



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

JOSÉ FAUSTINO MACÊDO DE SOUZA FERREIRA

**MUITO ALÉM DOS CÓDIGOS: uma abordagem ética da Inteligência Artificial no
Poder Judiciário**

Recife

2023

JOSÉ FAUSTINO MACÊDO DE SOUZA FERREIRA

**MUITO ALÉM DOS CÓDIGOS: uma abordagem ética da Inteligência Artificial no
Poder Judiciário**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de mestre em Ciência da Computação. Área de concentração: Computação Inteligente

Orientador (a): Prof. Dr. Adriano Lorena Inácio de Oliveira

Recife

2023

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

F383m Ferreira, José Faustino Macêdo de Souza
Muito além dos códigos: uma abordagem ética da inteligência artificial no poder judiciário / José Faustino Macêdo de Souza Ferreira. – 2023.
136 f.: il., fig., tab.

Orientador: Adriano Lorena Inácio de Oliveira.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2023.
Inclui referências e apêndices.

1. Inteligência artificial. 2. Ética. I. Oliveira, Adriano Lorena Inácio de (orientador). II. Título.

006.31

CDD (23. ed.)

UFPE - CCEN 2023-179

JOSÉ FAUSTINO MACÊDO DE SOUZA FERREIRA

**MUITO ALÉM DOS CÓDIGOS: uma abordagem ética da Inteligência Artificial no
Poder Judiciário**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Computação da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de mestre em Ciência da Computação. Área de concentração: Computação Inteligente

Aprovado em: 16/06/2023.

BANCA EXAMINADORA

Prof. Dr. Adriano Lorena Inácio de Oliveira (Orientador)
Universidade Federal de Pernambuco - UFPE

Prof. Dr. Geber Ramalho (Examinador Interno)
Universidade Federal de Pernambuco - UFPE

Prof. Dr. Samuel Meira Brasil (Examinador Externo)
Escola Nacional de Formação e Aperfeiçoamento de Magistrados - ENFAM

Dedico este trabalho à minha estrela Luísa.

AGRADECIMENTOS

A Deus, toda honra e glória, hoje e sempre.

A minha família, em especial aos meus pais Paulo e Glauceide, por serem esteio, presença e força.

A Maria, porque toda entrega e conquista minhas, são tuas.

As minhas filhas Nina, Luísa e Mariana, porque me enchem de amor e vontade de ser e fazer o melhor.

A Adriano Lorena, pela confiança, tranquilidade e generosidade com que me conduziu no CIn.

A Raphael D'Castro, pelo encorajamento e parceria, em nome de quem agradeço aos demais servidores do TJPE, que foram apoio, vibração e cooperação.

RESUMO

Com o desenvolvimento tecnológico, ferramentas de inteligência artificial têm sido cada vez mais empregadas na tomada de decisões. Os principais motivadores para o uso de ferramentas de IA pelos tribunais são aumentar a produtividade, buscar a inovação, melhorar a qualidade dos serviços judiciários e reduzir custos. Segundo levantamento do Conselho Nacional de Justiça, o Brasil possui uma situação de litigância única, com quase 80 milhões de processos judiciais em trâmite e ajuizamento anual em média de 30 milhões de novos casos, o que gera elevada taxa de congestionamento e morosidade. No mesmo cenário de alta demanda, o Programa das Nações Unidas para o Desenvolvimento (PNUD) identificou 111 projetos desenvolvidos ou em desenvolvimento nos tribunais nacionais que envolvam inteligência artificial. Partindo da preocupação de que a inteligência artificial, ao mesmo tempo em que traz benefícios substanciais para os indivíduos e para a sociedade, também apresenta erros, riscos e impactos negativos que podem ser de difícil antecipação, identificação e mensuração, o Conselho Nacional de Justiça, editou a Resolução nº 332/2020, que dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário. Embora relevante e comemorada a iniciativa, porque aponta para a preocupação do Judiciário com o desenvolvimento ético de IA em seus muros, o fato é que a norma não deixa claro como cada um dos servidores e magistrados, técnicos de informática, cientistas de dados e programadores devem se portar concretamente para que os produtos e serviços criados tenham conformidade ética. Conduzimos estudos exploratórios, através de entrevistas narrativas, que nos permitiram levantar hipóteses, compreender os sentidos e as percepções sobre o fenômeno do uso e desenvolvimento de IA pelos tribunais brasileiros, abordando questões de utilidade e adequação da tecnologia, benefícios e riscos, desafios concretos e abordagem ética. Os estudos constataram um vazio de referencial ético para uso e desenvolvimento concreto de IA pelo Judiciário, de modo que foram identificados, no cenário internacional, estândares éticos usualmente utilizados e, posteriormente, aplicada pesquisa com o objetivo de testar as hipóteses levantadas na entrevista narrativa e capturar, de forma mais ampla, a percepção dos tribunais nacionais quanto aos temas ética e IA pelo Judiciário, além de permitir classificação em escala de importância e pertencimento, sob a ótica do Judiciário, de cada um dos estândares identificados.

Palavras-chave: inteligência artificial; ética; judiciário.

ABSTRACT

With technological development, artificial intelligence tools are being increasingly used in decision-making. The main drivers for the use of AI tools by courts are to increase productivity, pursue innovation, improve the quality of court services and reduce costs. According to a survey by the National Council of Justice (Conselho Nacional de Justiça), Brazil has a unique litigation situation, with almost 80 million lawsuits in progress and an average of 30 million new cases being filed annually, which generates a high rate of congestion and delays. In the same high-demand scenario, the United Nations Development Program (UNDP) identified 111 projects developed or under development in national courts that involve the use of artificial intelligence. Based on the concern that artificial intelligence, while bringing substantial benefits to individuals and society, also presents errors, risks and negative impacts that can be difficult to anticipate, identify and measure, the National Council of Justice edited Resolution N. 332/2020, which regulates questions related to ethics, transparency and governance in the production and use of Artificial Intelligence in the Judiciary. Although the initiative is relevant and celebrated, because it points to the concern of the Judiciary with the ethical development of AI within its walls, the fact is that the norm does not make it clear how each of the servers and judges, IT technicians, data scientists and programmers must behave concretely so that the products and services created are ethically compliant. Henceforth, we conducted exploratory studies, through narrative interviews, which allowed us to raise hypothesis, understand the meanings and perceptions about the phenomenon of the use and development of AI by Brazilian courts, addressing issues of usefulness and adequacy of technology, benefits and risks, concrete challenges and ethical approach. The studies found a lack of ethical references for use and concrete development of AI by the Judiciary, so that ethical standards usually used were identified in the international scenario, and, later, research was applied with the objective of testing the hypothesis raised in the narrative interview and capture, in a broader way, the perception of national courts regarding ethics and AI issues by the Judiciary, in addition to allowing classification in scale of importance and belonging, from the perspective of the Judiciary, of each of the identified standards.

Keywords: artificial intelligence; ethics; judiciary.

LISTA DE ILUSTRAÇÕES

Figura 1 –	Incidência dos parâmetros éticos pesquisados	91
Gráfico 1 –	Qual órgão ou Instituição você integra?	95
Gráfico 2 –	Qual seu cargo/função?	96
Gráfico 3 –	Formação Acadêmica	97
Gráfico 4 –	Qual seu nível de conhecimento em IA?	98
Gráfico 5 –	Com qual frequência você utiliza ou desenvolve ferramentas de IA?	99
Gráfico 6 –	Seu Tribunal/órgão usou ou desenvolveu alguma ferramenta que utilize IA?	100
Gráfico 7 –	Qual seu nível de conhecimento sobre ética?	101
Gráfico 8 –	Você se sente apto/preparado a aplicar conceitos éticos no desempenho concreto do seu trabalho?	103
Gráfico 9 –	Na sua opinião, a IA é tecnologia aplicável, útil e adequada para resolução dos desafios enfrentados pelo Poder Judiciário?	104
Gráfico 10 –	A abordagem ética é importante para uso e desenvolvimento de ferramentas de IA pelo Poder Judiciário?	105
Gráfico 11 –	Avalie a importância de uma lista de verificação (checklist) para conferência de conformidade ética e legal dos modelos de IA utilizados pelo Judiciário	106
Gráfico 12 –	Critérios e parâmetros éticos	108

LISTA DE TABELAS

Tabela 1 –	Tabela de parâmetros éticos por documento analisado – parte 01	89
Tabela 2 –	Tabela de parâmetros éticos por documento analisado – parte 02	90
Tabela 3 –	Tabela de parâmetros éticos por documento analisado – parte 03	90
Tabela 4 –	Tabela de parâmetros éticos por documento analisado – parte 04	91

SUMÁRIO

1	INTRODUÇÃO	13
1.1	MOTIVAÇÃO E PROBLEMA DE PESQUISA	13
1.2	OBJETIVOS	16
1.3	METODOLOGIA	17
1.4	ESTRUTURA E ORGANIZAÇÃO DA DISSERTAÇÃO	18
2	FUNDAMENTOS	20
2.1	ÉTICA: CONCEITO, IMPACTO NO DESENVOLVIMENTO DE IA, NECESSIDADE DE ESTABELECIMENTO DE UM FRAMEWORK	20
2.2	IA – CONCEITOS GERAIS	23
3	ENTREVISTA NARRATIVA VIRTUAL – ERA UMA VEZ: IA E ÉTICA NO JUDICIÁRIO	27
3.1	CONTEXTO	27
3.2	ENTREVISTA NARRATIVA: ADEQUAÇÃO METODOLÓGICA	27
3.3	ESTRUTURA E ORGANIZAÇÃO	28
3.4	DESCRIÇÃO ENTREVISTA 01	29
3.5	DESCRIÇÃO ENTREVISTA 02	31
3.6	DESCRIÇÃO ENTREVISTA 03	33
3.7	SISTEMATIZAÇÃO E RESULTADOS DA ENTREVISTA NARRATIVA	35
4	ÉTICA E IA – MULTIFACETADA ABORDAGEM	39
4.1	INICIATIVAS GOVERNAMENTAIS E INTERGOVERNAMENTAIS	39
4.1.1	China	40
4.1.2	Reino Unido	43
4.1.3	Brasil	44
4.1.3.1	EBIA	44
4.1.3.2	Resolução 332 do CNJ	48
4.1.3.3	PL 21/2020	49
4.1.3.4	PL 2338/2023	52

4.1.4	Singapura	54
4.1.5	União Europeia	57
4.1.6	França	59
4.1.7	Índia	60
4.1.8	Estados Unidos da América	61
4.1.9	Austrália	62
4.1.10	OCDE – Organização para Cooperação e Desenvolvimento Econômico	63
4.1.11	Declaração Ministerial sobre Comércio e Economia Digital – Princípios para IA Centrada nos Humanos (2019)	65
4.2	INICIATIVAS MERCADO/INDÚSTRIA	66
4.2.1	Asilomar	66
4.2.2	Institute of Electrical and Electronics Engineers – IEEE	69
4.2.3	Associação para Maquinária da Computação (ACM)	70
4.2.4	Google	77
4.2.5	Microsoft	79
4.2.6	Diretrizes Universais para Inteligência Artificial (Public Voice Coalition, 2018)	79
4.3	INICIATIVAS ACADÊMICAS	81
4.3.1	Universidade de Montreal – Canadá	81
4.3.2	Princípios de Inteligência Artificial de Beijing	85
4.3.3	AI4PEOPLE - AI4PEOPLE’S Ethical Framework for a Good AI Society	86
4.4	QUADRO-RESUMO ÉTICA NA IA	88
5	SURVEY – IA E ÉTICA NO JUDICIÁRIO	93
5.1	SURVEY: ADEQUAÇÃO, METODOLOGIA, ORGANIZAÇÃO E CONDUÇÃO	93
5.2	SURVEY: RESULTADOS	95
5.2.1	Perfil do Entrevistado: Qual órgão ou Tribunal representa?	95
5.2.2	Perfil do Entrevistado: Qual cargo/função?	96
5.2.3	Perfil do Entrevistado: Formação Acadêmica	97
5.2.4	IA e eu: Nível de conhecimento em IA	98

5.2.5	IA e eu: Frequência do uso e desenvolvimento de ferramentas de IA	99
5.2.6	IA e eu: Frequência do uso e desenvolvimento de ferramentas de IA pelo órgão ou Tribunal	100
5.2.7	Ética e eu: Nível de conhecimento	101
5.2.8	Ética e eu: Aptidão e preparo para aplicar conceitos éticos no desempenho concreto do seu trabalho	102
5.2.9	IA e o Judiciário: Aplicação, utilidade e adequação para enfrentamento dos desafios atuais	103
5.2.10	IA e o Judiciário: Importância da abordagem ética para uso e desenvolvimento de ferramentas de IA pelo Judiciário	105
5.2.11	IA e o Judiciário: Importância de uma lista de verificação (checklist) para conferência de conformidade ética e legal nos modelos de IA pelo Judiciário	106
5.2.12	Critério e Parâmetros Éticos	107
6	CONCLUSÕES	112
6.1	FRAQUEZAS À VALIDADE, CONFIABILIDADE E FIDEDIGNIDADE	112
	REFERÊNCIAS	114
	APÊNDICE A – ENTREVISTA NARRATIVA	124
	APÊNDICE B – ENTREVISTA NARRATIVA	126

1 INTRODUÇÃO

1.1 MOTIVAÇÃO E PROBLEMA DE PESQUISA

Com o desenvolvimento tecnológico, ferramentas de inteligência computacional têm sido cada vez mais empregadas na tomada de decisões privadas.

Algoritmos de aprendizagem de máquina possibilitam a classificação de consumidores, a oferta de cartões de crédito, o valor de seguros e os escolhidos para vagas de emprego. A técnica de aprendizado profundo promete que computadores possam examinar milhares de imagens médicas para identificar padrões de doenças. Sistemas de recomendação sugerem não somente filmes e seriados adequados ao seu perfil de telespectador, mas possíveis amigos nas redes sociais, o que comprar e que rota escolher. Algoritmos prenunciam padrões relevantes para a indústria e o comércio, e até mesmo para questões associadas à saúde física e mental.

Vive-se desmedido entusiasmo pelo automatizado e por máquinas inteligentes, capazes de tomar decisões que envolvam severos riscos, como no campo dos veículos e armamentos autônomos.

Inobstante o uso massivo de tecnologia e inovação tragam diversos benefícios, a implementação de mecanismos de inteligência artificial apresenta diversos riscos. Vakkuri e Abrahamsson (2018) demonstram preocupação com a negligência de parâmetros éticos e legais na construção de sistemas inteligentes, concluindo pela necessidade de aprofundamento de estudos que expliquem como a ética de ser implementada na prática. Blight (2018) alerta as nefastas consequências, notadamente ligadas ao *machine bias*, isto é, reprodução de padrão discriminatório pela máquina.

Um sistema de reconhecimento facial criado pela Google identificou pessoas negras como gorilas; sistema de busca de contatos do aplicativo LinkedIn demonstrou uma preferência por nomes de pessoas do sexo masculino; Tay, mecanismo de IA lançado pela Microsoft para interagir com usuários do Twitter, passou a reproduzir mensagens xenofóbicas, racistas e antissemitas; o aplicativo de chat SimSimi, que utiliza inteligência artificial para conversar com os usuários, foi suspenso no Brasil por reproduzir ameaças, palavrões e conversas de teor sexual.

No setor público, a penetração de inteligência computacional também parece ser uma tendência. O Poder Judiciário, atento à realidade social que o circunda, vem utilizando mineração de dados para desenvolver soluções inteligentes.

Segundo Hartman e Zumblick (2019), três grandes campos da IA despertaram interesse na área do Direito, sendo elas o “raciocínio baseado em casos, sistemas especialistas e redes neurais”. A primeira delas utiliza os precedentes no raciocínio jurídico, enquanto a segunda área propõe a reprodução artificial da forma como o humano resolve problemas e, por fim, as redes neurais permitem que pontos de entrada sejam conectados aos pontos de saída por uma rede simulada. Salientam que, na advocacia, a IA tem especial atenção para apontar tendência, tratamento e análise de dados, auxílio em sistemas de planejamento jurídico, avaliação de cenários, bem como atividades conexas ao exercício profissional.

Bench-Capon, Prakken e Sartor (2009) registram que a IA também tem sido utilizada na argumentação jurídica, criando sistemas que podem armazenar interpretações conflitantes e

propor soluções alternativas para um caso com base nessas interpretações. Além disso, descrevem sistemas que podem usar precedentes legais para gerar argumentos, traçando analogias ou distinguindo precedentes, além de sistemas que podem argumentar por que uma regra não deve ser aplicada a um caso, mesmo que todas as suas condições sejam atendidas. Também reportam que há sistemas que podem atuar como mediadores entre as partes em litígio, estruturando e registrando os seus argumentos e respostas, como também sugerem mecanismos e táticas para a formação de argumentos jurídicos.

Já no Judiciário, identificam especial atenção da IA para uma Justiça Preditiva, para a concretização dos direitos, assim como para o papel instrumental de apoiar decisões judiciais, permitindo ao Poder Judiciário mais celeridade e qualidade na prestação jurisdicional, também como mecanismo de reprodução de atividades repetitivas e pouco estratégicas.

Segundo levantamento do Conselho Nacional de Justiça, China, Reino Unido, Singapura, Japão, Finlândia, Índia, Dinamarca, Canadá e Estados Unidos são exemplos de países que, de forma bem-sucedida, incorporaram tecnologia ao processo para as mais diversas finalidades. Para facilitar o acesso ao Poder Judiciário, por meio do peticionamento *on-line*, ou para entregar informação de qualidade às partes, diminuindo a assimetria informacional entre elas e, assim, ampliar a possibilidade de composição extrajudicial. Também está sendo utilizada para organizar o processo, reestruturando e racionalizando as suas etapas, tornando mais confortável a participação das partes com o uso de ferramentas como videoconferência. Mas não é só isso, extremamente promissor é o uso de IA para automatizar a execução de tarefas burocráticas, possibilitando diminuir o tempo empregado nessas atividades e liberar o valioso quadro de pessoal do Poder Judiciário para atividades em que são imprescindíveis a atuação humana. Mas não é só, esta tecnologia também se presta a dar suporte na análise de documentos legais, redação de decisões, pesquisa jurídica, previsão de resultados judiciais, detecção de fraudes e lavagem de dinheiro, mediação e resolução de disputas online e assistência virtual.

Se nesses países o emprego de tecnologia é muito precioso, com mais razão deve o Poder Judiciário brasileiro recorrer a ela. O Brasil possui uma situação de litigância única. Segundo dados do Conselho Nacional de Justiça, existem quase 80 milhões de processos judiciais em trâmite no Poder Judiciário brasileiro, com ajuizamento anual em média de 30 milhões de novos casos. Cada um dos 18.141 juízes brasileiros decide, em média, 1.877 processos por ano, o que corresponde a oito casos solucionados por dia útil. Esses números tornam o nosso Judiciário um dos mais produtivos do mundo. Entretanto, cada vez mais sentimos que isso não é o bastante.

Deste total de processos em tramitação, mais de 80% estão totalmente em meio eletrônico, ou seja, totalmente digital, o que torna possível considerar um horizonte favorável à utilização de IA em larga escala.

Em junho de 2022 o Programa das Nações Unidas para o Desenvolvimento (PNUD) apoiou o Conselho Nacional de Justiça (CNJ) em uma pesquisa sobre o uso da inteligência artificial (IA) pelos tribunais brasileiros. O mapeamento foi realizado no âmbito do Programa Justiça 4.0, iniciativa que tem o objetivo de acelerar a transformação digital do Poder Judiciário brasileiro.

A pesquisa identificou 111 projetos desenvolvidos ou em desenvolvimento nos tribunais. Com isso, o número de iniciativas cresceu 171% em relação ao levantamento realizado em 2021, quando foram registrados apenas 41 projetos.

Houve incremento também no número de órgãos com projetos de IA. Atualmente, 53 tribunais desenvolvem soluções com uso dessa tecnologia. Na pesquisa anterior, apenas 32 órgãos declararam ter iniciativas na área.

O estudo evidencia que mesmo os tribunais sem projetos nesse campo – em sua maioria tribunais do ramo eleitoral e do trabalho – já têm soluções implementadas ou sendo estudadas por seus Tribunais Superiores ou pelo respectivo conselho superior, o que implica serem também beneficiados por projetos nacionais.

Os principais motivadores para o uso de uma ferramenta de IA pelos tribunais são aumentar a produtividade, buscar a inovação, melhorar a qualidade dos serviços judiciários e reduzir custos.

Neste cenário em que se unem a necessidade e disponibilidade tecnológica com a possibilidade da urgente adoção de tecnologias como a inteligência artificial nos órgãos de Justiça, a utilização de inteligência computacional para apoiar o processo de tomada de decisão torna imperioso abordar, na criação de algoritmos inteligentes, questões de discriminação, representatividade, parcialidade, manipulação e preconceitos associados à máquina.

Floridi (2018) registra que:

“O uso de inteligência artificial no sistema de justiça cria oportunidades significativas para lidar com deficiências e falhas conhecidas na administração da justiça, mas também apresenta perigos sérios e únicos, não apenas para cidadãos individuais, mas para o ideal do estado de direito de forma mais geral. Qualquer tentativa de construir uma abordagem eticamente responsável para o desenvolvimento, implantação e monitoramento pós-implantação de IA no sistema judiciário e no setor de serviços jurídicos enfrenta, portanto, uma série de desafios específicos do domínio jurídico. Ao mesmo tempo, eventos recentes também evidenciaram que nem todos os aspectos das tradições jurídicas merecem ser preservados, e que muitos deles perpetuam profundas injustiças. A tecnologia pode desempenhar um papel positivo na resolução desses problemas.”

Nos Estados Unidos, o COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) – mecanismo utilizado nos EUA para avaliar o risco de reincidência dos acusados no país, usado para conceder fiança, sentenciar e determinar a liberdade condicional – revelou eminente conteúdo discriminatório, classificando erroneamente acusados negros como prováveis reincidentes, ao passo que brancos teriam menos propensão a delinquir.

Tais circunstâncias estão a revelar que, para além de desconformidade ética, é possível que os algoritmos violem normas jurídicas.

Esclarecendo que a primeira certeza é a de que o conhecimento em IA é estratégico, Hartmann (2019) acrescenta que, no que se refere à interseção entre IA e o Direito:

“A segunda certeza é a de que não há que se falar em robustez, solidez, confiança e competitividade sem se levar em conta a dimensão ética e a capacidade de impacto da IA no Direito. (...) A preocupação concreta com parâmetros éticos está no próprio fundamento de pesquisa e desenvolvimento de aplicações de IA para o Direito, especialmente as que destinam aos fluxos de gestão processual e o apoio à decisão, responsáveis pela segurança de concretização de direitos e mecanismos de reforço democrático e afirmação da cidadania”.

Muito se tem falado – e não somente nas ciências jurídicas e sociais – sobre a necessidade de estabelecimento de um framework ético para desenvolvimento de inteligência computacional.

Ximenes e Ramalho (2021) registram o crescente número de princípios, manuais e códigos de ética, emitidos pelos setores público, privado e terceiro setor, referentes a aplicações de IA. Segundo os autores, os princípios éticos de ML, que eram praticamente inexistentes em 2018, já em 2019 ultrapassavam 170 documentos, abrangendo milhares de páginas.

Em levantamento, Jobin, Ienca e Vayena (2019) pontuam que “vários estudos discutiram o tema da IA ética, notadamente em meta-avaliações ou em relação a riscos sistêmicos e consequências negativas não intencionais, como viés algorítmico ou discriminação”.

Partindo da preocupação de que a inteligência artificial, ao mesmo tempo em que traz benefícios substanciais para os indivíduos e para a sociedade, também apresenta erros, riscos e impactos negativos que podem ser de difícil antecipação, identificação e mensuração, a Comissão Europeia divulgou as Diretrizes Éticas para a Inteligência Artificial Confiável. Segundo o documento, a compreensão ética no desenvolvimento de soluções inteligentes passa por quatro princípios: (i) o respeito pela autonomia humana, (ii) prevenção de danos, (iii) a justiça e (iv) a explicabilidade.

Seguindo a mesma toada, o Poder Judiciário brasileiro, através do Conselho Nacional de Justiça, editou a Resolução nº 332/2020, que dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário.

Embora relevante e comemorada a iniciativa, porque aponta para a preocupação do Judiciário com o desenvolvimento ético de IA em seus muros, o fato é que a norma não deixa claro como cada um dos servidores e magistrados, técnicos de informática, cientistas de dados e programadores devem se portar concretamente para que os produtos e serviços criados tenham validação ético-legal.

Neste cenário, fácil se constatar que embora a inteligência artificial esteja em larga escala de utilização no Judiciário brasileiro, com mais de cem projetos mapeados no último ano, pouco se tem de referencial prático e concreto de standards éticos a perseguir, gerando uma severa dificuldade de implementação ética, justamente pelo desconhecimento de um passo-a-passo, roteiro ou manual de conformidade.

1.2 OBJETIVOS

Essa dissertação tem como objetivo principal investigar o uso e desenvolvimento da inteligência artificial pela Justiça brasileira, com especial enfoque no diagnóstico da abordagem ética.

Objetiva também:

a) Captar e Analisar as percepções de servidores e membros do Poder Judiciário quanto ao uso e desenvolvimento de IA pelo Judiciário;

b) Realizar uma revisão sistemática de literatura sobre parâmetros éticos para uso e desenvolvimento de ferramentas de IA;

c) Verificar a importância de uma lista de verificação, manual de rotinas e procedimentos (checklist) para conferência de conformidade ética dos modelos de IA utilizados pelo Judiciário;

d) Investigar a percepção dos membros do Judiciário quanto aos critérios éticos adequados para uma validação dos modelos de IA pelo Judiciário.

1.3 METODOLOGIA

Consideramos a metodologia relacionando-se a natureza, objetivos, procedimentos técnicos, tipo de instrumento, além dos procedimentos técnicos e abordagem utilizada.

O presente trabalho pode ser considerado uma pesquisa básica, pois possui o objetivo de produzir conhecimentos que sejam importantes para o avanço de uma área de estudo, mas que, no entanto, não possui aplicação prática em seu desenvolvimento.

Analisando os objetivos de pesquisa aqui desenvolvida, este trabalho se caracteriza por ser pesquisa descritiva, uma vez que os fatos observados na pesquisa são apenas registrados, descritos, analisados, classificados e interpretados, sem que sofram quaisquer interferências em decorrência da pesquisa ou da atuação do pesquisador.

O tipo de questões de pesquisa que orientam este estudo são, principalmente, exploratórias, notadamente nos estágios iniciais da pesquisa, quando da execução da entrevista narrativa, enquanto tentamos entender os fenômenos e identificar distinções úteis que esclarecem nosso entendimento (Easterbrook *et al.*, 2008).

Para elaborar as explicações sobre o fenômeno estudado, seguimos uma abordagem indutiva, onde o pesquisador visa descobrir o que está acontecendo em um determinado contexto, buscando insights e, partir daí, gera ideias e hipóteses para estudos futuros (Runeson; Host, 2009).

Utilizamos abordagem qualitativa na metodologia de pesquisa. Foram utilizados três métodos de coleta de dados: entrevista narrativa exploratória, revisão da literatura e *survey*. A entrevista narrativa inicial possibilitou o levantamento de hipóteses na forma de enfrentamento de ética e IA no Judiciário, com foco nos desenvolvedores e nos juízes com atuação na área de TI, mas levando em consideração preocupações de diferentes campos de estudo, como forma de ampliar a perspectiva. A revisão sistemática da literatura buscou referenciais éticos para IA e forneceu as bases para formulação do *survey*, especificamente quanto ao levantamento de parâmetros éticos possíveis, utilizados mundialmente para checagem ética de modelos de IA. O *survey* final testou as hipóteses levantadas nas entrevistas narrativas e possibilitou a comparação dos parâmetros éticos, identificados e listados na pesquisa bibliográfica e os classificou em graus de importância.

Na elaboração e interpretação do *survey*, utilizamos as escalas de concordância, frequência e importância de Likert, popularmente utilizada para medir percepções, atitudes, caráter e opiniões (Boone Jr; Boone, 2012), atributos qualitativos passíveis de transformação quantitativa (Joshi, 2015).

