a FDA informa que o banco de dados representa uma fonte valiosa de informações, que se soma a outras ferramentas de controle, mas

⁶³²U.S. FOOD & DRUG ADMINISTRATION (FDA). **MAUDE - Manufacturer and User Facility Device Experience.** Disponível em: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Acesso em: 18 jul. 2022.

⁶³³Ibidem.

não pode ser usado, por si só, para indicar categoricamente a prevalência de tal ou qual evento adverso⁶³⁴.

A respeito do canal pelo qual é feito o registro do evento adverso que ficará armazenado no banco de dados MAUDE, desde 1993 foi criado o sistema *MedWatch*, que atualmente está disponibilizado para registro *on-line*, a ser usado pelas pessoas físicas que não são obrigadas a reportar incidentes, mas que se dispõem voluntariamente a fazê-lo. Pelo formulário *on-line*, é preciso selecionar entre as opções disponíveis os detalhes que correspondem ao evento reportado, bem como digitar em até 4.000 caracteres as circunstâncias do que ocorreu⁶³⁵.

Figura 19 – Tela inicial de registro de evento adverso no *MedWatch Online*



Fonte: U.S Food & Drug Administration⁶³⁶.

Fazendo um paralelo do que poderia ser adaptado para registro de eventos adversos associados a modelos de IA utilizados nos tribunais, talvez fosse interessante a segmentação em dois grupos: (i) aqueles sujeitos a registro obrigatório, que são os usuários internos, e (ii) aqueles que podem fazer o registro voluntário, que são os usuários externos.

Além disso, entre os sujeitos à obrigatoriedade do registro, talvez fosse interessante segmentar entre área técnica (profissionais de TI) e área de negócio

⁶³⁴U.S. FOOD & DRUG ADMINISTRATION (FDA). **MAUDE - Manufacturer and User Facility Device Experience.** Disponível em: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Acesso em: 18 jul. 2022.

⁶³⁵U.S. FOOD & DRUG ADMINISTRATION (FDA). **MedWatch online voluntary reporting form.** Disponível em: <https://www.accessdata.fda.gov/scripts/medwatch/index.cfm>. Acesso em: 18 jul. 2022.

⁶³⁶Ibidem.

(servidores e magistrados). Para cada categoria de usuário, poderia ser disponibilizado um formulário com campos próprios, adequados ao perfil da categoria.

No que se refere à IA aplicada à Medicina, atualmente a FDA disponibiliza uma lista de dispositivos médicos habilitados à AI/ML (siglas em inglês para designar Inteligência Artificial/Aprendizado de Máquina)⁶³⁷.

Quanto ao valor do banco de dados MAUDE para o avanço da pesquisa científica em termos de avaliação do risco-benefício de dispositivos médicos, apresenta-se como exemplo um estudo altamente referenciado, publicado em 2016, localizado na base de dados da *U.S. National Library of Medicine*, e que foi apresentado pela primeira vez em 2014 como *paper* na área de cirurgia cardíaca, por ocasião da 50^a Reunião Anual da Sociedade de Cirurgiões Torácicos⁶³⁸.

O estudo analisou 14 anos de dados extraídos do MAUDE sobre eventos adversos ocorridos em cirurgias robóticas que usaram como dispositivo médico o sistema cirúrgico *Da Vinci*, fabricado pela empresa norte-americana *Intuitive Surgical*, que nos Estados Unidos obteve autorização da FDA para funcionar no ano 2000, e ao tempo do estudo permanecia sendo o único robô autorizado.

Os pesquisadores tiveram que desenvolver uma ferramenta de IA para analisar os 2,9 milhões de registros encontrados no MAUDE entre 01/01/2000 e 01/01/2013, aplicando processamento de linguagem natural para estudar as circunstâncias dos eventos adversos conforme haviam sido reportadas. Vale citar aqui os resultados, apenas para deixar registrado o nível de detalhamento de inferência que foi possível realizar (em tradução livre):

Durante o período do estudo, foram relatadas 144 mortes (1,4% dos 10.624 relatórios), 1.391 lesões em pacientes (13,1%) e 8.061 defeitos no dispositivo (75,9%). O número de lesões e eventos de morte por procedimento permaneceu relativamente constante (média = 83,4, intervalo de confiança de 95% (IC), 74,2-92,7 por 100.000 procedimentos) ao longo dos anos. **As especialidades cirúrgicas** para as quais os robôs são amplamente utilizados, como ginecologia e urologia, apresentaram menor número de lesões, mortes e conversões por procedimento do que as cirurgias mais complexas, como cardiotorácica e cabeça e pescoço (106,3 vs. 232,9 por 100.000

⁶³⁷U.S. FOOD & DRUG ADMINISTRATION (FDA). **Artificial Intelligence and Machine Learning (AI/ML) – enabled medical devices.** Disponível em: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices#resources>. Acesso em: 18 jul. 2022.

⁶³⁸ALEMZADEH, Homa *et al.* Adverse events in robotic surgery: a retrospective study of 14 years of FDA data. US National Library of Medicine. **PloS one**, v. 11, n. 4, p. e0151470, 2016. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4838256/>. Acesso em: 18 jul. 2022.

procedimentos, razão de risco = 2,2, IC 95%, 1,9-2,6). **Mau funcionamento do dispositivo e instrumento**, como queda de instrumentos queimados / quebrados no paciente (14,7%), arco elétrico dos instrumentos (10,5%), operação não intencional dos instrumentos (8,6%), erros do sistema (5%) e problemas de vídeo / imagem (2,6 %), constituiu grande parte dos relatórios. **O mau funcionamento do dispositivo afetou os pacientes** em termos de lesões ou interrupções do procedimento. Em 1.104 (10,4%) de todos os eventos, o procedimento foi interrompido para reiniciar o sistema (3,1%), para converter o procedimento em técnicas não-robóticas (7,3%) ou para reagendá-lo (2,5%)⁶³⁹.

As principais conclusões do artigo foram no sentido da necessidade de aprimorar o treinamento dos operadores da máquina, com simuladores e exposição a problemas técnicos para que saibam lidar com eles e melhorem o tempo de resposta; aprimorar o nível de detalhamento dos relatórios de eventos adversos para facilitar a investigação e permitir a criação mais rápida de solução; aprimorar o *design* do produto com mais requisitos de segurança, tudo para reduzir os incidentes futuros.

Observando bem, todas as conclusões desse artigo médico também podem ser aplicadas à interação entre operadores do Direito e os robôs dos tribunais.

Enfim, confirmando a hipótese inicial de pesquisa, depreendendo-se que o MAUDE da FDA pode ser uma ferramenta útil para servir de modelo para o registro de eventos adversos da IA aplicada nos tribunais, verificou-se que a *Algorithm Justice League*, organização não governamental criada pela cientista da computação que é uma referência mundial em IA ética, Joy Buolamwini (conforme mencionado no item 3.4.4 do capítulo anterior), publicou em 2020 um relatório no qual apresentou justamente todos os motivos pelos quais recomenda que os mecanismos da FDA sejam usados como referência para gerenciar os riscos da IA, tendo em vista a experiência da agência reguladora acumulada ao longo dos anos com novas tecnologias de alta complexidade no avanço da área médica⁶⁴⁰.

Esse relatório da *Algorithm Justice League* corrobora a hipótese de pesquisa de modo mais enfático e direto, em relação ao que já havia sido observado

⁶³⁹ALEMZADEH, Homa *et al.* Adverse events in robotic surgery: a retrospective study of 14 years of FDA data. US National Library of Medicine. *PloS one*, v. 11, n. 4, p. e0151470, 2016. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4838256/>. Acesso em: 18 jul. 2022.

⁶⁴⁰LEARNED-MILLER, Erik *et al.* **Facial recognition technologies in the wild: a call for a federal office.** Algorithm Justice League/MacArthur Foundation. 29 maio 2020. Disponível em: https://assets.website-files.com/5e027ca188c99e3515b404b7/5ed1145952bc185203f3d009_FRTsFederalOfficeMay2020.pdf. Acesso em: 18 jul.2022.

desde o final do primeiro capítulo desta dissertação, no item 2.3, quando se verificou que um documento divulgado em dezembro de 2021 pela Comissão Europeia para a Eficiência da Justiça (CEPEJ), o “Roteiro revisado para garantir um acompanhamento adequado da Carta Ética da CEPEJ sobre o uso de inteligência artificial em sistemas judiciais e seu ambiente”⁶⁴¹, continha um item de nº 17 que sugeria – ainda que de forma vaga – a criação de um banco de dados onde possa ser registrada, dentre outras informações, a experiência do usuário.

4.3.3 “AI Incident Database” como parâmetro

Quando o projeto desta pesquisa foi apresentado para concorrer à seleção no Mestrado Profissional da ENFAM, em julho de 2020, não foi cogitada outra hipótese de pesquisa para servir como parâmetro de registro de eventos adversos da IA ocorridos nos tribunais, que não a de investigar a possibilidade de adaptação das ferramentas norte-americanas MAUDE e *MedWatch* da FDA, visto que essas demonstraram ser as opções mais avançadas de registro de incidentes na área de regulação médica no cenário mundial, pelo que foi possível observar até aquele momento em pesquisa inicial exploratória.

Para chegar a tal percepção, tomou-se por base um artigo brasileiro na área de regulação e saúde, publicado em 2019 por Stela Melchior e William Waissmann, respectivamente vinculados à Agência Nacional de Vigilância Sanitária (Anvisa) e à Escola Nacional de Saúde Pública Sérgio Arouca, da Fundação Oswaldo Cruz (ENSP/Fiocruz).

Em tal artigo, os pesquisadores analisaram a “Regulação de dispositivos médicos: vigilância pós-mercado como estratégia de gerenciamento de riscos”, fazendo um quadro comparativo entre a estrutura de regulação do Brasil, dos Estados Unidos, da União Europeia e do Japão, tendo em vista que “a tecnovigilância é apresentada como ação do Estado e como parte do gerenciamento de risco em serviços de saúde”⁶⁴².

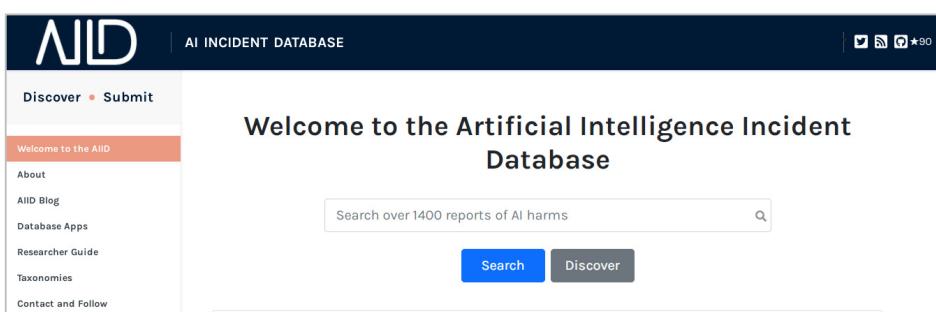
⁶⁴¹ CEPEJ. **Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment.** 9 dez. 2021. Disponível em: <https://rm.coe.int/cepej-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 6 jan. 2022.

⁶⁴² MELCHIOR, Stela Candioto; WAISSMANN, William. Regulação de dispositivos médicos: vigilância pós-mercado como estratégia de gerenciamento de riscos. **Vigilância Sanitária em Debate:**

Acontece que, alguns meses depois da apresentação do projeto desta pesquisa, foi lançada em novembro de 2020 uma iniciativa pioneira pela organização privada PAI – *Partnership on AI* (Parceria em IA)⁶⁴³, e que convergia com o objetivo do presente trabalho, sendo inclusive específico para a IA, razão pela qual se resolveu considerá-la como hipótese de pesquisa superveniente, a merecer também descrição e análise, até mesmo para compará-la com as ferramentas da FDA.

Trata-se do AIID - *AI Incident Database* (Banco de Dados de Incidentes de IA), que até fevereiro de 2022 já havia sido alimentado com mais de 1.200 incidentes de várias naturezas causados por aplicações de IA ao redor do mundo, sendo apresentado como “a única coleção de danos causados ou quase causados pelo uso de IA em todas as disciplinas, geografias e casos de uso”, o que fazia da iniciativa “um recurso tangível sobre o que pode acontecer se a IA não for construída ou implantada de uma maneira que considere as implicações do mundo real”⁶⁴⁴ (Figura 20).

Figura 20 – Tela inicial do *AI Incident Database*



Fonte: AIID⁶⁴⁵.

A inspiração para o *AI Incident Database* não foi um banco de dados semelhante da área médica, e sim de duas bases de dados de áreas distintas, aviação e cibersegurança⁶⁴⁶, a saber:

- O *Aviation Accident Database*, que documenta informações sobre acidentes ocorridos na aviação civil nos Estados Unidos, mantido

Sociedade, Ciência & Tecnologia, v. 7, n. 4, p. 67-76, 2019. Disponível em: <https://www.redalyc.org/journal/5705/570566202011/570566202011.pdf>. Acesso em: 18 jul. 2020.

⁶⁴³MCGRGOR, Sean. **When AI systems fail: introducing the AI incident database**. 18 nov. 2020. Disponível em: <https://partnershiponai.org/aiincidentdatabase/>. Acesso em: 6 jan. 2022.

⁶⁴⁴PARTNERSHIP ON AI. **AI incident database**. Disponível em: <https://incidentdatabase.ai/>. Acesso em: 6 jan. 2022.

⁶⁴⁵AI INCIDENT DATABASE. **Welcome do the artificial intelligence incident database**. Disponível em: <https://incidentdatabase.ai/>. Acesso em 18 jul. 2022.

⁶⁴⁶MCGRGOR, Sean. **When AI systems fail: introducing the AI incident database**. 18 nov. 2020. Disponível em: <https://partnershiponai.org/aiincidentdatabase/>. Acesso em: 6 jan. 2022.

desde 1962 pela agência norte-americana que regula a segurança de transportes⁶⁴⁷;

- O *Common Vulnerabilities and Exposures* (CVE), um banco de dados da área de cibersegurança existente desde 1999 para “identificar, definir e catalogar vulnerabilidades de segurança cibernética divulgadas publicamente”, patrocinado pela agência norte-americana que regula a segurança cibernética⁶⁴⁸.

Combinando os pontos fortes desses dois bancos de dados, surgiu o projeto de código aberto *AI Incident Database*, cuja arquitetura completa foi descrita em artigo⁶⁴⁹, apresentado em maio de 2021 por seu idealizador, Sean McGregor, na conferência anual da AAAI (*Association for the Advancement of Artificial Intelligence*), a importante organização mencionada no primeiro capítulo desta dissertação.

Em março de 2022, sobreveio a carta de fundação da *Responsible AI Collaborative*, um coletivo organizado por Sean McGregor para gerir o *AI Incident Database*, sob o ideal de atuação colaborativa para a IA responsável, com a missão de “identificar, definir e catalogar incidentes de Inteligência Artificial”, conforme relatório que detalhou todo o aprendizado desde que o banco de dados foi criado, dizendo: “Nós respondemos o aforismo de Santayana, ‘aqueles que não conseguem lembrar o passado estão condenados a repeti-lo’... **com dados**”⁶⁵⁰ (em tradução livre).

Assim, o relatório de fundação da *Responsible AI Collaborative* (i) reuniu reportagens e 16 artigos científicos que citaram a iniciativa pioneira, (ii) apresentou taxonomias para os incidentes, (iii) demonstrou as mudanças de interface do banco de dados para uma melhor experiência do usuário, (iv) descreveu as melhorias implementadas no sistema, (v) adiantou quais serão os próximos passos para o avanço do projeto e (vi) forneceu muitos outros detalhes importantes, incluindo o custo

⁶⁴⁷NATIONAL TRANSPORTATION SAFETY BOARD (NTSB). **Aviation Accident Database & Synopses**. Disponível em: <https://www.ntsb.gov/Pages/AviationQuery.aspx>. Acesso em: 18 jul. 2022.

⁶⁴⁸COMMON VULNERABILITIES AND EXPOSURES (CVE). **CVE® Program Mission**. Disponível em: <https://www.cve.org/>. Acesso em: 18 jul. 2022.

⁶⁴⁹MCGREGOR, Sean. Preventing repeated real world AI failures by cataloging incidents: The AI incident database. In: **Proceedings of the AAAI Conference on Artificial Intelligence**. 2021. p. 15458-15463. Disponível em: <https://ojs.aaai.org/index.php/AAAI/article/view/17817>. Acesso em: 18 jul. 2022.

⁶⁵⁰RESPONSIBLE AI COLLABORATIVE. **Founding report**. 28 mar. 2022. Disponível em: <https://docsend.com/view/47z2whznatd39vf9>. Acesso em: 18 jul. 2022.

total de manter *AI Incident Database* em 2021, que foi em torno de cem mil dólares (US\$ 105,073.66)⁶⁵¹.

Todas as evidências levantadas indicam que o *AI Incident Database*, por já ter sido desde o nascêdoiro projetado para a finalidade de registrar incidentes de IA, afigura-se o mais adequado para servir de parâmetro principal para o hipotético BIIAJud – Banco de Incidentes de Inteligência Artificial do Judiciário.

Inclusive utilizou-se nessa denominação hipotética a palavra “incidentes”, por soar mais adequada que a expressão “eventos adversos”, ao parecer permitir melhor compreensão pela população em geral, que é a destinatária dos serviços prestados pelo Judiciário.

Porém, por um lado, ao buscar inspiração no *AI Incident Database*, ainda serão necessárias muitas modificações para o *design* adequado às especificidades de uso pelo Poder Judiciário brasileiro, a serem sopesadas durante a metodologia de *design thinking*, caso o projeto venha a se realizar, no âmbito do Laboratório de Inovação e dos Objetivos de Desenvolvimento Sustentável do CNJ (LIODS/CNJ).