No mais, a pesquisa lançou mão de investigação teórica e foi baseada em revisão de literatura, a partir de abordagem analítica qualitativa e de natureza fundamentalmente interdisciplinar. Sempre foi buscada a interseção e indicação de pontos de contato, por exemplo, entre a literatura proveniente da ciência da computação, engenharia de software e aquela produzida pela dogmática jurídica, interseccionando comentários e contextualizações com discursos, relatórios e estudos produzidos por organizações governamentais e não-governamentais.

Foi reservado, sempre que possível, espaço para a análise crítica da documentação, sempre respaldando conclusões com as devidas referências.

1.4 ESTRUTURA E ORGANIZAÇÃO DA DISSERTAÇÃO

A presente dissertação divide-se em capítulos, que correspondem às etapas de sua execução.

O segundo capítulo, denominado de “FUNDAMENTOS” contém referencial teórico dos principais componentes deste trabalho, quais sejam, inteligência artificial e ética, apresentando fundamentação e referencial conceitual utilizados no trabalho desenvolvido. Também contempla contextualização quanto à abordagem ética no uso e desenvolvimento de inteligência artificial especificamente na Justiça.

Partindo do referencial teórico foi executada entrevista narrativa em formato virtual, retratada no terceiro capítulo, quando foram ouvidos servidores e magistrados de tribunais que usam ou desenvolvam projetos utilizando inteligência computacional, com a finalidade de levantar hipóteses, compreender os sentidos e as percepções sobre o fenômeno do uso e desenvolvimento de IA pelos tribunais brasileiros, abordando questões de utilidade e adequação da tecnologia, benefícios e riscos, desafios concretos e abordagem ética.

A partir dos insights gerados com as entrevistas, especificamente quanto à inexistência de parâmetros éticos concretos para uso e desenvolvimento de IA no Judiciário, o quarto capítulo apresenta um panorama atual do cenário internacional de discussões e preocupações com questões éticas associadas ao uso da IA, fazendo um levantamento concreto de padrões éticos preceituados não somente nas iniciativas governamentais, mas também enumerando e analisando proposições do mercado, da indústria e da academia. Encerra-se este tópico com sistematização, em quadro comparativo, das categorias éticas consideradas por cada um dos documentos analisados, identificando-se standards éticos.

A última etapa da pesquisa, constante do capítulo cinco, executa um *survey* com o objetivo de testar as hipóteses levantadas na entrevista narrativa, capturar, de forma mais ampla, a percepção dos tribunais nacionais quanto aos temas ética e IA pelo Judiciário, além de permitir classificação em escala de importância e pertencimento, sob a ótica do Judiciário, de cada um dos padrões levantados no capítulo quatro. Esse capítulo apresenta e interpreta os dados obtidos com a pesquisa com os tribunais entrevistados, identificando quais parâmetros foram mais pontuados, em escala de importância. Também realiza comparação entre os sentidos e significados que foram percebidos durante as entrevistas narrativas, contrapondo-as ao que foi apurado no *survey*, identificando consensos e discordâncias.

Busca-se, ao fim, coletar insumos para subsidiar trabalhos futuros, que busquem a construção de proposta de framework ético para uso e desenvolvimento de IA no Judiciário.

2 FUNDAMENTOS

2.1 ÉTICA: CONCEITO, IMPACTO NO DESENVOLVIMENTO DE IA, NECESSIDADE DE ESTABELECIMENTO DE UM FRAMEWORK

Antes de tudo, convém pontuar que este trabalho não tem por escopo discutir propriamente os conceitos, limites e construções teóricas sobre ética.

É bem sabido que a filosofia se ocupou com os estudos da ética desde a antiguidade, distinguindo-se da moral e se segmentando, adquirindo contemporaneamente um sentido amplo e outro mais estreito. Atualmente, existe uma ética da humanidade que pauta comportamentos pensando em pressupostos maiores; e outra que padroniza ações no interior de um grupo específico.

Fazendo-se uma passagem sobre o estudo da ética na história, o fato é que não há clareza quanto ao contorno ético, que varia conforme a linha teórica e filosófica adotada, bem assim com o momento histórico vivido.

A ética de Platão e Aristóteles, que buscava compreender o bem, a verdade e a felicidade, passou por contorno teocêntrico com Tomás de Aquino, abordagem racional e crítica com Hegel e Kant, chegando a uma ética utilitarista com Hume, Bentham, Mill e Rawls, confrontada pela filosofia da suspeita de Marx, Nietzsche e Freud.

Como se vê, é um conceito filosófico aberto e em evolução.

Nesse pisar, feito o registro de que não se fará nesse trabalho uma discussão de conceito ético, é patentemente adequado que se faça breve anotação, até para servir de referencial teórico, da concepção ética construída pela humanidade.

Etimologicamente, ética vem do grego *ethos*, que quer dizer caráter, modo de ser de uma pessoa.

Ética é um conjunto de valores morais e princípios que norteiam a conduta humana na sociedade (Valls, 1994).

A ética, hoje, é compreendida como parte da Filosofia, cuja teoria estuda o comportamento moral e relaciona a moral como uma prática, entendida por Cortella (2007, p. 103) como o “exercício das condutas”. Além disso, é entendida como um tipo ou qualidade de conduta que é esperada das pessoas como resultado do uso de regras morais no comportamento social.

É ético tudo que está em conformidade com os princípios do comportamento do ser humano, de acordo com o uso comum. É a busca pelo agir de forma moral, boa, certa, justa e honesta.

Segundo Cortella e Barros Filho (2014), “A ética é construída por uma sociedade com base nos valores históricos e culturais. Do ponto de vista da Filosofia, a Ética é uma ciência que estuda os valores e princípios morais de uma sociedade e seus grupos”.

A ética possui, portanto, conceito fluido – justamente porque conglera conjunto de princípios, valores e normas morais e de conduta de um indivíduo ou de um grupo social ou

de uma sociedade –, que varia no tempo (momento histórico), no espaço (geográfico e cultural) e a depender da corrente filosófica seguida pelo intérprete.

Esta fluidez conceitual, contudo, não pode ser um obstáculo à concretude necessária para que cada cidadão, em seu agir diuturno, possa pautar suas ações conforme o que se espera sob a perspectiva ética.

Dito isto, convém acrescentar o importante contributo do filósofo italiano Luciano Floridi quando se discute filosofia e ética da informação.

Dentro da sua Filosofia da Informação – que compõe área de pesquisa autônoma, definida num movimento constante de diálogo e confronto com as principais expressões das tradições filosóficas e do pensamento contemporâneo – Floridi (2006) constrói o que chama de ética da informação.

De Gómez (2013) explica que:

“A ética da informação de Luciano Floridi tem como um de seus pontos de partida a análise crítica das principais tradições ético-filosóficas, por entender que os critérios em que sustentam o reconhecimento e julgamento de dano moral não são sempre eficazes nem adequados aos problemas e situações do mundo contemporâneo. A tese principal do autor sustenta que os padrões éticos vigentes estariam vinculados a ontologias sociais cujas concepções paradigmáticas do agente não correspondem hoje a composições ontológicas que se manifestam como agências heterogêneas, capazes de ações moralmente qualificáveis”.

Nos ensinamentos do próprio autor, a ética da informação (EI) sustenta que o comportamento (relativo ao agente) e o status (relativo ao paciente) dos objetos informacionais enquanto objetos informacionais podem ter um significado moral que vai além da função instrumental que lhes é atribuída por outras perspectivas éticas e, portanto, concluir que podem contribuir para determinar de forma normativa (com caráter obrigatório) deveres éticos e direitos juridicamente exigíveis.

Conclui que:

“A posição da IE, como a de qualquer outra macroética, não é isenta de problemas, mas esta teoria pode interagir com outras teorias da macroética e contribuir para o desenvolvimento de uma perspectiva interessante e inovadora: um processo ou ação será moralmente bom ou ruim dependendo de como eles afetam a infosfera, independentemente de suas consequências, motivos, universalidade ou caráter virtuoso. Uma ética ontocêntrica nos fornece uma perspectiva de longo alcance. Sem a contribuição da IE, nosso entendimento dos fatos morais em geral, e não apenas dos problemas particulares relacionados às TIC, seria incompleto”.

Firmado o conceito de ética da informação, é de se registrar que até encontramos consenso no nível abstrato de ética na IA, no que se refere a princípios gerais.

Essa conclusão se extrai, por exemplo, quando se vê os resultados do extenso e recente trabalho de Jobin, Ienca e Vayena (2019), intitulado de “O cenário global das diretrizes de ética da IA”, em que identificaram e analisaram 84 documentos contendo princípios éticos ou diretrizes para IA. De acordo com estes autores, onze valores e princípios éticos abrangentes emergiram da análise de conteúdo. São eles, por frequência do número de fontes em que foram apresentados: transparência (73/84), justiça e equidade (68/84), não maleficência (60/84), responsabilidade (60/84), privacidade (47/84), beneficência (41/84), liberdade e autonomia (34/84), confiança (28/84), sustentabilidade (14/84), dignidade (13/84) e solidariedade (6/84).

Registram também que “o rápido aumento no número e variedade de documentos de orientação atesta o crescente interesse na ética da IA pela comunidade internacional e em diferentes tipos de organização”.

Ao verificarem proporção quase equivalente entre documentos emitidos pelo setor público e pelo setor privado, ressaltam os autores que “a ética da IA diz respeito tanto a entidades públicas quanto a empresas privadas”. No entanto, segundo eles, as soluções propostas para enfrentar os desafios éticos divergem significativamente.

Floridi et al. (2018) também jogam luzes na definição de um conteúdo ético mínimo na produção de IA. Partindo de 47 princípios, após “impressionante e tranquilizadora” coerência e sobreposição, utilizando-se de base na bioética, concluem por seis conjuntos de princípios: beneficência, não maleficência, autonomia, justiça e explicabilidade, entendida como incorporando tanto inteligibilidade quanto responsabilidade.

A beneficência incorpora conceitos de bem-estar, preservação da dignidade, sustentabilidade do planeta; a não maleficência abrange privacidade e segurança; autonomia garante poder de decisão; por outro lado, justiça abrange prosperidade e preservação da solidariedade; explicabilidade, por fim, traz valores de inteligibilidade, responsabilidade e prestação de contas.

Após apresentar este conteúdo ético mínimo, Floridi *et al.* (2018) afirmam contundentemente que conformidade com a lei é meramente necessário (é o mínimo que se exige), mas significativamente insuficiente (não é o máximo que pode e deve ser feito).

A partir daí, em uma metáfora relativa aos jogos, pontuam que jogar conforme as regras não garante que se vença o jogo. Conduzem, então ao raciocínio de que, mais que conformidade legal, é preciso abordagem ética para vencer a partida.

Para tanto, enaltecem a dupla vantagem de uma abordagem ética da IA:

“De um lado, a ética permite que as organizações obtenham vantagem no valor social que a IA permite. Esta é a vantagem de ser capaz de identificar e alavancar novas oportunidades socialmente aceitáveis ou preferíveis. Por outro lado, a ética permite que as organizações antecipem e evitem ou pelo menos minimizar erros dispendiosos. Esta é a vantagem da prevenção e mitigação de cursos de ação que acabam sendo socialmente inaceitáveis e, portanto, rejeitados, mesmo quando legalmente inquestionável”.

Registram os autores que essa dupla vantagem da ética só pode funcionar em um ambiente de confiança pública e responsabilidades claras e com ampla transparência.

Refletem que “a aceitação pública e adoção de tecnologias de IA ocorrerá apenas se os benefícios forem vistos como significativos e os riscos como potenciais, ainda evitável, minimizável, ou pelo menos algo contra o qual alguém pode ser protegido, através da gestão de risco (por exemplo, seguro) ou reparação”.

Essas atitudes vão depender por sua vez, do envolvimento do público com o desenvolvimento de tecnologias de IA, abertura sobre como eles operam e mecanismos compreensíveis e amplamente acessíveis de regulamentação e reparação.

Uma abordagem ética da IA, portanto, também pode ser vista como um sistema de alerta precoce contra riscos que possam pôr em perigo organizações inteiras. O valor claro para qualquer organização da dupla vantagem de uma abordagem ética à IA justifica amplamente a despesa de engajamento, abertura e contestabilidade que tal abordagem requer.

Jobin, Ienca e Vayena (2019) ressaltam que “traduzir os princípios para a prática e buscar a harmonização entre os códigos de ética da IA (*soft law*) e a legislação (*hard law*) são os próximos passos importantes para a comunidade global”.

Para Ximenes e Ramalho (2021):

“Embora haja muito mérito nesses trabalhos, pois eles são a base para a construção de aplicativos de ML mais éticos e responsáveis, os desenvolvedores não podem se beneficiar diretamente deles, principalmente por três razões: a grande quantidade de informações é esmagadora; a nomenclatura não é uniforme e os conceitos nem sempre claros; e as diretrizes são muito genéricas, mais preocupadas com o que é e o porquê de um aprendizado de máquinas ético, do que com o como, que é o aspecto mais relevante para os desenvolvedores”.

Em conclusão desde subtópico, especificamente no que se refere aos limites éticos da IA, definir parâmetros concretos éticos é demanda urgente para todas as áreas do conhecimento e notadamente no segmento da Justiça, mormente se considerado que operadores do direito estão em franco desenvolvimento de ferramentas que utilizam da inteligência computacional e a falta de parâmetros ou mesmo a ausência de clareza quanto à definição do contorno ético necessário pode gerar prejuízos e riscos à própria sociedade.

2.2 IA – CONCEITOS GERAIS

Boden (1996) afirma que a inteligência artificial procura preparar os computadores para fazer o tipo de coisas que a mente é capaz de fazer. Segundo a autora, da mesma maneira que a inteligência não é uma dimensão única, mas um espaço ricamente estruturado com diferentes habilidades de processar informações, a IA utiliza muitas técnicas diferentes e se dedica a muitas tarefas diferentes.

Para Michalski, Carbonell e Mitchell (1984) a expressão inteligência artificial se referia a funções cognitivas associadas à mente humana, como a solução de problemas e o aprendizado, de modo que o aprendizado e máquinas seria pré-requisito para o desenvolvimento da IA.

Russell e Norvig (2013) definem IA como o estudo dos métodos para fazer computadores se comportarem de forma inteligente, assim entendido na medida em que faz a coisa certa (ação que é mais propensa a atingir um objetivo). Ressalta que a IA inclui tarefas de aprendizagem, raciocínio, planejamento, percepção, compreensão de linguagem e robótica.

Definições mais recentes apontam que a expressão se refere à habilidade de um sistema interpretar corretamente dados externos, aprender a partir desses dados e usar o aprendizado para alcançar objetivos e tarefas específicos.

Apresentado conceito, convém mencionar que a literatura classifica a inteligência artificial em geral e específica, ou, de maneira mais vulgar e menos usual na atualidade, em IA forte e IA fraca, respectivamente.

Também conhecida como autoconsciente, a Inteligência Artificial Forte é aquela que emula o raciocínio humano com tamanha perfeição que é capaz de resolver situações de maneira mais rápida e assertiva que uma pessoa. É o que Nick Bostrom (2014) chama de superinteligência, pois teria o potencial não apenas de se igualar aos indivíduos, mas de superá-los.

Este tipo de tecnologia, costumeiramente vista nas películas de ficção científica, se um dia for alcançada, demandará profundos e variados debates, notadamente em torno das questões éticas e do centrismo humano da IA.

De outra banda, a inteligência artificial fraca ou específica não possui esse poder de imitar cognitivamente o raciocínio humano. Não há autoconsciência.

Trata-se da noção de como lidar com problemas não determinísticos. A IA fraca centra a sua investigação na criação de inteligência artificial que não é capaz de verdadeiramente raciocinar e resolver problemas.

Registram Vale e Pereira (2023) que:

“a inteligência artificial fraca ou específica vem ganhando fôlego e sua expansão não encontra precedentes na história. São inúmeros os demonstrativos de ferramentas direcionadas a determinado setores que desenvolvem, com grau de superioridade inquestionável, atividades antes exclusiva dos humanos. (...) No Direito, não é diferente, pois aplicações de Inteligência Artificial são desenvolvidas, diuturnamente, para viabilizar a tomada de decisões por máquinas, a análise documental (...) e a promoção de análises preditivas de pronunciamentos judiciais”.

Esta constatação é plenamente aplicável quando se pensa no uso e aplicação de IA nos serviços judiciários, justamente porque, com a grande gama de processos que tramitam em formato digital, a IA tem sido identificada como tecnologia útil, adequada e necessária para alavancar a prestação dos serviços prestados pelo Judiciário, segundo Villas Boas Cueva (2021).

Nessa ordem de ideias, com a massificação e repetição de demandas a IA fraca pode, e muito, ser fonte de ganho de eficiência e produtividade.

Pioneiros no conceito de máquinas que precisam compartilhar os mesmos valores das pessoas, Stuart Rusell e Peter Norvig (2013) afirmam que um dos métodos sugeridos para considerar a questão ética nas soluções de Inteligência Artificial consiste na aprendizagem inversa por reforço, que teria a função de uma consciência externa. Nessa abordagem, a conduta moral seria aprendida com base nos resultados manifestados por agentes externos, levando ao desenvolvimento de um comportamento moral ou ético por parte da solução.

Kong e Dietterich (1995) lembram que o termo *bias* foi introduzido por Mitchell (1980) para significar qualquer base para escolher uma generalização (hipótese) em detrimento de outra, além da estrita consistência com exemplos de treinamento observados. Apontam que todo algoritmo de aprendizado indutivo deve adotar um viés para generalizar dados de treinamento. Sem viés, todas as funções possíveis devem ser consideradas hipóteses e, juntas, essas funções preveem que todos os possíveis resultados futuros são igualmente prováveis, de modo que não podem fornecer uma base para generalização ou predição.

O viés de um algoritmo de aprendizado fornece uma especificação para o comportamento desejado do algoritmo e esclarece o *design* e a implementação de algoritmos de aprendizado de máquina. O problema reside quando, a despeito da necessidade de definição de um viés, o algoritmo incorpora e reproduz preconceitos humanos.

O preconceito de máquina, ou *machine bias*, é o efeito indesejado de suposições errôneas nos processos de aprendizado de máquina. O preconceito reflete problemas relacionados à coleta ou uso de dados, nos quais os sistemas tiram conclusões inadequadas sobre conjuntos de dados, seja por causa da intervenção humana ou como resultado da falta de avaliação cognitiva dos dados.

Em inteligência artificial, *algorithmic bias* pretende conhecer, identificar e resguardar os modelos matemáticos utilizados de consequências graves decorrentes da falta de transparência e conhecimento sobre essa tecnologia emergente.

Mehrabi, Ninareh et al. (2021) esclarecem que o preconceito pode existir de muitas formas, alguns dos quais podem levar à injustiça em diferentes aspectos e tarefas de aprendizagem. As mais diversas fontes de preconceito no aprendizado de máquina com suas respectivas categorizações e descrições precisam ser mapeadas, a fim de motivar soluções futuras para cada uma das fontes de preconceito. Partindo de diversos outros estudos, artigos e papers, os autores reiteram as fontes mais importantes de preconceito introduzidas, apresentando uma categorização diferente dessas definições no artigo de acordo com os dados, algoritmo e ciclo de interação do usuário.

O primeiro tipo são os vieses nos dados utilizados no algoritmo (Data to Algorithm), vieses nos dados, que, quando usados por ML, pode resultar em resultados algorítmicos tendenciosos. Os vieses podem ser de medição, oriundos da omissão de variável, da ausência de representatividade ou diversidade, incorreta agregação (por heterogeneidade de dados ou viés estatístico na análise geoespacial, erros na amostragem, falácia de dados longitudinais ou, por fim, falhas na vinculação de atributos de usuários.

A segunda classe de vieses reportada consiste na discriminação criada no comportamento dos usuários, modulada por algoritmos enviesados (Algorithm to User). Aqui enumeram: viés algorítmico (viés não está nos dados de entrada, mas adicionado pelo algoritmo), viés de interação do usuário (que surge de um viés de apresentação dos resultados ou de classificação), viés de popularidade, viés emergente (resultado do uso e da interação com usuários reais) e viés de avaliação do modelo.

Reportam, por fim, a categoria de vieses que surgem dos usuários para os dados (User to data), justamente porque muitas fontes de dados usadas para treinar modelos de ML são geradas pelo usuário, de modo que quaisquer preconceitos inerentes aos usuários podem ser refletidos nos dados que eles geram. Listam, então, os seguintes vieses: histórico, populacional, autosseleção, social, comportamental, temporal e de produção de conteúdo.

Segundo Floridi (2018), que a IA terá fortes impactos na sociedade não é mais uma questão. As questões pertinentes atuais são quem, como, onde e quando serão sentidos os impactos positivos e negativos da IA.

“Para enquadrar estas questões de uma forma mais substantiva e prática, apresentamos aqui o que consideramos as quatro principais oportunidades para a sociedade que a IA oferta. São quatro porque abordam os quatro pontos fundamentais na compreensão da dignidade humana e do florescimento: quem podemos nos tornar (autorrealização autônoma); o que podemos fazer (agência humana); o que podemos alcançar (individual e capacidades sociais); e como podemos interagir uns com os outros e com o mundo (coesão social). Em cada caso, a IA pode ser usada para promover a natureza humana e suas potencialidades, criando assim oportunidades; subutilizados, criando assim custos de oportunidade; ou usado em demasia e mal utilizados, criando assim riscos”.

Kaplan e Haenlein (2020) afirmam que em um contexto mais amplo, os governos precisarão criar regras e normas mais específicas para o funcionamento do ecossistema de gestores, empregados, máquinas, usuários e concorrentes.

Por fim, é preciso registrar que há meios para se controlar ou mitigar o viés algorítmico.

Mehrabi, Ninareh et al. (2021) registram que métodos que visam enfrentar vieses nos algoritmos se enquadram em três categorias: 1) técnicas de pré-processamento, que tentam transformar os dados para que os dados subjacentes à discriminação seja eliminada; 2) técnicas durante a fase de processamento, que tentam modificar e mudar a aprendizagem de última geração dos algoritmos, para remover a discriminação durante o processo de treinamento do modelo e; 3) no pós-processamento, em que após o treinamento, acessando um conjunto de validação que não esteve envolvido durante o treinamento do modelo.

Os autores então enumeram algumas técnicas: a primeira dela consiste na construção de um *dataset* com dados imparciais; depois propõem um Aprendizado de Máquina Justo, que se desdobra em justiça na classificação, na regressão, previsão estruturada, dentre outras técnicas; aprendizagem de representação justa, com codificadores automáticos variacionais, aprendizado adversarial e; processamento de linguagem natural justa, que precisa primar por incorporação de palavras, resolução de correferências, modelo de linguagem, codificador de frases, máquinas de tradução, além de outras.

3 ENTREVISTA NARRATIVA VIRTUAL – ERA UMA VEZ: IA E ÉTICA NO JUDICIÁRIO

3.1 CONTEXTO

Embora os principais acontecimentos determinantes para surgimento da inteligência artificial tenham se dado em meados do século passado, o fato é que somente nas últimas décadas é que houve adequado ambiente para um desenvolvimento e utilização exponenciais desta tecnologia.

Um mundo que gera 2,5 quintilhões de bytes de dados todos os dias, associado à evolução da qualidade de rede, dispositivos e softwares, aumento da capacidade de processamento de dados e investimento de grandes empresas tecnológicas, acompanhadas pelo interesse público e estatal, são alguns dos fatores que levaram a que o impacto da inteligência artificial em nossas vidas tenha subido de forma substancial.

Dentre os mais diversificados usos, tem se notabilizado a utilização da inteligência artificial no direito, seja nos escritórios de advocacia – com ferramentas que ajudam na conferência de contrato, acompanham tramitação de processos, comparam julgamentos, fazem predição de decisões e sugerem textos e modelos –, seja na Justiça. Os sistemas de Justiça estão se utilizando de inteligência artificial para analisar e comparar fluxos de processos, em assistentes virtuais, *chatbots*, triagem e classificação de petições e tarefas e até mesmo no apoio a decisões judiciais.

Segundo o Conselho Nacional de Justiça, a inteligência artificial está presente na maioria dos tribunais brasileiros, tendo o órgão administrativo verificado mais de cem projetos em desenvolvimento somente no Brasil.

O desenvolvimento e uso de IA no Judiciário é, portanto, uma realidade.

Diante deste fato é que surgem inúmeras questões: como se dá o interesse do Judiciário em desenvolver e aplicar IA? Quais usos, ganhos e benefícios esperados/auferidos? Foram identificados riscos ou ameaças? Quais dificuldades e limites no desenvolvimento de IA pelo Judiciário? Há alguma preocupação ética? Se sim, com qual conformação, seguindo quais critérios e em que fases de desenvolvimento e aplicação do modelo?

É neste oceano de desconhecimento da realidade e das experiências do Judiciário brasileiro que se impõe a necessidade de realização de uma entrevista, um estudo de percepção, tendente a capturar o sentimento dos integrantes dos tribunais nacionais para utilização de IA, notadamente com vistas para os aspectos éticos.

3.2 ENTREVISTA NARRATIVA: ADEQUAÇÃO METODOLÓGICA

Nesse ponto, reputamos relevante metodologicamente executar uma entrevista narrativa de forma virtual.

Segundo Jovchelovitch e Bauer (2017), “não há experiência humana que não possa ser expressa na forma de uma narrativa [...] contar histórias implica estados intencionais que

aliviam, ou ao menos tornam familiares, acontecimentos e sentimentos que confrontam a vida cotidiana normal”.

Idealizada por Fritz Schütze como um dispositivo para compreender os contextos em que as biografias foram construídas, os fatores que produzem mudanças e motivam as ações dos portadores da biografia, essa entrevista produz textos narrativos sobre as experiências das pessoas, expressando maneiras como os seres humanos vivem o mundo por meio de histórias pessoais, sociais e coletivas. Ela tem como objetivo incentivar a produção de uma narrativa pelo depoente.

Jovchelovitch e Bauer ressaltam que “A entrevista narrativa é classificada como um método de pesquisa qualitativa (Lamnek, 1989; Hatch & Wisniewski, 1995; Riesman, 1993; Flick, 1998). Ela é considerada uma forma de entrevista não estruturada, de profundidade, com características específicas”.

Jovchelovitch e Bauer (2002) afirmam que a questão central deve fazer parte da experiência do entrevistado e possuir relevância social, pessoal ou comunitária; ser ampla, para permitir ao narrador desenvolver uma história longa, com situações iniciais, e percorrer o passado até chegar à atual circunstância; e evitar formulações indexadas, as quais se caracterizam como perguntas sobre quem faz o que, quando, onde e o motivo.

Esse dispositivo de entrevista se constitui em uma maneira específica de produção de dados, proposta por Schütze (2011), por meio da reconstrução dos acontecimentos sociais, a partir da perspectiva dos sujeitos de uma investigação.

Schütze (1977) indica como útil a realização de entrevista narrativa nos casos de projetos que investigam acontecimentos específicos, especialmente assuntos “quentes”, tais como um projeto de desenvolvimento específico ou políticas locais, assim como e projetos onde variadas “versões” estão em jogo, de forma que grupos sociais diferentes constroem histórias diferentes e as maneiras como elas diferem são cruciais para se apreender a dinâmica plena dos acontecimentos.

Como visto, a entrevista narrativa se presta como forma eficiente de descrever e interpretar os hábitos, costumes, valores e práticas de uma comunidade específica (no caso, membros do Judiciário), investigando, em profundidade, comunidades e grupos humanos em suas particularidades.

Dito isto, resta claro que a opção metodológica assumida é adequada ao enfrentamento do recente fenômeno de uso e desenvolvimento da IA pelo Judiciário, com especial enfoque ético. A intenção de pesquisa narrativa é perceber o fenômeno do uso da inteligência artificial pelo Judiciário, colocando luzes sobre os aspectos éticos envolvidos neste tema, como etapa preparatória para execução desta dissertação, oportunidade em que se buscará capturar os elementos que podem interferir de alguma maneira no fenômeno pesquisado.

3.3 ESTRUTURA E ORGANIZAÇÃO

A entrevista narrativa foi organizada seguindo o roteiro de uma entrevista aberta semiestruturada, somente com questões abertas, que permitem maior apreensão de informações, elementos e achados de pesquisa.

Desenvolvida com prévia autorização de participação livre e esclarecida, a entrevistas foi disposta em três seções: a primeira cuida da definição do perfil pessoal e profissional do entrevistado, identificando formação acadêmica, experiência profissional; na segunda etapa, são levantados os conhecimentos teóricos e práticos do entrevistado com a IA, com ética, buscando especificamente a experiência com a operação do modelo, benefícios, riscos e ameaças, assim como as questões éticas, tais como a abordagem, os critérios e standards; a última parte da entrevista busca colher as percepções finais do entrevistado quanto à ideação, desenvolvimento e governança dos modelos de IA, questionando a avaliação do entrevistado quanto à verificação concreta de ética na IA, e seu entendimento quanto à proposição de uma lista de verificação ordenada para conferência de tarefas e de etapas para conformação ética e legal, e com qual contorno.

O recorte dos três entrevistados teve como objetivo obter a mais abrangente representatividade, abarcando profissionais da área de tecnologia, bem como juízes usuários dos sistemas informatizados à disposição, além de gestores responsáveis pelo setor de TI, incluindo tribunais federais e estaduais, de pequeno porte e de grande porte¹, de regiões diversas do País, além de que foram entrevistados representantes de tribunais que nunca tiveram qualquer experiência no uso e desenvolvimento de IA, bem como de tribunais já experientes nesta tecnologia, inclusive gestores nacionais de tecnologia da informação no Judiciário.

Com esses esclarecimentos, passemos, pois, à descrição das entrevistas.

3.4 DESCRIÇÃO ENTREVISTA 01

O primeiro entrevistado tem o seguinte perfil: magistrado integrante de tribunal estadual de grande porte, exercendo a função de juiz auxiliar gestor de TI; possui graduação e pós-graduação *latu sensu*, cursando mestrado; possui certa familiaridade com programação, fez cursos de ciência de dados e inteligência artificial com linguagem em *python*, pagou algumas cadeiras no mestrado que abordaram inteligência artificial; define-se como detentor de conhecimento moderado quanto à inteligência artificial, bem como a ética na IA.

Informou que não tem experiência em projetos de IA e que não há iniciativas em curso, no seu tribunal, de desenvolvimento de modelos de inteligência artificial. Inobstante isso, asseverou que seu tribunal tem interesse em desenvolver projetos com IA, porque “as aplicações que se buscam através da IA são de relevância para o negócio do tribunal”. Segundo o entrevistado, a IA foi definida como a tecnologia mais aderente para a proposta do projeto de gestão do tribunal, possibilitando fazer entregas e dar retornos no que se refere à predição em relação a grande volume de dados.

Embora não tenha conhecimentos mais aprofundados sobre ideação, desenvolvimento e operação de produtos que envolvam IA, o entrevistado afirmou vislumbrar riscos e ameaças no uso de IA pelo Judiciário. Destacou que ferramentas que envolvam vasta gama de tarefas complexas, essencialmente de cognição humana, como a tarefa de julgamento, existe o risco do “halo tecnológico”, que é confiança que o ser humano tem na máquina. Esclareceu que o

¹ Classificação utilizada pelo Conselho Nacional de Justiça, que agrupa, para fins de comparação, os tribunais nacionais, em tribunais de pequeno, médio e grande portes, valendo-se de critérios como a estrutura administrativa e de pessoal, além da população atendida.

usuário mais leigo acaba seguindo orientações do sistema. E “essa tendência que o ser humano tem de seguir orientações da máquina podem envolver um risco quando a resposta que a aplicação está te dando envolve consequências que ela mesma não consegue prever, não foi programada para isso”. Arremata que se o ser humano não estiver bem orientado quanto a este apoio da ferramenta é feito, quais são as limitações, pode ser que a pessoa seja induzida a erro.

Também registrou preocupação com a questão da discriminação algorítmica e enviesamento, afirmando que a “IA é um espelho nosso, da sociedade; se ela está enviesada, reflete nós mesmos, humanos. Quando as pessoas colocam o dedo no nariz da IA, estão colocando o dedo no seu próprio nariz”.

Por esta razão, em sua ótica, as questões da IA devem estar na pauta do desenvolvimento de IA pelo Poder Judiciário, dada a potencialidade da tecnologia.

Segundo o entrevistado, a Resolução nº 332 do CNJ é um documento robusto e, portanto, mais do que suficiente, no que se refere à normatização necessária para a segurança no desenvolvimento de IA pelo Judiciário.

A capacidade de os modelos captarem informações e darem respostas sobre essas informações gera projeção quase que infinita da IA, daí que um mínimo e controle é salutar para garantir que essa ferramenta seja utilizada de forma mais adequada possível.

Defende o entrevistado, sob o ponto de vista de fomento ao desenvolvimento de inovação, que o Poder Judiciário não precisa de uma regulação muito forte, notadamente pela dificuldade com burocracia administrativa e de pessoal. “Se a gente não tivesse nenhuma regulamentação como obstáculo ao desenvolvimento, já seria muito difícil de a gente desenvolver artefatos. Então quando você traz mais exigência, torna ainda mais difícil, o que é um entrave no desenvolvimento”. Diferente situação é identificada pelo entrevistado no que se refere ao setor privado, que tem desenvolvimento mais dinâmico e, portanto, necessita de maior controle e rígido acompanhamento ético. Conclui que uma regulamentação muito forte no desenvolvimento de IA pelo Judiciário provoca desaceleração de projetos, justamente pela burocracia.

Instado a identificar critérios éticos importantes no desenvolvimento de IA, citou: necessidade de uma equipe multidisciplinar, representatividade/pluralidade de equipes, tratamento de dados, seleção adequada de dados para evitar enviesamento, pluralidade, imparcialidade e, por fim, testes e homologação do modelo.

Disse que a abordagem ética no desenvolvimento de IA pelo Judiciário deve passar pelas fases de ideação, desenvolvimento e governança, com peso e importância, oportunidade em que citou o conceito de “*ethics by design*”, já mencionado em capítulo anterior deste trabalho. E que deve ser realizado um mapeamento prévio de riscos envolvidos, com as consequentes medidas de contenção e reversão.

Ressaltou que é importante um controle sobre os modelos e governança nacional, central, exercido pelo Conselho Nacional de Justiça, como já vem sendo feito.

Embora defenda que a regulamentação deva ser mínima o Judiciário, ressaltando que ela já existe, por questões práticas seria interessante um procedimento para essa validação ética no desenvolvimento de IA.

Ressaltado que não há hoje um procedimento bem especificado para verificação ética de IA no Judiciário, disse que a norma poderia trazer procedimentos dessa validação e curadoria

ética, com requisitos objetivos, lógica científica, o que permitiria aos tribunais ganhos expressivos, “porque facilita atividade das equipes, otimiza as atividades e simplifica e confere efetividade ao processo de verificação de conformidade”.

Nesse contexto avalia como sendo “ótima” a proposição de um checklist ordenado para conferência de tarefas e de etapas para conformação ética e legal dos modelos de IA produzidos pelo Judiciário.

Finalizando a entrevista sugeriu um validador automatizado de conformidade ética. O ideal seria uma aplicação/ferramenta que fizesse a curadoria e verificação ética do modelo e fornecesse um parecer, com critérios objetivos, sem burocracia. Um validador ético de modelos de IA, sistema indicaria se está válido ou não eticamente, com checklist pré-definido e indicação de pontos a melhorar.

3.5 DESCRIÇÃO ENTREVISTA 02

O segundo entrevistado é juiz, integra tribunal federal de médio porte, exerceu por quatro anos a função de juiz assessor especial do Conselho Nacional de Justiça responsável pelas políticas de TI na Justiça brasileira; possui especialização e mestrado em direito; define-se como um autodidata em ciência da computação e com bom conhecimento em relação à inteligência artificial, inclusive porque atuou no CNJ para regulamentar o uso ético de IA no Judiciário.

Reportou que seu interesse quanto ao uso da Inteligência Artificial no Judiciário se deu ainda em 2006 quando colaborou, com dados da sua unidade jurisdicional, em pesquisa científica que fez análise do tempo de duração de processos trabalhistas utilizando redes neurais artificiais como apoio à tomada de decisões. A partir daí passou a estudar sobre o tema.

Entusiasta de tecnologia da informação, o entrevistado ressaltou que o sistema de Justiça precisa melhorar, porque diante da sua complexidade, está se tornando disfuncional. A massa de processos que entram no Judiciário de hoje exige soluções adequadas e, segundo ele, a inteligência artificial tem condições de auxiliar organizações na tomada de decisões e promover melhorias de seus processos de trabalho, assim como na predição de eventos futuros e tendências.

Embora não acredite que o algoritmo possa substituir o juiz, o que seria uma distopia, identifica claramente três grandes benefícios do uso da IA pelo Judiciário: 1) como apoio à tomada de decisões judiciais; 2) melhoria de processos de trabalho, como, por exemplo, com a mineração de processos, oportunidade que mencionou convênio do CNJ com o CIn-UFPE para criação da ferramenta JUMP; 3) ferramentas que ajudem a entender a jurisdição em determinados temas e mostrar tendências.

Arremata que “se o mercado atua no atacado, a gente (Judiciário) não pode atuar no varejo”.

Destacou ter conhecimento amplo sobre inteligência artificial e seu uso no Judiciário, tendo participado, tanto em seu tribunal, como nas funções administrativas perante o CNJ, não somente nas fases de ideação, desenvolvimento e uso de modelos de IA, mas sobretudo na governança nacional do manejo da IA nos sistemas da Justiça.

Fazendo um esboço histórico, reportou que desde 2016 o CNJ identificou uma tendência nacional de uso de IA, assim como uma oportunidade de crescimento e desenvolvimento de IA na Justiça, a partir de uma experiência exitosa da Advocacia Geral da União (AGU), que desenvolveu o sistema SAPIENS, que usa técnicas de mineração de dados para auxiliar procuradores federais na tomada de decisões. A partir de então passaram a desenvolver pesquisa no Judiciário e encontraram padrões em modelos com uma resposta interessante, com boa perspectiva.

Registrou que somente em 2018, o CNJ, diante das potencialidades da tecnologia da IA, resolveu atuar em duas grandes frentes: a) conhecer, entender e estimular as soluções de IA já em desenvolvimento e uso, atuando de maneira a estabelecer política de gestão e governança que permitisse uma interconexão para todo o sistema de Justiça e não pontualmente por um ou outro tribunal, evitando uma multiplicidade de soluções desconectadas, sem considerar o Judiciário como um sistema uno; b) desenvolvimento de uma plataforma de IA que servisse a todo o Judiciário, que, no caso, foi o SINAPSES, desenvolvido em convênio com TJRO, criando um repositório de soluções de IA.

Ressaltou o papel do CNJ em capitanear a nacionalização e garantia de interoperabilidade de modelos de IA desenvolvidos em todo o Brasil, afirmando que o SINAPSE permite que a IA seja utilizada como plataforma, mencionando o conceito de “IA *as a service*”, ou seja, a inteligência artificial vista como um serviço à disposição de todo o Judiciário brasileiro.

Questões éticas no desenvolvimento de IA não estiveram na pauta do desenvolvimento do modelo, nem quando esteve à frente de projetos do seu tribunal, nem no próprio CNJ com o SINAPSE, mesmo porque, neste último caso, não se trata de um modelo em si, mas de uma plataforma, repositório de modelos de IA.

Paralelamente a essas duas iniciativas de desenvolvimento de IA, registrou que o CNJ iniciou estudos científicos, de informática e jurídicos quanto ao uso ético de IA no Judiciário, criando uma comissão de juristas e técnicos de computação, a fim de elaborar um documento que definisse abordagem ética para uso da IA nos sistemas de justiça, o que culminou com a Resolução CNJ nº 332/2020. Enalteceu que a resolução vem estabelecer parâmetros mínimos de modelo de IA, com observância dos direitos fundamentais. Registrou que buscaram trazer regulamentação mínima, sem impedir a inovação, de forma que o tribunal que esteja pesquisando e desenvolvendo IA apenas “declara” ao CNJ este fato, não sendo necessário “pedir autorização”.

Ao comemorar a edição da Resolução CNJ nº 332/2020, a primeira e única norma publicada pelo Judiciário em todo o mundo, também destacou a previsão de necessária participação multidisciplinar na elaboração de modelos de IA e a necessidade de retroalimentação e reaprendizagem sistemática dos modelos de IA, haja vista a dinamicidade dos fatos sociais.

Alerta que o uso da IA no Judiciário apresenta riscos, elencando-os: a) acomodação do ser humano à definição do algoritmo, o que demanda uma revisão constante para reaprendizagem do algoritmo, capaz de captar as mudanças e conflitos sociais; b) tendência de estabilizar as coisas do jeito em que estão no momento específico e estanque e, para resolver essa questão, a abordagem de reaprendizagem deve ser proativa, intensa, de forma a repetir o

ciclo de inovação, desde a ideação, prototipação e testes; c) necessária atividade de pesquisa judiciária, com centros de inteligências.

Afirmou ser fundamental catalogar e historiar riscos e ameaças previamente, quando for iniciar a ideação do modelo, pensando nos artefatos suficientes para neutralizá-los. Segundo o entrevistado, a resolução não traz essa obrigação, mas essa é uma evolução necessária. Para ele, “o Tribunal deveria responder a um questionário e a cada etapa do processo de desenvolvimento do modelo ele deveria olhar para o questionário novamente e ver se as respostas continuam fazendo sentido” e sair “ticando” cada obrigação ética.

Ressalta ser necessário e fundamental a criação de uma metodologia de verificação ética, com parâmetros concretos de validação. Sobre esse tema, disse que a proposição de uma lista de verificação (checklist) é muito bem-vinda, ao tempo em que enalteceu o trabalho produzido pelo TRF3 – SP, pelo juiz federal Caio Moyses. Dentre os critérios determinantes para validação ética, elencou como sendo importante olhar qual o problema que se quer resolver com aquele modelo, quais serão as pessoas atingidas por ele, o *dataset* utilizado (fonte primária de dados), entender como o modelo funciona e promover revisão constante. Ressaltou que o Poder Judiciário só deve ter o cuidado para que esse checklist não venha ser burocrático a inviabilizar os modelos de IA e inibir a inovação.

3.6 DESCRIÇÃO ENTREVISTA 03

O terceiro entrevistado é servidor público, integra tribunal estadual de pequeno porte, analista de sistemas e desenvolvedor, com formação acadêmica em ciência da computação e especialização em engenharia de software e também em ciência de dados.

Além de estudar inteligência artificial, desenvolveu, no seu tribunal e em cooperação com o Conselho Nacional de Justiça, dois projetos que se utilizam de inteligência artificial para o Judiciário. Por conta disso tem vasta experiência no uso desta tecnologia nos órgãos da Justiça, tanto em âmbito local, como em âmbito nacional, haja vista que atualmente ainda presta serviços em cooperação com o CNJ, no desenvolvimento e manutenção da plataforma SINAPSES.

Relembrou que o seu primeiro contato com IA se deu ainda em seu tribunal quando, em 2018, por direcionamento do órgão superior, especificamente o Presidente do tribunal, ficou definida a estruturação de uma equipe para pesquisar, estudar e implementar projetos de IA em âmbito local, oportunidade em que, por se tratar de uma novidade e, diante da inexperiência geral no uso de IA no Judiciário, tiveram 8 meses somente para estudar, prospectar e propor usos, com vistas a fazer o melhor uso da IA para os problemas da Justiça.

Identifica diversos benefícios no desenvolvimento e uso de IA na Justiça, notadamente para automatizar tarefas repetitivas, com modelos específicos: “cada atividade repetitiva é uma oportunidade de criar um modelo de IA”. Essa automatização promoverá sensível celeridade nos processos judiciais. Alerta, entretanto, que o intuito não é substituir humanos, mas agregar valor à decisão, tornando-a mais precisa, assertiva, segura e estável.

Quanto aos riscos e ameaças reportou que o principal deles é o viés algorítmico, que, por vezes está incorporado na própria natureza do desenvolvedor, não como um preconceito

em si, mas é a vivência dele e como ele trabalhava. Registrou que o tecnicismo de quem desenvolve muitas vezes mascara algum tipo de aspecto preconceituoso que possa estar envolvido.

Relatou que na maioria dos usos de IA na Justiça, o que se vê é a IA fraca, modelos simples de classificação de processos e situações, registrando que esse tem viés reduzido. Por outro lado, embora registre que o Judiciário ainda não se vale de IA forte, ressalta que quando se passa ao apoio na decisão judicial, que vai interferir no resultado do processo judicial, será recomendado uma auditoria com equipe multiétnica para validação.

As questões éticas sempre estiveram em mente quando desenvolveram IA. O principal ponto observado foi a responsabilidade das pessoas envolvidas.

Disse que na sua atuação perante o CNJ restou definido que não caberia a eles, servidores da área técnica, definir as questões éticas, sendo necessária uma equipe multidisciplinar para discussão e verificação. Esclarece que o desenvolvedor, por natureza, acaba buscando a via mais rápida e pode existir algum viés que ele não consegue identificar até mesmo por desconhecimento teórico ou jurídica em relação ao assunto. Daí que ficou definido que em casos de modelos que envolvessem classificações ou pudessem interferir de alguma forma no processo judicial, este modelo demandaria validação por uma equipe multidisciplinar, tanto no aspecto técnico, como no ético.

Lembrou o entrevistado que quando desenvolveram as duas ferramentas de IA, havia muito foco do Judiciário em usar IA, mas não havia uma norma com foco em ética e IA, de modo que não havia critérios ou standard para verificação, o que gerava um clima de medo entre os que operavam com modelos de IA, que se questionavam: “como é que a gente vai disponibilizar algo que não tem regramento?”.

Diante deste vácuo e da insegurança no desenvolvimento, sua equipe procurou o CNJ para pedir que fosse editada uma norma regulamentando o uso ético da IA. Foi nesse contexto que nasceu a Res. 332 do CNJ, ressaltando que é o único Poder da República que tem esse tipo de preocupação a ponto de normatizar o uso de IA ética.

Embora avalie que a mencionada resolução é extremamente genérica, relatou o entrevistado que parte dos seus colegas de profissão criticaram a mencionada resolução, por entenderem que ela engessava o processo de desenvolvimento.

O olhar ético deve se dar em todos os momentos, desde a concepção, extração, transformação e carga de dados (ETL), modelagem, validação e homologação. Mencionou que este enfoque ético tem que seguir no pós aplicação do modelo, lembrando norma francesa que garante às pessoas impactadas por modelo de IA o direito a ter acesso a todo o processo que envolveu o uso de IA na sentença na qual ela foi afetada.

Relatou que um sindicato de servidores instou o CNJ a especificar o que os modelos de IA pretendiam, porque estavam temendo que o uso de IA gerasse perda de cargos.

Durante o desenvolvimento dos modelos não chegaram a levantar e catalogar potenciais riscos e ameaças ou mesmo modos de correção e contenção.

Segundo o entrevistado, a abordagem ética na fase de ideação e de pensamento de um modelo de IA é superimportante. Para tanto, valendo-se de uma metáfora, compara o trabalho dos desenvolvedores ao de um minerador: “a gente é minerador, vê uma pedra e não sabe se é diamante ou não, cabe justamente à equipe que tem conhecimento negocial dizer: ‘isso aqui é

diamante, vamos lapidar isso aqui, bate dessa forma, refina dessa maneira'; nossa equipe sabe movimentar a ferramenta, mas não sabe dizer onde tem que fazer essa lapidação". Arremata afirmando que lapidação e curadoria são necessários e devem ser realizados por quem tem conhecimento de legislação, processos e fluxos, a fim de se chegar a um resultado.

De igual forma, na definição e desenvolvimento do algoritmo, será necessária validação ética, para verificar se aquilo que está sendo feito mediante análise estatística está de acordo com o que foi discutido na ideação.

Igualmente importante o estabelecimento de governança de modelos de IA, sob pena de se liberar modelos incorretos, que serão escalados e piorados ainda mais. Exemplificou uma situação em que "um modelo que foi utilizado por um tribunal, foi super bem aceito, está dando um ótimo resultado sob o ponto de vista da celeridade, mas contém uma falha na questão ética dele e antes de se descobrir essa falha ele foi replicado em 72 tribunais". Se lá no início, antes de entrar em produção, este modelo tivesse conferência e validação ética, essa escala do problema seria minimizada.

Avalia como superimportante uma proposição de metodologia para verificação e conferência de conformidade ética, registrando que do ponto de vista técnico não vê óbices, mas somente benefícios no estabelecimento de um padrão de desenvolvimento ético: "é até importante para quem trabalha no desenvolvimento ter esses parâmetros para poder nortear o trabalho". Seria interessante, segundo ele, a definição de um checklist, simples e concreto, até porque "eu não tenho hoje esse parâmetro".

Disse que, "como técnico, o primeiro pensamento nosso não é esse (verificação e validação ética), a gente parte mais para uma análise de como resolver o problema". Sabe-se que a norma exige conformidade ética, mas isso, na prática, só se dá numa etapa posterior: "desenvolve para poder validar". Concluir que esse processo de verificação e validação "deveria ser feito antes".

Reportou que é frustrante empreender esforços e trabalho para desenvolver algo e, ao final, embora tecnicamente perfeito, não ter conformidade ética e, portanto, não poder ser implementado. É uma perda de recursos e tempo, segundo ele.

O checklist deveria perseguir os seguintes valores: sociais, raciais, gênero.

3.7 SISTEMATIZAÇÃO E RESULTADOS DA ENTREVISTA NARRATIVA

Analisando sistematicamente os dados e descobertas coletados durante as entrevistas narrativas, podemos inferir que as contribuições e conclusões dos entrevistados circundam, basicamente, temas específicos:

1. Motivação e interesse – do entrevistado e da instituição pública em que atua – no tema inteligência artificial;
2. Apontamento de benefícios no uso da IA pelo Judiciário e os respectivos riscos;
3. Percepções sobre governança da IA, gestão e regulamentação deste tema nos órgãos da Justiça;

4. Desafios éticos no desenvolvimento de IA pelos tribunais nacionais – se existem, se são observados;
5. Fases em que a abordagem ética deve ser priorizada;
6. Definição de critérios éticos que precisam ser checados;
7. Necessidade de adoção de um framework ético objetivo e concreto.

Nesse contexto, para fins didáticos, faremos a análise do conteúdo das entrevistas em cada uma das categorias acima indicadas, enlaçando semelhanças e apontando eventuais dissensos.

A primeira conclusão extraída é o interesse comum e geral do Poder Judiciário brasileiro no estudo, desenvolvimento e uso da inteligência artificial. Esse fato ficou evidenciado em todas as entrevistas realizadas, havendo que se destacar que dois entrevistados reportaram que o interesse institucional na adoção da IA se deu por determinação da presidência da instituição, de onde se extrai que a administração superior do Judiciário brasileiro elegeu esta tecnologia como útil e adequada às suas soluções de negócio.

As razões indicadas pelos entrevistados residem justamente na antevisão dos resultados esperados na melhoria da prestação do serviço jurisdicional com adoção da tecnologia, daí que foi relatado por um dos entrevistados que a IA é a tecnologia mais aderente ao projeto de gestão do seu tribunal. Nesse mesmo sentido, outro entrevistado explicou que o grande número de ações já existentes, somadas às novas ações que chegam todos os dias, exigem do Judiciário “soluções adequadas” e, segundo ele, a inteligência artificial tem condições de auxiliar organizações na tomada de decisões e promover melhorias de seus processos de trabalho, assim como na predição de eventos futuros e tendências.

Por fim, ainda no que se refere aos interesses e motivações, ressaltou um dos entrevistados que a quantidade de tarefas repetitivas, em situações e processos semelhantes, é uma relevante oportunidade de utilização de IA para promover celeridade de julgamento.

O segundo ponto que merece destaque entre as conclusões das entrevistas diz respeito aos riscos inerentes ao emprego de IA nos órgãos da Justiça. Isto porque se, por um lado, ficou claro o interesse do Judiciário no uso de IA, por outro, também restou assente que o uso e desenvolvimento de IA envolve riscos e ameaças reais.

Todos os entrevistados elencaram o viés algorítmico e a discriminação que daí pode surgir como o mais relevante risco e impacto potencialmente negativo no uso e IA na Justiça. Isto porque, como guardião da Constituição Federal, o que se espera da atuação do Judiciário é justamente o respeito aos direitos humanos fundamentais, a igualdade e a justiça. Daí que, ao mesmo tempo em que se impõe investimento no uso da tecnologia, também se deve investir no desenvolvimento ético das soluções e modelos de IA.

Dois entrevistados reportaram que outro relevante risco no uso de IA pela Justiça, que reside no excesso de confiança que o ser humano possui no resultado da máquina. Segundo eles, quando a IA passa a ser utilizada como apoio na tomada de decisão judicial e na resolução de conflitos da sociedade, o “comodismo algorítmico”, entendido como a simples e imediata adoção do resultado indicado pela máquina, pode gerar severas distorções de justiça e equidade. Impõe-se, portanto, revisão humana das sugestões algorítmicas e uma revisão da base de conhecimento do modelo, o que deve se dar de modo constante e proativo, para reaprendizagem do algoritmo, capaz de captar as mudanças e conflitos sociais.

Quanto à governança, todos os entrevistados apontaram para necessidade de gestão, regulação, definição de uma política estatal e governanças unificados, o que já vem sendo feito pelo CNJ, pelas seguintes razões: obter aproveitamento nacional e integral da tecnologia implementada individualmente por quase uma centena de tribunais existentes no Brasil; estabelecimento de uma política pública direcionada à inovação e tecnologia no Judiciário, inclusive com uma mesma plataforma de compartilhamento (SINAPSES); controle e definição de parâmetros comuns, que permitam desenvolvimento ético e interoperabilidade de sistemas (“*AI as a service*”).

Nesse ponto, destacaram a Resolução CNJ nº 332/2020, que dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário, primeira e única norma publicada pelo Judiciário em todo o mundo para regular o desenvolvimento de IA na Justiça.

A respeito desta norma, um dos entrevistados registrou que a resolução é robusta e suficiente, no que dissentiu outro, que a vê como extremamente genérica e parca.

No entanto, todos registraram que, em se tratando de inovação e tecnologia de ponta, uma regulamentação minudente, complexa e burocratizada pode inibir, desacelerar, engessar ou mesmo impedir o uso da IA pelo Judiciário. Também consentem que se faz necessária uma participação multidisciplinar na elaboração de modelos de IA, além de que se deve cuidar para promover retroalimentação e reaprendizagem sistemática dos modelos de IA, haja vista a dinamicidade dos fatos sociais.

Quando indagados quanto aos critérios éticos ideais para desenvolvimento de IA pelo Judiciário houve patente descompasso nas entrevistas, de onde se pode extrair que não há entre os entrevistados um liame uníssono ou visão coesa entre eles. O levantamento de critérios foi, portanto, pulverizado, de onde se extrai que há pouca ou quase nenhuma definição concreta e precisa desse contorno ético ideal.

Enquanto um falou em necessidade de uma equipe multidisciplinar, representativa e plural, além de seleção adequada de dados para evitar enviesamento, testes e homologação de modelo, outro entrevistado pontuou que deveriam ser observados valores sociais, raciais e de gênero e, por fim, o terceiro abordou como premente a definição das pessoas atingidas pelo modelo, o *dataset* utilizado e a revisão constante.

Embora não haja um consenso, o que identificamos é que os critérios elencados apontam para a preocupação dos entrevistados no sentido de que a classificação algorítmica seja justa, correta, equânime e não discriminatória.

Em relação às fases, os entrevistados fizeram voz comum ao ressaltarem que o enfoque ético deve se dar em todos os momentos, desde a concepção, extração, transformação e carga de dados (ETL), modelagem, validação e homologação, o que se convencionou chamar de “*ethics by design*”.

Mais que isso, um deles mencionou que abordagem ética deve ser priorizada no pós aplicação do modelo, garantindo-se o acesso a toda a documentação e processo que gerou o conhecimento artificial utilizado no algoritmo que definiu a regra de IA aplicada na sentença que especificamente impactou aquele usuário interessado.

Preocupante, mas realista, foi a afirmação de dois dos entrevistados, que informaram que, diante da inexistência de parâmetro legais concretos, já desenvolveram modelos de IA em

seus tribunais ou órgãos, sem, contudo, qualquer etapa formatada de validação ética do modelo. Segundo um deles, quando muito, exerceram controle ético por intuição, seguindo-se o que os desenvolvedores pensavam como sendo correto e ético.

Nesse contexto, é que entramos no último ponto de análise da entrevista narrativa, qual seja, necessidade ou não da definição de um framework ético.