Por outro lado, não se pode deixar de levar em consideração toda a experiência acumulada com o MAUDE e o *MedWacth*, que existem desde os anos 1990 e foram se aprimorando ao longo de mais de três décadas, enquanto o *AI Incident Database* sequer completou dois anos de atividade.

Assim, a solução que se apresenta como a mais recomendável é refletir, durante o processo de criação do hipotético BIIAJud por *design thinking*, sobre o que cada parâmetro tem de melhor, mesclando características e funcionalidades de cada sistema aqui apresentado, de modo a customizar uma solução adequada ao Poder Judiciário brasileiro.

Em se tratando de um *wicked problem*, essa construção requer a participação de vários atores em sinergia de conhecimentos multidisciplinares, o que não é possível fazer nesta pesquisa acadêmica individual, limitada às contribuições que se empenhou em entregar para o avanço do debate ético sobre a Inteligência Artificial aplicada ao sistema judicial de seu país.

⁶⁵¹Ibidem.

5 CONCLUSÃO

Esta pesquisa procurou passar dos princípios à ação, em busca de concretizar preceitos éticos a serem observados na utilização da Inteligência Artificial pelo Poder Judiciário brasileiro, tendo por maior baliza o princípio da precaução.

Ao enfrentar o desafio, evidenciou a dificuldade de equilibrar na prática, em um lado da balança, o salto de eficiência, produtividade e qualidade da prestação jurisdicional que se espera com uso da IA, e do outro lado, o gerenciamento efetivo dos riscos potenciais que essa tecnologia produz a direitos humanos fundamentais dos usuários do sistema de justiça.

A interdisciplinaridade foi a tônica do estudo, dando uma dimensão da complexidade do tema, pois até atingir o objetivo final foi necessário buscar referenciais em diferentes áreas do conhecimento humano, como Filosofia, Bioética, História, Sociologia, Ciência Política, Administração Pública, Direito, Design, Tecnologia da Informação e Comunicação, Medicina e até Aviação.

Para amalgamar informações obtidas em searas tão díspares, foi preciso acrescentar-lhes contexto histórico-cronológico à medida que iam sendo apresentadas, permitindo evidenciar como e por que se relacionavam.

Tendo em vista que o problema de pesquisa estava ligado à última etapa do ciclo de gerenciamento de riscos – tratamento de eventos adversos – foi necessário ter uma noção prévia do todo: princípios éticos e regulação da IA, levantamento dos tipos de riscos em diversos aspectos, aplicações reais de IA nos tribunais, mecanismos de controle durante o *design* do modelo de IA, até chegar no ponto em que foi preciso refletir sobre qual seria a providência mais eficaz a tomar diante de um evento adverso.

Durante o caminho percorrido para ter uma visão holística do problema, foram encontrados alguns achados que podem ser apresentados como contribuições adicionais desta pesquisa, inclusive para agenda de pesquisas futuras, conforme será visto a seguir.

No primeiro capítulo, demonstrou-se que, embora a IA represente um avanço tecnológico sem precedentes pela rapidez de sua evolução, é preciso ter em mente que não é a primeira vez que a humanidade se depara com criações tecnológicas poderosas e arriscadas.

Há muito o que aprender com as lições do passado, principalmente as que se relacionam com os dilemas bioéticos, quando a engenharia genética foi capaz de produzir o DNA recombinante, porque há evidências de que a História está se repetindo em alguns aspectos no tocante à IA, mesmo sendo esta uma tecnologia tão disruptiva.

O primeiro aspecto é sobre a classificação de risco, que em 1975 foi o ponto-chave para o consenso obtido entre os cientistas para suspender a moratória da pesquisa em DNA recombinante, passando-se à criação de mecanismos de maior ou menor controle a partir do risco das aplicações.

Demonstrou-se que a História está se repetindo porque atualmente o debate sobre a elaboração do *AI Act* na Europa também tem como ponto-chave justamente a **classificação de risco das aplicações de IA** – que é uma agenda que convém ao Poder Judiciário brasileiro acompanhar, refletindo sobre o grau de risco dos diferentes modelos de IA em uso nos tribunais.

O segundo aspecto diz respeito à **importância da pesquisa acadêmica independente**, capitaneada por organizações públicas e instituições de ensino e de pesquisa científica, e não por empresas. Enquanto estas são movidas pelo lucro e por isso podem acabar abdicando da ética, aquelas buscam o progresso desde que seja benéfico para a humanidade.

Em 1975 a liberdade de pensamento crítico com compromisso ético dos cientistas revelou-se fundamental para encontrar as melhores soluções no sentido de controlar os riscos da tecnologia de DNA recombinante. No que se refere à IA, como se demonstrou ao longo do primeiro capítulo, a História está se repetindo porque as principais iniciativas em torno de diretrizes éticas e uso benéfico têm sido construídas de modo colaborativo por milhares de pesquisadores vinculados às mais prestigiadas instituições do mundo.

Entretanto, conforme foi visto no segundo capítulo, acendendo um sinal de alerta, nos últimos anos duas polêmicas vieram à tona, envolvendo gigantes empresas de tecnologia: uma foi a demissão de uma das líderes da equipe de Ética em IA do Google motivada porque ela escreveu artigo científico incômodo para a empresa, a outra foi o ataque da Amazon à honra da pesquisadora que verificou viés algorítmico discriminatório em *software* de reconhecimento facial.

De tais episódios, puderam-se extrair duas lições úteis ao Poder Judiciário: (i) convém adotar cautelas para garantir que o **investimento em pesquisa**

acadêmica sobre IA seja direcionado a instituições livres de conflitos de interesses, bem como (ii) convém estar preparado para receber críticas da comunidade acadêmica ou da sociedade civil às aplicações de IA em uso nos tribunais, tendo um **plano de gerenciamento de crise de imagem** e um **plano para averiguação técnico-científica da crítica**, cuidando de comunicar os resultados.

Demonstrou-se que nos Estados Unidos e no Reino Unido há organizações não governamentais de defesa dos direitos humanos e uso benéfico da tecnologia que costumam realizar ou patrocinar estudos sobre aplicações reais a fim de verificar vieses, com grande repercussão quando são encontrados resultados, afetando a credibilidade da instituição pesquisada e a confiança da população na nova tecnologia. Talvez seja uma questão de tempo começar a acontecer esse tipo de verificação também no Brasil em relação a aplicações de IA usadas no Poder Judiciário, razão pela qual convém estar preparado para enfrentar eventual situação semelhante.

Aliás, registrou-se que o valor da **adequada comunicação institucional**, informando à sociedade sobre uma nova tecnologia, já havia sido notado pelos cientistas desde 1975 quanto ao DNA recombinante, como medida necessária para tranquilizar a população sobre a novidade, conquistando e mantendo sua confiança.

A respeito do tema, no segundo capítulo deste trabalho foi visto que um Judiciário que inspire confiança é o principal anseio da sociedade em relação a tal Poder, tratando-se de atributo essencial ao Estado Democrático de Direito.

Em se tratando de um Judiciário cada vez mais digital, destacou-se a importância da **confiança da população** nos serviços tecnológicos prestados, principalmente em relação à IA, quando o imaginário popular e até o meio acadêmico são influenciados por notícias sensacionalistas sobre “robô-juiz” – algo que simplesmente não existe no Brasil, por não ser permitida pelo CNJ qualquer decisão automatizada sem controle humano, como ficou demonstrado.

Contudo, verificou-se que não há muitas informações disponíveis nas páginas dos tribunais na Internet sobre seus projetos de IA, encontrando-se apenas algumas boas práticas isoladas de disponibilização de documentos, tutoriais e vídeos explicativos. Foi visto que há proposta no CNJ de reformulação dos sítios eletrônicos dos tribunais para que contenham informações essenciais de forma padronizada, o que permitiu ter-se considerado relevante incluir **informações sobre IA** entre as essenciais, em se tratando de atividade de risco.

A respeito dos riscos, foi visto no segundo capítulo que no topo da lista estão aqueles relacionados à **cibersegurança**, com frequentes ataques *hackers* aos tribunais, figurando o Brasil como país mais visado da América do Sul. A pesquisa demonstrou, de um lado, o **déficit de profissionais de TI** em geral e de segurança cibernética em especial, e do outro, a ausência de um plano concreto por parte do Poder Judiciário para enfrentar o prognóstico de escassez cada vez maior de especialistas. Foram apontados alguns instrumentos para incrementar a resiliência cibernética do Poder Judiciário, conforme o estado da arte sobre o tema.

Da mesma maneira, quanto aos riscos de **discriminação algorítmica**, também ficou constatada a **ausência de um plano concreto para garantir a real diversidade** nas equipes de desenvolvimento da IA nos tribunais, sobretudo quanto a pessoas negras e mulheres, segmentos mais atingidos por vieses preconceituosos e que estão sub-representados na força de trabalho de TI, embora sejam maioria na população brasileira, conforme se evidenciou. Foram apresentadas diversas estratégias para implementação da diversidade, de acordo com fontes variadas de pesquisa.

No tocante aos riscos relacionados à **falta de transparência algorítmica**, tratando-se de tema correlato à **discriminação algorítmica**, demonstrou-se a dimensão do risco de aplicações de IA para predição e avaliação de risco em matéria penal, bem como para reconhecimento facial. Foi possível encontrar estudos promissores no sentido da **IA interpretável**, em vez da IA explicável, assim como foram localizados instrumentos avançados para **documentação**, de modo a colaborar com o suprimento de uma deficiência notada em número significativo de projetos dos tribunais.

Quanto aos riscos relacionados à falta de **controle do usuário**, demonstrou-se com base em artigo elaborado à luz de pesquisa quantitativa que a esmagadora maioria dos magistrados brasileiros não conhece conceitos básicos de IA, não conhece a fundo o instrumento regulatório sobre o tema e não se sente apta a exercer efetivo controle da IA, necessitando de **capacitação**.

No terceiro capítulo, foi apresentado um panorama dos projetos de IA nos tribunais de acordo com dados oficiais, constatando que a quase totalidade dos modelos se destina à **área-meio**, para atividades burocráticas das secretarias e cartórios. Foi visto que não há projetos de predição de reincidência e de avaliação de

risco em matéria penal, e foi feita uma análise crítica de projetos de reconhecimento facial para levantar o debate sobre questões éticas envolvidas.

Ao analisar a ferramenta auxiliar mais próxima da atividade-fim, que seleciona modelos de decisão à escolha do magistrado, criada em São Paulo pelo Tribunal Regional Federal da 3^a Região e premiada pelo CNJ, também foi possível observar a notável qualidade do **manual de diretrizes** que os laboratórios de inovação do referido tribunal criaram para validação ético-jurídica de seus modelos de IA.

Ficou evidenciada a necessidade de **criação de um Comitê de Ética em IA** ou órgão equivalente no âmbito do CNJ e de todos os tribunais, haja vista as inúmeras atividades necessárias para verificar continuamente o risco-benefício dos modelos de IA desenvolvidos.

Para tratar das falhas, erros, eventos adversos e incidentes de IA em geral, foi necessário compreender a regulação e o controle da IA como um **wicked problem**, ou seja, um problema complexo para a formulação de políticas públicas, situado no final do ciclo de gerenciamento de risco. Observou-se que a metodologia de **design thinking** se afigura como a mais adequada para a criação da solução centrada no usuário e mediante a colaboração multidisciplinar de vários atores, que esta pesquisa acadêmica individual não tem como suprir, até porque se trata de assunto institucional para ficar a cargo do Laboratório de Inovação e dos Objetivos de Desenvolvimento Sustentável do CNJ (LIODS/CNJ).

Demonstrou-se que o regramento atual para uso ético da IA no Judiciário quanto ao tema, previsto no art. 27 da Resolução 332/2020, limitando-se a determinar a informação e registro de eventos adversos mas sem providenciar uma estrutura para tanto, não é suficiente para gerenciar de maneira eficaz os riscos dessa tecnologia, do ponto de vista dos princípios da precaução e da prevenção, os quais pressupõem evitar riscos futuros e incertos a partir dos riscos conhecidos.

Ficou esclarecido que os dados relativos ao registro dos eventos adversos, se não ficarem organizados em estrutura adequada, poderão ser desperdiçados ou subutilizados, em vez de terem grande proveito para auxiliar o Judiciário a se precaver de novos eventos.

Atingindo o objetivo final desta pesquisa, **como principal contribuição, apresentaram-se dois parâmetros de bancos de dados (MAUDE/MedWatch e AI Incident Database)** que podem ser combinados em seus detalhes mais funcionais e

adaptados para as especificidades do Poder Judiciário, por meio de *design thinking* a se realizar no LIODS/CNJ, para criar um **Banco de Incidentes de Inteligência Artificial do Judiciário (BIIAJud)**, contribuindo para o progresso da ciência e a proteção dos direitos fundamentais dos jurisdicionados, em prol da inovação responsável na Justiça brasileira.

REFERÊNCIAS

- AAAI. **AAAI Presidential Panel on Long-Term AI Futures: 2008-2009 Study**. Disponível em: <https://www.aaai.org/Organization/presidential-panel.php>. Acesso em: 16 nov. 2021.
- AAAI. **Association for the Advancement of Artificial Intelligence**. Disponível em: <https://www.aaai.org/>. Acesso em: 16 nov. 2021.
- AAAI. **Duke computer scientist wins \$1 million artificial intelligence prize, a “new Nobel”**. Disponível em: <https://aaai.org/Pressroom/Releases/release-21-1012.php>. Acesso em: 27 jun. 2022.
- AAAI. **Past AAAI Officials**. Disponível em: <https://aaai.org/Organization/past-officers.php>. Acesso em: 16 nov. 2021.
- AAAI. **AAAI Squirrel AI Award for Artificial Intelligence for the Benefit of Humanity**. Disponível em: <https://aaai.org/Awards/squirrel-ai-award.php>. Acesso em: 27 jun. 2022.
- ACAT. **Assistive Context-Aware Toolkit (ACAT) is a free, open source software to enable people with severe disabilities to communicate**. Disponível em: <https://01.org/acat/>. Acesso em: 21 nov. 2021.
- ACLU. **ACLU History**. Disponível em: <https://www.aclu.org/about/aclu-history>. Acesso em: 15 jul. 2022.
- AI INCIDENT DATABASE. **Welcome do the artificial intelligence incident database**. Disponível em: <https://incidentdatabase.ai/>. Acesso em 18 jul. 2022.
- AJL. **Algorithmic Justice League**. Disponível em: <https://www.ajl.org/>. Acesso em: 13 jul. 2022.
- ALEMZADEH, Homa *et al.* Adverse events in robotic surgery: a retrospective study of 14 years of FDA data. US National Library of Medicine. **PloS one**, v. 11, n. 4, p. e0151470, 2016. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4838256/>. Acesso em: 18 jul. 2022.
- ALMEIDA, Romário César de. **Falta de profissionais de TI é prelúdio para a catástrofe iminente de Cibersegurança**. ANATI Notícias, 30 nov. 2021. Disponível em: <https://www.anati.org.br/noticias/falta-de-profissionais-de-ti-e-preludio-para-a-catastrofe-iminente-de-ciberseguranca>. Acesso em: 4 jun. 2022.
- ALTARES, Guillermo. **O russo que criou o antivírus do seu computador**. El País. 1 out. 2017. Disponível em: https://brasil.elpais.com/brasil/2017/09/28/internacional/1506610335_514029.html. Acesso em: 18 jul. 2022.
- AMB; FGV; IPESPE. **Estudo da Imagem do Judiciário Brasileiro**. Dez. 2019. Disponível em: https://www.amb.com.br/wp-content/uploads/2020/04/ESTUDO_DA_IMAGEM_DO_JUDICIARIO_BRASILEIRO_COMPLETO.pdf. Acesso em: 5 jun. 2022.
- AREFEEN, Mohammed Shamsul; SCHILLER, Michael. Continuous Integration Using Gitlab. **Undergraduate Research in Natural and Clinical Science and Technology Journal**, v. 3, p. 1-6, 2019.
- ASIMOV, Isaac. **Eu, robô**. Tradução Aline Storto Pereira. 1. ed. São Paulo: Aleph, 2014. Título original: I, robot.
- ASSOCIATED PRESS. **John Lewis é o primeiro parlamentar negro a ser velado na Rotunda do Capitólio, em Washington**. 27 jul. 2020. Disponível em: <https://g1.globo.com/mundo/noticia/2020/07/27/john-lewis-e-o-primeiro-parlamentar-negro-a-ser-velado-na-rotunda-do-capitolio-em-washington.ghtml>. Acesso em: 15 jul. 2022.

BASS, Dina. **Amazon schooled on AI facial technology by Turing Award winner.** Bloomberg. 3 abr. 2019. Disponível em: <https://www.bloomberg.com/news/articles/2019-04-03/amazon-schooled-on-ai-facial-technology-by-turing-award-winner#xj4y7vzkg>. Acesso em: 15 jul. 2022.

BBC. **Caso George Floyd: morte de homem negro filmado com policial branco com joelhos em seu pescoço causa indignação nos EUA.** 27 maio 2020. Disponível em: <https://g1.globo.com/mundo/noticia/2020/05/27/caso-george-floyd-morte-de-homem-negro-filmado-com-policial-branco-com-joelhos-em-seu-pescoco-causa-indignacao-nos-eua.ghtml>. Acesso em: 13 jul. 2022.

BBC. **Margaret Mitchell: Google fires AI ethics founder.** BBC News. 20 fev. 2021. Disponível em: <https://www.bbc.com/news/technology-56135817>. Acesso em: 13 jul. 2022.

BBC News. **SolarWinds Orion: more US government agencies hacked.** 15 dez. 2020. Disponível em: <https://www.bbc.com/news/technology-55318815>. Acesso em: 25 jun. 2022.

BECK, Ulrich. **Sociedade de risco: rumo a uma outra modernidade.** Tradução de Sebastião Nascimento. São Paulo: Editora 34, 2011 (2ª edição). 384 p.

BENDER, Emily M. et al. On the dangers of stochastic parrots: can language models be too big?  In: **Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.** 2021. p. 610-623. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3442188.3445922>. Acesso em: 13 jul. 2022.

BERG, Paul. Asilomar 1975: DNA modification secured. **Nature**, v. 455, n. 7211, p. 290-291, 2008. Disponível em: <https://www.nature.com/articles/455290a>. Acesso em: 16 nov. 2021.

BERG, Paul et al. Summary statement of the Asilomar conference on recombinant DNA molecules. **Proceedings of the National Academy of Sciences of the United States of America**, v. 72, n. 6, p. 1981, 1975. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC432675/>. Acesso em: 16 nov. 2021.

BERNAL, Daniel; HAGAN, Margaret. Redesigning Justice Innovation: a standardized methodology. **Stanford Journal of Civil Rights & Civil Liberties**, vol. 16, n. 2, p. 335-384, 2020. Hein Online.

BERTUZZI, Luca. **AI regulation filled with thousands of amendments in the European Parliament.** Euractiv, 2 jun. 2022. Disponível em: <https://www.euractiv.com/section/digital/news/ai-regulation-filled-with-thousands-of-amendments-in-the-european-parliament/>. Acesso em: 23 jun. 2022.

BERTUZZI, Luca. **Czech presidency sets out path for AI Act discussions.** Euractiv, 22 jun. 2022. Disponível em: <https://www.euractiv.com/section/digital/news/czech-presidency-sets-out-path-for-ai-act-discussions/>. Acesso em: 24 jun. 2022.

BIG BROTHER WATCH. **A closer look at Experian big data and artificial intelligence in Durham Police.** 6 abr. 2018. Disponível em: <https://bigbrotherwatch.org.uk/2018/04/a-closer-look-at-experian-big-data-and-artificial-intelligence-in-durham-police/>. Acesso em: 11 jun. 2022.

BLACK IN AI. **Black in AI.** Disponível em: <https://blackinai.github.io/#/about>. Acesso em: 13 jul. 2022.

BRASIL. **30 anos da Constituição.** Disponível em: <https://www.camara.leg.br/internet/agencia/infograficos-html5/constituinte/index.html>. Acesso em: 22 maio 2022.

BRASIL. **Anteprojeto de Lei de Proteção de Dados para segurança pública e persecução penal.** Disponível em: <https://www2.camara.leg.br/atividade-legislativa/comissoes/grupos-de-trabalho/56a-legislatura/comissao-de-juristas-dados-pessoais-seguranca-publica/documentos/outros-documentos/DADOSAnteprojetocomissaoprotecaodadossegurancapersecucaoFINAL.pdf>. Acesso em: 2 jul. 2022.

BRASIL. Ato do Presidente de 17/12/2020 – Câmara dos Deputados. Disponível em: <https://www2.camara.leg.br/atividade-legislativa/comissoes/grupos-de-trabalho/56a-legislatura/comissao-de-juristas-combate-ao-racismo-no-brasil/conheca-a-comissao/criacao-e-constitucionalizacao/ato-de-criacao>. Acesso em: 20 jun. 2022.

BRASIL. Ato do Presidente do Senado Federal nº 4, de 2022. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9087218&ts=1652881168283&disposition=inline>. Acesso em: 26 maio 2022.

BRASIL. CJSUBIA - Comissão de juristas responsável por subsidiar elaboração de substitutivo sobre inteligência artificial. Senado. Disponível em: <https://legis.senado.leg.br/comissoes/comissao?codcol=2504>. Acesso em: 24 jun. 2022.

BRASIL. Constituição da República Federativa do Brasil de 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 2 jan. 2022.

BRASIL. Consulta pública – Estratégia Brasileira de Inteligência Artificial. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-consulta-publica.pdf>. Acesso em: 27 maio 2022.

BRASIL. Decreto nº 4.377, de 13 de setembro de 2002. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/2002/d4377.htm. Acesso em: 21 jun. 2022.

BRASIL. Decreto nº 6.949, de 25 de agosto de 2009. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2007-2010/2009/decreto/d6949.htm. Acesso em: 21 jun. 2022.

BRASIL. Decreto nº 9.920, de 18 de julho de 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2019-2022/2019/decreto/D9920.htm. Acesso em: 2 jan. 2022.

BRASIL. Decreto nº 10.088, de 5 de novembro de 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/Ato2019-2022/2019/Decreto/D10088.htm#art5. Acesso em: 21 jun. 2022.

BRASIL. Decreto nº 10.109, de 7 de novembro de 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2019-2022/2019/decreto/D10109.htm#:~:text=DECRETO%20N%C2%BA%2010.109%2C%20DE%207,3%20de%20junho%20de%202015. Acesso em: 2 jan. 2022.

BRASIL. Decreto 10.932, de 10 de janeiro de 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/Ato2019-2022/2022/Decreto/D10932.htm. Acesso em: 21 jun. 2022.

BRASIL. Emenda Constitucional nº 49, de 8 de fevereiro de 2006. Disponível em: <https://www2.camara.leg.br/legin/fed/emecon/2006/emendaconstitucional-49-8-fevereiro-2006-540999-publicacaooriginal-42678-pl.html>. Acesso em: 22 maio 2022.

BRASIL. Emenda Constitucional nº 85, de 26 de fevereiro de 2015. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc85.htm#art1. Acesso em: 22 maio 2022.

BRASIL. Emenda Constitucional nº 115, de 10 de fevereiro de 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc115.htm. Acesso em: 2 jul. 2022.

BRASIL. Emenda Constitucional nº 117, de 5 de abril de 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc117.htm. Acesso em: 20 jun. 2022.

BRASIL. Emenda Constitucional nº 118, de 26 de abril de 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc118.htm. Acesso em: 22 maio 2022.

BRASIL. Estratégia Brasileira de Inteligência Artificial - EBIA. Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-diagramacao_4-979_2021.pdf. Acesso em: 27 maio 2022.

BRASIL. Inteligência Artificial. Ministério da Ciência, Tecnologia e Inovações. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligencia-artificial>. Acesso em: 27 maio. 2022.

BRASIL. Lei nº 8.213, de 24 de julho de 1991. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8213cons.htm. Acesso em: 20 jun. 2022.

BRASIL. Lei nº 8.906, de 4 de julho de 1994. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8906.htm. Acesso em: 15 jul. 2022.

BRASIL. Lei nº 9.100, de 29 de setembro de 1995. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9100.htm. Acesso em: 20 jun. 2022.

BRASIL. Lei nº 9.504, de 30 de setembro de 1997. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9504.htm. Acesso em: 20 jun. 2022.

BRASIL. Lei nº 10.406, de 10 de janeiro de 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm. Acesso em: 26 maio 2022.

BRASIL. Lei nº 12.288, de 20 de julho de 2010. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2007-2010/2010/lei/l12288.htm. Acesso em: 20 jun. 2022.

BRASIL. Lei nº 12.711, de 29 de agosto de 2012. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2011-2014/2012/lei/l12711.htm. Acesso em: 20 jun. 2022.

BRASIL. Lei nº 12.990, de 9 de junho de 2014. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2011-2014/2014/lei/l12990.htm. Acesso em: 20 jun. 2022.

BRASIL. Lei nº 13.105, de 16 de março de 2015. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2015-2018/2015/lei/l13105.htm. Acesso em: 9 jul. 2022.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/L13709.htm. Acesso em: 4 jun. 2022.

BRASIL. Lei nº 14.226, de 20 de outubro de 2021. Disponível em: <https://www.in.gov.br/en/web/dou/-/lei-n-14.226-de-20-de-outubro-de-2021-353673576>. Acesso em: 8 jul. 2022.

BRASIL. Lei Complementar nº 35, de 14 de março de 1979. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp35.htm. Acesso em: 3 jun. 2022.

BRASIL. Levantamento de Pessoal de TI. Tribunal de Contas da União; Relator Ministro Raimundo Carreiro. – Brasília: TCU, Secretaria de Fiscalização de Tecnologia da Informação (Sefti), 2015. 52 p. – (Sumário Executivo. Tecnologia da Informação). Disponível em: <https://portal.tcu.gov.br/biblioteca-digital/levantamento-de-pessoal-de-ti-sumario.htm>. Acesso em: 10 jun. 2022.

BRASIL. Levantamento nacional de informações penitenciárias – período de julho a dezembro de 2021. Disponível em: <https://www.gov.br/depes/pt-br/servicos/sisdepen/sisdepen>. Acesso em: 15 jun. 2022.

BRASIL. MCTI tem reunião sobre ciência e tecnologia com representantes do ICTP.Br. Ministério da Ciência, Tecnologia e Inovações. 12 maio 2022. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2022/05/mcti-tem-reuniao-sobre-ciencia-e-tecnologia-com-representantes-do-ictp.br>. Acesso em: 22 maio 2022.

BRASIL. **Nota à imprensa nº 170/2019.** Ministério das Relações Exteriores. Disponível em: https://www.gov.br/mre/pt-br/canais_atendimento/imprensa/notas-a-imprensa/2019/declaracao-de-osaka-dos-lideres-do-g20. Acesso em: 2 jan. 2022.

BRASIL. **Parecer de Plenário pelas Comissões de Ciência e Tecnologia, Comunicação e Informática, Trabalho, de Administração e Serviço Público e Constituição e Justiça e de Cidadania ao Projeto de Lei nº 21, de 2020.** Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarIntegra?codteor=2082139. Acesso em: 26 maio 2022.

BRASIL. **Portaria MCTI nº 4.167, de 6 de abril de 2021.** Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-portaria_mcti_4-617_2021.pdf. Acesso em: 27 maio 2022.

BRASIL. **Portaria MCTI nº 4.979, 13 de julho de 2021.** Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-portaria_mcti_4-979_2021_anexo1.pdf. Acesso em: 27 maio 2022.

BRASIL. **Projeto de Lei nº 21/2020.** Disponível em: <https://www.congressonacional.leg.br/materias/materias-bicamerais/-/ver/pl-21-2020>. Acesso em: 26 maio 2022.

BRASIL. **Projeto de Lei nº 21-A de 2020.** Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarIntegra?codteor=2129459. Acesso em: 15 jul. 2022.

BRASIL. **Projeto de lei nº 5.051, de 2019.** Disponível em: <https://www.congressonacional.leg.br/materias/materias-bicamerais/-/ver/pl-5051-2019>. Acesso em: 31 maio 2022.

BRASIL. **Relatório final – Comissão de juristas destinada a avaliar e propor estratégias normativas com vistas ao aperfeiçoamento da legislação de combate ao racismo estrutural e institucional no país.** 30 nov. 2021. Disponível em: <https://www2.camara.leg.br/atividade-legislativa/comissoes/grupos-de-trabalho/56a-legislatura/comissao-de-juristas-combate-ao-racismo-no-brasil/documentos/outros-documentos/relatorio-final>. Acesso em: 20 jun. 2022.

BRASIL. **Rodrigo Maia recebe anteprojeto para controle de dados de investigações criminais.** Câmara dos Deputados - Agência Câmara de Notícias. 5 nov. 2020. Disponível em: <https://www.camara.leg.br/noticias/705293-rodrigo-maia-recebe-anteprojeto-para-controle-de-dados-de-investigacoes-criminais/>. Acesso em: 2 jul. 2022.

BRASIL JR, Samuel Meira. Rules and Principles in Legal Reasoning. A Study of Vagueness and Collisions in Artificial Intelligence and Law. **Information & Communications Technology Law**, v. 10, n. 1, p. 67-77, 2001. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/13600830124910>. Acesso em: 31 maio 2022.

BASS, Dina. **Amazon schooled on AI facial technology by Turing Award winner.** Bloomberg. 3 abr. 2019. Disponível em: <https://www.bloomberg.com/news/articles/2019-04-03/amazon-schooled-on-ai-facial-technology-by-turing-award-winner#xj4y7vzkg>. Acesso em: 15 jul. 2022.

BRASSCOM. **Demanda de talentos em TIC e estratégia Σ TECEM 2021.** São Paulo, dez. 2021. Disponível em: <https://brasscom.org.br/demandade-talentos-em-tic-e-estrategia-%cf%83-tcem/>. Acesso em: 4 jun. 2022.

BUOLAMWINI, Joy Adowaa. **Gender shades: intersectional phenotypic and demographic evaluation of face datasets and gender classifiers.** 2017. Tese de Doutorado. Massachusetts Institute of Technology. Disponível em: <https://dspace.mit.edu/handle/1721.1/114068>. Acesso em: 12 jul. 2022.

BUOLAMWINI, Joy. **Response: racial and gender bias in Amazon Rekognition - commercial AI system for analyzing faces.** Medium. 25 jan. 2019. Disponível em: <https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a28922eeecd>. Acesso em: 16 jul. 2022.

BUOLAMWINI, Joy; GEBRU, Timnit Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In: **Conference on fairness, accountability and transparency**. PMLR, 2018. p. 77-91. Disponível em: http://proceedings.mlr.press/v81/buolamwini18a.html?mod=article_inline. Acesso em: 12 jul. 2022.

BURGESS, Matt. **UK police are using AI to inform custodial decisions – but it could be discriminating against the poor.** Wired. 1 mar. 2017. Disponível em: <https://www.wired.co.uk/article/police-ai-uk-durham-hart-checkpoint-algorithm-edit>. Acesso em: 11 jun. 2022.

CASTRO, Paulo Alexandre Batista de. Confiança pública e legitimidade do Poder Judiciário: a agenda de pesquisa pendente no Brasil. **Revista da CGU**, v. 11, p. 1284-1292, 2019. Disponível em: https://repositorio.cgu.gov.br/bitstream/1/44358/11/V11.N20_Confian%C3%A7a%20p%C3%ABblica.pdf. Acesso em: 5 jun. 2022.

CAVALCANTE, Pedro; MENDONÇA, Letícia; BRANDALISE, Isabella. Políticas Públicas e Design Thinking: interações para enfrentar desafios contemporâneos. In: CAVALCANTE, Pedro (Org.). **Inovação e Políticas Públicas: superando o mito da ideia**. Brasília: Ipea, 2019, p. 29-52. Disponível em: https://www.ipea.gov.br/portal/images/stories/PDFs/livros/livros/190910_livro_inovacao_e_politica_s_publicas_cap01.pdf. Acesso em: 18 jul. 2022.

CEPEJ. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente.** Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 10 jul. 2021.

CEPEJ. **Council of Europe European Commission for the efficiency of justice (CEPEJ).** Disponível em: <https://www.coe.int/en/web/cepej/home>. Acesso em: 5 jan. 2022.

CEPEJ. **Cyberjustice and artificial intelligence used in the field of justice.** Disponível em: <https://www.coe.int/en/web/cepej/cepej-working-group-cyber-just>. Acesso em: 6 jan. 2022.

CEPEJ. **European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment.** Disponível em: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>. Acesso em: 5 jan. 2022.

CEPEJ. **Possible introduction of a mechanism for certifying artificial intelligence tools and services in the sphere of justice and the judiciary: Feasibility Study.** Dez. 2020. Disponível em: <https://rm.coe.int/feasibility-study-en-cepej-2020-15/1680a0adf4>. Acesso em: 5 jan. 2022.

CEPEJ. **Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment.** Dez. 2021. Disponível em: <https://rm.coe.int/cepej-2021-16-en-revised-roadmap-follow-up-charter/1680a4cf2f>. Acesso em: 5 jan. 2022.

CIMPANU, Catalin. **EU: No evidence of Kaspersky spying despite 'confirmed malicious' classification.** ZDNet, 16 abr. 2019. Disponível em: <https://www.zdnet.com/article/eu-no-evidence-of-kaspersky-spying-despite-confirmed-malicious-classification/>. Acesso em: 18 jul. 2022.

CNJ. **Estatísticas do Poder Judiciário.** Disponível em: <https://painelestatistica.stg.cloud.cnj.jus.br/estatisticas.html>. Acesso em: 9 jul. 2022.

CNJ. **Guia da Estratégia Nacional de Tecnologia da Informação e Comunicação do Poder Judiciário ENTIC-JUD 2021-2026.** Versão 1.1. Ago. 2021. Disponível em: <https://atos.cnj.jus.br/files/compilado1841452021102661784be9efedd.pdf>. Acesso em: 10 jun. 2022.

CNJ. **Igualdade racial no judiciário - Relatório de atividade. Grupo de trabalho Políticas Judiciárias sobre a Igualdade Racial no âmbito do Poder Judiciário (Portaria n. 108, de 8/7/2020).** Out. 2020. Disponível em: https://www.cnj.jus.br/wp-content/uploads/2020/10/Relatorio_Igualdade-Racial_2020-10-02_v3-2.pdf. Acesso em: 20 jun. 2022, p. 14.

CNJ. **Justiça em Números 2020.** Brasília: CNJ, 2020. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2020/08/WEB-V3-Justi%C3%A7a-em-N%C3%BAmeros-2020-atualizado-em-25-08-2020.pdf>. Acesso em: 8 jul. 2022.

CNJ. **Justiça em Números 2021.** Brasília: CNJ, 2021. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2021/09/relatorio-justica-em-numeros2021-12.pdf>. Acesso em: 28 jun. 2022.

CNJ. **Padronizar portais no judiciário é facilitar acesso a informações da justiça.** 26 jan. 2022. Disponível em: <https://www.cnj.jus.br/padronizar-portais-no-judiciario-e-facilitar-acesso-a-informacoes-da-justica/>. Acesso em: 3 jul. 2022.

CNJ. **Painéis e publicações - Plataforma Sinapses/Inteligência Artificial.** Disponível em: <https://www.cnj.jus.br/sistemas/plataforma-sinapses/paineis-e-publicacoes>. Acesso em: 2 jul. 2022.

CNJ. **Painel dá transparência a projetos de inteligência artificial no Judiciário.** Agência CNJ de Notícias, 23 dez. 2020. Disponível em: <https://www.cnj.jus.br/painel-da-transparencia-a-projetos-de-inteligencia-artificial-no-judiciario/>. Acesso em: 8 jul. 2022.

CNJ. **Portaria nº 25, de 19 de fevereiro de 2019.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2829>. Acesso em: 16 jul. 2022.