Primeiro registro – e relevante – é que não há hoje um modelo ou procedimento para validação ética no desenvolvimento de IA, o que foi indicado pelos entrevistados como um dos problemas enfrentados, que causa insegurança, até mesmo um “clima de medo” entre os que operam modelos de IA. Segundo um dos entrevistados, que já desenvolveu duas ferramentas de IA, havia muito foco no Judiciário em usar IA, mas não havia nenhuma indicação de critério ou standard ético em IA.

Segundo um dos entrevistados, da área técnica, esse vácuo leva com que os programadores primeiro desenvolvam, para depois, se for o caso, validarem o modelo em parâmetros éticos. Ocorre que, na ausência desses critérios, nem mesmo validação posterior é possível. No seu sentir, o processo de verificação e validação deveria ser feito antes, concomitante e posterior.

Daí que, segundo todos eles, urge a proposição de uma metodologia para verificação e conferência de conformidade ética, com requisitos objetivos, de simples e concreta checagem, lógica científica, o que permitiria aos tribunais ganhos expressivos, “porque facilita atividade das equipes, otimiza as atividades e simplifica e confere efetividade ao processo de verificação de conformidade ética”.

Também refletiram sobre como esse framework deveria ser implementado: para um dos entrevistados o Tribunal deveria responder a um questionário a cada etapa do processo de desenvolvimento do modelo, “ticando” cada obrigação ética, de maneira que esta conformidade ética seria pré-requisito para avançar à etapa seguinte; para outro, deveria existir validador automatizado de conformidade ética, uma aplicação/ferramenta que fizesse a curadoria e verificação ética do modelo e fornecesse um parecer, com critérios objetivos, sem burocracia.

Em conclusão, é de se registrar que essas entrevistas tiveram como perspectiva, a partir de indução, gerar insights que subsidiem a construção de uma entrevista mais ampla sobre o fenômeno do desenvolvimento da IA pelo Judiciário, com enfoque ético.

Diante desse vazio prático e concreto de parâmetros éticos a nortear o Judiciário, partirmos, então, a navegar pelos referenciais teóricos, normativos, manuais e relatórios, nacionais e transnacionais, produzidos pela academia, pelos governos e pelo mercado, que evidenciem algum contorno de checagem ética ideal, não somente a aplicação da IA no Direito, mas também nos demais ramos da atividade e conhecimento humanos.

4 ÉTICA E IA – MULTIFACETADA ABORDAGEM

Faz-se necessário, para melhor compreensão do tema e para traçar linhas teóricas mínimas a fundar o presente trabalho, realizar análise da ética na IA sob multifacetada abordagem.

É que essa discussão não se circunscreve a determinados espaços: a academia tem focado esforços em pesquisas com enfoque ético; governos e organizações intergovernamentais internacionais dedicam-se e estão construindo ambientes para discussão sobre inovação, ética e responsabilidade em IA; a indústria e o mercado, representados por grandes empresas de tecnologia, de igual forma, buscam coadunar seus interesses econômicos a uma conformidade ética em produtos e serviços de IA, justamente porque já vislumbram que este seja um valor importante na escolha e contratação por seus clientes.

Desta forma, iniciaremos por traçar um panorama atual do cenário internacional de discussões e preocupações com questões éticas associadas ao uso da IA, passando pelas iniciativas governamentais, de mercado/indústria e, por fim, da academia.

4.1 INICIATIVAS GOVERNAMENTAIS E INTERGOVERNAMENTAIS

A recente evolução nos algoritmos e nas técnicas de IA, combinada com um enorme aumento do poder de computação e uma explosão na quantidade de dados disponíveis, tem promovido avanços significativos e tangíveis, que prometem gerar valor para as pessoas, economias e sociedade no geral.

Diante deste cenário, o que se percebeu foi uma corrida internacional não somente para protagonizar o desenvolvimento desta tecnologia, mas também para regular a pesquisa e os projetos, de forma a promover uma rápida e inteligente captura dos seus benefícios, mas também prevenir e mitigar os riscos inerentes.

Países e organismos internacionais multiestatais passaram não somente a estudar os conceitos, alcances e usos desta tecnologia, mas também constituíram grupos de trabalhos, pesquisas e estudos para trazer um ambiente favorável ao desenvolvimento sustentável da IA.

Nota-se, portanto, uma insipiente e tímida regulação, marcada, por um lado, pela preocupação com valores como bem-estar, justiça, equidade, não discriminação, responsabilidade e segurança, e, por outro, por uma difusa e disforme compreensão dos parâmetros éticos e legais que devem ser almejados em projetos e produtos que utilizem IA.

A presente seção objetiva justamente expor como alguns dos entes políticos mundiais estão enfrentando o tema, de forma a subsidiar a discussão final de conformidade ética no Judiciário.

Registre-se, por oportuno, que não houve nessa etapa a pretensão de avaliar se algum dos entes políticos listados abaixo é eficiente na verificação ou validação ética na produção de IA. A escolha se deu pela existência ou não de documentos oficiais sobre o tema IA e ética, assim como pela importância dos países na geopolítica mundial, buscando-se, se possível, representação em todos os continentes.

Aliás, importante, desde já, destacar que as pesquisas realizadas revelam que apenas dois órgãos produziram documentos com enfoque no desenvolvimento ético da IA

especificamente no Judiciário e, portanto, merecem especial atenção, são eles: a Comissão Europeia para a Eficiência da Justiça (CEPEJ), pertencente ao Conselho da Europa, e o Poder Judiciário brasileiro, através do Conselho Nacional de Justiça.

Passemos, pois, à análise de cada uma dessas manifestações.

4.1.1 China

A China editou um plano de desenvolvimento da inteligência artificial, denominado de “A Next Generation Artificial Intelligence Development Plan”.

O documento elege três objetivos estratégicos, a serem buscados em etapas evolutivas, com previsões e diretrizes a serem observadas até, 2020, até 2025 e até 2030.

Segundo o plano, até 2020, a tecnologia geral e a aplicação da IA estarão em sintonia com os níveis globalmente avançados, a indústria da IA terá se tornado um novo e importante ponto de crescimento econômico e as aplicações da tecnologia da IA terão se tornado uma nova maneira de melhorar a vida das pessoas.

Em um segundo momento, até 2025, a China alcançará grandes avanços nas teorias básicas de IA, de modo que algumas tecnologias e aplicações alcançarão um nível de liderança mundial e a IA se tornará a principal força motriz para a modernização industrial e a transformação econômica da China, enquanto a construção social inteligente progresso. Ainda até 2025, uma nova geração de teoria e sistema de tecnologia de IA será inicialmente estabelecida, pois a IA com capacidade de aprendizado autônomo alcança avanços em muitas áreas para obter resultados de pesquisa líderes. A indústria de IA entrará na cadeia de valor global de ponta.

Registra o plano que até 2025, a China terá visto o estabelecimento inicial de leis e regulamentos de IA, normas éticas e sistemas de políticas e a formação de capacidades de avaliação e controle de segurança de IA.

Finalizando os objetivos estratégicos, define o plano que até 2030, as teorias, tecnologias e aplicativos de IA da China devem atingir níveis de liderança mundial, tornando a China o principal centro de inovação de IA do mundo, alcançando resultados visíveis em aplicações de economia inteligente e sociedade inteligente e estabelecendo uma base importante para se tornar uma inovação líder estilo de nação e uma potência econômica. A China terá estabelecido uma série de centros (ou bases) de treinamento de pessoal e inovação em tecnologia de IA líderes mundiais e terá construído leis e regulamentos de IA mais abrangentes e normas e políticas éticas.

Numa análise mais aprofundada do plano, encontra-se o objetivo estratégico de **promover a inteligência da governança social**, que abrange a promoção da aplicação da tecnologia de IA para gestão administrativa, gestão judicial, gestão urbana, proteção ambiental e outras questões relevantes e difíceis na governança social, para promover a modernização da governança social, desdobrando-se em governo inteligente, tribunais inteligentes, cidades inteligentes e transportes inteligentes.

Neste tópico, no que se requiere ao Governo Inteligente, enuncia ao plano que deve a China: desenvolver uma plataforma de IA para serviços governamentais e tomada de decisões;

desenvolver um mecanismo de tomada de decisão para o ambiente aberto; promover aplicações em pesquisa sobre problemas sociais complexos, avaliação de políticas, alerta de risco, resposta a emergências e outros assuntos importantes de tomada de decisão estratégica; fortalecer a integração dos recursos de informação do governo e a previsão precisa das demandas públicas e facilitar os canais de comunicação entre o governo e o público.

Destacamos, até pelo objeto deste trabalho, as previsões de diretrizes para instituição de Tribunais Inteligentes. Neste aspecto, prevê o plano que a China deve buscar: construir um conjunto de julgamento, pessoal, aplicativos de dados, divulgação judicial e monitoramento dinâmico em uma plataforma integrada de dados do tribunal; promova aplicativos de IA para aplicativos, incluindo coleta de evidências, análise de casos e leitura e análise de documentos legais; alcançar a inteligência dos tribunais e sistemas de julgamento e capacidade de julgamento.

Outro ponto de destaque, para finalizar nossa análise do documento, é o levantamento das “Medidas de garantia”. Visando os requisitos realistas de promover o desenvolvimento saudável e rápido da IA na China, é necessário lidar com os possíveis desafios da IA, formar um arranjo institucional para se adaptar ao desenvolvimento da IA, construir um ambiente internacional aberto e inclusivo e reforçar a base social do desenvolvimento da IA.

O primeiro desafio apontado é “Desenvolver leis, regulamentos e normas éticas que promovam o desenvolvimento da IA”, sugerindo as seguintes ações:

1. Fortalecer a pesquisa sobre questões legais, éticas e sociais relacionadas à IA e estabeleça leis, regulamentos e estruturas éticas para garantir o desenvolvimento saudável da IA;
2. Realizar pesquisas sobre questões legais, como confirmação de responsabilidade civil e criminal, proteção de privacidade e propriedade e utilização de segurança da informação relacionada a aplicativos de IA;
3. Estabelecer um sistema de rastreabilidade e responsabilidade e esclareça o corpo principal da IA e os direitos, obrigações e responsabilidades relacionados;
4. Concentrar-se em direção autônoma, robôs de serviço e outros subsectores de aplicativos com uma base de uso comparativamente boa e acelere o estudo e o desenvolvimento de leis e regulamentos relevantes de gerenciamento de segurança, para estabelecer uma base legal para a rápida aplicação de novas tecnologias;
5. Lançar pesquisas sobre ciência e ética do comportamento da IA e outras questões, estabelecer uma estrutura de julgamento de vários níveis éticos e morais e uma estrutura ética de colaboração humano-computador;
6. Desenvolver um código de conduta ética e design de P&D para produtos de IA, fortaleça a avaliação dos riscos e benefícios potenciais da IA e crie soluções para emergências em cenários complexos de IA

O plano se encerra afirmando que a China “participará ativamente da governança global da IA, fortalecerá o estudo dos principais problemas comuns internacionais, como alienação de robôs e supervisão de segurança, aprofundará a cooperação internacional em leis e regulamentos de IA, regras internacionais e assim por diante, e lidará em conjunto com os desafios globais”.

Como se pode perceber dos excertos acima apontados, o plano da próxima geração de IA na China, é documento nitidamente com conteúdo programático, uma vez que estabelece

diretrizes e programas a serem cumpridos pelo legislador, órgãos públicos ou pela indústria da IA, tentando identificar uma trilha, um caminho que o país deve seguir nos próximos anos, sem explicitar, contudo, qualquer conteúdo mais concreto.

Não há sequer definição de princípios propriamente ditos, mas mera indicação de caminho que deve ser percorrido.

Interessante que, embora o país esteja há anos em pleno desenvolvimento de IA nos mais diversos setores da vida social (inclusive na Justiça – o documento cita expressamente), o “estabelecimento inicial de leis e regulamentos de IA, normas éticas e sistemas de políticas e a formação de capacidades de avaliação e controle de segurança de IA” **somente se dará até 2025**. Mais que isso, o plano também deixa claro que **somente até 2030 é que terá construído leis e regulamentos de IA mais abrangentes e normas e políticas éticas**.

Embora seja de conhecimento amplo, divulgado em toda a imprensa internacional, que a China esteja efetivamente utilizando IA para sistemas de Justiça desde 2016 – o que é corroborado pelo plano, quando afirma o estímulo à instituição de Tribunais Inteligentes – o que observamos é que há pouca clareza na indicação objetiva de arcabouço ético. Mais que isso, não há qualquer regulamentação específica para utilização de IA na Justiça.

Outro documento produzido pelo governo Chinês é denominado de “Princípios de Governança para a Nova Geração de Inteligência Artificial - Desenvolvendo Inteligência Artificial Responsável”.

Elaborado pelo Comitê Nacional de Governança para a Nova Geração de Inteligência Artificial, a fim de promover o desenvolvimento saudável da nova geração de IA, melhor equilíbrio entre desenvolvimento e governança, garantir a segurança, confiabilidade e controlabilidade da IA, apoiar os pilares econômico, social e ambiental dos objetivos de desenvolvimento sustentável da ONU e, em conjunto, construir uma comunidade humana com um futuro compartilhado, todas as partes interessadas no desenvolvimento da IA devem observar os seguintes princípios:

1. Harmônica e amigável aos humanos - O objetivo do desenvolvimento da IA deve ser promover o bem-estar da humanidade. Deve estar em conformidade com os valores humanos e princípios éticos, promover a harmonia homem-máquina e servir ao progresso da civilização humana. O desenvolvimento da IA deve basear-se na premissa de garantir a segurança social e respeitar os direitos humanos, e deve evitar o uso indevido, abuso e mau uso da tecnologia de IA por todos os meios.

2. Equidade e Justiça - O desenvolvimento da IA deve promover equidade e justiça, proteger os direitos e interesses de todas as partes interessadas e promover a igualdade de oportunidades. Por meio do avanço da tecnologia e da melhoria da gestão, preconceitos e discriminações devem ser eliminados o máximo possível no processo de aquisição de dados, design de algoritmos, desenvolvimento de tecnologia e desenvolvimento e aplicação de produtos.

3. Inclusão e Compartilhamento - A IA deve promover o desenvolvimento verde para atender aos requisitos de respeito ao meio ambiente e conservação de recursos; A IA deve promover o desenvolvimento coordenado promovendo a transformação e modernização de todas as indústrias e reduzindo as disparidades regionais; A IA deve promover o desenvolvimento inclusivo por meio de melhor educação e treinamento, apoio aos grupos vulneráveis para se adaptarem e esforços para

eliminar a exclusão digital; A IA deve promover o desenvolvimento compartilhado, evitando monopólios de dados e plataformas e promovendo uma concorrência aberta e justa.

4. Respeito à Privacidade - O desenvolvimento da IA deve respeitar e proteger a privacidade dos indivíduos e proteger totalmente os direitos individuais de conhecer e escolher. Limites e regras devem ser estabelecidos para a coleta, armazenamento, processamento e uso de informações pessoais. Os mecanismos de autorização e revogação de privacidade pessoal devem ser estabelecidos e atualizados. Roubar, fazer malabarismos, vazar e outras formas de coleta e uso ilegal de informações pessoais devem ser estritamente proibidos.

5. Segurança e Controlabilidade - A transparência, interpretabilidade, confiabilidade e controlabilidade dos sistemas de IA devem ser aprimoradas continuamente para tornar os sistemas mais rastreáveis, confiáveis e fáceis de auditar e monitorar. A segurança da IA em diferentes níveis dos sistemas deve ser garantida, a robustez da IA e o desempenho anti-interferência devem ser melhorados e as capacidades de avaliação e controle da segurança da IA devem ser desenvolvidas.

6. Responsabilidade Compartilhada - Desenvolvedores de IA, usuários e outras partes interessadas relacionadas devem ter um alto senso de responsabilidade social e autodisciplina, e devem cumprir estritamente as leis, regulamentos, princípios éticos, padrões técnicos e normas sociais. Mecanismos de responsabilidade de IA devem ser estabelecidos para esclarecer as responsabilidades de pesquisadores, desenvolvedores, usuários e partes relevantes. Os usuários de produtos e serviços de IA e outras partes interessadas devem ser informados antecipadamente sobre os possíveis riscos e impactos. O uso de IA para atividades ilegais deve ser estritamente proibido.

7. Abertura e Colaboração - Intercâmbios e cooperação transdisciplinares e transfronteiriços devem ser incentivados no desenvolvimento da IA. Interações coordenadas devem ser fomentadas entre organizações internacionais, agências governamentais, instituições de pesquisa, instituições educacionais, indústrias, organizações sociais e o público em geral no desenvolvimento e governança da IA. Com total respeito pelos princípios e práticas de desenvolvimento da IA em vários países, os diálogos e a cooperação internacionais devem ser incentivados para promover a formação de uma estrutura internacional de governança da IA com amplo consenso.

8. Governança Ágil - A governança da IA deve respeitar os princípios subjacentes ao desenvolvimento da IA. Ao promover o desenvolvimento inovador e saudável da IA, deve-se manter alta vigilância para detectar e resolver possíveis problemas em tempo hábil. A governança da IA deve ser adaptativa e inclusiva, atualizando constantemente o nível de inteligência.

Diferentemente do plano estabelecido pelo governo Chinês, este último documento, ao eleger claramente princípios, dá passos importantes na concretização de princípios éticos aplicáveis a todos os atores envolvidos com IA, alinhando-se aos documentos produzidos pelo Ocidente.

4.1.2 Reino Unido

No Reino Unido, a Câmara dos Lordes, câmara alta do parlamento, por considerar as implicações econômicas, éticas e sociais dos avanços na inteligência artificial, nomeou o Comitê Seletor de Inteligência Artificial, para elaborar estudos e propor medidas.

Publicaram um relatório e estudo intitulado de “IA no Reino Unido: pronta, disposta e capaz?”. Registrando que, por enquanto, ainda falta clareza sobre como a IA pode ser melhor usada para beneficiar os indivíduos e a sociedade, a Câmara dos Lordes propõe cinco princípios que podem se tornar a base para uma estrutura ética compartilhada em IA.

Embora a regulamentação específica de IA não seja apropriada nesta fase, tal estrutura fornece clareza a curto prazo e pode sustentar a regulamentação, se for necessária, no futuro. Os reguladores existentes estão mais indicados a regular a IA em seus respectivos setores. Devem ser dotados de recursos adequados e poderes para tal.

Estão no centro do relatório cinco recomendações de princípios éticos que podem ser aplicados a todos os setores, nacional e internacionalmente:

1. a inteligência artificial deve ser desenvolvida para o bem comum e o benefício da Humanidade;
2. devem operar nos princípios da inteligibilidade e justiça;
3. não devem ser usadas para diminuir a proteção de dados ou a privacidade dos indivíduos, famílias e comunidades;
4. todos os cidadãos devem ter o direito a serem educados para que possam florescer mentalmente, emocionalmente e economicamente, junto com a inteligência artificial;
5. forças autônomas capazes de ferir, destruir ou enganar os seres humanos não devem receber investimentos da inteligência artificial.

Escrito em formato de relatório, em cada ponto enumera conclusões e indica recomendações direcionadas para cada ator (órgãos governamentais ou não), porém o documento não identifica explicitamente e de modo objetivo os princípios éticos.

O texto, que parece ser mais introdutório aos conceitos da ciência de dados, é em sua maior parte descritivo, relatando os trabalhos da comissão e o que cada testemunha trouxe de informação ao levantamento de dados para o documento.

Pouco se fala em Judiciário, em lei ou em juízes, trazendo meramente o exemplo do COMPASS e do Big Brother Watch (polícia de Durham – UK).

4.1.3 Brasil

4.1.3.1 EBIA

Instituída pela Portaria MCTI nº 4.617, de 6 de abril de 2021, alterada pela Portaria MCTI nº 4.979, de 13 de julho de 2021, a Estratégia Brasileira de Inteligência Artificial - EBIA assume o papel de nortear as ações do Estado brasileiro em prol do desenvolvimento das ações, em suas várias vertentes, que estimulem a pesquisa, inovação e desenvolvimento de soluções em Inteligência Artificial, bem como, seu uso consciente, ético e em prol de um futuro melhor.

Alinhada às diretrizes da OCDE endossadas pelo Brasil, a EBIA fundamenta-se nos cinco princípios definidos pela Organização para uma gestão responsável dos sistemas de IA, quais sejam: (i) crescimento inclusivo, o desenvolvimento sustentável e o bem-estar; (ii) valores

centrados no ser humano e na equidade; (iii) transparência e explicabilidade; (iv) robustez, segurança e proteção e; (v) a responsabilização ou a prestação de contas (*accountability*).

A Estratégia foi construída em três etapas: (i) contratação de consultoria especializada em IA; (ii) benchmarking nacional e internacional; e (iii) processo de consulta pública.

A Consulta Pública à sociedade por realizada por meio de plataforma eletrônica do Governo Federal, entre 12 de dezembro de 2019 e 3 de março de 2020. Foram recebidas cerca de 1.000 contribuições serviram como base para a construção da presente estratégia.

A Estratégia Brasileira de Inteligência Artificial (EBIA) tem por objetivo potencializar o desenvolvimento e a utilização da tecnologia com vistas a promover o avanço científico e solucionar problemas concretos do País, identificando áreas prioritárias nas quais há maior potencial de obtenção de benefícios. Espera-se que a IA possa trazer ganhos na promoção da competitividade e no aumento da produtividade brasileira, na prestação de serviços públicos, na melhoria da qualidade de vida das pessoas e na redução das desigualdades sociais, entre outros.

A EBIA estabelece nove eixos temáticos, caracterizados como os pilares do documento, apresenta um diagnóstico da situação atual da IA no mundo e no Brasil, destaca os desafios a serem enfrentados, oferece uma visão de futuro e apresenta um conjunto de ações estratégicas que nos aproximam dessa visão.

A EBIA tem como objetivos:

1. Contribuir para a elaboração de princípios éticos para o desenvolvimento e uso de IA responsáveis;
2. Promover investimentos sustentados em pesquisa e desenvolvimento em IA;
3. Remover barreiras à inovação em IA;
4. Capacitar e formar profissionais para o ecossistema da IA;
5. Estimular a inovação e o desenvolvimento da IA brasileira em ambiente internacional;
6. Promover ambiente de cooperação entre os entes públicos e privados, a indústria e os centros de pesquisas para o desenvolvimento da Inteligência Artificial;

Para tanto, a Estratégia estabelece nove eixos temáticos, caracterizados como os pilares do documento; apresenta um diagnóstico da situação atual da IA no mundo e no Brasil; destaca os desafios a serem enfrentados; oferece uma visão de futuro; e apresenta um conjunto de ações estratégicas que nos aproximam dessa visão.

Eixos Transversais

1. Legislação, regulação e uso ético.
2. Governança de Inteligência Artificial.
3. Aspectos Internacionais.

Eixos Verticais

1. Qualificações para um futuro digital.
2. Força de trabalho e capacitação.
3. Pesquisa, desenvolvimento, inovação e empreendedorismo.

4. Aplicação nos setores produtivos.
5. Aplicação no poder público.
6. Segurança pública.

Conforme Portaria MCTI nº 4.617, de 6 de abril de 2021, caberá ao Ministério de Ciência, Tecnologia e Inovações, no âmbito de suas atribuições:

I - Criar instâncias e práticas de governança para priorizar, implantar, monitorar e atualizar as ações estratégicas estabelecidas na Estratégia Brasileira de Inteligência Artificial;

II - Coordenar e estabelecer as ações que possibilitem a implementação da Estratégia Brasileira de Inteligência Artificial;

III - convidar instituições do setor público, privado e da academia para subsidiá-lo nas ações estratégicas definidas na Estratégia Brasileira de Inteligência Artificial;

IV - Elaborar relatórios de avaliação da implementação das ações estratégicas definidas na Estratégia Brasileira de Inteligência Artificial;

V - Dar publicidade aos relatórios de que trata o inciso III do caput deste artigo no sítio eletrônico deste Ministério; e

VI - Avaliar a periodicidade de atualização da Estratégia Brasileira de Inteligência Artificial.

A Governança da EBIA é composta por:

- I - Ministério da Ciência, Tecnologia e Inovações – MCTI;
- II - Rede MCTI/EMBRAPII de Tecnologias e Inovação Digital; e
- III - Instituições Convidadas.

A Estratégia, como é chamada, traz um importante farol para o desenvolvimento de IA no Brasil. Embora não se dedique à aplicação da IA no Judiciário especificamente, traz uma série de contribuições para o desenvolvimento ético em IA.

É de se ressaltar o conceito de *ethics by design* – segundo o qual a ética deve ser considerada em todas as etapas do desenvolvimento e no uso da IA. Revela-se, assim, a importância de criação de rotinas de gestão de riscos, de monitoramento e de supervisão quanto ao uso de sistemas de IA ao longo de todo o seu ciclo de vida, assim como para a necessidade de que tais princípios sejam incorporados desde o momento da concepção do sistema (*privacy by design, security by design, human rights by design, ethics by design*).

Interessantes as ações estratégicas identificadas no decorrer de todo o texto, aqui pinçadas:

1. estabelecer valores éticos para uso da IA na Administração Pública Federal;
2. estimular a produção de IA ética financiando projetos de pesquisa que visem aplicar soluções éticas, principalmente nos campos de equidade/ não-discriminação / Justiça (*fairness*), responsabilidade/prestação de contas (*accountability*) e transparência (*transparency*), conhecidas como a matriz FAT;

3. estimular parcerias com corporações que estejam pesquisando soluções comerciais dessas tecnologias de IA ética;
4. estabelecer como requisito técnico em licitações que os proponentes ofereçam soluções compatíveis com a promoção de uma IA ética (por exemplo, estabelecer que soluções de tecnologia de reconhecimento facial adquiridas por órgãos públicos possuam um percentual de falso positivo abaixo de determinado limiar);
5. estabelecer, de maneira multissetorial, espaços para a discussão e definição de princípios éticos a serem observados na pesquisa, no desenvolvimento e no uso da IA;
6. estimular que as organizações criem conselhos de revisão de dados ou comitês de ética em relação à IA;
7. estimular ações de transparência e de divulgação responsável quanto ao uso de sistemas de IA, e promover a observância, por tais sistemas, de direitos humanos, de valores democráticos e da diversidade;
8. desenvolver técnicas para identificar e tratar o risco de viés algorítmico;
9. elaborar política de controle de qualidade de dados para o treinamento de sistemas de IA;
10. promover abordagens inovadoras para a supervisão regulatória (por exemplo, *sandboxes* e hubs regulatórios).

O item 8 da Estratégia se dedica à aplicação de IA no Setor Público e, portanto, tem especial atenção, dado o objeto deste estudo.

Registra a Estratégia que, “não por acaso, dos 50 países com estratégias nacionais de IA, 36 possuem estratégias específicas para a transformação do setor público por meio da IA ou têm foco no setor público dentro de uma estratégia mais ampla”.

Algumas questões centrais contidas em tais estratégias são:

1. Colaboração entre diferentes setores, inclusive por meio de parcerias público-privadas, facilitada por hubs e por laboratórios de inovação.
2. Criação de conselhos, redes e comunidades envolvendo diferentes áreas do governo.
3. Automação de processos rotineiros para aumentar a eficiência.
4. Uso de IA para apoiar processos de tomada de decisão.
5. Gestão estratégica e abertura de dados governamentais, inclusive para alavancar IA no setor privado.
- 6. Orientações quanto ao uso transparente e ético de IA no setor público”.**

Analisando a utilização da inteligência artificial como um motor propulsor para efetiva entrega de serviços públicos eficientes, enaltece o documento que “No campo do Poder Público, a Inteligência Artificial representa importante oportunidade para melhorias na qualidade de atendimento e nos índices de satisfação dos cidadãos. A inovação na gestão pública, ao modernizar os processos administrativos, possibilita que o Estado supere obstáculos burocráticos e restrições orçamentárias para oferecer serviços novos, melhores e mais eficientes à população”.

Demonstra, entretanto, preocupação com a gestão e tratamento dos riscos envolvidos com o uso da IA pelos Estado, de forma que “ao mesmo tempo em que a IA traz inúmeros

benefícios, como o aumento de eficiência dos serviços públicos, é fundamental que os valores éticos de uma IA justa e responsável também estejam refletidas nas soluções adquiridas e/ou desenvolvidas pelo Poder Público”.

A Estratégia enumera aplicações de IA no setor público, enaltecendo trabalhos em diversos instituições e órgãos públicos, do executivo, legislativo de judiciário, tais como no TCU, CGU e, na Justiça, menciona STF (Victor), e TJPE (Elis).

4.1.3.2 Resolução 332 do CNJ

A preocupação do Poder Judiciário brasileiro com a pesquisa, uso e desenvolvimento de IA pelos órgãos da Justiça remonta ao início de 2019, quando foi editada a Portaria nº 25/2019 onde se obrigou que os projetos de IA deveriam observar certas premissas definidas no item 4 do Anexo à Portaria nº 25/2019, dentre as quais incluía-se a “validação jurídica e ética dos modelos”, assim definida no item 4-c: “Os modelos de IA que forem utilizados na tomada de decisões ou produção de artefatos deverão ser passíveis de auditoria para análise dos resultados a partir de critérios éticos jurídicos. O processo de auditoria será definido pelo CNJ”.

Mas somente em meados de 2020 é que importante passo foi dado. Com protagonismo e ineditismo, a Justiça brasileira, através do Conselho Nacional de Justiça (CNJ), órgão do Poder Judiciário, editou Resolução nº 332/2020, que dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário.

A resolução baseia-se na Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes, conta com 31 artigos e 11 páginas, e é dividida nos seguintes capítulos: Respeito a direitos fundamentais; não discriminação; Publicidade e transparência; Governança e qualidade; Segurança; Controle do usuário; Pesquisa, desenvolvimento, implantação de serviços de IA; Prestação de contas e responsabilização.

A própria resolução, no 11º considerando, registra a ausência, no Brasil, de normas específicas quanto à governança e aos parâmetros éticos para o desenvolvimento e o uso de IA.

Embora a resolução se destine à ética na produção e no uso de inteligência artificial, não há clara definição de quais os critérios seriam considerados efetivamente numa validação ética.

Daí que, da leitura do da norma, surgem algumas dúvidas: É necessário ao usuário percorrer e atender a todos os standards listados no decorrer da resolução? Cada um dos capítulos listados acima pode/deve ser considerado como um dos tipos de verificação ética a se atentar? Ou é possível deduzir que o critério ético se restringe a “transparência, previsibilidade, possibilidade de auditoria e garantia de imparcialidade e justiça substancial” (conforme considerando)?

Não se permite ao leitor ou ao usuário compreender exatamente qual o contorno da ética no uso de IA no Judiciário, nem tampouco há clareza nos elementos que precisam ser verificados para ter validação ética.

Aliás, o nome “ética”, utilizado no título da norma, tem mais cinco inserções no texto, quatro nos “considerandos” e uma no art. 22, §1º, que dispõe:

“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”.

Também importa acrescentar que após a edição da Res. 332/2021-CNJ, o mesmo Conselho editou a Portaria Nº 271 de 04/12/2020, que regulamenta o uso de Inteligência Artificial no âmbito do Poder Judiciário.

De acordo com o Art. 3º, a pesquisa e desenvolvimento em matéria de inteligência artificial observará:

- I – economicidade, evitando-se a realização de pesquisas e projetos com conteúdo semelhante em diferentes órgãos, sem colaboração e interação;
- II –promoção da interoperabilidade tecnológica dos sistemas processuais eletrônicos do Poder Judiciário;
- III –adoção de tecnologias, padrões e formatos abertos e livres;
- IV –acesso à informação;
- V –transparência;
- VI –capacitação humana e sua preparação para a reestruturação dos fluxos processuais e de trabalho, à medida que a inteligência artificial é implantada;
- VII –foco na celeridade processual; e
- VIII – estabelecimento de mecanismos de governança colaborativa e democrática, com a participação do Poder Judiciário, daqueles que exercem funções essenciais à justiça, da comunidade acadêmica e da sociedade civil.

Também se definiu que o uso de inteligência artificial no âmbito do Poder Judiciário se dará em plataforma comum, acessível por todos, que incentive a colaboração, a transparência, o aprimoramento e a divulgação dos projetos. Estabeleceu que a plataforma de inteligência artificial do Poder Judiciário Nacional é o Sinapses, disponibilizada pelo CNJ em parceria com o Tribunal de Justiça do Estado de Rondônia.

Com essas novas regras, tornou-se obrigatório o depósito, na Plataforma Sinapses, de todos os modelos de IA produzidos no âmbito do Poder Judiciário (cf. art. 10, inciso III, da Resolução CNJ).

Nesse contexto, embora mereça comemoração que o Judiciário brasileiro tenha se dedicado a regular o uso ético da IA por seus tribunais e juízes, o fato é que há importante lacuna aberta, no que concerne à definição do contorno ético esperado, o que dificulta sobremaneira a aplicação da resolução e portarias no cotidiano dos desenvolvedores, operadores do direito e do cidadão.

4.1.3.3 PL 21/2020

Outro importante registro, no que se refere ao caso brasileiro, é o Projeto de Lei nº 21/2020, de autoria do Deputado Eduardo Bismarck (PDT-CE), que já fora aprovada na Câmara dos Deputados e segue em tramitação no Senado Federal.

O projeto estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil.

De acordo com a proposta, a aplicação de inteligência artificial no Brasil tem por objetivo o desenvolvimento científico e tecnológico, bem como:

1. A promoção do desenvolvimento econômico sustentável e inclusivo e do bem-estar da sociedade;
2. O aumento da competitividade e da produtividade brasileira;
3. A inserção competitiva do Brasil nas cadeias globais de valor;
4. A melhoria na prestação de serviços públicos e na implementação de políticas públicas;
5. A promoção da pesquisa e desenvolvimento com a finalidade de estimular a inovação nos setores produtivos; e
6. A proteção e a preservação do meio ambiente.

Dentre os objetivos elencados no art. 3º da norma, é de se destacar o foco na prestação de serviços públicos e implementação de políticas públicas, haja vista que o legislador brasileiro optou por escolher a tecnologia da IA como aderente à definição de políticas públicas adequadas, bem assim para implementar e melhorar os serviços públicos prestados pelo Estado brasileiro, seja ele pelo Executivo, Legislativo ou Judiciário.

Dáí se extrai importante reconhecimento da adoção da inteligência computacional como instrumento para consecução do interesse público.

Além disso, estabelece o projeto que o respeito à ética, aos direitos humanos e aos valores democráticos é um dos fundamentos do desenvolvimento e aplicação de IA no Brasil (art. 4º, III).

Interessante notar que a proposição legislativa enumera diretrizes para observância do poder público quando da sua atividade regulatória, mas também quando da sua atuação quanto ao uso e fomento de sistemas de IA. Dada a relevância e a didática do texto, permitimo-nos transcrevê-lo:

“Art. 6º Constituem diretrizes para a atuação da União, dos Estados, do Distrito Federal e dos Municípios em relação ao uso e ao fomento dos sistemas de inteligência artificial no Brasil:

I – Intervenção subsidiária: regras específicas deverão ser desenvolvidas para os usos de sistemas de inteligência artificial apenas quando absolutamente necessárias para a garantia do atendimento ao disposto na legislação vigente;

II – Atuação setorial: a atuação do poder público deverá ocorrer pelo órgão ou entidade competente, considerados o contexto e o arcabouço regulatório específicos de cada setor;

III – gestão baseada em risco: o desenvolvimento e o uso dos sistemas de inteligência artificial deverão considerar os riscos concretos, e as definições sobre a necessidade de regulação dos sistemas de inteligência artificial e sobre o respectivo grau de intervenção deverão ser sempre proporcionais aos riscos concretos oferecidos por cada sistema e à probabilidade de ocorrência desses riscos, avaliados sempre em comparação com: a) os potenciais benefícios sociais e

econômicos oferecidos pelo sistema de inteligência artificial; e b) os riscos apresentados por sistemas similares que não envolvam inteligência artificial, nos termos do inciso V deste caput;

IV – Participação social e interdisciplinar: a adoção de normas que impactem o desenvolvimento e a operação de sistemas de inteligência artificial será baseada em evidências e precedida de consulta pública, realizada preferencialmente pela internet e com ampla divulgação prévia, de modo a possibilitar a participação de todos os interessados e as diversas especialidades envolvidas;

V – Análise de impacto regulatório: a adoção de normas que impactem o desenvolvimento e a operação de sistemas de inteligência artificial será precedida de análise de impacto regulatório, nos termos do Decreto nº 10.411, de 30 de junho de 2020, e da Lei nº 13.874, de 20 de setembro de 2019; e

VI – responsabilidade: as normas sobre responsabilidade dos agentes que atuam na cadeia de desenvolvimento e operação de sistemas de inteligência artificial deverão, salvo disposição legal em contrário, pautar-se na responsabilidade subjetiva e levar em consideração a efetiva participação desses agentes, os danos específicos que se deseja evitar ou remediar e a forma como esses agentes podem demonstrar adequação às normas aplicáveis, por meio de esforços razoáveis compatíveis com os padrões internacionais e as melhores práticas de mercado.

§ 1º Na gestão com base em risco a que se refere o inciso III do caput deste artigo, a administração pública, nos casos de baixo risco, deverá incentivar a inovação responsável com a utilização de técnicas regulatórias flexíveis.

§ 2º Na gestão com base em risco a que se refere o inciso III do caput deste artigo, a administração pública, nos casos concretos em que se constatar alto risco, poderá, no âmbito da sua competência, requerer informações sobre as medidas de segurança e prevenção enumeradas no inciso VI do caput do art. 5º desta Lei, e respectivas salvaguardas, nos termos e nos limites de transparência estabelecidos por esta Lei, observados os segredos comercial e industrial.

§ 3º Quando a utilização do sistema de inteligência artificial envolver relações de consumo, o agente responderá independentemente de culpa pela reparação dos danos causados aos consumidores, no limite de sua participação efetiva no evento danoso, observada a Lei nº 8.078 de 11 de setembro de 1990 (Código de Defesa do Consumidor).

§ 4º As pessoas jurídicas de direito público e as de direito privado prestadoras de serviços públicos responderão pelos danos que seus agentes, nessa qualidade, causarem a terceiros, assegurado o direito de regresso contra o responsável nos casos de dolo ou culpa.

Art. 7º Constituem diretrizes para a atuação da União, dos Estados, do Distrito Federal e dos Municípios em relação ao uso e ao fomento dos sistemas de inteligência artificial no Brasil:

I – Promoção da confiança nas tecnologias de inteligência artificial, com disseminação de informações e de conhecimento sobre seus usos éticos e responsáveis;

II – Incentivo a investimentos em pesquisa e desenvolvimento de inteligência artificial;

III – promoção da interoperabilidade tecnológica dos sistemas de inteligência artificial utilizados pelo poder público, de modo a permitir o intercâmbio de informações e a celeridade de procedimentos;

IV – Incentivo ao desenvolvimento e à adoção de sistemas de inteligência artificial nos setores público e privado;

V – Estímulo à capacitação e à preparação das pessoas para a reestruturação do mercado de trabalho;

VI – Estímulo a práticas pedagógicas inovadoras, com visão multidisciplinar, e ênfase da importância de ressignificação dos processos de formação de professores para lidar com os desafios decorrentes da inserção da inteligência artificial como ferramenta pedagógica em sala de aula;

VII – estímulo à adoção de instrumentos regulatórios que promovam a inovação, como ambientes regulatórios experimentais (sandboxes regulatórios), análises de impacto regulatório e autorregulações setoriais;

VIII – estímulo à criação de mecanismos de governança transparente e colaborativa, com a participação de representantes do poder público, do setor empresarial, da sociedade civil e da comunidade científica; e

IX – Promoção da cooperação internacional, mediante estímulo ao compartilhamento do conhecimento sobre sistemas de inteligência artificial e à negociação de tratados, acordos e padrões técnicos globais que facilitem a interoperabilidade entre os sistemas e a harmonização da legislação a esse respeito.

Parágrafo único. Para fins deste artigo, o poder público federal promoverá a gestão estratégica e as orientações quanto ao uso transparente e ético de sistemas de inteligência artificial no setor público, conforme as políticas públicas estratégicas para o setor”.

Assim como tem se observado em diversos diplomas legais e em propostas de normatização, embora se eleja a abordagem ética como relevante, não se faz descrição, conceituação, categorização ou qualquer especificação quanto ao seu conteúdo.

É de se destacar, por fim, que a proposta de lei estabeleceu estratégia regulatória interessante, dando-se primazia à autorregulação, mediante adoção de códigos de conduta e de guias de boas práticas, que podem servir como elementos indicativos de conformidade, além do necessário impacto regulatório, o que conferem simplicidade e plasticidade necessários em ambiente de inovação tecnológica. Afora isso, diretrizes como o da intervenção subsidiária e atuação setorial permitem escolhas diferentes, flexíveis e adequadas, a depender do ramo ou setor que se utiliza da IA. Ainda, a noção de gestão baseada em riscos, mas também com olhar nos benefícios trazem maior segurança e o encorajamento necessário aos gestores públicos. Por derradeiro, a definição de responsabilidade civil subjetiva em regra – ressalvadas apenas as hipóteses de aplicação do Código de Defesa do Consumidor e de responsabilidade das pessoas jurídicas de direito público – também é tema que confere maior segurança no ambiente de desenvolvimento.

4.1.3.4 PL 2338/2023

O projeto foi concebido por comissão de juristas criada em 30 de março de 2022, instaurada no Senado Federal, denominada de CJSUBIA, Comissão de Juristas responsável pela elaboração de um projeto substitutivo para regulamentação da IA no Brasil.

Partindo das referências dos Projetos de Lei n°s 5.051, de 2019, 21, de 2020, e 872, de 2021, a CJSUBIA que teve como objetivo estabelecer princípios, regras, diretrizes e fundamentos para regular o desenvolvimento e a aplicação da inteligência artificial no Brasil.

Em maio deste ano fora apresentado novo projeto de lei, que segue a linha da proposta pela União Europeia (AI Act) para regulação do tema, ao prever o detalhamento de obrigações e responsabilidades aos fornecedores e operadores de IA no próprio texto da lei, o que difere completamente da abordagem principiológica do PL 21/20, oriundo da Câmara dos Deputados, e dá ênfase na regulação setorial e forte investimento privado, nos moldes preconizados pelos EUA.

O projeto dispõe sobre o uso da inteligência artificial tem como objetivos proteger os direitos fundamentais e garantir a implementação de sistemas seguros e confiáveis, em benefício da pessoa humana, do regime democrático e do desenvolvimento científico e tecnológico (art. 1º).

O artigo 2º da proposta traz fundamentos O desenvolvimento, a implementação e o uso de sistemas de inteligência artificial no Brasil, dentre eles:

- I – a centralidade da pessoa humana;
- II – o respeito aos direitos humanos e aos valores democráticos;
- III – o livre desenvolvimento da personalidade;
- IV – a proteção ao meio ambiente e o desenvolvimento sustentável;
- V – a igualdade, a não discriminação, a pluralidade e o respeito aos direitos trabalhistas;
- VI – o desenvolvimento tecnológico e a inovação;
- VII – a livre iniciativa, a livre concorrência e a defesa do consumidor;
- VIII – a privacidade, a proteção de dados e a autodeterminação informativa;
- IX – a promoção da pesquisa e do desenvolvimento com a finalidade de estimular a inovação nos setores produtivos e no poder público; e
- X – o acesso à informação e à educação, e a conscientização sobre os sistemas de inteligência artificial e suas aplicações.

Também elenca princípios, a seguir discriminados:

- I – crescimento inclusivo, desenvolvimento sustentável e bem-estar;
- II – autodeterminação e liberdade de decisão e de escolha;
- III – participação humana no ciclo da inteligência artificial e supervisão humana efetiva;
- IV – não discriminação;
- V – justiça, equidade e inclusão;
- VI – transparência, explicabilidade, inteligibilidade e auditabilidade;
- VII – confiabilidade e robustez dos sistemas de inteligência artificial e segurança da informação;
- VIII – devido processo legal, contestabilidade e contraditório;
- IX – rastreabilidade das decisões durante o ciclo de vida de sistemas de inteligência artificial como meio de prestação de contas e atribuição de responsabilidades a uma pessoa natural ou jurídica;
- X – prestação de contas, responsabilização e reparação integral de danos;
- XI – prevenção, precaução e mitigação de riscos sistêmicos derivados de usos intencionais ou não intencionais e de efeitos não previstos de sistemas de inteligência artificial; e

XII – não maleficência e proporcionalidade entre os métodos empregados e as finalidades determinadas e legítimas dos sistemas de inteligência artificial.

Afora esse núcleo de objetivos, fundamentos e princípios, o projeto também traz regramentos sobre direitos dos usuários, avaliação preliminar e classificação de riscos governança dos sistemas de IA, avaliação de impacto algorítmico, boas práticas de governança, comunicação de incidentes, responsabilidade civil, supervisão, fiscalização e sanções, além de medidas para fomentar a inovação.

4.1.4 Singapura

No discurso global sobre ética e governança da IA, Singapura acredita que sua abordagem equilibrada pode facilitar a inovação, proteger os interesses do consumidor e servir como um ponto de referência global comum.

Em 25/05/2022 foi criado o AI VERIFY - Estrutura e kit de ferramentas de teste de governança de IA, que está em MVP, ou seja, o produto viável mínimo, que é a versão mais simples de um produto que pode ser lançada com uma quantidade mínima de esforço e desenvolvimento.

A primeira estrutura e kit de ferramentas de teste de governança de IA do mundo para empresas que desejam demonstrar IA responsável de maneira objetiva e verificável, visa promover a transparência entre as empresas e seus stakeholders.

O Verify empacota um conjunto de soluções de teste de código aberto, incluindo um conjunto de verificações de processo em um Toolkit para uma autoavaliação conveniente. O Toolkit gerará relatórios para desenvolvedores, gerentes e parceiros de negócios, cobrindo as principais áreas que afetam o desempenho da IA.

Diversas empresas de setores diferentes foram convidadas a testar e dar feedback.

O AI VERIFY tem os seguintes princípios orientadores:

1. As organizações que usam IA na tomada de decisões devem garantir que o processo de tomada de decisão seja explicável, transparente e justo;
2. As soluções de IA devem ser centradas no ser humano.

As organizações devem detalhar um conjunto de princípios éticos quando embarcam na implantação de IA em escala dentro de seus processos ou para capacitar seus produtos e/ou serviços.

O governo de Singapura lançou a primeira edição do “Quadro Modelo para Governança de IA (Model AI Governance Framework)”, que busca traduzir princípios éticos em práticas implementáveis no processo de desenvolvimento de IA.

É um guia completo que permite às mais diversas organizações e pessoas (desenvolvedores ou não) que compreendam os impactos da aplicação da IA, trazendo diversas seções com conteúdo prático, “encorajamentos”.

Em texto extremamente didático, repleto de exemplos, além de ilustrações de aplicações dos conceitos em organizações públicas e privadas, o documento de Singapura traz uma visão mais prática e sistemática, elegendo standards éticos para aplicação de IA.

O principal contributo do documento de Singapura é justamente a concretude. Isto porque estabelece, no seu Anexo A, uma compilação de princípios éticos de IA existentes, elegendo-os como de necessária observância para obtenção de validação ética.

Afirma o texto que, onde necessário, organizações podem querer consultar a compilação de princípios éticos de IA princípios do Anexo A. Na medida do possível, as organizações devem também rever seus valores corporativos existentes e incorporar a princípios éticos por eles articulados. Algumas das questões éticas princípios (por exemplo, segurança) podem ser articulados como riscos que podem ser incorporados à estrutura de gerenciamento de riscos corporativos.

O Modelo de Estrutura foi projetado para ajudar as organizações, incorporando princípios éticos em práticas familiares e pré-existentes, estruturas de governança corporativa e, assim, ajudar a orientar a adoção da IA em uma organização.

O “ANEXO A - PARA REFERÊNCIA: UMA COMPILAÇÃO DE PRINCÍPIOS ÉTICOS DE IA EXISTENTES”, compreende uma coleção de princípios éticos fundamentais de IA, provenientes de várias fontes (IEEE, SIIA, FATML, Declaração Europeia e OCDE), escolhidos como necessários para validação ética de um projeto de IA, quais sejam:

1. Controle/prestação de contas/Responsabilidade (*accountability*): Garantir que os atores de IA sejam responsáveis e prestem contas pelo funcionamento dos sistemas de IA e pelo respeito à ética e aos princípios da IA, com base em seus papéis, o contexto e a consistência com o estado da arte;
2. Precisão (*accuracy*): Identifique, registre e articule fontes de erro e incerteza em todo o algoritmo e suas fontes de dados para que as implicações esperadas e de pior caso podem ser compreendidos e podem informar os procedimentos de mitigação;
3. Auditabilidade (*auditability*): Permitir que terceiros interessados investiguem, compreendam e revisem o comportamento do algoritmo por meio da divulgação de informações que possibilitam o monitoramento, verificação ou crítica;
4. Explicação (*explainability*): Garantir que decisões automatizadas e algorítmicas e quaisquer os dados que orientam essas decisões podem ser explicados aos usuários finais e outras partes interessadas em termos não técnicos
5. Equidade/Imparcialidade/Justiça (*fairness*):
 - a) garantir que as decisões algorítmicas não criem impactos discriminatórios ou injustos em diferentes linhas demográficas (por exemplo, raça, sexo, etc.);
 - b) desenvolver e incluir mecanismos de monitoramento e contabilidade para evitar discriminação não intencional na implementação de sistemas de tomada de decisão;
 - c) para consultar uma diversidade de vozes e dados demográficos ao desenvolver sistemas, aplicações e algoritmos.

6. Centricidade Humana e Bem-estar (*human centricity and well-being*);
 - a. Visar uma distribuição equitativa dos benefícios das práticas de dados e evitar que o uso dos dados prejudique desproporcionalmente os grupos vulneráveis;
 - b. Visar criar o maior benefício possível do uso de dados e técnicas de modelagem avançadas;
 - c. Envolver-se em práticas de dados que incentivem a prática de virtudes que contribuem para florescimento humano, dignidade humana e autonomia humana;
 - d. Para dar peso aos julgamentos considerados de pessoas ou comunidades afetadas pelas práticas de dados e estar alinhado com os valores e princípios éticos das pessoas ou comunidades afetadas;
 - e. Para tomar decisões que não devem causar danos previsíveis ao indivíduo, ou deve pelo menos minimizar esse dano (em circunstâncias necessárias, quando pesado contra o bem maior);
 - f. Para permitir que os usuários mantenham o controle sobre os dados que estão sendo usados, o contexto em que estão sendo usados e a capacidade de modificar esse uso e contexto;
 - g. Para garantir que o bem-estar geral do usuário seja central para a funcionalidade do sistema de IA;
7. Alinhamento dos direitos humanos: Garantir que a concepção, desenvolvimento e implementação de tecnologias não infringem os direitos humanos internacionalmente reconhecidos;
8. Inclusão (*inclusivity*): Garanta que a IA seja acessível a todos;
9. Progressividade (*progressiveness*): Favorecer implementações onde o valor criado é materialmente melhor do que não se envolver nesse projeto;
10. Responsabilidade, prestação de contas e transparência (*responsability, accountability and transparency*):
 - a. Construir confiança, garantindo que designers e operadores sejam responsáveis e prestem contas de seus sistemas, aplicativos e algoritmos, e para garantir que tais sistemas, aplicativos e algoritmos operem de forma transparente e justa;
 - b. Disponibilizar vias de reparação de danos externamente visíveis e imparciais para efeitos individuais ou sociais de um sistema de decisão algorítmico, e para designar a função a uma pessoa ou escritório responsável pela solução oportuna de tais questões;
 - c. Incorporar medidas e processos a jusante para que usuários ou partes interessadas verifiquem como e quando a tecnologia de IA está sendo aplicada.
 - d. Para manter registros detalhados dos processos de design e tomada de decisão.

11. Robustez e Segurança (*robustness and security*): Os sistemas de IA devem ser seguros e protegidos, não vulneráveis a adulteração ou comprometimento dos dados em que são treinados;

12. Sustentabilidade (*sustainability*): Favorecer implementações que efetivamente prevejam o comportamento futuro e gerar insights benéficos durante um período de tempo razoável.

Esclarece o texto que o Anexo A pode ser utilizado como um quadro-base, sem prejuízo que organizações considerem incorporar esses princípios em seus próprios princípios corporativos, onde relevante e desejado.

Em seção específica, o IA VERIFY busca ajudar organizações a tomar passos para construir confiança nas estratégias de relacionamento com as partes interessadas ao implantar a IA, que passa pelo fornecimento de informações gerais sobre se a IA é usada em seus produtos e serviços, definição de políticas, explicação a serem fornecidas, explicabilidade e transparência, opção de exclusão, teste de interface, etc.

Essa provocação também precisa ser feita no âmbito do Judiciário, a fim de se indagar em qual medida todos os atores envolvidos no processo judicial participam ou podem participar da construção das ferramentas de IA. Mais que isso, qual a importância dessa participação democrática para construção de inovação em IA.

4.1.5 União Europeia

A Comissão Europeia para a Eficiência da Justiça (CEPEJ), pertencente ao Conselho da Europa adotou o primeiro texto europeu que estabelece princípios éticos relativos ao uso de inteligência artificial (IA) em sistemas judiciais.

A Comissão Europeia Pela Eficiência da Justiça (CEPEJ, do francês *Commission européenne pour l'efficacité de la justice*) faz parte do Conselho da Europa, sediado em Estrasburgo, na França. O Conselho tem como propósito a defesa dos Direitos Humanos, da Democracia e a Justiça no continente europeu. Foi fundada em 1949 e hoje tem 47 Estados filiados, incluindo os 28 integrantes da União Europeia.

A Carta fornece uma estrutura de princípios que podem orientar os formuladores de políticas, legisladores e profissionais da justiça quando lidam com o rápido desenvolvimento da IA nos processos judiciais nacionais.

Serve de embasamento a diversos outros documentos citados neste estudo, produzidos pelas mais variadas instituições governamentais ou não.

A visão do CEPEJ, conforme estabelecido na Carta, é que a aplicação da IA no domínio da justiça pode contribuir para melhorar a eficiência e a qualidade e deve ser implementada de forma responsável e em conformidade com os direitos fundamentais garantidos, nomeadamente na Convenção Europeia de Direitos Humanos (CEDH) e a Convenção do Conselho da Europa sobre a Proteção de Dados Pessoais.

Para o CEPEJ, é fundamental garantir que a IA continue sendo uma ferramenta a serviço do interesse geral e que seu uso respeite os direitos individuais.

O CEPEJ identificou os seguintes princípios fundamentais a serem respeitados no campo da IA e da justiça:

1. Princípio do respeito pelos direitos fundamentais: garantir que a concepção e implementação de ferramentas e serviços de inteligência artificial sejam compatíveis com os direitos fundamentais;
2. Princípio da não discriminação: prevenir especificamente o desenvolvimento ou intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos;
3. Princípio da qualidade e segurança: no que diz respeito ao tratamento de decisões e dados judiciais, utilizando fontes certificadas e dados intangíveis com modelos concebidos de forma multidisciplinar, em ambiente tecnológico seguro;
4. Princípio da transparência, imparcialidade e equidade: tornar os métodos de tratamento de dados acessíveis e compreensíveis, autorizando auditorias externas;
5. Princípio “sob controle do usuário”: impedir uma abordagem prescritiva e garantir que os usuários sejam atores informados e no controle de suas escolhas.

Para o CEPEJ, o cumprimento desses princípios deve ser assegurado no processamento das decisões judiciais e dos dados por algoritmos e na utilização dos mesmos.

A Carta do CEPEJ é acompanhada de um estudo aprofundado sobre o uso da IA em sistemas judiciais, notadamente aplicativos de IA que processam decisões e dados judiciais.

Ao enumerar os cinco princípios que compõem a abordagem ética de uso de IA nos sistemas de Justiça, com explicação breve de cada um deles, a Carta Ética fornece importante contributo: trazer um pouco mais de objetividade e identificação do delineado ético de IA no Judiciário.

Igualmente relevantes são os estudos anexos.

No primeiro, há estudos aprofundados sobre uso da IA em sistemas judiciais, enumerando-se garantias a serem reafirmadas no contexto de processos civis, comerciais e administrativos.

O ANEXO II – QUAIS OS USOS DA IA NOS SISTEMAS JUDICIAIS EUROPEUS – separa os usos de IA, agrupando-os em “usos a serem encorajados”, “possíveis usos, que exigem precauções metodológicas”, “Usos a serem considerados após estudos científicos adicionais” e “Usos a considerar com reservas mais extremas”.