CNJ. **Portaria nº 46, de 10 de fevereiro de 2022.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4366>. Acesso em: 5 jun. 2022.

CNJ. **Portaria n. 78, de 19 de maio de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3325>. Acesso em: 10 jul. 2021.

CNJ. **Portaria nº 112, de 11 de julho de 2013.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/1806>. Acesso em: 5 jun. 2022.

CNJ. **Portaria nº 170, de 20 de maio de 2022.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4549>. Acesso em: 20 jun. 2022.

CNJ. **Portaria n. 197, de 22 de novembro de 2019.** Disponível em: <https://atos.cnj.jus.br/files/original190722201911265ddd77ea7b46e.pdf>. Acesso em: 10 jul. 2021.

CNJ. **Portaria nº 242, de 10 de novembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3566>. Acesso em: 5 jun. 2022.

CNJ. **Portaria nº 271, de 4 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3613>. Acesso em: 4 jul. 2022.

CNJ. **Procedimento de Ato Normativo nº 0005432-29.2020.2.00.0000,** Plenário Virtual, 14 ago. 2020, votação unânime.

CNJ. **Projeto BRA/20/015 CNJ-PNUD.** Disponível em: https://www.cnj.jus.br/wp-content/uploads/2020/12/TCOT_051_2020-atualizado.pdf. Acesso em: 4 jul. 2022.

CNJ. **Projetos com Inteligência Artificial no Poder Judiciário.** Disponível em: <https://paineisanalytics.cnj.jus.br/single/?appid=29d710f7-8d8f-47be-8af8-a9152545b771&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&opt=ctxmenu,currsel>. Acesso em: 8 jul. 2022.

CNJ. **Protocolo para Julgamento com Perspectiva de Gênero** [recurso eletrônico] — Brasília: Conselho Nacional de Justiça – CNJ; Escola Nacional de Formação e Aperfeiçoamento de

Magistrados — Enfam, 2021. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2021/10/protocolo-18-10-2021-final.pdf>. Acesso em: 21 jun. 2022.

CNJ. **Recomendação nº 128, de 15 de fevereiro de 2022.** Disponível em: <https://atos.cnj.jus.br/files/original18063720220217620e8ead8fae2.pdf>. Acesso em: 21 jun. 2022.

CNJ. **Repositório de projetos do CNJ.** Disponível em: https://git.cnj.jus.br/users/sign_in. Acesso em: 4 jul. 2022.

CNJ. **Resolução nº 203, de 23 de junho de 2015.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2203>. Acesso em: 20 jun. 2022.

CNJ. **Resolução nº 230, de 22 de junho de 2016.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2301>. Acesso em: 21 jun. 2021.

CNJ. **Resolução nº 325, de 29 de junho de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3365>. Acesso em: 10 jun. 2022.

CNJ. **Resolução nº 331, de 20 de agosto de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3428>. Acesso em: 20 jun. 2022.

CNJ. **Resolução nº 332, de 21 de agosto de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 10 jul. 2021.

CNJ. **Resolução nº 335, de 29 de setembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3496>. Acesso em: 7 jul. 2022.

CNJ. **Resolução nº 336, de 29 de setembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3497>. Acesso em: 20 jun. 2022.

CNJ. **Resolução nº 343, de 9 de setembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3459>. Acesso em: 21 jun. 2020.

CNJ. **Resolução nº 345, de 9 de outubro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3512>. Acesso em: 7 jul. 2022.

CNJ. **Resolução nº 349, de 23 de outubro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3547>. Acesso em: 7 jul. 2022.

CNJ. **Resolução nº 350, de 27 de outubro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3556>. Acesso em: 7 jul. 2022.

CNJ. **Resolução nº 351, de 28 de outubro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3557>. Acesso em: 21 jun. 2022.

CNJ. **Resolução nº 360, de 17 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3633>. Acesso em: 5 jun. 2022.

CNJ. **Resolução nº 361, de 17 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3634>. Acesso em: 5 jun. 2022.

CNJ. **Resolução nº 362, de 17 de dezembro de 2020.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3635>. Acesso em: 5 jun. 2022.

CNJ. **Resolução nº 370, de 28 de janeiro de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3706>. Acesso em: 25 jun. 2022.

CNJ. **Resolução nº 372, de 12 de fevereiro de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3742>. Acesso em: 7 jul. 2022.

CNJ. **Resolução nº 385, de 6 de abril de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3843>. Acesso em: 7 jul. 2022.

CNJ. **Resolução nº 395, de 7 de junho de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3973>. Acesso em: 7 jul. 2022.

CNJ. **Resolução nº 396, de 7 de junho de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3975>. Acesso em: 5 jun. 2022.

CNJ. **Resolução nº 401, de 16 de junho de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3987>. Acesso em: 21 jun. 2022.

CNJ. **Resolução nº 423, de 5 de outubro de 2021.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4147>. Acesso em: 29 jun. 2022.

CNJ. **Resolução nº 446, de 14 de março de 2022.** Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4417>. Acesso em: 7 jul. 2022.

CNJ. **Resultados pesquisa IA no Poder Judiciário – 2022.** Disponível em: https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu.curssel&select=language.BR. Acesso em: 2 jul. 2022.

CNJ. Webinar **A Resolução CNJ n. 332: ética e governança para produção e uso da Inteligência Artificial no Poder Judiciário.** 04 set. 2020. Disponível em: <https://www.cnj.jus.br/agendas/webinar-a-resolucao-n-332-e-a-etica-na-producao-e-uso-da-i-a-no-poder-judiciario/>. Acesso em: 10 jul. 2021.

CÓBE, Raphael M. O. et al. Rumo a uma política de Estado para inteligência artificial. **Revista USP**, n. 124, p. 37-48, 2020. Disponível em: <https://www.revistas.usp.br/revusp/article/view/167914/159995>. Acesso em: 27 maio 2022.

COE. **About us.** Disponível em: <https://www.coe.int/en/web/about-us/who-we-are>. Acesso em: 5 jan. 2022.

COE. **Values.** Disponível em: <https://www.coe.int/en/web/about-us/values>. Acesso em: 5 jan. 2022.

COMISSÃO EUROPEIA. **Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União.** Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>. Acesso em: 26 maio 2022.

COMMON VULNERABILITIES AND EXPOSURES (CVE). **CVE® Program Mission.** Disponível em: <https://www.cve.org/>. Acesso em: 18 jul. 2022.

CONCERNED RESEARCHERS. **On recent research auditing commercial facial analysis technology.** Medium. 26 mar. 2019. Disponível em: <https://medium.com/@bu64dcjrytwitb8/on-recent-research-auditing-commercial-facial-analysis-technology-19148bda1832>. Acesso em: 16 jul. 2022.

COSTA, Eduarda; REIS, Carolina. **Histórico da LGPD Penal: o que foi feito até aqui e quais são os próximos passos?** Laboratório de Políticas Públicas e Internet. 16 abr. 2021. Disponível em: <https://lapin.org.br/2021/04/16/lgpd-penal-o-que-foi-feito-ate-aqui-e-quais-sao-os-proximos-passos/>. Acesso em: 2 jul. 2022.

CREESE, Sadie et al. **Future Series: cybersecurity, emerging technology and systemic risk.** In: World Economic Forum. 2020. Disponível em: https://www3.weforum.org/docs/WEF_Future_Series_Cybersecurity_emerging_technology_and_systemic_risk_2020.pdf. Acesso em: 7 jun. 2022.

CROWLEY, Kate; HEAD, Brian W. The enduring challenge of 'wicked problems': revisiting Rittel and Webber. **Policy Sciences**, v. 50, n. 4, p. 539-547, 2017. Disponível em: https://www.researchgate.net/publication/320883976_The_enduring_challenge_of'_wicked_problems'_revisiting_Rittel_and_Webber. Acesso em: 17 jul. 2022.

CSER. **Centre for the Study of Existential Risk**. Disponível em: <https://www.cser.ac.uk/>. Acesso em: 20 dez. 2021.

DAIR. **Distributed AI Research Institute**. Disponível em: <https://www.dair-institute.org/>. Acesso em: 13 jul. 2022.

DE BAEY, M E. The National Library of Medicine: evolution of a premier information center. **JAMA**, 1991. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/1870251/>. Acesso em: 12 jul. 2021.

DELOITTE. **Building successful AI teams**. 2021. Disponível em: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-ai-institute-building-successful-AI-teams.pdf>. Acesso em: 16 jun. 2022.

DESIGN RESEARCH SOCIETY. **About the DRS**. Disponível em: <https://www.designresearchsociety.org/cpages/about>. Acesso em: 17 jul. 2022.

DURHAM CONSTABULARY. **AI can predict reoffending, university study finds**. 24 jan. 2022. Disponível em: <https://durham.police.uk/News/News-Articles/2022/January/AI-can-predict-reoffending-university-study-finds.aspx>. Acesso em: 11 jun. 2022.

DUKE UNIVERSITY. **Cynthia Rudin - Interpretable ML Lab**. Disponível em: <https://users.cs.duke.edu/~cynthia/home.html>. Acesso em: 27 jun. 2022.

DUPUY, Jean-Pierre. A catástrofe de Chernobyl vinte anos depois. **Estudos Avançados**, [S. I.], v. 21, n. 59, p. 243-252, 2007. Disponível em: <https://www.revistas.usp.br/eav/article/view/10219>. Acesso em: 23 maio 2022.

EEOC. **Diversity in high tech**. U.S. Equal Employment Opportunity Comission. Disponível em: <https://www.eeoc.gov/special-report/diversity-high-tech>. Acesso em: 10 jun. 2022.

EMERON. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente**. Tradução livre de Teresa Germana Lopes de Azevedo. Disponível em: <https://emeron.tjro.jus.br/images/biblioteca/revistas/Avulso/CartaEticaEuropeia.pdf>. Acesso em: 31 maio 2022.

ENFAM. **Linhos de pesquisa**. Disponível em: <https://www.enfam.jus.br/programa-de-pos-graduacao-profissional-em-direito/mestrado-profissional/pesquisa/linhas-de-pesquisa/>. Acesso em: 6 jan. 2022.

ENFAM. **Sobre a Enfam**. Disponível em: <https://www.enfam.jus.br/institucional/sobre-a-escola/>. Acesso em: 2 jan. 2022.

EPIC. **About us**. Disponível em: <https://epic.org/about/>. Acesso em: 1 jul. 2022.

EPIC. **Brief of Amicus Curiae Electronic Privacy Information Center (EPIC) in support of appellant**. SJC-13197, 14 fev. 2022. Disponível em: <https://epic.org/wp-content/uploads/2022/02/EPIC-Amicus-Rodriguez-Parole-Board.pdf>. Acesso em: 1 jul. 2022.

EPIC. **Brief of appellant**. José Rodriguez v. Massachusetts Parole Board. Disponível em: https://epic.org/wp-content/uploads/2022/02/SJC-13197_01_Appellant_Rodriguez_Brief.pdf. Acesso em: 1 jul. 2022.

EPIC. **Liberty at risk: pre-trial risk assessment tools in the U.S. Set. 2020**. Disponível em: <https://epic.org/wp-content/uploads/2022/02/Liberty-At-Risk-Report-FALL-2020-UPDATE.pdf>. Acesso em: 2 jul. 2022.

EPIC. **Rodriguez v. Massachusetts Parole Board. SJC-13197 (2021)**. Disponível em: <https://epic.org/documents/rodriguez-v-massachusetts-parole-board/#epic-s-brief>. Acesso em: 1 jul. 2022.

EUROPEAN COMISSION. Orientações éticas para uma IA de confiança. Directorate-General for Communications Networks, Content and Technology. **Publications Office**, 2019. Disponível em: <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-pt/format-PDF>. Acesso em: 16 jul. 2022.

FAGONE, Jason. Exclusive: The Silicon Valley quest to preserve Stephen....: How a Silicon Valley team helped rebuild his distinctive robotic sound. **San Francisco Chronicle**, 18 mar. 2018. Disponível em: <https://www.sfgate.com/bayarea/article/The-Silicon-Valley-quest-to-preserve-Stephen-12759775.php>. Acesso em: 14 nov. 2021.

FARRIER, Ellie. **É seguro usar o Kaspersky em 2022?** Avast. 8 abr. 2022. Disponível em: <https://www.avast.com/pt-br/c-is-kaspersky-safe>. Acesso em: 18 jul. 2022.

FAST COMPANY. **The 10 most innovative companies in Artificial Intelligence.** 2021. Disponível em: <https://www.fastcompany.com/90600124/artificial-intelligence-most-innovative-companies-2021>. Acesso em: 13 jul. 2022.

FBSP. **Anuário Brasileiro de Segurança Pública.** 2022. p. 456. Disponível em: <https://forumseguranca.org.br/wp-content/uploads/2022/06/anuario-2022.pdf?v=4>. Acesso em: 1 jul. 2022.

FBSP. **A violência contra pessoas negras no Brasil 2021.** 19 nov. 2021. Disponível em: https://forumseguranca.org.br/publicacoes_posts/a-violencia-contra-pessoas-negras-no-brasil-2021/. Acesso em: 15 jun. 2022.

FDA. **Artificial Intelligence and Machine Learning (AI/ML) Software as a Medical Device Action Plan.** 12 jan. 2021. Disponível em: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>. Acesso em: 10 jul. 2021.

FDA. **MAUDE - Manufacturer and User Facility Device Experience.** Disponível em: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Acesso em: 10 jul. 2021.

FDA. **Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD): Discussion Paper and Request for Feedback.** Regulations.gov. 2 abr. 2019. Disponível em: <https://www.regulations.gov/document/FDA-2019-N-1185-0001>. Acesso em: 10 jul. 2021.

FHI. **Future of Humanity Institute.** Disponível em: <https://www.fhi.ox.ac.uk/>. Acesso em: 20 dez. 2021.

FINDLAW. **State v. Loomis.** Disponível em: <https://caselaw.findlaw.com/wi-supreme-court/1742124.html>. Acesso em: 26 jun. 2022.

FLI. **About us.** Disponível em: <https://futureoflife.org/team/>. Acesso em: 20 dez. 2021.

FLI. **AI Safety Research.** Disponível em: <https://futureoflife.org/ai-safety-research/>. Acesso em: 21 dez. 2021.

FLI. **A principled AI discussion in Asilomar.** Disponível em: <https://futureoflife.org/2017/01/17/principled-ai-discussion-asilomar/>. Acesso em: 20 dez. 2021.

FLI. **Asilomar AI Principles.** Disponível em: <https://futureoflife.org/2017/08/11/ai-principles/>. Acesso em: 20 dez. 2021.

FLI. **Beneficial AI 2017.** Future of Life Institute. Disponível em: <https://futureoflife.org/bai-2017/>. Acesso em: 20 dez. 2021.

FLI. **FLI Value Alignment Research Landscape**. Disponível em: <https://futureoflife.org/valuealignmentmap/>. Acesso em: 21 dez. 2021.

FLI. **Grant Programs**. Disponível em: <https://futureoflife.org/grant-programs/>. Acesso em: 21 dez. 2021.

FLI. **FLI Policy Work**. Disponível em: <https://futureoflife.org/policy-work/>. Acesso em: 21 dez. 2021.

FLI. **Future of Life Institute**. Disponível em: <https://futureoflife.org/>. Acesso em: 20 dez. 2021.

FLI. **Principles signatories**. Disponível em: <https://futureoflife.org/2017/01/11/principles-signatories/>. Acesso em: 21 dez. 2021.

FLI. **The AI Act**. Disponível em: <https://artificialintelligenceact.eu/>. Acesso em: 26 maio 2022.

FLORIDI, Luciano et al. **capAI: a procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act**. 23 mar. 2022. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4064091. Acesso em: 16 jul. 2022.

FONAJE. **Enunciados - Fórum Nacional de Juizados Especiais**. Disponível em: <https://fonaje.amb.com.br/enunciados/>. Acesso em: 9 jul. 2022.

FORTUNE. **50 world's greatest leaders**. 2019. Disponível em: <https://fortune.com/worlds-greatest-leaders/2019/joy-buolamwini/>. Acesso em: 13 jul. 2022.

FORTUNE. **50 world's greatest leaders**. 2021. Disponível em: <https://fortune.com/worlds-greatest-leaders/2021/timnit-gebru/>. Acesso em: 13 jul. 2022.

FRAGOSO, Roberto. **Comissão da inteligência artificial prorroga o prazo trabalhos por 120 dias**. Senado Notícias, 9 jun. 2022. <https://www12.senado.leg.br/noticias/audios/2022/06/comissao-da-inteligencia-artificial-prorroga-trabalhos-por-120-dias>. Acesso em: 26 maio. 2022.

FUNDAÇÃO GETÚLIO VARGAS (FGV). **Inteligência Artificial: Tecnologia aplicada à gestão de conflitos no âmbito do Poder Judiciário**. Coordenação Luis Felipe Salomão. 2020. Disponível em: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/estudos_e_pesquisas_ia_1afase.pdf. Acesso em: 18 jul. 2022.

FUNDAÇÃO GETÚLIO VARGAS (FGV). **Inteligência Artificial: Tecnologia aplicada à gestão de conflitos no âmbito do Poder Judiciário**. Coordenação Luis Felipe Salomão. 2ª edição. 2022. Disponível em: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/relatorio_ia_2fase.pdf. Acesso em: 18 jul. 2022.

FUNDAÇÃO GETÚLIO VARGAS (FGV). Publicações. **Centro de Inovação, Administração e Pesquisa do Judiciário**. 29 abr. 2022. Disponível em: <https://ciapj.fgv.br/publicacoes>. Acesso em: 18 jul. 2022.

G1. **Pioneiro na luta pelos direitos civis, John Lewis morre aos 80 anos nos EUA**. 18 jul. 2020. <https://g1.globo.com/mundo/noticia/2020/07/18/pioneiro-na-luta-pelos-direitos-civis-deputado-john-lewis-morre-nos-eua.ghtml>. Acesso em: 15 jul. 2022.

G20. **About the G20**. Disponível em: <https://g20.org/about-the-g20/>. Acesso em: 2 jan. 2022.

G20 Insights. **G20 AI Principles**. Disponível em: <https://www.g20-insights.org/wp-content/uploads/2019/07/G20-Japan-AI-Principles.pdf>. Acesso em: 2 jan. 2022.

GASPAR, Walter Britto; MENDONÇA, Yasmin Curzi de. A Inteligência Artificial no Brasil ainda precisa de uma estratégia. FGV Direito Rio: CTS: **Papers e textos para discussões**. 2021. Disponível em: <https://bibliotecadigital.fgv.br/dspace;handle/10438/30500>. Acesso em: 27 maio 2022.

GEBRU, Timnit et al. Datasheets for datasets. **Communications of the ACM**, v. 64, n. 12, p. 86-92, 2021. Disponível em: <https://dl.acm.org/doi/10.1145/3458723>. Acesso em: 13 jul. 2022.

GOLDIM, J.R. **Conferência de Asilomar**. Disponível em: <https://www.ufrgs.br/bioetica/asilomar.htm>. Acesso em: 16 nov. 2021.

GOOGLE. **2021 Diversity Annual Report**. Disponível em: https://static.googleusercontent.com/media/diversity.google/pt-BR//annual-report/static/pdfs/google_2021_diversity_annual_report.pdf?cachebust=2e13d07. Acesso em: 18 jun. 2022.

GOOGLE. **Model cards**. Disponível em: <https://modelcards.withgoogle.com/about>. Acesso em: 13 jul. 2022.

GOOGLE WALKOUT FOR REAL CHANGE. **Standing with Dr. Timnit Gebru** - #ISupportTimnit #BelieveBlackWomen. Medium. 3 dez. 2020. Disponível em: <https://googlewalkout.medium.com/standing-with-dr-timnit-gebru-isupporttimnit-believeblackwomen-6dadcd382>. Acesso em: 13 jul. 2022.

GOV.UK. **Risk potential assessment form**. Disponível em: <https://www.gov.uk/government/publications/risk-potential-assessment-form>. Acesso em: 11 jun. 2022.

GRANGEIA, Marcos Alaor. **Desafios éticos da inteligência artificial**. Judiciário Exponencial. Youtube, 15 set. 2020. Disponível em: <https://www.youtube.com/watch?v=HwLGZ69R6cM&t=12s>. Acesso em: 31 maio 2022.

GRCI. **Global Catastrophic Risk Institute**. Disponível em: <https://gcrinstitute.org/>. Acesso em: 20 dez. 2021.

HAO, Karen. **Congress wants answers from Google about Timnit Gebru's firing**. MIT Technology Review. 17 dez. 2020. Disponível em: <https://www.technologyreview.com/2020/12/17/1014994/congress-wants-answers-from-google-about-timnit-gebrus-firing/>. Acesso em: 13 jul. 2022.

HAO, Karen. **We read the paper that forced Timnit Gebru out of Google. Here's what it says**. MIT Technology Review. 4 dez. 2020. Disponível em: <https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/>. Acesso em: 13 jul. 2022.

HARVARD LAW REVIEW. State vs. Loomis. **Harvard Law Review**, v. 130, n. 5, p. 1530-7, 2017. Disponível em: <https://harvardlawreview.org/2017/03/state-v-loomis/>. Acesso em: 26 jun. 2022.

HAWKING. **Biography**. Disponível em: <https://www.hawking.org.uk/biography>. Acesso em: 14 nov. 2021.

HAWKING, Stephen. **Breves respostas para grandes questões**. Rio de Janeiro: Editora Intrínseca, 2018. p. 94.

HORVITZ, Eric; SELMAN, Bart. Interim Report from the Panel Chairs: AAAI Presidential Panel on Long-Term AI Futures. In: **Singularity Hypotheses**. Springer, Berlin, Heidelberg, 2012. p. 301-308. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-32560-1_15. Acesso em: 16 nov. 2021.

IBM. **IBM CEO's Letter to Congress on Racial Justice Reform**. 8 jun. 2020. Disponível em: <https://www.ibm.com/policy/facial-recognition-sunset-racial-justice-reforms/>. Acesso em 13 jul. 2022.

IBM. **X-Force Threat Intelligence Index 2022**. IBM Security. Disponível em: <https://www.ibm.com/downloads/cas/ADLMLAZ>. Acesso em: 4 jun. 2022.

IEEE. **History of IEEE**. Disponível em: <https://www.ieee.org/about/ieee-history.html>. Acesso em: 6 jan. 2022.

IEEE. **Mission & Vision**. Disponível em: <https://www.ieee.org/about/vision-mission.html>. Acesso em: 6 jan. 2022.

IEEE. **The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems**. Disponível em: <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>. Acesso em: 6 jan. 2021.

IEEE. **The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems Key Information, Milestones, and FAQs about The Initiative**. Disponível em: <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/faqs/gieais-faq-11.22.2020.pdf>. Acesso em: 6 jan. 2022.

INSTITUTO INNOVARE. **Premiada CNJ/Tecnologia: Iniciativa do TRF 3ª Região que facilita produção de minutas é premiada pelo Innovare e CNJ**. 21 out. 2021. Disponível em: <https://www.premioinnovare.com.br/noticias/premiada-cnjtecnologia:-iniciativa-do-trf-3a.-regiao-que-facilita-producao-de-minutas-e-premiada-pelo-innovare-e-cnj/108>. Acesso em: 16 jul. 2022.

ISA. **Books of the XX century**. Disponível em: <https://www.isa-sociology.org/en/about-isa/history-of-isa/books-of-the-xx-century>. Acesso em: 23 maio 2022.

ISACA. **About us**. Disponível em: <https://www.isaca.org/why-isaca/about-us>. Acesso em: 24 jun. 2022.

ISACA. **Cyberrisk quantification**. 2021. Disponível em: <https://store.isaca.org/s/store#/store/browse/detail/a2S4w000004Ko4jEAC>. Acesso em: 24 jun. 2022.

(ISC)². **(ISC)² Cybersecurity workforce study 2021**. Disponível em: <https://www.isc2.org/-/media/ISC2/Research/2021/ISC2-Cybersecurity-Workforce-Study-2021.ashx>. Acesso em: 1 jul. 2022.

(ISC)². **(ISC)²: The world's leading cybersecurity professional organization**. Disponível em: <https://www.isc2.org/about>. Acesso em: 4 jun. 2022.

JABBARI, Ramtin et al. What is DevOps? A systematic mapping study on definitions and practices. In: **Proceedings of the Scientific Workshop Proceedings of XP2016**. 2016. p. 1-11.

JACOB, Snow. **Amazon's face recognition falsely matched 28 members of Congress with mugshots**. ACLU. 26 jul. 2018. Disponível em: <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>. Acesso em: 15 jul. 2022.

JONAS, Hans. **O princípio responsabilidade: ensaio de uma ética para a civilização tecnológica**. Tradução de Marijane Lisboa, Luiz Barros Montez. Rio de Janeiro: Contraponto: Ed. PUC-Rio, 2006.

JUSTEN FILHO, Marçal. **Curso de Direito Administrativo**. 10. ed. São Paulo: Revista dos Tribunais, 2014, p. 1.335.

KAHNEMAN, Daniel et al. **Ruído: uma falha no julgamento humano**. Trad. Cássio de Arantes Leite. – 1^a ed. – Rio de Janeiro: Objetiva, 2021.

KASPERSKY, Eugene. **Aprenda a usar as regras YARA – como prever císnos negros**. Kaspersky Daily. 3 set. 2020. Disponível em: <https://www.kaspersky.com.br/blog/cybersecurity-expert-training/15992/>. Acesso em: 18 jul. 2022.

KEHL, Danielle Leah; KESSLER, Samuel Ari. Algorithms in the criminal justice system: Assessing the use of risk assessments in sentencing. Responsive Communities Initiative, Berkman Klein Center for Internet & Society, **Harvard Law School**. 2017. Disponível em: <https://dash.harvard.edu/handle/1/33746041>. Acesso em: 7 jun. 2022.

KEMPF, Hervé. **Asilomar: a ciência aos cientistas.** Transgênicos, terapia genética e células-tronco: questões para a ciência e para a sociedade. Brasília: Núcleo de Estudos Agrários e Desenvolvimento Rural, Instituto Interamericano de Cooperação para a Agricultura, 2004. Disponível em: http://www.ensp.fiocruz.br/portal-ensp/uploads/documentos-pessoais/documento-pessoal_11531.pdf. Acesso em: 16 nov. 2021.

KIRKPATRICK, Keith. It's not the algorithm, it's the data. **Communications of the ACM**, v. 60, n. 2, p. 21-23, 2017. Disponível em: <https://cacm.acm.org/magazines/2017/2/212422-its-not-the-algorithm-its-the-data/fulltext>. Acesso em: 26 jun. 2022.

KURZWEIL, Ray. **A singularidade está próxima: quando os humanos transcendem a biologia.** Tradução Ana Goldberger. Itaú Cultural, 2019. Título original: The singularity is near.

KUSTERS, Remy et al. Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities. **Frontiers in Big Data**, v. 3, p. 45, 2020. Disponível em: <https://www.frontiersin.org/articles/10.3389/fdata.2020.577974/full>. Acesso em: 11 out. 2021.

LANZELLOTTI, Andressa; GUEDES, Maria Rita. A Inteligência Artificial no Judiciário. **AMB Informa – Especial VII ENAJE**, p. 15. Disponível em: https://www.amb.com.br/wp-content/uploads/2019/06/Especial_VII_Enaje_PARA_PROVA-1-1.pdf. Acesso em: 31 maio 2022.

LARSON, Jeff et al. **How we analyzed the COMPAS recidivism algorithm.** ProPublica, 23 maio de 2016. Disponível em: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Acesso em: 26 jun. 2022.

LEARNED-MILLER, Erik et al. **Facial recognition technologies in the wild: a call for a federal office.** Algorithm Justice League/MacArthur Foundation. 29 maio 2020. Disponível em: https://assets.website-files.com/5e027ca188c99e3515b404b7/5ed1145952bc185203f3d009_FRTsFederalOfficeMay2020.pdf. Acesso em: 18 jul. 2022.

LEE, Nicol Turner; RESNICK, Paul; BARTON, Genie. **Algorithmic bias detection and mitigation: best practices and policies to reduce consumer harms.** Brookings. 22 maio 2019. Disponível em: <https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/>. Acesso em: 7 jun. 2022.

LEMOS, Ronaldo. **Estratégia de IA brasileira é patética.** Folha de S. Paulo. 11 abr. 2021. Disponível em: <https://www1.folha.uol.com.br/colunas/ronaldolemos/2021/04/estrategia-de-ia-brasileira-e-patetica.shtml>. Acesso em: 27 maio 2022.

LIAA-3R. **Manual de Diretrizes de Auditabilidade e Conformidade no Desenvolvimento e Testes de Soluções de IA no Âmbito do LIAA-3R/ Grupo de Validação Ético-Jurídica (GVEJ) do LIAA-3R, iLabTRF3, iJuspLab.** 2. ed., rev. e atual. São Paulo: LIAA-3R, 2022.

LIMA, Eduardo Jorge da Fonseca; ALMEIDA, Amalia Mapurunga; KFOURI, Renato de Ávila. Vacinas para COVID-19 - o estado da arte. **Revista Brasileira de Saúde Materno Infantil**, v. 21, p. 13-19, 2021. Disponível em: <https://doi.org/10.1590/1806-9304202100S100002>. Acesso em: 16 nov. 2021.

LINCOLN COLLEGE. **Abraham Lincoln's namesake college set to close after 157 years.** Disponível em: <https://lincolncollege.edu/home>. Acesso em: 25 jun. 2022.

LOPEZ, Teresa Ancona. **Princípio da precaução e evolução da responsabilidade civil.** São Paulo: Quartier Latin, 2010, p. 110.

MACHADO, Jorge; LINO, Daniel. Gerencialismo e justiça aberta. **Acervo - Revista do Arquivo Nacional**, v. 34, n. 3, p. 1-18, 2021. Disponível em: <https://www.brapci.inf.br/index.php/res/v/163601>. Acesso em: 20 jun. 2022.

MANDIA, Kevin. **FireEye shares details of recent cyber attack, actions to protect community.** FireEye Stories Blog, 8 dez. 2020. Disponível em: <https://www.fireeye.com/blog/products-and-services/2020/12/fireeye-shares-details-of-recent-cyber-attack-actions-to-protect-community.html>. Acesso em: 25 jun. 2022.

MARIANO JÚNIOR, Raul; MARTINS, Claudio R. N. **Parecer nº GVEJ.** Tribunal Regional Federal da 3ª Região. 17 dez. 2021. Disponível em: https://www.trf3.jus.br/documentos/adeq/Inova/LIAA-3R/PARECER_TECNICO/SEI_8354929_Parecer_N.I._GVEJ_1.pdf. Acesso em: 16 jul. 2022.

MARTINELLI, Julieta. **As nation changes stance on life sentences for minors, Tennessee's life juvenile lifers are left behind.** Latino Rebels. 6 maio 2022. Disponível em: <https://www.latinorebels.com/2022/05/06/tennesseejuvencelifers/>. Acesso em: 1 jul. 2022.

MCCARTHY, J.; MINSKY, M. L.; ROCHESTER, N. **A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.** 1955. Disponível em: <http://jmc.stanford.edu/articles/dartmouth.html>. Acesso em: 16 nov. 2021.

MCGREGOR, Sean. Preventing repeated real world AI failures by cataloging incidents: The AI incident database. In: **Proceedings of the AAAI Conference on Artificial Intelligence 2021.** p. 15458-15463. Disponível em: <https://ojs.aaai.org/index.php/AAAI/article/view/17817>. Acesso em: 18 jul. 2022.

MCGREGOR, Sean. **When AI systems fail: introducing the AI incident database.** 18 nov. 2020. Disponível em: <https://partnershiponai.org/aiincidentdatabase/>. Acesso em: 6 jan. 2022.

MELCHIOR, Stela Candioto; WAISSMANN, William. Regulação de dispositivos médicos: vigilância pós-mercado como estratégia de gerenciamento de riscos. Vigilância Sanitária em Debate: **Sociedade, Ciência & Tecnologia**, v. 7, n. 4, p. 67-76, 2019. Disponível em: <https://www.redalyc.org/journal/5705/570566202011/570566202011.pdf>. Acesso em: 18 jul. 2020.

MELO, Jairo Simão Santana; NEVES, Thiago Arruda; NETO, Celso Oliveira. AMON: controle de acesso do jurisdicionado no TJDFT a partir de técnicas de reconhecimento facial. **Sistema e-Revista CNJ**, v. 5, n. 1, p. 129-140, 2021. Disponível em: <https://www.cnj.jus.br/ojs/index.php/revista-cnj/article/view/157/110>. Acesso em: 10 jan. 2022.

MELLO, Vitor Tadeu Carramão. A repercussão geral e o writ of certiorari: breve diferenciação. **Revista da SJRJ**, Rio de Janeiro, n. 26, p. 139-146. Disponível em: <https://www.jfrj.jus.br/sites/default/files/revista-sjrj/arquivo/32-149-1-pb.pdf>. Acesso em: 26 jun. 2022.

MEYER, David. **IBM pulls out of facial recognition, fearing racial profiling and mass surveillance.** Fortune. 9 jun. 2020. Disponível em: <https://fortune.com/2020/06/09/george-floyd-ibm-exits-facial-recognition-bias-human-rights/>. Acesso em: 13 jul. 2022.

MICROSOFT. **Aether dimension.** Disponível em: <https://www.minecraft.net/pt-br/marketplace/pdp?id=bedf5379-7c46-4f26-8428-475179761d46>. Acesso em: 13 jul. 2022.

MIRI. **Machine Intelligence Research Institute.** Disponível em: <https://intelligence.org/>. Acesso em: 20 dez. 2021.

MIT MEDIA LAB. **Gender shades.** 2018. Disponível em: <http://gendershades.org/>. Acesso em: 12 jul. 2022.

MITCHELL, Margaret et al. Model cards for model reporting. In: **Proceedings of the conference on fairness, accountability, and transparency.** 2019. p. 220-229. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3287560.3287596>. Acesso em: 13 jul. 2022.

MORAES, Alexandre de. **Direito Constitucional.** 17 ed. São Paulo: Atlas, 2005. p. 30.

MOREIRA, Adilson José. **Tratado de Direito Antidiscriminatório** – São Paulo: Editora Contracorrente, 2020, p. 513 e 514.

MORLEY, Jessica et al. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. In: **Ethics, Governance, and Policies in Artificial Intelligence**. Springer, Cham, 2021. p. 153-183.

MOURA, Rafael Moraes; BORGES, Laryssa. **A impunidade dos hackers que colocaram o Judiciário de joelhos.** Veja, 27 mar. 2022. Disponível em: <https://veja.abril.com.br/politica/a-impunidade-dos-hackers-que-colocaram-o-judiciario-de-joelhos/>. Acesso em: 5 jun. 2022.

MUÑOZ, Antonio et al. P2ISE: preserving project integrity in CI/CD based on secure elements. **Information**, v. 12, n. 9, p. 357, 2021. Disponível em: <https://www.mdpi.com/2078-2489/12/9/357>. Acesso em: 18 jul. 2022.

NATIONAL TRANSPORTATION SAFETY BOARD (NTSB). **Aviation Accident Database & Synopses.** Disponível em: <https://www.ntsb.gov/Pages/AviationQuery.aspx>. Acesso em: 18 jul. 2022.

NETFLIX. **Coded bias.** 2020. Disponível em: <https://www.netflix.com/br/title/81328723>. Acesso em: 12 jul. 2022.

NEVES, Isabel Clara. Contribuição de Horst Rittel para a abordagem científica ao projecto no início da era computacional. **PARC Pesquisa em Arquitetura e Construção**, SP, v. 6, n. 1, p. 39–55, 2015. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/parc/article/view/8635025/4273>. Acesso em: 17 jul. 2022.