O ANEXO IV propõe um “CHECKLIST PARA INTEGRAR OS PRINCÍPIOS DA CARTA NOS SEUS TRATAMENTOS”.

Segundo o texto, para avaliar a compatibilidade do seu método de processamento com a Carta, existe uma escala de autoavaliação disponível para cada um dos princípios listados.

Deixando claro que esta avaliação é puramente informativa e de forma alguma equivale a qualquer certificação, indica o checklist que, para cada princípio, deve-se marcar a caixa que corresponde aos seus métodos de processamento, de forma que a uma marcação mais à esquerda indica integração completa, enquanto uma marcação mais à direita indica que não há integração. A coluna com a maior pontuação indica o nível de compatibilidade do método de processamento com o regulamento proposto.

Após a checagem de cada um dos cinco princípios, apresenta um CHECKLIST PARA AVALIAR SEUS MÉTODOS DE PROCESSAMENTO, seguindo-se uma lista de verificação, para cada um dos princípios, conforme segue:

1. Princípio do respeito pelos direitos fundamentais:

Assegurar que a elaboração e a implementação de ferramentas e serviços de Inteligência Artificial sejam compatíveis com os direitos fundamentais

2. Princípio da não discriminação:

Prevenir o desenvolvimento ou intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos

3. Princípio da qualidade e segurança:

No processamento de decisões e dados judiciais, utilizar fontes certificadas e dados intangíveis, com modelos elaborados de forma multidisciplinar, num ambiente tecnológico seguro

4. Princípio da transparência, imparcialidade e justiça:

Tornar os métodos de processamento de dados acessíveis e compreensíveis, autorizar auditorias externas

5. Princípio “sob controle do usuário”: impedir uma abordagem prescritiva e garantir que os usuários sejam informados e controlem suas escolhas

Ressalte-se, por fim, que o item 10 do anexo trata a necessidade de um debate público aprofundado sobre estas ferramentas antes da implementação de políticas públicas para o seu desenvolvimento. A necessidade urgente da **ciberética** para o desenvolvimento de algoritmos de inteligência artificial, respeitando os direitos fundamentais.

Conclui-se o documento destacando-se a necessidade de “estabelecimento de um quadro ético”.

4.1.6 França

A Comissão Nacional de Computação e Liberdades da França (CNIL - *Commission Nationale de l'Informatique et des Libertés*), apresentou relatório sobre as questões éticas que permeiam os algoritmos e a inteligência artificial.

Na França, o relatório surgiu como uma resposta ao Projeto de Lei da República Digital, que conferiu uma atribuição única à Autoridade Francesa de Proteção de Dados (CNIL): liderar uma reflexão sobre as questões éticas e sociais levantadas pelo rápido desenvolvimento das tecnologias digitais.

Segundo o documento, há dois princípios fundantes: 1) princípio da justiça (fairness) aplicado a todos os tipos de algoritmos, que leva em consideração não apenas seus resultados pessoais, mas também seus resultados coletivos; 2) princípio de atenção e vigilância contínua, de modo que todas as partes interessadas estejam em estado de alerta permanente que nossas sociedades precisam adotar em relação aos objetos sociotécnicos complexos e mutáveis que os sistemas algorítmicos representam.

A partir destes princípios, enumera seis recomendações de natureza política, destinados tanto ao poder público, como à sociedade civil: 1) Fomentar a formação de todos os atores envolvidos em “cadeias algorítmicas” (designers, profissionais, cidadãos) em matéria de ética; 2) Tornar os sistemas algorítmicos compreensíveis, fortalecendo os direitos existentes e organizando a mediação com os usuários; 3) melhorar a concepção de sistemas algorítmicos no interesse da liberdade humana; 4) Criação de uma plataforma nacional de algoritmos de auditoria; 5) Aumentar os incentivos para pesquisa em IA ética e lançar uma causa nacional participativa em um projeto de pesquisa de interesse geral; 6) Fortalecimento da ética nas empresas.

Além de princípios e de recomendações, também identifica seis preocupações com questões éticas substancialmente associadas ao uso de algoritmos e IA, a saber: 1) Máquinas autônomas: uma ameaça ao livre arbítrio e à responsabilidade? 2) Preconceitos, discriminação e exclusão; 3) Perfil algorítmico: personalização versus benefícios coletivos; 4) Prevenindo arquivos massivos e aprimorando a IA: buscando um novo equilíbrio; 5) Qualidade, quantidade, relevância: o desafio da curadoria de dados para IA; 6) A identidade humana diante do desafio da inteligência artificial.

4.1.7 Índia

Atualmente, a Índia não possui uma estrutura de orientação abrangente e genérica para o uso de sistemas de IA. Estabelecer tal estrutura seria crucial para fornecer orientação a várias partes interessadas na gestão responsável da Inteligência Artificial na Índia.

Existem certas estruturas específicas de alguns setores que foram identificadas para o desenvolvimento e uso da IA.

Em finanças, o SEBI (*Securities and Exchange Board of India*) emitiu uma circular em janeiro de 2019 para Corretores de Valores, Participantes Depositários, Bolsas de Valores Reconhecidas e Depositários e, em maio de 2019, para todos os Fundos Mútuos, Empresas Gestoras de Ativos, Conselho de Curadores de Fundos Mútuos sobre requisitos de relatórios para aplicativos e sistemas de Inteligência Artificial (AI) e *Machine Learning* (ML) oferecidos e usados. O relatório visa criar um inventário de sistemas de IA no mercado e orientar políticas futuras.

No campo da saúde, a estratégia da National Digital Health Mission (NDHM) identifica a necessidade de criação de orientações e padrões para garantir a confiabilidade dos sistemas de IA na saúde.

A *Data Empowerment and Protection Architecture* (DEPA) do NITI Aayog apresenta uma estrutura técnica para que as pessoas mantenham o controle de seus dados pessoais e os meios para aproveitá-los para obter serviços e benefícios.

Atualmente, a Índia não possui uma legislação abrangente e específica para IA. O mais próximo disso é o projeto de Lei de Proteção de Dados Pessoais (2019).

O Projeto de Lei de Privacidade de 2019 foi criado devido à crescente preocupação do governo da Índia com o uso indevido de dados pessoais confidenciais nos últimos anos. Com a

Lei, empresas teriam que pedir uma solicitação explícita para utilizar ou armazenar os dados dos cidadãos.

O projeto de lei gerou preocupação em no mercado tecnológico e nas *big-techs*, pois há anos as companhias têm procurado novas soluções para operar na Índia devido às regulamentações do país em relação a tecnologia.

Neste cenário de preocupações e incertezas, embora promulgada, em 03/08/2022 o governo da Índia revogou a lei.

4.1.8 Estados Unidos da América

Em fevereiro de 2019 foi emitida a Ordem Executiva nº 13859 por parte do presidente Donald Trump, estabelecendo categoricamente IA como prioridade para os investimentos não somente militares, como também dos diversos setores do Estado, em pesquisa e desenvolvimento – P&D. Este documento, direcionado às agências reguladoras, embora tímido, demonstra a atenção que a Casa Branca tem dado ao desenvolvimento de IA.

O texto igualmente advoga que “à luz das escolhas feitas por nossos competidores estratégicos, os EUA também devem examinar IA sob lentes militares, inclusive os conceitos de IA habilitada para realizar operações autônomas”. Ou seja, o emprego de armas de IA sem o controle humano.

Este projeto de memorando estabelece considerações de política que devem orientar, na medida em que permitido por lei, supervisão regulatória e não regulatória de aplicativos de IA desenvolvidos e implantado fora do governo federal.

Embora as agências federais atualmente usem IA em muitas maneiras de realizar suas missões, o uso governamental de IA está fora do escopo deste MEMORANDO.

De acordo com a lei, as agências devem levar em consideração os seguintes princípios ao formular abordagens regulatórias e não regulatórias para o projeto, desenvolvimento, implantação e operação de aplicativos de IA, gerais e específicos do setor.

Esses princípios, muitos dos quais estão inter-relacionados, refletem os objetivos e princípios da Ordem Executiva 13859. As agências devem calibrar as abordagens de acordo com princípios e considerar fatores específicos para otimizar a rede de benefícios.

1. Confiança pública na IA
2. Participação pública
3. Integridade Científica e Qualidade da Informação
4. Avaliação e Gestão de Riscos
5. Benefícios e custos
6. Flexibilidade
7. Equidade/Justiça e não discriminação
8. Divulgação e transparência

9. Segurança e proteção
10. Coordenação interagências

A seguir, em junho de 2019, foi emitida a atualização do plano estratégico nacional de pesquisa e desenvolvimento em inteligência artificial. O documento estabelece mais uma vez que a liderança mundial em IA é uma questão de segurança nacional para os EUA. Nesse sentido, a prevalência na área é vista como aspecto primordial no tocante à manutenção da hegemonia militar global.

4.1.9 Austrália

Baseada na *Ethically Aligned Design report by IEEE*, a Austrália definiu um Framework Ético para Inteligência Artificial.

Os 8 Princípios de Ética de Inteligência Artificial (IA) da Austrália são projetados para garantir que a IA seja segura, protegida e confiável.

Eles ajudarão a atender os principais objetivos da Austrália, quais sejam: a) alcançar resultados mais seguros, confiáveis e justos para todos os australianos; b) reduzir o risco de impacto negativo sobre as pessoas afetadas por aplicativos de IA e; c) empresas e governos a praticar os mais altos padrões éticos ao projetar, desenvolver e implementar a IA.

Os princípios são voluntários, de modo que sejam aspiracionais e complementem – não substituam – os regulamentos e práticas de IA existentes.

O documento australiano elenca os benefícios da adoção de um framework ético em IA. Segundo o texto, ao aplicar os princípios e se comprometer com práticas éticas de IA, o órgão público ou a empresa pode: a) construir a confiança do público em seu produto ou organização; b) promover a fidelidade do consumidor em seus serviços habilitados para IA; c) influenciar positivamente os resultados da IA; d) garantir que todos os australianos se beneficiem dessa tecnologia transformadora.

Enumera, então, os seguintes princípios:

- 1) Bem-estar humano, social e ambiental: os sistemas de IA devem beneficiar os indivíduos, a sociedade e o meio ambiente.
- 2) Valores centrados no ser humano: os sistemas de IA devem respeitar os direitos humanos, a diversidade e a autonomia dos indivíduos.
- 3) Justiça/Equidade: os sistemas de IA devem ser inclusivos e acessíveis e não devem envolver ou resultar em discriminação injusta contra indivíduos, comunidades ou grupos.
- 4) Proteção e segurança da privacidade: os sistemas de IA devem respeitar e defender os direitos de privacidade e proteção de dados e garantir a segurança dos dados.
- 5) Confiabilidade e segurança: Os sistemas de IA devem operar de forma confiável de acordo com a finalidade pretendida.
- 6) Transparência e explicabilidade: deve haver transparência e divulgação responsável para que as pessoas possam entender quando estão sendo

significativamente impactadas pela IA e possam descobrir quando um sistema de IA está se envolvendo com elas.

7) **Contestabilidade:** Quando um sistema de IA impacta significativamente uma pessoa, comunidade, grupo ou ambiente, deve haver um processo oportuno para permitir que as pessoas desafiem o uso ou os resultados do sistema de IA.

8) **Responsabilidade (accountability):** As pessoas responsáveis pelas diferentes fases do ciclo de vida do sistema de IA devem ser identificáveis e responsáveis pelos resultados dos sistemas de IA, e a supervisão humana dos sistemas de IA deve ser habilitada.

4.1.10 OCDE - Organização para a Cooperação e Desenvolvimento Econômico

A Organização para Cooperação e Desenvolvimento Econômico (OCDE) aprovou, em maio de 2019, seus Princípios sobre Inteligência Artificial, aos quais 42 países, inclusive o Brasil, já aderiram.

A Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), partindo de uma perspectiva humanística, tem promovido estudos e conferências para tratar de diferentes dimensões de políticas públicas para IA, com discussão de questões como desenvolvimento sustentável, gênero, educação e ética da ciência.

O documento estabelece os 5 princípios baseados em valores e 5 recomendações os seus respectivos governos.

Os Princípios para a Inteligência Artificial segundo a OCDE são:

- 1) **Crescimento inclusivo, desenvolvimento sustentável e bem-estar;**
- 2) **Valores centrados nos humanos e justiça - Os sistemas de IA devem ser projetados de forma a respeitar o estado de direito, os direitos humanos, os valores democráticos e a diversidade, e devem incluir salvaguardas apropriadas para garantir uma sociedade justa e igualitária;**
- 3) **Transparência e explicabilidade.** Esse princípio é sobre transparência e divulgação responsável em torno dos sistemas de IA para garantir que as pessoas entendam quando estão se envolvendo com eles e possam desafiar os resultados.
- 4) **Robustez, segurança e proteção.** Os sistemas de IA devem funcionar de maneira robusta, segura e protegida ao longo de suas vidas, e os riscos potenciais devem ser continuamente avaliados e gerenciados.
- 5) **Responsabilidade (ou prestação de contas).** Organizações e indivíduos que desenvolvem, implantam ou operam sistemas de IA devem ser responsabilizados por seu funcionamento adequado de acordo com os princípios baseados em valores da OCDE para IA.

Com a finalidade de viabilizar e instrumentalizar os princípios elencados, a OCDE elaborou cinco recomendações que devem pautar a atuação dos países membros e aqueles que aderiram ao documento. São aconselhamentos de com foco em políticas nacionais e na cooperação internacional, elaborando objetivos estratégicos que podem ser adotados pelos atores estatais globais.

1) Investir em pesquisa e desenvolvimento de IA

Os governos devem facilitar o investimento público e privado em pesquisa e desenvolvimento para estimular a inovação em IA confiável.

Os governos devem considerar o investimento público de longo prazo e incentivar o investimento privado, em pesquisa e desenvolvimento, incluindo esforços interdisciplinares, para estimular a inovação em IA confiável que se concentre em questões técnicas desafiadoras **e nas implicações e políticas sociais, legais e éticas relacionadas à IA questões.**

Os governos também devem considerar o investimento público e incentivar o investimento privado em conjuntos de dados abertos que sejam representativos e respeitem a privacidade e a proteção de dados para apoiar um ambiente de pesquisa e desenvolvimento de IA livre de preconceitos inadequados e para melhorar a interoperabilidade e o uso de padrões.

2) Promover um ecossistema digital para IA

Os governos devem promover ecossistemas de IA acessíveis com infraestrutura e tecnologias digitais e mecanismos para compartilhar dados e conhecimento.

Os governos devem promover o desenvolvimento e o acesso a um ecossistema digital para IA confiável. Esse ecossistema inclui, em particular, tecnologias e infraestruturas digitais e mecanismos para compartilhar conhecimento de IA, conforme apropriado. Nesse sentido, os governos devem considerar a promoção de mecanismos, como confiança de dados, para apoiar o compartilhamento seguro, justo, legal e ético de dados.

3) Criar ambiente de política favorável para IA

Os governos devem criar um ambiente político que abra caminho para a implantação de sistemas confiáveis de IA.

Os governos devem promover um ambiente de políticas que apoie uma transição ágil do estágio de pesquisa e desenvolvimento para o estágio de implantação e operação de sistemas de IA confiáveis. Para esse efeito, eles devem considerar o uso de experimentação para fornecer um ambiente controlado no qual os sistemas de IA possam ser testados e ampliados, conforme apropriado.

Os governos devem revisar e adaptar, conforme apropriado, suas políticas e estruturas regulatórias e mecanismos de avaliação à medida que se aplicam aos sistemas de IA para incentivar a inovação e a competição por IA confiável.

4) Capacitação humana e preparação para a transformação do mercado de trabalho

Os governos devem equipar as pessoas com as habilidades de IA e apoiar os trabalhadores para garantir uma transição justa.

Os governos devem trabalhar em estreita colaboração com as partes interessadas para se preparar para a transformação do mundo do trabalho e da sociedade. Eles devem capacitar as pessoas a usar e interagir efetivamente com os sistemas de IA em toda a variedade de aplicativos, inclusive equipando-os com as habilidades necessárias.

Os governos devem tomar medidas, inclusive por meio do diálogo social, para garantir uma transição justa para os trabalhadores à medida que a IA é implantada, como por meio de programas de treinamento ao longo da vida profissional, apoio às pessoas afetadas pelo deslocamento e acesso a novas oportunidades no mercado de trabalho.

Os governos também devem trabalhar em estreita colaboração com as partes interessadas para promover o uso responsável da IA no trabalho, melhorar a segurança dos trabalhadores e a qualidade dos empregos, promover o empreendedorismo e a produtividade e ter como objetivo garantir que os benefícios da IA sejam compartilhados de maneira ampla e justa.

5) Cooperação internacional para IA confiável

Os governos devem cooperar além das fronteiras e setores para compartilhar informações, desenvolver padrões e trabalhar para uma administração responsável da IA.

Os governos, incluindo os países em desenvolvimento e as partes interessadas, devem cooperar ativamente para promover esses princípios e progredir na administração responsável da IA confiável.

Os governos devem trabalhar juntos na OCDE e em outros fóruns globais e regionais para promover o compartilhamento de conhecimento de IA, conforme apropriado. Eles devem incentivar iniciativas internacionais, intersetoriais e abertas de várias partes interessadas para obter conhecimento de longo prazo em IA.

Os governos devem promover o desenvolvimento de padrões técnicos globais multissetoriais e orientados por consenso para IA interoperável e confiável.

Os governos também devem incentivar o desenvolvimento e seu próprio uso de métricas comparáveis internacionalmente para medir a pesquisa, o desenvolvimento e a implantação de IA e reunir a base de evidências para avaliar o progresso na implementação desses princípios.

Apesar de não vincularem juridicamente os países membros da OCDE e os aderentes ao documento, os princípios e as recomendações fornecem perspectivas a respeito de legislações e regulamentos nacionais que estão em fase de discussão ou elaboração.

Aliás, este documento foi importante referencial na elaboração da Estratégia Brasileira de Inteligência Artificial – EBIA, instituída pela Portaria MCTI nº 4.617, de 6 de abril de 2021, alterada pela Portaria MCTI nº 4.979, de 13 de julho de 2021, como já salientado em tópico anterior.

Além disso, importante destacar que, embora o documento não preveja claramente um quadro mínimo e concreto para conferir validação ética e legal às pesquisas e desenvolvimento com IA, o documento serve de importante referencial para que cada estado signatário, valendo-se de suas respectivas autonomias políticas, possam definir parâmetros éticos adequados.

4.1.11 G20 – Declaração Ministerial sobre Comércio e Economia Digital – Princípios para IA Centrada nos Humanos (2019).

Outro importante documento, que demonstra a preocupação dos governantes globais quanto ao enfoque ético no uso da inteligência artificial é a Declaração Ministerial sobre Comércio e Economia Digital, elaborada pelos Líderes do G20, em reunião realizada em Osaka, Japão, em 28 e 29 de junho de 2019. Em união de esforços conjuntos para tratar dos principais desafios econômicos globais, reconheceram o papel da inteligência artificial para promover o crescimento econômico global, aproveitando o poder da inovação tecnológica, em particular a digitalização e seu uso em benefício de todos.

Estabeleceram que, a fim de promover ainda mais a inovação na economia digital, apoiam o intercâmbio de boas práticas sobre políticas eficientes e enfoques e marcos regulatórios que sejam inovadores, ágeis, flexíveis e adaptadas à era digital, inclusive por meio de regras experimentais (“regulatory sandboxes”).

Também destacaram que o desenvolvimento e o uso responsável da Inteligência Artificial (IA) podem ser uma força motora para ajudar a promover os objetivos de desenvolvimento sustentável definidos na Agenda 2030 da ONU e a criar uma sociedade sustentável e inclusiva.

Para promover a confiança do público em tecnologias de IA e realizar plenamente seu potencial, comprometeram-se a buscar um enfoque de IA centrado no ser humano. Neste ponto, anunciaram os Princípios sobre IA do G20, extraídos da Recomendação da OCDE sobre IA, que foram elencados no tópico anterior.

Adicionalmente, reconheceram a crescente importância de promover a segurança na economia digital e tratar de lacunas e vulnerabilidades. Afirmaram a importância da proteção da propriedade intelectual, juntamente com a rápida expansão de tecnologias emergentes, inclusive a “Internet das Coisas (IoT, em inglês), cresce o valor de uma permanente discussão sobre a segurança na economia digital.

Os membros do G20 afirmam também a necessidade de continuar tratando desses desafios urgentes e reafirmam a importância de superar o hiato digital e promover a adoção da digitalização pelas pequenas e médias empresas (PMEs) e por todos os indivíduos, particularmente os grupos vulneráveis, e também encorajam a comunicação e a troca de experiências entre cidades, para o desenvolvimento de cidades inteligentes.

4.2 INICIATIVAS MERCADO/INDÚSTRIA

4.2.1 Asilomar

A Conferência Asilomar ocorreu em um cenário de crescente interesse da sociedade mundial em geral no potencial da inteligência artificial (IA) e uma sensação de que aqueles que participam de seu desenvolvimento têm a responsabilidade e a oportunidade de moldá-la para o melhor.

O objetivo da Conferência, que reuniu líderes da academia e da indústria, era discernir a visão compartilhada da comunidade de IA para a IA. Entre seus membros identificamos nomes como Stephen Hawking, Morgan Freeman, Elon Musk e Stuart J. Russell.

Foram analisados relatórios sobre as oportunidades e ameaças criadas pela IA e compilados uma longa lista das diversas visões sobre como a tecnologia deve ser gerenciada. Em seguida, utilizando-se de sucessivas rodadas de discussões, com os mais diversos especialistas, chegaram em Asilomar a uma lista final, que consistiu em 23 princípios, cada um dos quais recebeu apoio de pelo menos 90% dos participantes da conferência.

Esses “Princípios Asilomar” se tornaram um dos conjuntos de princípios de governança mais influentes e servem para orientar trabalho em IA em todo o mundo. Foram agrupados em blocos, denominados de: a) questões sobre a pesquisa; b) éticas e valores e; c) questões de longo prazo.

Os princípios são descritos a seguir:

a) Questões sobre a pesquisa

1) Objetivo da Pesquisa: O objetivo da pesquisa em IA não deve ser criar inteligência não direcionada, mas inteligência benéfica.

2) Financiamento da Pesquisa: Investimentos em IA devem ser acompanhados de financiamentos em pesquisa para assegurar seu uso benéfico, incluindo pesquisas espinhosas em Ciência da Computação, Economia, Leis, Ética, e Estudos Sociais, tais como: Como podemos tornar futuros sistemas com IA altamente robustos, para que eles façam o que queremos sem mal funcionamento ou serem hackeados? Como podemos aumentar nossa prosperidade por meio da automação, mantendo os recursos e a finalidade das pessoas? Como podemos atualizar nossos sistemas jurídicos para sermos mais justos e eficientes, para manter o ritmo da IA e para gerenciar os riscos associados à IA? Com que conjunto de valores a IA deve estar alinhada, e que posicionamento legal e ético ela deve ter?

3) Link 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 fomentada entre pesquisadores e desenvolvedores de IA.

5) Prevenção de Corrida: Equipes que desenvolvem sistemas de inteligência artificial devem cooperar ativamente para evitar cortes nas normas de segurança.

b) Éticas e Valores

6) Segurança: Os sistemas com IA devem ser seguros e protegidos durante toda a sua vida útil operacional, e verificáveis, quando aplicável e viável.

7) Transparência de falha: Se um sistema com IA causar dano, deve ser possível determinar o motivo.

8) Transparência Judicial: Qualquer envolvimento de um sistema autônomo na tomada de decisões judiciais deve fornecer uma explicação satisfatória passível de auditoria por uma autoridade humana competente.

9) Responsabilidade: Designers e construtores de sistemas avançados com IA são partes interessadas nas implicações morais de seu uso, abuso e ações, com responsabilidade e oportunidade de moldar essas implicações.

10) Alinhamento de valor: Sistemas com IA altamente autônomos devem ser projetados de modo que seja assegurado que seus objetivos e comportamentos serão alinhados com os valores humanos durante toda a operação.

11) Valores Humanos: Os sistemas com IA devem ser projetados e operados de modo a serem compatíveis com os ideais da 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 com AI de analisar e utilizar esses dados.

13) Liberdade e Privacidade: A aplicação de IA aos dados pessoais não deve restringir de forma injustificável a liberdade real ou percebida das pessoas.

14) Benefício compartilhado: Tecnologias com IA devem beneficiar e capacitar o maior número de pessoas possível.

15) Prosperidade compartilhada: A prosperidade econômica criada pela IA deve ser compartilhada amplamente, para beneficiar toda a humanidade.

16) Controle Humano: Os seres humanos devem escolher como e se devem delegar decisões aos sistemas com IA, para realizar os objetivos escolhidos pelo homem.

17) Não-subversão: O poder conferido pelo controle de sistemas com IA altamente avançada deve respeitar e melhorar, ao invés de subverter, os processos sociais e cívicos dos quais depende a saúde da sociedade.

18) Corrida Armada com IA: Deve ser evitada uma corrida armamentista com armas autônomas letais.

c) Questões de Longo Prazo

19) Atenção na Capacidade: Não havendo consenso, devemos evitar fortes suposições sobre os limites superiores em futuras capacidades de IA.

20) Importância: IA avançada poderia representar uma mudança profunda na história da vida na Terra, e deveria ser planejada e administrada com cuidado e recursos proporcionais.

21) Riscos: Os riscos colocados pelos sistemas com IA, especialmente os riscos catastróficos ou existenciais, devem estar sujeitos a esforços de planejamento e mitigação proporcionais ao impacto esperado.

22) Autoaprimoramento Recursivo: Sistemas com IA projetados para melhorar ou autorreplicar-se recursivamente de uma maneira que poderia levar a um aumento rápido da qualidade ou quantidade, devem estar sujeitos a rígidas medidas 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.

Conforme se vê da leitura dos princípios e Asilomar para ética e IA, além de ter forte valorativa destinada à busca do bem-estar humano, do bem comum, dos direitos humanos e do desenvolvimento sustentável, há dois destaques no documento, que o diferencia dos demais: primeiro a forte preocupação com questões relacionadas à pesquisa de IA e, depois, o estabelecimento de princípios que norteiam a geopolítica global no que se refere ao desenvolvimento da IA. Daí que vemos princípios como o de benefício e prosperidade

compartilhados, além de não subversão de processos cívicos e sociais, até chegar a um princípio que busca ordenar a corrida armada em torno da inteligência artificial.

Denota-se, portanto, preocupação do documento com a proteção política internacional para crescimento da IA, evitando-se problemas concorrenciais e até conflito entre nações.

4.2.2 Institute of Electrical and Electronics Engineers – IEEE

O IEEE (Instituto de Engenheiros Eletricistas e Eletrônicos) é uma organização profissional fundada em 1963 nos Estados Unidos, e atualmente é conhecida como a maior organização profissional do mundo em número de membros. A organização não tem fins lucrativos e sua meta é fomentar os conhecimentos na área da engenharia elétrica, eletrônica e da computação.

O Instituto é o responsável pelas normas que são implementadas internacionalmente nas áreas da engenharia elétrica e informática. IEEE – *Institute of Electrical and Electronic Engineers* colabora no incremento da prosperidade mundial, promovendo a engenharia de criação, desenvolvimento, integração, compartilhamento e o conhecimento aplicado no que se refere à ciência e tecnologias da eletricidade e da informação, em benefício da humanidade e da profissão.

Em junho de 2020, reconhecendo a importância das tecnologias disponíveis para afetar a qualidade de vida em todo o mundo e aceitando uma obrigação pessoal dos seus membros para com suas profissões e com as comunidades mundiais, comprometeram-se com a mais alta conduta ética e profissional, estabelecendo 10 princípios, alocados em 3 blocos.