NLM. **PubMed.** Disponível em: <https://pubmed.ncbi.nlm.nih.gov/>. Acesso em: 12 jul. 2021.

NOBEL PRIZE. **The Nobel Prize in Chemistry 1980.** Disponível em: <https://www.nobelprize.org/prizes/chemistry/1980/summary/>. Acesso em: 16 nov. 2021.

NOGUEIRA, Pablo. Projeto de marco legal da IA no Brasil é pouco consistente e pode ser inútil, dizem especialistas. **Jornal da Unesp.** 29 jul 2021. Disponível em: <https://jornal.unesp.br/2021/07/29/projeto-de-marco-legal-da-ia-no-brasil-e-pouco-consistente-e-pode-ser-inutil-dizem-especialistas/>. Acesso em: 26 maio 2022.

OBSERVADOR. **Tudo o que Stephen Hawking disse à Web Summit.** Youtube, 14 mar. 2018. Disponível em: <https://www.youtube.com/watch?v=nVgoWh7Wang>. Acesso em: 14 nov. 2021.

OCDE. OECD Framework for the classification of AI systems. **OECD Digital Economy Papers**, n. 323, OECD Publishing: Paris, 2022. Disponível em: <https://doi.org/10.1787/cb6d9eca-en>. Acesso em: 15 jul. 2022.

OCDE. AI: Intelligent Machines, Smart Policies. Conference Summary. **OECD Digital Economy Papers n. 270.** Paris: OECD Publishing, 2018. Disponível em: <https://doi.org/10.1787/f1a650d9-en>. Acesso em: 02 jan. 2022.

OCDE. **A OCDE e o Brasil: uma relação mutuamente benéfica.** Disponível em: <https://www.oecd.org/latin-america/paises/brasil-portugues/>. Acesso em: 2 jan. 2022.

OCDE. Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. **OECD Legal Instruments.** 21 maio 2019. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 02 jan. 2022.

OCDE. Scoping the OECD AI sources: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO). **OECD Digital Economy Papers** n. 291. Paris: OECD Publishing, 2019. Disponível em: https://www.oecd-ilibrary.org/science-and-technology/scoping-the-oecd-ai-principles_d62f618a-en. Acesso em: 2 jan. 2022.

OCDE. State of implementation oh the OECD Principles: insights from national AI policies. **OECD Digital Economy Papers** n. 311. Paris: OECD Publishing, 2021. Disponível em: <https://www.oecd-ilibrary.org/science-and-technology/state-of-implementation-of-the-oecd-principles-979a-en>

https://library.org/science-and-technology/state-of-implementation-of-the-oecd-ai-principles_1cd40c44-en. Acesso em: 2 jan. 2022

OCDE. **Summary of the CDEP Technology Foresight Forum Economic And Social Implications Of Artificial Intelligence**. Paris, 10 fev. 2017. Disponível em: [https://www.oecd.org/digital/ieconomy/DSTI-CDEP\(2016\)17-ENG.pdf](https://www.oecd.org/digital/ieconomy/DSTI-CDEP(2016)17-ENG.pdf) . Acesso em: 02 jan. 2022.

OCDE. **Where: Global reach**. Disponível em: <https://www.oecd.org/about/members-and-partners/>. Acesso em: 2 jan. 2022.

OECDAI. **About OECD.AI**. Disponível em: <https://oecd.ai/en/about>. Acesso em: 2 jan. 2022.

OECDAI. **Compute & climate**. Disponível em: <https://oecd.ai/en/network-of-experts/working-group/1138>. Acesso em: 2 jan. 2022.

OECDAI. **List of participants in the OECD Expert Group on AI (AIGO)**. Disponível em: <https://oecd.ai/en/list-of-participants-oecd-expert-group-on-ai>. Acesso em: 2 jan. 2022.

OECDAI. **OCDE Network of Experts on AI (ONE AI)**. Disponível em: <https://oecd.ai/en/network-of-experts>. Acesso em: 2 jan. 2022.

O'CONNOR, Niamh. **Leading from the Centre: Web Summit's hall-of-famers**. Web Summit. 6 mar. 2019. Disponível em: <https://websummit.com/blog/hall-of-famers>. Acesso em: 14 nov. 2021.

ONU. **Chefe de agência atômica da ONU visita Chernobil 36 anos após desastre nuclear**. ONU News. 26 abr. 2022. Disponível em: <https://news.un.org/pt/story/2022/04/1787232>. Acesso em: 22 maio 2022.

ONU. **Comentários aos Princípios de Bangalore de Conduta Judicial/Escritório Contra Drogas e Crime**. Tradução Marlon da Silva Malha, Ariane Emílio Kloth. Brasília: Conselho da Justiça Federal, 2008. Tradução de: Commentary on The Bangalore Principles of Judicial Conduct. Disponível em: https://www.unodc.org/documents/lpo-brazil/Topics_corruption/Publicacoes/2008_Comentarios_aos_Principios_de_Bangalore.pdf. Acesso em: 29 jun. 2022.

ONU. **Os objetivos de desenvolvimento sustentável no Brasil**. Nações Unidas - Brasil. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 10 jun. 2022.

ONU. **Report of the Secretary General. Roadmap for Digital Cooperation**. Jun. 2020. Disponível em: https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf#page=32. Acesso em: 21 dez. 2021.

ONU News. **Organização Mundial da Saúde declara novo coronavírus uma pandemia**. Disponível em: <https://news.un.org/pt/story/2020/03/1706881#:~:text=A%20Organiza%C3%A7%C3%A3o%20Mundial%20da%20Sa%C3%BAde,pre-sente%20agora%20em%20114%20pa%C3%ADses>. Acesso em: 10 jul. 2021.

OSWALD, Marion et al. Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality. **Information & Communications Technology Law**, v. 27, n. 2, p. 223-250, 2018. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/13600834.2018.1458455>. Acesso em: 11 jun. 2022.

PALAZOLO, Gustavo; TEIXEIRA, Ialle; DUARTE, Felipe. **RansomEXX: Análise do Ransomware utilizado no ataque ao STJ**. Medium, 15 nov. 2020. Disponível em: <https://gustavopalazolo.medium.com/ransomexx-an%C3%A1lise-do-ransomware-utilizado-no-ataque-ao-stj-918001ec8195>. Acesso em: 4 jun. 2022.

PARTNERSHIP ON AI. **AI incident database**. Disponível em: <https://incidentdatabase.ai/>. Acesso em: 6 jan. 2022.

PEREIRA, João Carlos Murta; RODRIGUES, Marcus Vinícius Jardim. A plataforma Sinapses e a continuidade dos modelos de IA no judiciário. In: ARAÚJO, Valter Shuenquener de; GOMES, Marcus Lívio (Coord.); CANEN, Doris (Org.) **Inteligência Artificial e aplicabilidade prática no Direito**. CNJ. Brasília, 2022. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/04/inteligencia-artificial-e-a-aplicabilidade-pratica-web-2022-03-11.pdf>. Acesso em: Acesso em: 3 jun. 2022.

PEREIRA, S. Tavares. **Machine learning nas decisões: o uso jurídico dos algoritmos aprendizes**. 1ª edição. Santa Catarina/Florianópolis, 2021, p. 786-787.

PERRIGO, Billy. **Why Timnit Gebru isn't waiting for Big Tech to fix AI's problems**. Time. 18 jan. 2022. Disponível em: <https://time.com/6132399/timnit-gebru-ai-google/>. Acesso em: 13 jul. 2022.

PORTO, Fábio R. A “corrida maluca” da Inteligência Artificial no Poder Judiciário. In: ARAÚJO, Valter Shuenquener de; GOMES, Marcus Lívio (Coord.); CANEN, Doris (Org.) **Inteligência Artificial e aplicabilidade prática no Direito**. CNJ. Brasília, 2022. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2022/04/inteligencia-artificial-e-a-aplicabilidade-pratica-web-2022-03-11.pdf>. Acesso em: Acesso em: 18 jul. 2022.

PRADO, Eunice M B. **Responsabilidade civil por atos de inteligência artificial nas relações de consumo**: uma análise dos primeiros casos concretos de falhas em aplicações ao redor do mundo e dos rumos da legislação brasileira sobre o tema. Monografia (Especialização em Direito do Consumidor e Responsabilidade Civil) - Centro Universitário Mauricio de Nassau/ESMAPE. Recife, 2020.

PRADO, E. M. B.; MÜNCH, L. A.C.; VILLARROEL, M. A. C. U. “Sob controle do usuário”: formação dos juízes brasileiros para o uso ético da IA no judiciário. **Direito Público**, v. 18, n. 100, 2022. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/6021>. Acesso em: 28 jun. 2022.

PRETALAB. **Quem coda o Brasil?** - #quemcodab. PretaLab/ThoughtWorks. Disponível em: https://www.thoughtworks.com/content/dam/thoughtworks/documents/report/tw_report_quem_coda_br.pdf. Acesso em: 15 jun. 2022.

PRINCÍPIOS de **Yogyakarta**: princípios sobre a aplicação da legislação internacional de direitos humanos em relação à orientação sexual e identidade de gênero. Tradução Jones de Freitas. Disponível em: http://www.dhnet.org.br/direitos/sos/gays/principios_de_yogyakarta.pdf. Acesso em: 21 jun. 2022.

PUNKE, Michael. **Some thoughts on facial recognition legislation**. AWS Machine Learning Blog. 7 fev. 2019. Disponível em: <https://aws.amazon.com/pt/blogs/machine-learning/some-thoughts-on-facial-recognition-legislation/>. Acesso em: 16 jul. 2022.

PWC. **O abismo digital no Brasil**. PwC/Instituto Locomotiva. 2022. Disponível em: https://www.pwc.com.br/pt/estudos/preocupacoes-ceos/mais-temas/2022/O_Abismo_Digital.pdf. Acesso em: 15 jun. 2022.

PYBUS, Robert. **The uncanny mind that build Ethereum**. Wired. 13 jun. 2016. Disponível em: <https://www.wired.com/2016/06/the-uncanny-mind-that-built-ethereum/#.u88ip6pb1>. Acesso em: 21 dez. 2021.

QUACH, Katyanna. **We listened to more than 3 hours of US Congress testimony on facial recognition so you didn't have to go through it**. The Register. 22 maio 2019. https://www.theregister.com/2019/05/22/congress_facial_recognition/. Acesso em: 13 jul. 2022.

QUESTIONPRO. **Calculadora de amostra de pesquisa**. Disponível em: <https://www.questionpro.com/pt-br/mobile-diaries.html>. Acesso em: 28 jun. 2022.

RAINEY, Hal G. **Understanding and managing public organizations** - 5th Edition. San Francisco: Jossey-Bass, 2014, p. 9.

RAJI, Inioluwa Deborah; BUOLAMWINI, Joy. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In: **Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. 2019.** p. 429-435. Disponível em: <https://doi.org/10.1145/3306618.3314244>. Acesso em: 16 jul. 2022.

REINA, Eduardo. **Em 18 meses, hackers violaram sistemas de tribunais no Brasil a cada 41 dias.** Revista Consultor Jurídico, 15 abr. 2022. Disponível em: <https://www.conjur.com.br/2022-abr-15/onda-invasoes-hackers-estruturas-tecnologicas-tribunais>. Acesso em: 4 jun. 2022.

RESPONSIBLE AI COLLABORATIVE. **Founding report.** 28 mar. 2022. Disponível em: <https://docsend.com/view/47z2whznatd39vf9>. Acesso em: 18 jul. 2022.

REUTERS. **Europe's biggest tech conference to be held in-person in November.** 11 ago. 2021. Disponível em: <https://www.reuters.com/technology/europe-s-biggest-tech-conference-be-held-in-person-november-2021-08-11/>. Acesso em: 14 nov. 2021.

RITTEL, H.; WEBBER, M. Dilemmas in a general theory of planning. **Policy Sciences** 4, p. 155-169, 1973. Disponível em: <https://www.semanticscholar.org/paper/Dilemmas-in-a-general-theory-of-planning-Rittel-Webber/3667a56bd911445a6c0fc9447771d964e6831146>. Acesso em: 17 jul. 2022.

ROSA, Alexandre Morais da. **Inteligência artificial e Direito: ensinando um robô a julgar.** Consultor Jurídico. 4 set. 2020. Disponível em: <https://www.conjur.com.br/2020-set-04/limite-penal-inteligencia-artificial-direito-ensinando-robo-julgar>. Acesso em: 16 jul. 2022.

RUDIN, Cynthia. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. **Nature Machine Intelligence**, v. 1, n. 5, p. 206-215, 2019.

RUDIN, Cynthia et. al. Interpretable machine learning: fundamental principles and 10 grand challenges. **Statistics Surveys**, v. 16, p. 1-85, 2022.

RUDIN, Cynthia et. al. The age of secrecy and unfairness in recidivism prediction. **Harvard Data Science Review**, 2 (1), 3 2020.

SALVADOR, João Pedro Favaretto; GUIMARÃES, Tatiane. **O ataque ao STJ é mais um grito de socorro da segurança cibernética no Brasil.** FVG Direito. 9 nov. 2020. Disponível em: <https://portal.fvg.br/artigos/ataque-ao-stj-e-mais-grito-socorro-seguranca-cibernetica-brasil>. Acesso em: 4 jun. 2022.

SCHREIBER, Anderson. **PL da Inteligência Artificial cria fratura no ordenamento jurídico brasileiro.** Jota. 2 nov 2021. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/coluna-do-anderson-schreiber/pl-inteligencia-artificial-cria-fratura-no-ordenamento-juridico-02112021>. Acesso em: 26 maio 2022.

SCHWAB, Klaus. **A quarta revolução industrial.** Tradução Daniel Moreira Miranda. São Paulo: Edipro, 2016. Tradução de: *The Fourth Industrial Revolution*.

SCOTUS BLOG. **Loomis v. Wisconsin.** Disponível em: <https://www.scotusblog.com/case-files/cases/loomis-v-wisconsin/>. Acesso em: 26 jun. 2022.

SGARIONI, Mariana. **IA: especialistas alertam que proposta de regulação da UE não serve de modelo para o Brasil.** Mobile Time, 30 maio 2022. Disponível em: <https://www.mobiletme.com.br/noticias/30/05/2022/ia-especialistas-alertam-que-proposta-de-regulacao-da-ue-nao-serves-de-modelo-para-o-brasil/>. Acesso em: 24 jun. 2022.

SILVA, Tarcízio. **Audiência pública – inteligência artificial e riscos: vieses e discriminação.** 30 maio 2022. Disponível em: <https://tarciziosilva.com.br/blog/audiencia-publica-inteligencia-artificial-e-riscos-vieses-e-discriminacao/>. Acesso em: 10 jun. 2022.

SILVA, Tarcízio. **Linha do tempo do racismo algorítmico: casos, dados e reações.** 2019. Disponível em: <https://tarciziosilva.com.br/blog/destaques/posts/racismo-algoritmico-linha-do-tempo/>. Acesso em: 10 jun. 2022.

STF. **Ação Direta de Inconstitucionalidade nº 41 – Distrito Federal.** Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=13375729>. Acesso em: 20 jun. 2022.

STF. **Arguição de Descumprimento de Preceito Fundamental nº 186 – Distrito Federal.** Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=6984693>. Acesso em: 20 jun. 2022.

STJ. **Comunicado da presidência do STJ.** 19 nov. 2020. Disponível em: <https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/19112020-Comunicado-da-Presidencia-do-STJ.aspx>. Acesso em: 4 jun. 2020.

STJ. ENFAM/PPGPD. **Edital n. 1, de 08 de julho de 2020.** DJe/STJ, 9 jul. 2020. Disponível em: <https://bdjur.stj.jus.br/jspui/handle/2011/144345>. Acesso em: 10 jul. 2021.

STJ. **Sessão para escolha dos desembargadores do TRF6 será presencial, em 1º de agosto, com votação secreta.** 13 jun. 2022. Disponível em: <https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/13062022-Sessao-para-escolha-de-desembargadores-do-TRF6-sera-presencial-em-1-de-agosto-com-votacao-secreta.aspx>. Acesso em: 8 jul. 2022.

STRECK, Lenio Luiz. **Um robô pode julgar? Quem programa o robô?** Consultor Jurídico. 3 set. 2020. Disponível em: <https://www.conjur.com.br/2020-set-03/senso-incomum-robo-julgar-quem-programa-robo>. Acesso em: 16 jul. 2022.

TADEU, Erivelto. **Ataque ao STJ visava servidores de backup, nova tendência entre hackers.** CISO Advisor, 5 fev. 2021. Disponível em: <https://www.cisoadvisor.com.br/ataque-ao-stj-visou-servidores-de-backup-nova-tendencia-entre-hackers/>. Acesso em: 4 jun. 2022.

TALEB, Nassim Nicholas. **A lógica do cisne negro: o impacto do altamente improvável.** Editora Best Business, 2018.

TALEB, Nassim Nicholas. **Official research biography of Nassim Nicholas Taleb.** Disponível em: <https://www.fooledbyrandomness.com/CV.htm>. Acesso em: 18 jul. 2022.

TEGMARK, Max. **Stephen Hawking in Memoriam.** 14 mar. 2018. Disponível em: <https://futureoflife.org/2018/03/14/stephen-hawking-memoriam/>. Acesso em: 20 dez. 2021.

THALER, Richard H.; SUSTEIN, Cass R. **Nudge: como tomar melhores decisões sobre saúde, dinheiro e felicidade.** Tradução Ângelo Lessa. 1ª ed. Rio de Janeiro: Objetiva, 2019.

THE GUARDIAN. **US cybersecurity firm FireEye says it was hacked by foreign government.** Guardian staff and agencies, 9 dez. 2020. Disponível em: <https://www.theguardian.com/technology/2020/dec/08/fireeye-hack-cybersecurity-theft>. Acesso em: 25 jun. 2022.