O primeiro deles, recomenda manter os mais altos padrões de integridade, comportamento responsável e conduta ética nas atividades profissionais, especificamente:

- 1) Manter a segurança, a saúde e o bem-estar do público em primeiro lugar, esforçar-se para cumprir o design ético e as práticas de desenvolvimento sustentável, proteger a privacidade de outras pessoas e divulgar prontamente fatores que possam colocar em risco o público ou o meio ambiente;
- 2) Melhorar a compreensão dos indivíduos e da sociedade sobre as capacidades e implicações sociais das tecnologias convencionais e emergentes, incluindo sistemas inteligentes;
- 3) Evitar conflitos de interesse reais ou percebidos sempre que possível e divulgá-los às partes afetadas quando existirem;
- 4) Evitar condutas ilícitas nas atividades profissionais e rejeitar o suborno em todas as suas formas;
- 5) Buscar, aceitar e oferecer críticas honestas ao trabalho técnico, reconhecer e corrigir erros, ser honesto e realista ao declarar reivindicações ou estimativas com base em dados disponíveis e creditar adequadamente as contribuições de outros;
- 6) Manter e melhorar nossa competência técnica e realizar tarefas tecnológicas para terceiros somente se qualificado por treinamento ou experiência, ou após a divulgação completa das limitações pertinentes;

No que se refere ao tratamento de dados, elenca a necessidade de tratar todas as pessoas com justiça e respeito, não se envolver em assédio ou discriminação e evitar ferir outras pessoas, enaltecendo:

- 7) Tratar todas as pessoas com justiça e respeito e não se envolver em discriminação com base em características como raça, religião, gênero, deficiência, idade, nacionalidade, orientação sexual, identidade de gênero ou expressão de gênero;
- 8) Não se envolver em assédio de qualquer tipo, incluindo assédio sexual ou comportamento de bullying;
- 9) Evitar ferir outras pessoas, suas propriedades, reputação ou emprego por meio de ações falsas ou maliciosas, boatos ou quaisquer outros abusos verbais ou físicos;

Conclui com o dever de esforçar-se para garantir que este código seja mantido por colegas e colegas de trabalho e, para tanto:

- 10) Apoiar colegas e colegas de trabalho no cumprimento deste código de ética, esforçar-se para garantir que o código seja mantido e não retaliar contra indivíduos que relatam uma violação.

4.2.3 Associação para Maquinária da Computação (ACM)

A Associação para Maquinaria da Computação foi fundada em 1947 como a primeira sociedade científica e educacional dedicada à computação, contando com mais de 100 mil associados em todo o mundo.

Em junho de 2018 publicou Código de Ética e Condutas profissionais da ACM, destinado aos profissionais de computação e redigidos em forma de comandos, subdivide-se em: a) princípios éticos gerais; b) responsabilidades profissionais; c) princípios de liderança profissional; d) conformidade com o código.

Dentre os **princípios éticos gerais**, elencam:

- 1) Contribua para a sociedade e para o bem-estar humano, reconhecendo que todas as pessoas são partes interessadas na computação.

Esse princípio, que diz respeito à qualidade de vida de todas as pessoas, afirma a obrigação dos profissionais de computação, individual e coletivamente, de usar suas habilidades em benefício da sociedade, de seus membros e do ambiente que os cerca. Esta obrigação inclui a promoção dos direitos humanos fundamentais e a proteção do direito à autonomia de cada indivíduo. Um objetivo essencial dos profissionais de computação é minimizar as consequências negativas da computação, incluindo ameaças à saúde, segurança, segurança pessoal e privacidade. Quando os interesses de vários grupos entram em conflito, as necessidades dos menos favorecidos devem receber maior atenção e prioridade.

Os profissionais de computação devem considerar se os resultados de seus esforços respeitarão a diversidade, serão usados de maneira socialmente responsável, atenderão às necessidades sociais e serão amplamente acessíveis.

Além de um ambiente social seguro, o bem-estar humano requer um ambiente natural seguro. Portanto, os profissionais de computação devem promover a sustentabilidade do meio ambiente local e globalmente.

2) Evite danos.

Neste documento, "dano" significa consequências negativas, especialmente quando essas consequências são significativas e injustas.

Evitar danos começa com uma consideração cuidadosa dos impactos potenciais em todos os afetados pelas decisões. Quando o dano é uma parte intencional do sistema, os responsáveis são obrigados a garantir que o dano seja eticamente justificado. Em ambos os casos, certifique-se de que todos os danos sejam minimizados.

Para minimizar a possibilidade de prejudicar outras pessoas de forma indireta ou não intencional, os profissionais de computação devem seguir as melhores práticas geralmente aceitas, a menos que haja uma razão ética convincente para agir de outra forma.

Um profissional de computação tem a obrigação adicional de relatar quaisquer sinais de riscos do sistema que possam resultar em danos. Se os líderes não agirem para reduzir ou mitigar tais riscos, pode ser necessário "denunciar" para reduzir possíveis danos. No entanto, relatórios caprichosos ou equivocados de riscos podem ser prejudiciais. Antes de relatar riscos, um profissional de computação deve avaliar cuidadosamente os aspectos relevantes da situação.

3) Seja honesto e confiável.

A honestidade é um componente essencial da confiabilidade. Um profissional de computação deve ser transparente e fornecer divulgação completa de todos os recursos, limitações e problemas potenciais do sistema para as partes apropriadas. Fazer declarações deliberadamente falsas ou enganosas, fabricar ou falsificar dados, oferecer ou aceitar subornos e outras condutas desonestas são violações do Código.

4) Seja justo e tome medidas para não discriminar.

Os valores de igualdade, tolerância, respeito ao próximo e justiça regem este princípio. A justiça exige que mesmo os processos de decisão cuidadosos forneçam algum caminho para a reparação de queixas.

Os profissionais de computação devem promover a participação justa de todas as pessoas, incluindo aquelas de grupos sub-representados. A discriminação preconceituosa com base em idade, cor, deficiência, etnia, situação familiar, identidade de gênero, filiação sindical, situação militar, nacionalidade, raça, religião ou crença, sexo, orientação sexual ou qualquer outro fator inadequado é uma violação explícita de o código. O assédio, incluindo assédio sexual, bullying e outros abusos de poder e autoridade, é uma forma de discriminação que, entre outros danos, limita o acesso justo aos espaços virtuais e físicos onde ocorre o assédio.

O uso da informação e da tecnologia pode causar novas desigualdades ou aumentar as já existentes. As tecnologias e práticas devem ser tão inclusivas e acessíveis quanto possível e os profissionais de computação devem tomar medidas para evitar a criação de sistemas ou tecnologias que desprivilegiem ou oprimam as pessoas. A falha em projetar para inclusão e acessibilidade pode constituir discriminação injusta.

5) Respeite o trabalho necessário para produzir novas ideias, invenções, trabalhos criativos e artefatos de computação.

Desenvolver novas ideias, invenções, trabalhos criativos e artefatos de computação cria valor para a sociedade, e aqueles que despendem esse esforço devem esperar obter valor de seu trabalho. Os profissionais de computação devem, portanto, dar crédito aos criadores de ideias, invenções, trabalhos e artefatos e respeitar direitos autorais, patentes, segredos comerciais, contratos de licença e outros métodos de proteção das obras dos autores.

Tanto os costumes quanto a lei reconhecem que algumas exceções ao controle de um criador sobre uma obra são necessárias para o bem público. Os profissionais de computação não devem se opor indevidamente a usos razoáveis de suas obras intelectuais. Os esforços para ajudar os outros, contribuindo com tempo e energia para projetos que ajudam a sociedade, ilustram um aspecto positivo desse princípio. Tais esforços incluem software livre e de código aberto e trabalhos colocados em domínio público. Os profissionais de computação não devem reivindicar a propriedade privada do trabalho que eles ou outros compartilharam como recursos públicos.

6) Respeite a privacidade.

A responsabilidade de respeitar a privacidade se aplica aos profissionais de computação de maneira particularmente profunda. A tecnologia permite a coleta, monitoramento e troca de informações pessoais de forma rápida, barata e muitas vezes sem o conhecimento das pessoas afetadas. Portanto, um profissional de computação deve estar familiarizado com as várias definições e formas de privacidade e deve entender os direitos e responsabilidades associados à coleta e uso de informações pessoais.

Os profissionais de computação só devem usar informações pessoais para fins legítimos e sem violar os direitos de indivíduos e grupos. Isso requer a tomada de precauções para evitar a reidentificação de dados anônimos ou coleta de dados não autorizada, garantindo a precisão dos dados, entendendo a proveniência dos dados e protegendo-os contra acesso não autorizado e divulgação acidental. Os profissionais de computação devem estabelecer políticas e procedimentos transparentes que permitam aos indivíduos entender quais dados estão sendo coletados e como estão sendo usados, dar consentimento informado para coleta automática de dados e revisar, obter, corrigir imprecisões e excluir seus dados pessoais.

7) Honre a confidencialidade.

Os profissionais de computação geralmente recebem informações confidenciais, como segredos comerciais, dados de clientes, estratégias de negócios não públicas, informações financeiras, dados de pesquisa, artigos acadêmicos pré-publicação e pedidos de patentes. Os profissionais de computação devem proteger a confidencialidade, exceto nos casos em que haja evidência de violação da lei, dos

regulamentos organizacionais ou do Código. Nesses casos, a natureza ou o conteúdo dessas informações não devem ser divulgados, exceto às autoridades competentes. Um profissional de computação deve considerar cuidadosamente se tais divulgações são consistentes com o Código.

Ao enumerar **responsabilidades profissionais**, destacam:

1) Esforçar-se para alcançar alta qualidade tanto nos processos quanto nos produtos do trabalho profissional.

Os profissionais de computação devem insistir e apoiar o trabalho de alta qualidade de si mesmos e dos colegas. A dignidade de empregadores, empregados, colegas, clientes, usuários e qualquer outra pessoa afetada direta ou indiretamente pelo trabalho deve ser respeitada durante todo o processo. Os profissionais de computação devem respeitar o direito dos envolvidos à comunicação transparente sobre o projeto. Os profissionais devem estar cientes de quaisquer consequências negativas graves que afetem qualquer parte interessada que possam resultar de um trabalho de baixa qualidade e devem resistir aos incentivos para negligenciar essa responsabilidade.

2) Manter altos padrões de competência profissional, conduta e prática ética.

A computação de alta qualidade depende de indivíduos e equipes que assumem a responsabilidade pessoal e de grupo pela aquisição e manutenção da competência profissional. A competência profissional começa com o conhecimento técnico e com a consciência do contexto social em que seu trabalho pode ser implantado. A competência profissional também requer habilidade em comunicação, em análise reflexiva e em reconhecer e enfrentar desafios éticos. A atualização de habilidades deve ser um processo contínuo e pode incluir estudo independente, participação em conferências ou seminários e outra educação formal ou informal. As organizações profissionais e os empregadores devem encorajar e facilitar essas atividades.

3) Conhecer e respeitar as regras existentes relativas ao trabalho profissional.

"Regras" aqui incluem leis e regulamentos locais, regionais, nacionais e internacionais, bem como quaisquer políticas e procedimentos das organizações às quais o profissional pertence.

Os profissionais de computação devem cumprir essas regras, a menos que haja uma justificativa ética convincente para fazer o contrário. Um profissional de computação deve considerar desafiar a regra por meio dos canais existentes antes de infringir a regra. Um profissional de computação que decida violar uma regra por ser antiético, ou por qualquer outro motivo, deve considerar as possíveis consequências e aceitar a responsabilidade por essa ação.

4) Aceitar e fornecer revisão profissional apropriada.

O trabalho profissional de alta qualidade em computação depende de revisão profissional em todas as etapas. Sempre que apropriado, os profissionais de computação devem buscar e utilizar a revisão por pares e partes interessadas. Os profissionais de computação também devem fornecer revisões críticas e construtivas do trabalho de outras pessoas.

5) Faça avaliações abrangentes e completas dos sistemas de informática e seus impactos, incluindo a análise de possíveis riscos.

Os profissionais de computação estão em uma posição de confiança e, portanto, têm uma responsabilidade especial de fornecer avaliações e testemunhos objetivos e confiáveis a empregadores, funcionários, clientes, usuários e ao público. Os profissionais de computação devem se esforçar para serem perspicazes, minuciosos e objetivos ao avaliar, recomendar e apresentar descrições e alternativas do sistema. Cuidados extraordinários devem ser tomados para identificar e mitigar riscos potenciais em sistemas de aprendizado de máquina. Um sistema para o qual os riscos futuros não podem ser previstos de forma confiável requer reavaliação frequente do risco à medida que o sistema evolui em uso, ou não deve ser implantado. Quaisquer problemas que possam resultar em grandes riscos devem ser relatados às partes apropriadas.

6) Executar trabalhos apenas em áreas de competência.

Um profissional de computação é responsável por avaliar possíveis atribuições de trabalho. Isso inclui avaliar a viabilidade e conveniência do trabalho e fazer um julgamento sobre se a atribuição do trabalho está dentro das áreas de competência do profissional. Se a qualquer momento antes ou durante o trabalho o profissional identificar a falta de uma perícia necessária, deve comunicá-la ao empregador ou cliente. O cliente ou empregador pode decidir prosseguir a atribuição com o profissional após um tempo adicional para adquirir as competências necessárias, prosseguir a atribuição com outra pessoa que tenha a experiência necessária ou renunciar à atribuição. O julgamento ético de um profissional de computação deve ser o guia final para decidir se deve ou não trabalhar na tarefa.

7) Promover a conscientização e a compreensão do público sobre computação, tecnologias relacionadas e suas consequências.

Conforme apropriado ao contexto e às habilidades de cada um, os profissionais de computação devem compartilhar conhecimento técnico com o público, promover a conscientização sobre a computação e incentivar a compreensão da computação. Essas comunicações com o público devem ser claras, respeitadas e acolhedoras. Questões importantes incluem os impactos dos sistemas de computador, suas limitações, suas vulnerabilidades e as oportunidades que eles apresentam. Além disso, um profissional de computação deve abordar respeitadamente informações imprecisas ou enganosas relacionadas à computação.

8) Acessar recursos de informática e comunicação somente quando autorizado ou compelido pelo bem público.

Indivíduos e organizações têm o direito de restringir o acesso a seus sistemas e dados, desde que as restrições sejam consistentes com outros princípios do Código. Consequentemente, os profissionais de computação não devem acessar o sistema de computador, software ou dados de outra pessoa sem uma crença razoável de que tal ação seria autorizada ou uma crença convincente de que é consistente com o bem público. Um sistema acessível ao público não é, por si só, motivo suficiente para implicar autorização. Em circunstâncias excepcionais, um profissional de computação pode usar acesso não autorizado para interromper ou inibir o funcionamento de sistemas

maliciosos; precauções extraordinárias devem ser tomadas nesses casos para evitar danos a outras pessoas.

9) Projete e implemente sistemas que sejam robustos e seguros para uso.

As violações de segurança do computador causam danos. A segurança robusta deve ser uma consideração primária ao projetar e implementar sistemas. Os profissionais de computação devem realizar a devida diligência para garantir que o sistema funcione conforme pretendido e tomar as medidas adequadas para proteger os recursos contra uso indevido, modificação e negação de serviço acidental e intencional. Como as ameaças podem surgir e mudar após a implantação de um sistema, os profissionais de computação devem integrar técnicas e políticas de mitigação, como monitoramento, aplicação de patches e relatórios de vulnerabilidade. Os profissionais de computação também devem tomar medidas para garantir que as partes afetadas por violações de dados sejam notificadas de maneira oportuna e clara, fornecendo orientação e remediação apropriadas.

Para garantir que o sistema atinja o objetivo pretendido, os recursos de segurança devem ser projetados para serem os mais intuitivos e fáceis de usar possível. Os profissionais de computação devem desencorajar as precauções de segurança que são muito confusas, são inadequadas à situação ou inibem o uso legítimo.

Nos casos em que o uso indevido ou danos são previsíveis ou inevitáveis, a melhor opção pode ser não implementar o sistema.

Ao tratar de **princípios de liderança profissional**, enaltecem o papel do líder (amplamente considerado) em defender e promover os princípios estabelecidos do documento, estabelecendo os seguintes deveres:

1) Assegurar que o bem público seja a preocupação central durante todo o trabalho de computação profissional.

As pessoas — incluindo usuários, clientes, colegas e outros afetados direta ou indiretamente — sempre devem ser a preocupação central da computação. O bem público deve sempre ser uma consideração explícita ao avaliar tarefas associadas à pesquisa, análise de requisitos, projeto, implementação, teste, validação, implantação, manutenção, retirada e descarte. Os profissionais de computação devem manter esse foco independentemente das metodologias ou técnicas que utilizam em sua prática.

2) Articular, estimular a aceitação e avaliar o cumprimento das responsabilidades sociais pelos membros da organização ou grupo.

Organizações e grupos técnicos afetam a sociedade em geral, e seus líderes devem aceitar as responsabilidades associadas. As organizações — por meio de procedimentos e atitudes orientadas para a qualidade, transparência e bem-estar da sociedade — reduzem os danos ao público e aumentam a conscientização sobre a influência da tecnologia em nossas vidas. Portanto, os líderes devem incentivar a plena participação dos profissionais de computação no cumprimento de responsabilidades sociais relevantes e desencorajar tendências de fazer o contrário.

3) Gerenciar pessoal e recursos para melhorar a qualidade de vida no trabalho.

Os líderes devem garantir que eles melhorem, não degradem, a qualidade de vida no trabalho. Os líderes devem considerar o desenvolvimento pessoal e profissional, requisitos de acessibilidade, segurança física, bem-estar psicológico e dignidade humana de todos os trabalhadores. Padrões apropriados de ergonomia humano-computador devem ser usados no local de trabalho.

4) Articular, aplicar e apoiar políticas e processos que reflitam os princípios do Código.

Os líderes devem buscar políticas organizacionais claramente definidas que sejam consistentes com o Código e comunicá-las efetivamente às partes interessadas relevantes. Além disso, os líderes devem incentivar e recompensar o cumprimento dessas políticas e tomar as medidas adequadas quando as políticas forem violadas. Projetar ou implementar processos que deliberadamente ou negligentemente violem ou tendam a permitir a violação dos princípios do Código é eticamente inaceitável.

5) Criar oportunidades para que os membros da organização ou grupo cresçam como profissionais.

As oportunidades educacionais são essenciais para todos os membros da organização e do grupo. Os líderes devem garantir que os profissionais de computação tenham oportunidades para ajudá-los a melhorar seus conhecimentos e habilidades no profissionalismo, na prática da ética e em suas especialidades técnicas. Essas oportunidades devem incluir experiências que familiarizem os profissionais de computação com as consequências e limitações de tipos específicos de sistemas. Os profissionais de computação devem estar plenamente cientes dos perigos de abordagens simplificadas, da improbabilidade de antecipar todas as condições operacionais possíveis, da inevitabilidade de erros de software, das interações dos sistemas e seus contextos e outras questões relacionadas à complexidade de sua profissão - e, assim, ser confiantes em assumir responsabilidades pelo trabalho que realizam.

6) Tenha cuidado ao modificar ou retirar sistemas.

Alterações de interface, remoção de recursos e até atualizações de software têm impacto na produtividade dos usuários e na qualidade de seu trabalho. Os líderes devem tomar cuidado ao alterar ou descontinuar o suporte para recursos do sistema dos quais as pessoas ainda dependem. Os líderes devem investigar minuciosamente alternativas viáveis para remover o suporte para um sistema legado. Se essas alternativas forem inaceitavelmente arriscadas ou impraticáveis, o desenvolvedor deve ajudar na migração tranquila das partes interessadas do sistema para uma alternativa. Os usuários devem ser notificados sobre os riscos do uso contínuo do sistema sem suporte muito antes do término do suporte. Os profissionais de computação devem ajudar os usuários do sistema a monitorar a viabilidade operacional de seus sistemas de computação e ajudá-los a entender que a substituição oportuna de recursos inadequados ou desatualizados ou sistemas inteiros pode ser necessária.

7) Reconhecer e cuidar especialmente dos sistemas que se integram na infraestrutura da sociedade.

Mesmo os sistemas de computador mais simples têm o potencial de impactar todos os aspectos da sociedade quando integrados às atividades cotidianas, como comércio, viagens, governo, saúde e educação. Quando organizações e grupos

desenvolvem sistemas que se tornam uma parte importante da infraestrutura da sociedade, seus líderes têm a responsabilidade adicional de serem bons administradores desses sistemas. Parte dessa gestão requer o estabelecimento de políticas para acesso justo ao sistema, inclusive para aqueles que podem ter sido excluídos. Essa gestão também exige que os profissionais de computação monitorem o nível de integração de seus sistemas na infraestrutura da sociedade. À medida que o nível de adoção muda, as responsabilidades éticas da organização ou grupo provavelmente também mudarão. O monitoramento contínuo de como a sociedade está usando um sistema permitirá que a organização ou grupo para permanecerem consistentes com suas obrigações éticas descritas no Código. Quando não existem padrões adequados de atendimento, os profissionais de computação têm o dever de garantir que sejam desenvolvidos.

Por fim, ao especificar princípios relacionados com a **conformidade com o Código**, o documento destaca:

- 1) Defender, promover e respeitar os princípios do Código.

O futuro da computação depende da excelência técnica e ética. Os profissionais de informática devem aderir aos princípios do Código e contribuir para aprimorá-los. Os profissionais de computação que reconhecem violações do Código devem tomar medidas para resolver as questões éticas que reconhecem, incluindo, quando razoável, expressar sua preocupação à pessoa ou pessoas que estejam violando o Código.

- 2) Tratar as violações do Código como inconsistentes com a adesão ao ACM.

Cada membro do ACM deve incentivar e apoiar a adesão de todos os profissionais de computação, independentemente da associação ao ACM. Os membros da ACM que reconhecem uma violação do Código devem considerar relatar a violação à ACM, o que pode resultar em ação corretiva conforme especificado no Código de Ética e Política de Aplicação de Conduta Profissional da ACM.

4.2.4 Google

A Google – empresa de tecnologia multinacional americana que se concentra em tecnologia de mecanismos de pesquisa, publicidade online, computação em nuvem, software de computador, computação quântica, comércio eletrônico, inteligência artificial e eletrônicos de consumo – enaltece que o potencial de crescimento e uso da IA vem atrelado a dúvidas igualmente importantes.

Por isso anunciaram em junho 2018 sete princípios que vão orientar o de desenvolvedores e pesquisadores. Segundo a empresa, não se trata de conceitos teóricos e abstratos: são padrões concretos que vão guiar, na prática, nosso trabalho de pesquisa e desenvolvimento de produtos. Esses princípios serão usados para tomar decisões reais de negócios.

Segundo a Google, a IA deve:

- 1) Trazer benefícios para a sociedade

O crescente alcance das novas tecnologias atinge porções cada vez maiores da sociedade. O avanço da inteligência artificial causará impactos transformadores numa série de campos, como saúde, segurança, energia, transporte, manufatura e entretenimento. Quando pensarmos no desenvolvimento de possíveis usos dessa tecnologia, levaremos em conta uma série de fatores sociais e econômicos, e só seguiremos adiante nos casos em que, na nossa avaliação, os benefícios gerais superarem de longe os riscos e as desvantagens previsíveis.

A IA também aumenta nossa capacidade de compreender, em grande escala, o significado de diversos conteúdos. Trabalharemos para garantir que informações precisas e de qualidade estejam sempre disponíveis graças à IA, respeitando normas culturais, sociais e jurídicas dos países onde atuamos. Continuaremos avaliando cuidadosamente as situações em que nossas tecnologias devem estar disponíveis de forma não-comercial.

2) Evitar a criação e a afirmação de ideias e comportamentos injustos e tendenciosos

Algoritmos e conjuntos de dados alimentados por IA podem refletir, reforçar e até reduzir ideias e comportamentos injustos e tendenciosos. Sabemos que nem sempre é fácil diferenciar o justo do injusto, pois os conceitos variam entre culturas e sociedades. Buscaremos sempre evitar impactos injustos sobre as pessoas, principalmente aqueles relacionados a questões delicadas como raça, etnia, gênero, nacionalidade, renda, orientação sexual, habilidades, crenças religiosas e políticas.

3) Ser criada e testada para garantir segurança

Vamos continuar desenvolvendo e aplicando práticas sólidas e fortes de segurança, para evitar efeitos não-desejados que apresentem algum risco de serem prejudiciais. Nossos sistemas de IA serão projetados para agir com cautela, e trabalharemos para desenvolvê-los de acordo com as melhores práticas da pesquisa segura em inteligência artificial. Nos casos em que considerarmos adequado, testaremos tecnologias de AI em ambientes restritos, monitorando as operações e o posterior desenvolvimento.

4) Ser responsável com os usuários

Vamos projetar sistemas de IA que garantam oportunidades adequadas de receber comentários e sugestões, explicações relevantes e de interesse. Nossas tecnologias de IA estarão sujeitas ao devido direcionamento e controle humano.

5) Ser projetadas de acordo com princípios de privacidade

O desenvolvimento e uso da inteligência artificial irão respeitar os princípios de privacidade do Google. Haverá avisos e solicitações de consentimento; incentivaremos arquiteturas que protejam a privacidade; garantiremos a transparência e o controle adequados sobre o uso dos dados.

6) Observar padrões elevados de excelência científica

O método científico e o compromisso com investigação aberta, rigor intelectual, integridade e colaboração são fundamentais para a inovação tecnológica. As ferramentas de IA são capazes de abrir novas portas para pesquisa científica e conhecimento em áreas fundamentais, como biologia, química, medicina e ciências ambientais. Buscaremos os mais altos padrões de excelência científica em nosso trabalho pelo progresso da IA.

Temos também o compromisso de compartilhar os conhecimentos da IA de forma responsável, publicando materiais educativos, melhores práticas e pesquisas que levem para mais pessoas a possibilidade de desenvolver aplicações úteis usando IA.

7) Estar disponível para outras pessoas, para utilizações que respeitem os princípios acima

Tecnologias podem ser usadas para diferentes fins. Trabalharemos para limitar aplicações com potencial de causar danos ou levar a abusos. Nos reservamos o direito de evitar e impedir o uso da nossa tecnologia nos casos em que soubermos de aplicações que não respeitem nossos princípios.

Segundo o documento do Google, à medida que desenvolverem e utilizarem a inteligência artificial, buscaremos responder às seguintes perguntas:

“Objetivo e uso primário: o propósito principal e a utilização provável de tecnologias e aplicações podem levar a um fim prejudicial? A tecnologia oferece possibilidade de ser adaptada para objetivos danosos?”

Natureza e diferenciação: a tecnologia que tornamos disponível – seja por publicações, open-sourcing, distribuição geral ou acordos específicos – é única e diferente? Ou está amplamente disponível?

Escala: o uso dessa tecnologia terá um impacto considerável?

Natureza do envolvimento do Google: estamos fornecendo ferramentas de uso geral, integrando ferramentas para clientes ou desenvolvendo soluções sob medida? Aplicações de IA que não farão parte da nossa estratégia”.

4.2.4 Microsoft

A Microsoft, empresa transnacional dos Estados Unidos que desenvolve, fabrica, licencia, apoia e vende softwares de computador, produtos eletrônicos, computadores e serviços pessoais, elaboraram um conceito chamado “IA RESPONSÁVEL”, com vídeos explicativos, em que se comprometem com o avanço da AI impulsionados por princípios éticos que colocam as pessoas em primeiro lugar, especificamente:

- 1) Justiça;
- 2) Inclusão;
- 3) Confiabilidade e segurança;
- 4) Transparência;
- 5) Segurança e privacidade;
- 6) Responsabilidade.

4.2.5 Diretrizes Universais para Inteligência Artificial (*Public Voice Coalition, 2018*)

A coalizão Public Voice foi estabelecida em 1996 pelo *Electronic Privacy Information Center* (EPIC) para promover a participação pública nas decisões relativas ao futuro da Internet.

O Public Voice tem buscado questões que vão desde privacidade e liberdade de expressão até proteção ao consumidor e governança da Internet. Por meio de conferências internacionais, relatórios e financiamento de viagens, o projeto Public Voice busca aumentar a presença de ONGs em reuniões em todo o mundo. Em cooperação com o ICDPPC, a OCDE, a UNESCO e outras organizações internacionais, o projeto Public Voice coloca os líderes da sociedade civil frente a frente com funcionários do governo para um engajamento construtivo sobre questões políticas atuais.

Em 23 de outubro de 2018, em Bruxelas, na Bélgica, publicaram Diretrizes Universais para informar e melhorar o design e o uso da IA.

As Diretrizes visam maximizar os benefícios da IA, minimizar o risco e garantir a proteção dos direitos humanos. Essas Diretrizes devem ser incorporadas às normas éticas, adotadas na legislação nacional e nos acordos internacionais e incorporadas ao projeto de sistemas. Afirmamos claramente que a principal responsabilidade pelos sistemas de IA deve residir nas instituições que financiam, desenvolvem e implantam esses sistemas.