TJDFT. **Diagnóstico organizacional de equidade e inclusão da diversidade.** Disponível em: https://www.tjdft.jus.br/institucional/governanca/integridade-e-etica/programa-pro-equidade-e-diversidade/relatorio_diversidade.pdf. Acesso em: 10 jun. 2022.

TJDFT. **Juiz do TJDFT vence Prêmio Desafio Lideranças Públicas Negras.** Disponível em: <https://www.tjdft.jus.br/institucional/imprensa/noticias/2021/fevereiro/juiz-do-tjdft-vence-premio-desafio-liderancas-publicas-negras>. Acesso em: 10 jun. 2022.

TJDFT. **Sistema de apresentação e reconhecimento facial - SAREF.** Disponível em: <https://saref.tjdft.jus.br/>. Acesso em: 15 jul. 2022.

TJRO. **Sinapses: termo de cooperação técnica com o CNJ completa 1 ano.** Disponível em: <https://www.tjro.jus.br/noticias/item/11953-sinapses-termo-de-cooperacao-tecnica-com-o-cnj-completa-1-ano>. Acesso em: 3 jun. 2022.

TJRS. **Iniciativas de IA no TJRS-2022.** Disponível em: <https://datastudio.google.com/reporting/2201f416-9182-4f8a-9638-1cf7d047aa3/page/6sHKC>. Acesso em: 3 jul. 2022.

TJRS. **Projetos com Inteligência Artificial no TJRS.** Disponível em: <https://www.tjrs.jus.br/novo/institucional/direcoes/ditic/projetos-com-inteligencia-artificial-no-tjrs/>. Acesso em: 3 jul. 2022.

TRF3. **Laboratório de Inteligência Artificial Aplicada LIAA-3R.** Disponível em: https://www.trf3.jus.br/adeg/governanca-e-estrategia/laboratorio-de-inovacao-do-trf3a-regiao/laboratorio-de-inteligencia-artificial-aplicada-liaa-3r?sword_list%5B0%5D=sigma&no_cache=1. Acesso em: 16 jul. 2022.

UK PARLIAMENT. **Preparing for extreme risks:** building a resilient society. Select Committee on Risk Assessment and Risk Planning. Report of Session 2021–22. HL Paper 110. 3 dez. 2021. Disponível em: <https://committees.parliament.uk/publications/8082/documents/83124/default/>. Acesso em: 11 jun. 2022.

UK PARLIAMENT. **Risk assessment and risk planning committee.** Disponível em: <https://committees.parliament.uk/committee/483/risk-assessment-and-risk-planning-committee/>. Acesso em: 11 jun. 2022.

U.S. FOOD & DRUG ADMINISTRATION (FDA). **About manufacturer and user facility device experience (MAUDE).** Disponível em: <https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities/about-manufacturer-and-user-facility-device-experience-maude>. Acesso em: 18 jul. 2022.

U.S. FOOD & DRUG ADMINISTRATION (FDA). **Artificial Intelligence and Machine Learning (AI/ML) – enabled medical devices.** Disponível em: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices#resources>. Acesso em: 18 jul. 2022.

U.S. FOOD & DRUG ADMINISTRATION (FDA). **MAUDE - Manufacturer and User Facility Device Experience.** Disponível em: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Acesso em: 18 jul. 2022.

U.S. FOOD & DRUG ADMINISTRATION (FDA). **MedWatch online voluntary reporting form.** Disponível em: <https://www.accessdata.fda.gov/scripts/medwatch/index.cfm>. Acesso em: 18 jul. 2022.

VIEIRA, Luciano J. M.; PINHEIRO, Ivan A. Contribuições do Conselho Nacional de Justiça para a gestão do Poder Judiciário. In: **XXXII Encontro da ANPAD**, Rio de Janeiro, 2008. Disponível em: http://www.anpad.org.br/abrir_pdf.php?e=ODcxNQ==. Acesso em: 11 jul. 2022.

WANG, Kathy. **Inside the GitLab public bug bounty program.** GitLab, 29 abr. 2019. Disponível em: <https://about.gitlab.com/blog/2019/04/29/inside-the-gitlab-public-bug-bounty-program/>. Acesso em: 18 jul. 2022.

WEF. **Global Cybersecurity Outlook 2022:** insight report. Jan. 2022. Disponível em: https://www3.weforum.org/docs/WEF_Global_Cybersecurity_Outlook_2022.pdf. Acesso em: 6 jun. 2022.

WEF. **The Global Risks Report 2022**, 17th edition. World Economic Forum. Disponível em: https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2022.pdf. Acesso em: 5 jun. 2022.

WOOD, Matt. **Thoughts on recent research paper and associated article on Amazon Rekognition.** AWS Machine Learning Blog. 26 jan. 2019. Disponível em: <https://aws.amazon.com/pt/blogs/machine-learning/thoughts-on-recent-research-paper-and-associated-article-on-amazon-rekognition/>. Acesso em: 16 jul. 2022.

WORDCLOUDS.COM. **Free online word cloud generator and tag cloud creator.** Disponível em: <https://www.wordclouds.com/>. Acesso em: 12 jul. 2022.

YARA. **Yara: the pattern matching swiss knife for malware researchers (and everyone else).** Disponível em: <https://virustotal.github.io/yara/>. Acesso em: 5 jun. 2022.

ZUCKERMAN, Ethan. **The narrative of public criticism.** MIT Media Lab. 7 fev. 2019. Disponível em: <https://civic.mit.edu/index.html%3Fp=2440.html>. Acesso em: 16 jul. 2022.

Este relatório foi gerado automaticamente pela plataforma OriginalMy.

Relatório requisitado por
EUNICE MARIA BATISTA PRADO

Data
03/10/2022

Hora
10:35

A plataforma OriginalMy assina digitalmente o hash do documento original e os endereços da(s) rede(s) blockchain onde o mesmo foi registrado.

Identidade Digital
1oriGMy27TK1YyrJuYxPvRparcf5HKuxs

Assinatura Digital
HPDQIwh5dadBcc5XvPg4N/MNFFucN8AtBiirR4r5KxpAgmTVBrgt9s/tFbiXoHXBC
aREiEVWLcpWdJukx5hR1HQ=

Conteúdo assinado digitalmente (codificado em base64)

```
LS0tLS0gaW5pY2lvIGFzc2luYXR1cmEgZGlnaXRhbCBkb2N1bWVudG8gKHNoYTI1N
ikgLS0tLS0KZjc1YWQ2NTA0NmY4YWI5M2ZhOTJkNzE0ZjE0YTUwOTY1ZTNiODAyMj
k1ODEyYzhmM2ZiOWRkYjRkNzg0YWFhNQotLS0tLSBmaW0gYXNzaW5hdHVyYSBkaWd
pdGfsIGRvY3VtZW50byAoc2hhMjU2KSAtLS0tLQotLS0tLSBpbmIjaW8gY29kaWdv
IGRhIHRyYW5zYWNhbyBibG9ja2NoYWIulC0tLS0tCkVUQyBUWCAweDA1ZGQxN2M1Z
WRkMGVkmGY5MDdmMTVjNTJkNmNjZjNjZGNjOTk5MmZkMzhjYWQxZjFkZTVkMDlyNj
ExZjZmMjYKRENSIFRYIDNkZWZhN2JmMmJIMDFkNWU1MWMzZGFkZWExNDIkZmVjZGM
4Y2YzM2M5MDMxMzdiZDZkYWZlZjY1ZDgyN2Q3ZGMKRENSIE1FUktMRSBmNzVhZDY1
MDQ2ZjhhYjkzZmE5MmQ3MTRmMTRhNTA5NjVIM2I4MDlyOTU4MTJjOGYzZml5ZGRIN
GQ3ODRhYWE1Ci0tLS0tIGZpbSBjb2RpZ28gZGEgdHJhbnNhY2FvIGJsb2NrY2hhaW
4gLS0tLS0=
```

Link para verificação

Esta assinatura digital pode ser verificada em [https://originalmy.com/verify?address=MW9yaUdNeT13VEsxWXlySnVZeFB2UnBhcmNmNUhLdXhz&signature=SFBEUWx3aDVkYWRCY2M1WHZQZzR0L010RkZ1Y044QXRCAwlyUjRyNUt4cEFnbVRWQnJndDlZ3RGYmlYb0hYQkNhUkVpRVZXTGNwV2RKdWt4NWhSMUhRPQ==&message=LS0tLS0gaW5pY2lvIGFzc2luYXR1cmEgZGlnaXRhbCBkb2N1bWVudG8gKHNoYTI1NikglS0tLS0KZjc1YWQ2NTA0NmY4YWI5M2ZhOTJkNzE0ZjE0YTUwOTY1ZTNiODAyMjk1ODEyYz hmM2ZiOWRkYjRkNzg0YWFhNQotLS0tLSBmaW0gYXNzaW5hdHVyYSBkaWdpdGFsIGRvY3VtZW50byAoc2hhMjU2KSAtLS0tLQotLS0tLSBpbmIjaW8gY29kaWdvIGRhIHRyYW5zYWNhbyBibG9ja2NoYWIulC0tLS0tCkVUQyBUWCAweDA1ZGQxN2M1ZWRkMGVkmGY5MDdmMTVjNTJkNmNjZjNjZGNjOTk5MmZkMzhjYWQxZjFkZTVkMDlyNjExZjZmMjYKRENSIFRYIDNkZWZhN2JmMmJIMDFkNWU1MWMzZGFkZWExNDIkZmVjZGM4Y2YzM2M5MDMxMzdiZDZkYWZlZjY1ZDgyN2Q3ZGMKRENSIE1FUktMRSBmNzVhZDY1MDQ2ZjhhYjkzZmE5MmQ3MTRmMTRhNTA5NjVIM2I4MDlyOTU4MTJjOGYzZml5ZGRINQ3ODRhYWE1Ci0tLS0tIGZpbSBjb2RpZ28gZGEgdHJhbnNhY2FvIGJsb2NrY2hhaW4gLS0tLS0="](https://originalmy.com/verify?address=MW9yaUdNeT13VEsxWXlySnVZeFB2UnBhcmNmNUhLdXhz&signature=SFBEUWx3aDVkYWRCY2M1WHZQZzR0L010RkZ1Y044QXRCAwlyUjRyNUt4cEFnbVRWQnJndDlZ3RGYmlYb0hYQkNhUkVpRVZXTGNwV2RKdWt4NWhSMUhRPQ==&message=LS0tLS0gaW5pY2lvIGFzc2luYXR1cmEgZGlnaXRhbCBkb2N1bWVudG8gKHNoYTI1NikglS0tLS0KZjc1YWQ2NTA0NmY4YWI5M2ZhOTJkNzE0ZjE0YTUwOTY1ZTNiODAyMjk1ODEyYz hmM2ZiOWRkYjRkNzg0YWFhNQotLS0tLSBmaW0gYXNzaW5hdHVyYSBkaWdpdGFsIGRvY3VtZW50byAoc2hhMjU2KSAtLS0tLQotLS0tLSBpbmIjaW8gY29kaWdvIGRhIHRyYW5zYWNhbyBibG9ja2NoYWIulC0tLS0tCkVUQyBUWCAweDA1ZGQxN2M1ZWRkMGVkmGY5MDdmMTVjNTJkNmNjZjNjZGNjOTk5MmZkMzhjYWQxZjFkZTVkMDlyNjExZjZmMjYKRENSIFRYIDNkZWZhN2JmMmJIMDFkNWU1MWMzZGFkZWExNDIkZmVjZGM4Y2YzM2M5MDMxMzdiZDZkYWZlZjY1ZDgyN2Q3ZGMKRENSIE1FUktMRSBmNzVhZDY1MDQ2ZjhhYjkzZmE5MmQ3MTRmMTRhNTA5NjVIM2I4MDlyOTU4MTJjOGYzZml5ZGRINQ3ODRhYWE1Ci0tLS0tIGZpbSBjb2RpZ28gZGEgdHJhbnNhY2FvIGJsb2NrY2hhaW4gLS0tLS0=)



**SUPERIOR TRIBUNAL DE JUSTIÇA
Programa de Pós-Graduação - ENFAM**

ATA

ESCOLA NACIONAL DE FORMAÇÃO E APERFEIÇOAMENTO DE MAGISTRADOS - ENFAM
PROGRAMA DE PÓS-GRADUAÇÃO PROFISSIONAL EM DIREITO
MESTRADO PROFISSIONAL EM DIREITO E PODER JUDICIÁRIO - COORDENAÇÃO ACADÊMICA

DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO

Em 22/08/2022, na sede da Enfam em Brasília/DF, às 10:00 horas, reuniu-se a Banca Examinadora, designada para avaliação do **Trabalho de Conclusão de Curso**, sob a forma de **dissertação**, da aluna **Eunice Maria Batista Prado** do curso de Mestrado Profissional em Direito e Poder Judiciário da **ENFAM – ESCOLA NACIONAL DE FORMAÇÃO E APERFEIÇOAMENTO DE MAGISTRADOS**, intitulada “**Banco de Incidentes de Inteligência Artificial do Judiciário**”.

RESULTADO

X	Aprovada com louvor
	Aprovada
	Reprovada

MEMBROS DA BANCA

Prof. Dr. Samuel Meira Brasil Jr.– Presidente (Orientador/PPGPD-ENFAM)
Prof. Dr. Richard Paulro Pae Kim– (UNISA)
Prof. Dr. Fabrício Castagna Lunardi– Membro (PPGPD-ENFAM)

Observações:

Após a defesa, o aluno terá o prazo de trinta dias para a entrega da versão definitiva do trabalho de conclusão do curso.

Em caso de não aprovação pela banca examinadora, o aluno terá o prazo assinalado pela Banca Examinadora para realizar as mudanças sugeridas e reencaminhar o trabalho para nova avaliação.

Este documento deverá ser assinado, imediatamente após a conclusão dos trabalhos, pelo presidente da Banca.



31/08/2022, às 12:28, conforme art. 1º, III, "b", da Lei 11.419/2006.



A autenticidade do documento pode ser conferida no site
[https://sei.stj.jus.br/sei/controlador_externo.php?
acao=documento_conferir&id_orgao_acesso_externo=0](https://sei.stj.jus.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0) informando o código verificador **3084912** e o
código CRC **E46093EC**.

006492/2022

3084912v5

Presidência

PORTARIA N° 197, DE 22 DE NOVEMBRO DE 2019.

Institui Grupo de Trabalho destinado à elaboração de estudos e propostas voltadas à ética na produção e uso da inteligência artificial no poder judiciário e dá outras providências.

O PRESIDENTE DO CONSELHO NACIONAL DE JUSTIÇA, no uso de suas atribuições legais e regimentais,

RESOLVE:

Art. 1º Fica instituído Grupo de Trabalho destinado à elaboração de estudos e propostas voltadas à ética na produção e uso da Inteligência Artificial no poder judiciário.

Art 2º Integram o Grupo de Trabalho:

- I – Rubens de Mendonça Canuto Neto, Conselheiro do Conselheiro Nacional de Justiça, responsável pela Coordenação;
- II – Richard Pae Kim, Secretário Especial de Programas, Pesquisas e Gestão Estratégica do Conselho Nacional de Justiça;
- II – Bráulio Gabriel Gusmão, Juiz Auxiliar da Presidência do Conselho Nacional de Justiça;
- III – Marcos Alaor Diniz Grangeia, Desembargador do Tribunal de Justiça do Estado de Rondônia;
- IV – Eunice Prado, Juíza de Direito do Tribunal de Justiça do Estado de Pernambuco;
- V – Camila Miranda de Moraes, Juíza do Trabalho do Tribunal Regional do Trabalho da 7ª Região;
- VI – Tarciso Dal Maso Jardim, Consultor Legislativo do Senado Federal do Brasil;
- VII - Fabiano Hartmann Peixoto, Professor Adjunto da Universidade de Brasília;
- VIII – Luciane Gomes, Assessora-Chefe do Gabinete do Conselheiro Rubens Canuto;

Art. 3º O Grupo de Trabalho encerrará suas atividades com a apresentação de relatório à Comissão Permanente de Tecnologia e da Informação e Inovação, no prazo máximo de noventa dias, a contar a partir da publicação desta Portaria.

Art. 4º Para os objetivos da presente Portaria, o Grupo de Trabalho poderá propor a realização de audiências públicas, consultas públicas, palestras ou seminários com representantes de órgãos públicos e de entidades da sociedade civil, além de especialistas e operadores do Direito e em Tecnologia da Informação, para colher subsídios.

Parágrafo único. O Grupo de Trabalho contará com o apoio do Departamento de Tecnologia da Informação e da Secretaria Especial de Programas, Pesquisas e Gestão Estratégica do CNJ.

Art. 5º Esta Portaria entra em vigor na data de sua publicação.

Ministro **DIAS TOFFOLI**

Presidente

Secretaria Geral

Secretaria Processual

PJE

INTIMAÇÃO

N. 0008023-32.2018.2.00.0000 - PEDIDO DE PROVIDÊNCIAS - A: JOSÉ JOÃO DA SILVA. Adv(s): Nao Consta Advogado. R: CONSELHO NACIONAL DE JUSTIÇA - CNJ. Adv(s): Nao Consta Advogado. Conselho Nacional de Justiça Autos: PEDIDO DE PROVIDÊNCIAS - 0008023-32.2018.2.00.0000 Requerente: JOSÉ JOÃO DA SILVA Requerido: CONSELHO NACIONAL DE JUSTIÇA - CNJ DECISÃO Cuida-se de pedido de providências formulado por JOSÉ JOÃO DA SILVA em desfavor de CONSELHO NACIONAL DE JUSTIÇA - CNJ, no qual alega a prática de supostas irregularidades no âmbito do Juízo da 20ª zona Eleitoral, com sede no Município de São João do Piauí - PI. Aduz o representante "(...) a existência de dolo no indeferimento de seu Registro de Candidatura (eleições 2016), porque fundamentado na ausência de filiação partidária, muito embora tenha instruído o pedido com certidão atestando sua condição de Presidente do Partido da Mobilização Nacional (PMN) no Município de Lagoa do Barro do Piauí" (Id. 3624799). A Corregedoria local arquivou o processo por entender que não houve falta funcional, ou inépcia profissional, in verbis: "Analisando detidamente os autos, verifica-se que, em 3 de setembro de 2016, foi proferida decisão judicial, transitada em julgado, onde foi indeferido o pedido de registro de candidatura do ora requerente ao cargo de Prefeito no Município de Lagoa do Barro do Piauí, nas Eleições 2016, consoante Registro de Candidatura n.º 173-95.2016.6.18.0020. Constatata-se a absoluta regularidade em seu trâmite, consentânea ao ordenamento jurídico vigente. Além disso, averigua-se que não houve sequer interposição de recurso contra



Poder Judiciário
Conselho Nacional de Justiça

RESOLUÇÃO Nº 332, DE 21 DE AGOSTO DE 2020.

Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências.

O PRESIDENTE DO CONSELHO NACIONAL DE JUSTIÇA, no uso de suas atribuições legais e regimentais;

CONSIDERANDO que a Inteligência Artificial, ao ser aplicada no Poder Judiciário, pode contribuir com a agilidade e coerência do processo de tomada de decisão;

CONSIDERANDO que, no desenvolvimento e na implantação da Inteligência Artificial, os tribunais deverão observar sua compatibilidade com os Direitos Fundamentais;

CONSIDERANDO que a Inteligência Artificial aplicada nos processos de tomada de decisão deve atender a critérios éticos de transparência, previsibilidade, possibilidade de auditoria e garantia de imparcialidade e justiça substancial;

CONSIDERANDO que as decisões judiciais apoiadas pela Inteligência Artificial devem preservar a igualdade, a não discriminação, a pluralidade, a solidariedade e o julgamento justo, com a viabilização de meios destinados a eliminar ou minimizar a opressão, a marginalização do ser humano e os erros de julgamento decorrentes de preconceitos;





Poder Judiciário
Conselho Nacional de Justiça

CONSIDERANDO que os dados utilizados no processo de aprendizado de máquina deverão ser provenientes de fontes seguras, preferencialmente governamentais, passíveis de serem rastreados e auditados;

CONSIDERANDO que, no seu processo de tratamento, os dados utilizados devem ser eficazmente protegidos contra riscos de destruição, modificação, extravio, acessos e transmissões não autorizadas;

CONSIDERANDO que o uso da Inteligência Artificial deve respeitar a privacidade dos usuários, cabendo-lhes ciência e controle sobre o uso de dados pessoais;

CONSIDERANDO que os dados coletados pela Inteligência Artificial devem ser utilizados de forma responsável para proteção do usuário;

CONSIDERANDO que a utilização da Inteligência Artificial deve se desenvolver com vistas à promoção da igualdade, da liberdade e da justiça, bem como para garantir e fomentar a dignidade humana;

CONSIDERANDO o contido na Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes;

CONSIDERANDO a ausência, no Brasil, de normas específicas quanto à governança e aos parâmetros éticos para o desenvolvimento e uso da Inteligência Artificial;

CONSIDERANDO as inúmeras iniciativas envolvendo Inteligência Artificial no âmbito do Poder Judiciário e a necessidade de observância de parâmetros para sua governança e desenvolvimento e uso éticos;





Poder Judiciário
Conselho Nacional de Justiça

CONSIDERANDO a competência do Conselho Nacional de Justiça para zelar pelo cumprimento dos princípios da administração pública no âmbito do Poder Judiciário, à exceção do Supremo Tribunal Federal, conforme art. 103- B, § 4º, II, da Constituição da República Federativa do Brasil;

CONSIDERANDO a decisão proferida pelo Plenário do Conselho Nacional de Justiça no julgamento do Procedimento de Ato Normativo nº 0005432-29.2020.2.00.0000, na 71ª Sessão Virtual, realizada em 14 de agosto de 2020;

RESOLVE:

CAPÍTULO I
DAS DISPOSIÇÕES GERAIS

Art. 1º O conhecimento associado à Inteligência Artificial e a sua implementação estarão à disposição da Justiça, no sentido de promover e aprofundar maior compreensão entre a lei e o agir humano, entre a liberdade e as instituições judiciais.

Art. 2º A Inteligência Artificial, no âmbito do Poder Judiciário, visa promover o bem-estar dos jurisdicionados e a prestação equitativa da jurisdição, bem como descobrir métodos e práticas que possibilitem a consecução desses objetivos.

Art. 3º Para o disposto nesta Resolução, considera-se:

I – Algoritmo: sequência finita de instruções executadas por um programa de computador, com o objetivo de processar informações para um fim específico;

II – Modelo de Inteligência Artificial: conjunto de dados e algoritmos computacionais, concebidos a partir de modelos matemáticos, cujo objetivo é oferecer resultados inteligentes, associados ou comparáveis a determinados aspectos do pensamento, do saber ou da atividade humana;





Poder Judiciário
Conselho Nacional de Justiça

III – Sinapses: solução computacional, mantida pelo Conselho Nacional de Justiça, com o objetivo de armazenar, testar, treinar, distribuir e auditar modelos de Inteligência Artificial;

IV – Usuário: pessoa que utiliza o sistema inteligente e que tem direito ao seu controle, conforme sua posição endógena ou exógena ao Poder Judiciário, pode ser um usuário interno ou um usuário externo;

V – Usuário interno: membro, servidor ou colaborador do Poder Judiciário que desenvolva ou utilize o sistema inteligente;

VI – Usuário externo: pessoa que, mesmo sem ser membro, servidor ou colaborador do Poder Judiciário, utiliza ou mantém qualquer espécie de contato com o sistema inteligente, notadamente jurisdicionados, advogados, defensores públicos, procuradores, membros do Ministério Público, peritos, assistentes técnicos, entre outros.

CAPÍTULO II DO RESPEITO AOS DIREITOS FUNDAMENTAIS

Art. 4º No desenvolvimento, na implantação e no uso da Inteligência Artificial, os tribunais observarão sua compatibilidade com os Direitos Fundamentais, especialmente aqueles previstos na Constituição ou em tratados de que a República Federativa do Brasil seja parte.

Art. 5º A utilização de modelos de Inteligência Artificial deve buscar garantir a segurança jurídica e colaborar para que o Poder Judiciário respeite a igualdade de tratamento aos casos absolutamente iguais.

Art. 6º Quando o desenvolvimento e treinamento de modelos de Inteligência exigir a utilização de dados, as amostras devem ser representativas e observar as cautelas necessárias quanto aos dados pessoais sensíveis e ao segredo de justiça.

Parágrafo único. Para fins desta Resolução, são dados pessoais sensíveis aqueles assim considerados pela Lei nº 13.709/2018, e seus atos regulamentares.





CAPÍTULO III DA NÃO DISCRIMINAÇÃO

Art. 7º As decisões judiciais apoiadas em ferramentas de Inteligência Artificial devem preservar a igualdade, a não discriminação, a pluralidade e a solidariedade, auxiliando no julgamento justo, com criação de condições que visem eliminar ou minimizar a opressão, a marginalização do ser humano e os erros de julgamento decorrentes de preconceitos.

§ 1º Antes de ser colocado em produção, o modelo de Inteligência Artificial deverá ser homologado de forma a identificar se preconceitos ou generalizações influenciaram seu desenvolvimento, acarretando tendências discriminatórias no seu funcionamento.

§ 2º Verificado viés discriminatório de qualquer natureza ou incompatibilidade do modelo de Inteligência Artificial com os princípios previstos nesta Resolução, deverão ser adotadas medidas corretivas.

§ 3º A impossibilidade de eliminação do viés discriminatório do modelo de Inteligência Artificial implicará na descontinuidade de sua utilização, com o consequente registro de seu projeto e as razões que levaram a tal decisão.

CAPÍTULO IV DA PUBLICIDADE E TRANSPARÊNCIA

Art. 8º Para os efeitos da presente Resolução, transparência consiste em:

I – divulgação responsável, considerando a sensibilidade própria dos dados judiciais;

II – indicação dos objetivos e resultados pretendidos pelo uso do modelo de Inteligência Artificial;

III – documentação dos riscos identificados e indicação dos instrumentos de segurança da informação e controle para seu enfrentamento;

IV – possibilidade de identificação do motivo em caso de dano causado pela ferramenta de Inteligência Artificial;





V – apresentação dos mecanismos de auditoria e certificação de boas práticas;

VI – fornecimento de explicação satisfatória e passível de auditoria por autoridade humana quanto a qualquer proposta de decisão apresentada pelo modelo de Inteligência Artificial, especialmente quando essa for de natureza judicial.

CAPÍTULO V DA GOVERNANÇA E DA QUALIDADE

Art. 9º Qualquer modelo de Inteligência Artificial que venha a ser adotado pelos órgãos do Poder Judiciário deverá observar as regras de governança de dados aplicáveis aos seus próprios sistemas computacionais, as Resoluções e as Recomendações do Conselho Nacional de Justiça, a Lei nº 13.709/2018, e o segredo de justiça.

Art. 10. Os órgãos do Poder Judiciário envolvidos em projeto de Inteligência Artificial deverão:

I – informar ao Conselho Nacional de Justiça a pesquisa, o desenvolvimento, a implantação ou o uso da Inteligência Artificial, bem como os respectivos objetivos e os resultados que se pretende alcançar;

II – promover esforços para atuação em modelo comunitário, com vedação a desenvolvimento paralelo quando a iniciativa possuir objetivos e resultados alcançados idênticos a modelo de Inteligência Artificial já existente ou com projeto em andamento;

III – depositar o modelo de Inteligência Artificial no Sinapses.

Art. 11. O Conselho Nacional de Justiça publicará, em área própria de seu sítio na rede mundial de computadores, a relação dos modelos de Inteligência Artificial desenvolvidos ou utilizados pelos órgãos do Poder Judiciário.

Art. 12. Os modelos de Inteligência Artificial desenvolvidos pelos órgãos do Poder Judiciário deverão possuir interface de programação de aplicativos (API) que permitam sua utilização por outros sistemas.





Parágrafo único. O Conselho Nacional de Justiça estabelecerá o padrão de interface de programação de aplicativos (API) mencionado no *caput* deste artigo.

CAPÍTULO VI DA SEGURANÇA

Art. 13. Os dados utilizados no processo de treinamento de modelos de Inteligência Artificial deverão ser provenientes de fontes seguras, preferencialmente governamentais.

Art. 14. O sistema deverá impedir que os dados recebidos sejam alterados antes de sua utilização nos treinamentos dos modelos, bem como seja mantida sua cópia (dataset) para cada versão de modelo desenvolvida.

Art. 15. Os dados utilizados no processo devem ser eficazmente protegidos contra os riscos de destruição, modificação, extravio ou acessos e transmissões não autorizados.

Art. 16. O armazenamento e a execução dos modelos de Inteligência Artificial deverão ocorrer em ambientes aderentes a padrões consolidados de segurança da informação.

CAPÍTULO VII DO CONTROLE DO USUÁRIO

Art. 17. O sistema inteligente deverá assegurar a autonomia dos usuários internos, com uso de modelos que:

- I – proporcione incremento, e não restrição;
- II – possibilite a revisão da proposta de decisão e dos dados utilizados para sua elaboração, sem que haja qualquer espécie de vinculação à solução apresentada pela Inteligência Artificial.





Art. 18. Os usuários externos devem ser informados, em linguagem clara e precisa, quanto à utilização de sistema inteligente nos serviços que lhes forem prestados.

Parágrafo único. A informação prevista no *caput* deve destacar o caráter não vinculante da proposta de solução apresentada pela Inteligência Artificial, a qual sempre é submetida à análise da autoridade competente.

Art. 19. Os sistemas computacionais que utilizem modelos de Inteligência Artificial como ferramenta auxiliar para a elaboração de decisão judicial observarão, como critério preponderante para definir a técnica utilizada, a explicação dos passos que conduziram ao resultado.

Parágrafo único. Os sistemas computacionais com atuação indicada no *caput* deste artigo deverão permitir a supervisão do magistrado competente.

CAPÍTULO VIII **DA PESQUISA, DO DESENVOLVIMENTO E DA IMPLANTAÇÃO DE** **SERVIÇOS DE INTELIGÊNCIA ARTIFICIAL**

Art. 20. A composição de equipes para pesquisa, desenvolvimento e implantação das soluções computacionais que se utilizem de Inteligência Artificial será orientada pela busca da diversidade em seu mais amplo espectro, incluindo gênero, raça, etnia, cor, orientação sexual, pessoas com deficiência, geração e demais características individuais.

§ 1º A participação representativa deverá existir em todas as etapas do processo, tais como planejamento, coleta e processamento de dados, construção, verificação, validação e implementação dos modelos, tanto nas áreas técnicas como negociais.

§ 2º A diversidade na participação prevista no *caput* deste artigo apenas será dispensada mediante decisão fundamentada, dentre outros motivos, pela ausência de profissionais no quadro de pessoal dos tribunais.

§ 3º As vagas destinadas à capacitação na área de Inteligência Artificial serão, sempre que possível, distribuídas com observância à diversidade.





Poder Judiciário
Conselho Nacional de Justiça

§ 4º A formação das equipes mencionadas no *caput* deverá considerar seu caráter interdisciplinar, incluindo profissionais de Tecnologia da Informação e de outras áreas cujo conhecimento científico possa contribuir para pesquisa, desenvolvimento ou implantação do sistema inteligente.

Art. 21. A realização de estudos, pesquisas, ensino e treinamentos de Inteligência Artificial deve ser livre de preconceitos, sendo vedado:

I – desrespeitar a dignidade e a liberdade de pessoas ou grupos envolvidos em seus trabalhos;

II – promover atividades que envolvam qualquer espécie de risco ou prejuízo aos seres humanos e à equidade das decisões;

III – subordinar investigações a sectarismo capaz de direcionar o curso da pesquisa ou seus resultados.

Art. 22. Iniciada pesquisa, desenvolvimento ou implantação de modelos de Inteligência Artificial, os tribunais deverão comunicar imediatamente ao Conselho Nacional de Justiça e velar por sua continuidade.

§ 1º As atividades descritas no *caput* deste artigo serão encerradas quando, mediante manifestação fundamentada, for reconhecida sua desconformidade com os preceitos éticos estabelecidos nesta Resolução ou em outros atos normativos aplicáveis ao Poder Judiciário e for inviável sua readequação.

§ 2º Não se enquadram no *caput* deste artigo a utilização de modelos de Inteligência Artificial que utilizem técnicas de reconhecimento facial, os quais exigirão prévia autorização do Conselho Nacional de Justiça para implementação.

Art. 23. A utilização de modelos de Inteligência Artificial em matéria penal não deve ser estimulada, sobretudo com relação à sugestão de modelos de decisões preditivas.

§ 1º Não se aplica o disposto no *caput* quando se tratar de utilização de soluções computacionais destinadas à automação e ao oferecimento de subsídios destinados ao cálculo de penas, prescrição, verificação de reincidência, mapeamentos, classificações e triagem dos autos para fins de gerenciamento de acervo.





§ 2º Os modelos de Inteligência Artificial destinados à verificação de reincidência penal não devem indicar conclusão mais prejudicial ao réu do que aquela a que o magistrado chegaria sem sua utilização.

Art. 24. Os modelos de Inteligência Artificial utilizarão preferencialmente software de código aberto que:

I – facilite sua integração ou interoperabilidade entre os sistemas utilizados pelos órgãos do Poder Judiciário;

II – possibilite um ambiente de desenvolvimento colaborativo;

III – permita maior transparência;

IV – proporcione cooperação entre outros segmentos e áreas do setor público e a sociedade civil.

CAPÍTULO IX DA PRESTAÇÃO DE CONTAS E DA RESPONSABILIZAÇÃO

Art. 25. Qualquer solução computacional do Poder Judiciário que utilizar modelos de Inteligência Artificial deverá assegurar total transparência na prestação de contas, com o fim de garantir o impacto positivo para os usuários finais e para a sociedade.

Parágrafo único. A prestação de contas compreenderá:

I – os nomes dos responsáveis pela execução das ações e pela prestação de contas;

II – os custos envolvidos na pesquisa, desenvolvimento, implantação, comunicação e treinamento;

III – a existência de ações de colaboração e cooperação entre os agentes do setor público ou desses com a iniciativa privada ou a sociedade civil;

IV – os resultados pretendidos e os que foram efetivamente alcançados;

V – a demonstração de efetiva publicidade quanto à natureza do serviço oferecido, técnicas utilizadas, desempenho do sistema e riscos de erros.





Poder Judiciário
Conselho Nacional de Justiça

Art. 26. O desenvolvimento ou a utilização de sistema inteligente em desconformidade aos princípios e regras estabelecidos nesta Resolução será objeto de apuração e, sendo o caso, punição dos responsáveis.

Art. 27. Os órgãos do Poder Judiciário informarão ao Conselho Nacional de Justiça todos os registros de eventos adversos no uso da Inteligência Artificial.

**CAPÍTULO X
DAS DISPOSIÇÕES FINAIS**

Art. 28. Os órgãos do Poder Judiciário poderão realizar cooperação técnica com outras instituições, públicas ou privadas, ou sociedade civil, para o desenvolvimento colaborativo de modelos de Inteligência Artificial, observadas as disposições contidas nesta Resolução, bem como a proteção dos dados que venham a ser utilizados.

Art. 29. As normas previstas nesta Resolução não excluem a aplicação de outras integrantes do ordenamento jurídico pátrio, inclusive por incorporação de tratado ou convenção internacional de que a República Federativa do Brasil seja parte.

Art. 30. As disposições desta Resolução aplicam-se inclusive aos projetos e modelos de Inteligência Artificial já em desenvolvimento ou implantados nos tribunais, respeitados os atos já aperfeiçoados.

Art. 31. Esta Resolução entra em vigor na data da sua publicação.

Ministro **DIAS TOFFOLI**