O guia enumera os seguintes princípios:

1) Direito à Transparência.

Todos os indivíduos têm o direito de conhecer a base de uma decisão de IA que lhes diz respeito. Isso inclui acesso aos fatores, à lógica e às técnicas que produziram o resultado.

2) Direito à Determinação Humana.

Todos os indivíduos têm direito a uma determinação final feita por uma pessoa.

3) Obrigação de Identificação.

A instituição responsável por um sistema de IA deve ser divulgada ao público.

4) Obrigação de Justiça.

As instituições devem garantir que os sistemas de IA não reflitam preconceitos injustos ou tomem decisões discriminatórias inadmissíveis.

5) Obrigação de Avaliação e Responsabilidade.

Um sistema de IA deve ser implantado somente após uma avaliação adequada de sua finalidade e objetivos, seus benefícios, bem como seus riscos. As instituições devem ser responsáveis pelas decisões tomadas por um sistema de IA.

6) Obrigações de Exatidão, Confiabilidade e Validade.

As instituições devem garantir a precisão, confiabilidade e validade das decisões.

7) Obrigação de Qualidade de Dados.

As instituições devem estabelecer a proveniência dos dados e garantir a qualidade e relevância da entrada dos dados nos algoritmos.

8) Obrigação de Segurança Pública.

As instituições devem avaliar os riscos de segurança pública que surgem da implantação de sistemas de IA que direcionam ou controlam dispositivos físicos e implementam controles de segurança.

9) Obrigação de segurança cibernética.

As instituições devem proteger os sistemas de IA contra ameaças de segurança cibernética.

10) Proibição de Perfil Secreto.

Nenhuma instituição deve estabelecer ou manter um sistema secreto de perfis.

11) Proibição de pontuação unitária. Nenhum governo nacional deve estabelecer ou manter uma pontuação de propósito geral para seus cidadãos ou residentes.

12) Obrigação de Rescisão.

Uma instituição que estabeleceu um sistema de IA tem a obrigação afirmativa de encerrar o sistema se o controle humano do sistema não for mais possível.

4.3 INICIATIVAS ACADÊMICAS

4.3.1 Universidade de Montreal – Canadá

A Declaração de Montreal para o Desenvolvimento Responsável da Inteligência Artificial foi elaborada pela Universidade de Montreal com o apoio do Instituto Quebec de inteligência artificial.

A Declaração de Montreal pelo Desenvolvimento Responsável da Inteligência Artificial persegue 3 objetivos:

1. Elaborar um quadro ético para o desenvolvimento e a implementação da IA;
2. Orientar a transição digital para que todos possam se beneficiar dessa revolução tecnológica;
3. Abrir um espaço para o diálogo nacional e internacional de modo a obter sucesso coletivo no desenvolvimento inclusivo, equitativo e ecologicamente sustentável da IA.

O objetivo principal desta Declaração é identificar princípios e valores éticos que promovam os interesses fundamentais de indivíduos e grupos. Estes princípios aplicados aos domínios digital e da inteligência artificial permanecem gerais e abstratos.

A partir desses princípios foram elaboradas recomendações cujo objetivo é fornecer diretrizes para realizar a transição digital dentro do marco ético da Declaração.

Elas cobrem alguns temas intersetoriais fundamentais para pensar sobre a transição para uma sociedade na qual a IA promova o bem comum: governança algorítmica, alfabetização digital, inclusão digital da diversidade e sustentabilidade ecológica.

O primeiro ponto que merece evidência é que o documento, elaborado em meio acadêmico, mas em cocriação com diversos setores da sociedade, não se adstringe ao Judiciário, diferentemente da Carta da União Europeia e da Res 332 do CNJ.

A fim de se obter um desenvolvimento responsável de IA, especificamente identificar princípios e valores éticos que promovam os interesses fundamentais de indivíduos e grupos, eles definem princípios, que compõem o arcabouço mínimo do quadro ético para o desenvolvimento e a implementação da IA

1) PRINCÍPIO DO BEM-ESTAR.

O desenvolvimento e o uso de Sistemas de Inteligência Artificial (SIAs) devem permitir aumentar o bem-estar de todos os seres sencientes.

Os SIAs devem permitir que os indivíduos melhorem suas condições de vida, de saúde e de trabalho. Os SIAs devem permitir que os indivíduos satisfaçam suas preferências, dentro dos limites do que não cause danos a outro ser senciente. Os SIAs devem permitir que os indivíduos exerçam suas capacidades físicas e intelectuais. Os SIAs não devem ser fonte de desconforto, a menos que este último possa gerar um maior bem-estar que não possamos alcançar de outra forma. O uso de SIAs não deve contribuir para o aumento do estresse, da ansiedade e de sentimentos de assédio ligados ao ambiente digital.

2) PRINCÍPIO DO RESPEITO À AUTONOMIA

Os SIAs devem ser desenvolvidos e usados respeitando-se a autonomia das pessoas e com o objetivo de aumentar o controle, pelos indivíduos, de sua vida e de seu meio ambiente.

3) PRINCÍPIOS DE PROTEÇÃO DA INTIMIDADE E DA VIDA PRIVADA

A privacidade e a intimidade devem ser protegidas de intrusão por SIAs e por Sistemas de Aquisição e Arquivamento de Dados Pessoais (SAADs).

A intimidade do pensamento e das emoções deve ser estritamente protegida contra o uso de SIAs e SAADs que poderiam lhe causar danos, em particular quando usados para julgar moralmente pessoas ou suas escolhas de vida.

As pessoas devem sempre ter a opção de desconexão digital em relação à sua vida privada, além de possuir amplo controle sobre as informações relativas às suas preferências. Os SIAs não devem criar perfis de preferências individuais que visem influenciar o comportamento dos envolvidos sem seu consentimento livre e informado.

Os SAADs devem garantir a confidencialidade dos dados e o anonimato de perfis pessoais.

Toda pessoa deve ser capaz de manter um amplo controle sobre seus dados pessoais, em particular no que diz respeito à sua coleta, uso e disseminação. O uso de SIAs e de serviços digitais não pode estar condicionado à renúncia aos direitos de propriedade sobre seus dados pessoais.

Qualquer pessoa pode doar seus dados pessoais a organizações de pesquisa para contribuir para o progresso do conhecimento.

A integridade da identidade individual deve ser garantida.

4) PRINCÍPIO DA SOLIDARIEDADE

O desenvolvimento de SIAs deve ser compatível com a manutenção de relações solidárias entre pessoas e gerações.

Os SIAs não devem prejudicar a preservação de relações afetivas e morais que floresçam entre as pessoas, e devem ser desenvolvidos com o objetivo de fomentá-las de modo a reduzir a vulnerabilidade e o isolamento das pessoas.

Os SIAs devem ser desenvolvidos para colaborar com os seres humanos em tarefas complexas e devem fomentar o trabalho colaborativo entre humanos.

Os SIAs não devem ser implementados para substituir pessoas em tarefas que exigem relacionamento humano de qualidade, mas devem ser desenvolvidos para facilitar essas relações.

O desenvolvimento de SIAs não deve estimular comportamentos cruéis com robôs que se apresentam como seres humanos ou animais e que pareçam agir como eles.

5) PRINCÍPIO DA PARTICIPAÇÃO DEMOCRÁTICA

Os SIAs devem atender aos critérios de inteligibilidade, justificabilidade e acessibilidade, e devem estar sujeitos a escrutínios, debates e controles democráticos.

O funcionamento dos SIAs que tomam decisões que afetam a vida, a qualidade de vida ou a reputação dos indivíduos deve ser inteligível para seus desenvolvedores.

As decisões dos SIAs que afetam a vida, a qualidade de vida ou a reputação dos indivíduos devem sempre ser justificadas em linguagem compreensível para aqueles que os utilizam ou que sofrem as consequências de seu uso.

O código dos algoritmos, públicos ou privados, deve estar sempre acessível às autoridades públicas competentes e às partes interessadas para fins de verificação e controle.

A descoberta de erros operacionais nos SIAs, de seus efeitos imprevistos ou indesejados, de violações de segurança e vazamentos de dados, deve ser obrigatoriamente relatada às autoridades públicas, às partes interessadas e às pessoas afetadas pela situação.

De acordo com a exigência de transparência nas decisões públicas, o código de algoritmos de decisão utilizado pelas autoridades públicas deve ser acessível a todos, com exceção dos algoritmos que apresentam, em caso de uso indevido, alta probabilidade de perigo sério.

Para que os SIAs públicos tenham um impacto significativo na vida dos cidadãos, estes devem ter a oportunidade e a competência para deliberar sobre os seus parâmetros sociais, objetivos e limites de uso.

Todo usuário de um serviço deve saber se uma decisão que lhe diz respeito ou que o afete foi tomada por um SIA.

6) PRINCÍPIO DA EQUIDADE

O desenvolvimento e o uso dos SIAs devem contribuir para a obtenção de uma sociedade justa e equitativa.

Os SIAs devem ser projetados e treinados de forma a não criar, reforçar ou reproduzir discriminação baseada em diferenças sociais, sexuais, étnicas, culturais e religiosas, entre outras.

O desenvolvimento de SIAs deve beneficiar econômica e socialmente a todos, de modo a reduzir a precariedade e as desigualdades sociais.

7) PRINCÍPIO DA INCLUSÃO DA DIVERSIDADE

O desenvolvimento e o uso de SIAs devem ser compatíveis com a manutenção da diversidade social e cultural e não devem restringir o leque de escolhas de vida e experiências pessoais.

O desenvolvimento e o uso de SIAs não devem levar a uma padronização da sociedade através da normalização de comportamentos e opiniões.

O desenvolvimento e a implantação de SIAs devem levar em conta as múltiplas expressões de diversidade social e cultural, e isso deve ocorrer desde a concepção dos algoritmos.

Os ambientes de pesquisa em IA, tanto na investigação científica quanto na indústria, devem ser inclusivos e refletir a diversidade de indivíduos e grupos na sociedade.

8) PRINCÍPIO DA PRUDÊNCIA

Todas as pessoas envolvidas no desenvolvimento de SIAs devem ser cautelosas, antecipando, tanto quanto possível, as consequências adversas do uso dos SIAs, e tomando as medidas apropriadas para evitá-las.

É necessário desenvolver mecanismos que levem em conta o potencial de uso duplo (benéfico e prejudicial) da pesquisa em IA (tanto pública quanto privada) e também a possibilidade de desenvolvimento de SIAs para limitar seu uso prejudicial.

Quando o uso inapropriado de um SIA puder representar uma séria ameaça à segurança ou saúde pública, com alta probabilidade, é prudente restringir a disseminação pública ou o acesso aberto ao seu algoritmo.

Antes de serem colocados no mercado, sejam eles pagos ou gratuitos, os SIAs devem atender a critérios rigorosos de confiabilidade, segurança e integridade, e estar sujeitos a testes que não ponham em risco a vida das pessoas, não afetem sua qualidade de vida, nem prejudiquem sua reputação ou integridade psicológica. Estes testes devem estar abertos às autoridades públicas competentes e às partes interessadas.

O desenvolvimento de SIAs deve prevenir os riscos de uso nocivo dos dados do usuário e proteger a integridade e a confidencialidade dos dados pessoais.

9) PRINCÍPIO DA RESPONSABILIDADE

O desenvolvimento e o uso de SIAs não devem contribuir para a desresponsabilização dos seres humanos quando uma decisão vier a ser tomada.

Somente os seres humanos podem ser responsabilizados por decisões decorrentes de recomendações feitas por SIAs e pelas ações decorrentes delas.

A decisão de matar deve sempre ser tomada por seres humanos e a responsabilidade por esta decisão não pode ser transferida para um SIA.

Pessoas que autorizem um SIA a cometer um crime ou delito, ou que sejam negligentes ao permiti-los, são responsáveis por eles.

No caso de um erro ter sido infligido por um SIA, e o SAI se provar confiável e tiver sido usado de maneira normal, não é razoável culpar as pessoas envolvidas em seu desenvolvimento ou uso.

10) PRINCÍPIO DO DESENVOLVIMENTO SUSTENTÁVEL

O desenvolvimento e o uso de SIAs devem ser realizados de forma a garantir uma forte sustentabilidade ecológica do planeta.

4.3.2 Princípios de Inteligência Artificial de Beijing

O documento, produzido no âmbito científico, reúne a Academia de Inteligência Artificial de Pequim (BAAI), a Universidade de Pequim, a Universidade de Tsinghua, o Instituto de Automação, Academia Chinesa de Ciências, o Instituto de Tecnologia da Computação, a Academia Chinesa de Ciências e a Aliança de Estratégia de Inovação da Indústria de Inteligência Artificial (AITISA).

Segundo o texto, firmado em 25 de maio de 2019, o desenvolvimento da Inteligência Artificial (IA) diz respeito ao futuro de toda a sociedade, de toda a humanidade e do meio ambiente.

Os princípios abaixo são propostos como uma iniciativa para a pesquisa, desenvolvimento, uso, governança e planejamento de longo prazo da IA, exigindo seu desenvolvimento saudável para apoiar a construção de uma comunidade de destino comum e a realização de IA benéfica para a humanidade e natureza.

No que se refere à pesquisa e desenvolvimento (P&D) de IA deve observar os seguintes princípios:

- 1) Fazer o bem: A IA deve ser projetada e desenvolvida para promover o progresso da sociedade e da civilização humana, para promover o desenvolvimento sustentável da natureza e da sociedade, para beneficiar toda a humanidade e o meio ambiente e para melhorar o bem-estar da sociedade e da ecologia.
- 2) Para a Humanidade: A P&D da IA deve servir à humanidade e estar em conformidade com os valores humanos, bem como com os interesses gerais da humanidade. A privacidade, a dignidade, a liberdade, a autonomia e os direitos humanos devem ser suficientemente respeitados. A IA não deve ser usada para contra, utilizar ou prejudicar seres humanos.
- 3) Seja responsável: Pesquisadores e desenvolvedores de IA devem ter considerações suficientes sobre os possíveis impactos e riscos éticos, legais e sociais trazidos por seus produtos e tomar ações concretas para reduzi-los e evitá-los.
- 4) Riscos de controle: Esforços contínuos devem ser feitos para melhorar a maturidade, robustez, confiabilidade e controlabilidade dos sistemas de IA, de modo a garantir a segurança dos dados, a segurança do próprio sistema de IA e a segurança do ambiente externo onde o sistema de IA é implantado.
- 5) Seja ético: A P&D de IA deve adotar abordagens de design éticas para tornar o sistema confiável. Isso pode incluir, mas não se limita a: tornar o sistema o mais justo possível, reduzir possíveis discriminações e preconceitos, melhorar sua

transparência, explicabilidade e previsibilidade e tornar o sistema mais rastreável, auditável e responsável.

6) **Seja diverso e inclusivo:** o desenvolvimento da IA deve refletir a diversidade e a inclusão e ser projetado para beneficiar o maior número possível de pessoas, especialmente aquelas que seriam facilmente negligenciadas ou sub-representadas nas aplicações de IA.

7) **Abrir e compartilhar:** É incentivado o estabelecimento de plataformas abertas de IA para evitar monopólios de dados/plataforma, para compartilhar ao máximo os benefícios do desenvolvimento de IA e promover oportunidades iguais de desenvolvimento para diferentes regiões e indústrias.

Quanto aos usos da IA deve observar os seguintes princípios:

8) **Use de forma sábia e adequada:** Os usuários de sistemas de IA devem ter o conhecimento e a capacidade necessários para fazer o sistema operar de acordo com seu projeto e ter compreensão suficiente dos possíveis impactos para evitar possíveis usos indevidos e abusos, de modo a maximizar seus benefícios e minimizar os riscos.

9) **Consentimento informado:** devem ser tomadas medidas para garantir que as partes interessadas dos sistemas de IA tenham consentimento informado suficiente sobre o impacto do sistema em seus direitos e interesses. Quando ocorrerem circunstâncias inesperadas, mecanismos razoáveis de revogação de dados e serviços devem ser estabelecidos para garantir que os direitos e interesses dos próprios usuários não sejam infringidos.

10) **Educação e treinamento:** as partes interessadas dos sistemas de IA devem poder receber educação e treinamento para ajudá-los a se adaptar ao impacto do desenvolvimento da IA nos aspectos psicológicos, emocionais e técnicos.

4.3.3. AI4PEOPLE - AI4PEOPLE'S Ethical Framework for a Good AI Society

O AI4People é o primeiro fórum europeu que reúne as principais partes interessadas, como academia, indústria, organizações da sociedade civil e o Parlamento Europeu, para estabelecer as bases de uma “Boa Sociedade de IA”, moldando o impacto da Inteligência Artificial

O AI4People's Ethical Framework for a Good AI Society é o resultado do esforço colaborativo do AI4People para propor uma série de recomendações para o desenvolvimento de uma Good AI Society. Ele sintetiza três coisas: as oportunidades e os riscos associados que as tecnologias de IA oferecem para fomentar a dignidade humana e promover o florescimento humano; os princípios que devem fundamentar a adoção da IA; e vinte recomendações específicas que, se adotadas, permitirão a todas as partes interessadas aproveitar as oportunidades, evitar ou pelo menos minimizar e contrabalançar os riscos, respeitar os princípios e, portanto, desenvolver uma Good AI Society.

Em novembro de 2018, sob a liderança do professor Luciano Floridi, os pesquisadores apresentaram no Parlamento Europeu o “AI4People's Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations”, resultado do primeiro ano de atividade da AI4People. Este importante resultado serviu de inspiração à Comissão Europeia

e orientou a identificação dos 7 Requisitos Chave para uma IA Confiável apresentados pela Comissão em abril de 2019.

Segundo o texto,

“Vários grupos de múltiplas partes interessadas criaram declarações de princípios éticos que devem orientar o desenvolvimento e a adoção da IA. Em vez de repetir o mesmo processo aqui, apresentamos uma análise comparativa de vários desses conjuntos de princípios. Cada princípio expresso em cada um dos documentos que analisamos é encapsulado por um dos cinco princípios abrangentes. Quatro deles – beneficência, não maleficência, autonomia e justiça – são princípios estabelecidos de ética médica, mas um quinto – explicabilidade – também é necessário para capturar os novos desafios éticos colocados pela IA”.

O framework AI4people, inspirado nos princípios da Bioética, propõe cinco princípios abrangentes como sendo os mais importantes a serem levados em consideração:

- 1) Beneficência, referente à promoção do bem-estar geral, preservação da dignidade e sustentabilidade do planeta;
- 2) Não maleficência, estabelecendo que a ação (mesmo benéfica) deve causar o menor dano;
- 3) Justiça, estabelecendo a equidade como condição fundamental;
- 4) Autonomia, exigindo que os agentes tenham habilidades e competências para tomar decisões, e promovendo a autonomia dos seres humanos; e
- 5) Explicabilidade, que é a necessidade de entender e responsabilizar os processos de tomada de decisão da IA.

Posteriormente, já em 2020, por ocasião da AI4People Summit, o grupo de pesquisadores, seguindo seu trabalho anterior sobre ética de IA, identificou sete setores estratégicos (Automotivo, Bancário e Financeiro, Energia, Saúde, Seguros, Serviços Jurídicos, Mídia e Tecnologia) para a implantação de IA ética, nomeando 7 diferentes comitês para analisar como a IA confiável pode ser implementada nesses setores: os 7 AI Global Frameworks do AI4People é o resultado deste esforço.

Merece registro que um dos setores elencados é justamente o de serviços jurídicos, de forma que o documento estabelece princípios fundamentais para o uso responsável da IA pelo direito, desenvolvendo uma série de princípios que concretizam as preocupações éticas específicas para este setor, que devem informar tanto o desenvolvimento quanto o uso da IA na lei em um alto nível de abstração que se aplica ao todo variedade de tecnologias e casos de uso.

Segundo os autores, “como dizem respeito aos fundamentos do direito e ao estado de direito, eles não podem ser reduzidos à ética; na verdade, eles dizem respeito à natureza do direito positivo moderno conforme informado pelo estado de direito”.

Os princípios elencados são:

- A) Princípios fundamentais para o uso responsável da IA no direito
 - 1) Princípio do Respeito à Integridade da Lei e ao Estado de Direito
 - 2) Princípio da intencionalidade
 - 3) Princípio do respeito (natureza situada) do estado de direito

Princípio do respeito pela natureza situada do estado de direito:

4) Princípios de distribuição justa de impacto em nível individual e social

Princípio da distribuição aérea de riscos e benefícios:

B) Princípios para o Desenvolvimento Responsável da IA no Direito

1) Princípio da transparência processual

2) Princípio do respeito ao processo legislativo

3) Princípio da participação pública irrestrita

Princípio da compensação transparente e adequada

Princípio da diversidade e representação

Princípio da não discriminação

Princípio da contestabilidade temporal

C) Princípios para o emprego responsável da IA no Direito

1) *Accountability*: Princípios para uma IA Legal Responsável na Administração Pública

Princípio da validação empírica contínua

Princípio da contestabilidade

Princípios de justificação e explicação

Princípio da reparação prática e eficaz

2) *Justiciabilidade*: Princípios para uma IA Responsável no Judiciário

Princípio da justiça pública

Princípio do acesso igualitário à justiça

3) *Responsabilidade*: Princípios da IA Responsável na Prática Jurídica

Princípio da aplicação equivalente das normas deontológicas profissionais

Princípio da responsabilidade final

Princípio da responsabilidade coletiva

Princípio da neutralidade tecnológica

4.4 QUADRO-RESUMO ÉTICA NA IA

Com a perspectiva de possibilitar visão sistêmica, apresentamos abaixo o quadro-resumo da incidência de parâmetros éticos por cada documento analisado neste capítulo, indicando quais estândares éticos foram observados em quais textos, permitindo-se uma análise comparativa e global.

Para montagem do quadro-resumo optou-se pela leitura dos textos, verificação do parâmetro indicado e checagem de incidência de cada um desses parâmetros éticos em cada um

dos documentos elencados no capítulo anterior, identificando-se em quais textos foram citados ou não.

Partiu-se, portanto, de uma análise meramente textual, da nomenclatura ou título dado ao parâmetro pelo respectivo documento, sem interpretações de abrangência, ampliativa ou restritiva, nem tampouco exclusão de conteúdo.

Registre-se, por oportuno, que os conceitos atribuídos a cada um dos parâmetros éticos podem ser, a depender do documento, mais extenso ou com menor acepção, justamente porque não há globalmente um mesmo padrão, de forma que, por vezes, dois parâmetros são tratados, em documentos distintos, como espécie e gênero, ou gênero e espécie.

Não se optou pela clusterização ou agrupamento de similares, porquanto importaria agregar a este trabalho objeto diverso, qual seja, análise de conteúdo, contorno e interpretação dos estândaes, o que envolveria semântica e semiótica.

A lista dos parâmetros éticos identificados decorre da citação expressa, sem ulterior edição, exclusão ou validação do pesquisador, em qualquer dos documentos estudados e listados anteriormente, sendo certo que estudos futuros poderão contemplar maior aprofundamento e verticalização de conteúdo, assim como clusterização, fusão, criação de supercategorias, além de construção de parâmetros éticos que contenham mais intrínseca conexão com os serviços judiciários.

Dito isso, como última observação, é de se pontuar que a nomenclatura aqui adotada será de “parâmetros” éticos, haja vista que, fazendo-se um cotejo entre os critérios adotados, verificamos que parte deles é tratado pela literatura como princípios, outros como valores ou mecanismos. Dada a intenção metodológica de evitar edições do pesquisador, optamos por transpor essas categorias, aqui chamadas de “parâmetros” para os quadros abaixo.

Tabela 1 - Tabela de parâmetros éticos por documento analisado - parte 01

	Bem-estar	Meio-ambiente	Auditabilidade	Explicabilidade	Transparência	Publicidade	Confidencialidade	Justiça	Imparcialidade	Equidade / igualdade	Solidariedade	Liberdade	Honestidade
CHINA - plano													
CHINA - princípios	X	X		X	X	X		X		X			
REINO UNIDO	X			X				X					
BRASIL - EBI	X			X	X	X				X			
BRASIL - RES 332 CNJ					X	X		CONSIDERANDOS	CONSIDERANDOS				
BRASIL - PL 21/2020	X	X			x								
BRASIL - PL 2.338/2023	X	X	X	X	X	X		X	X	X		X	
UNIÃO EUROPEIA - CEPEJ					X				X	X			
ÍNDIA													
EUA ordem executiva 13859					X	X		X		X			
EUA - DoD			X		X					X			
FRANÇA - CNIL								X					
AUSTRÁLIA	X	X		X	X			X		X			
SINGAPURA	X		X	X	X	X		X	X	X			
OCDE	X			X	X	X		X					
G20	X			X	X	X		X					
IEEE	X					X		X					X
ACM	X	X			X	X	X	X		X			X
ASIOMAR	X			X	X							X	
GOOGLE	X							X					
MICROSOFT					X			X					
Public Voice Coalition			X		X	X		X					
UNIVERSIDADE DE MONTREAL	X									X	X		
Beijing IA Principles	X	X	X	X	X	X		X				X	
A4People	X	X	X	X	X	X		X	X	X			

Fonte: elaborada pelo autor.

Tabela 2 - Tabela de parâmetros éticos por documento analisado - parte 02

	Controle humano	Controle usuário	Contestabilidade	Qualidade	Robustez	Segurança	Confiabilidade	Prudência	Responsabilidade	Flexibilidade	Cooperação	Prestação de contas
CHINA - plano									X			
CHINA - princípios	X	X			X	X	X		X		X	
REINO UNIDO												
BRASIL - EBIA					X				X			X
BRASIL – RES 332 CNJ		X				x			X			X
BRASIL - PL 21/2020						x						
BRASIL - PL 2.338/2023	X	X	X		X	X	X		X	X		X
UNIÃO EUROPEIA - CEPEJ		X		X		X						
ÍNDIA												
EUA ordem executiva 13859				X		x				X		
EUA - DoD	X				X	X	X		X			
FRANÇA - CNIL												
AUSTRÁLIA			X			X	x		X			
SINGAPURA		X			x	X			X			X
OCDE					X	X			X			X
G20					X	X			X			X
IEEE						X						
ACM		X		X		X	X		X			X
ASILOMAR	X	X			X	x	X		X		X	X
GOOGLE	X			X		X			X		X	
MICROSOFT						X	X		X			
Public Voice Coalition	X	X		X		X	X		X			
UNIVERSIDADE DE MONTREAL		X						X	X			
Beijing IA Principles	X	X			X	X	X		X			
A4People	X	X							X			X

Fonte: elaborada pelo autor.

Tabela 3 - Tabela de parâmetros éticos por documento analisado - parte 03

	Proteção dos dados / privacidade	Participação pública / democrática	Ser humano no centro / Educação e evolução dos seres humanos	Proteção da integridade humana	Desenvolvimento sustentável / sustentabilidade	bem comum	Crescimento inclusivo / não discriminação	Direitos HUMANOS fundamentais	Qualidade de vida no trabalho	Governança / coordenação	Atenção / Vigilância contínua
CHINA - plano										X	
CHINA - princípios	X		X	X			X	X	X	X	X
REINO UNIDO	X		X	X							
BRASIL - EBIA	X		X	X	X		X				
BRASIL – RES 332 CNJ							X	X		X	
BRASIL - PL 21/2020			X		X	X	X	X			
BRASIL - PL 2.338/2023	X	X	X	X	X	X	X	X			
UNIÃO EUROPEIA - CEPEJ							X	X			
ÍNDIA											
EUA ordem executiva 13859		X		X			X			X	
EUA - DoD				X			X				X
FRANÇA - CNIL											X
AUSTRÁLIA	X		X								
SINGAPURA			X		X		X	X			
OCDE			X	X	X		X				
G20			X	X	X		X				
IEEE	X		X	X	X		X				
ACM	X	x	X	X	X	X	X	X	X		X
ASILOMAR	X		X			X	X	X		X	X
GOOGLE	X						X				
MICROSOFT	X		X				X				
Public Voice Coalition	X										
UNIVERSIDADE DE MONTREAL	X	X			X		X				
Beijing IA Principles	X	x	X	X	X	X	X	X			X
A4People	X	X	X	X	X	X	X	X			X

Fonte: elaborada pelo autor.

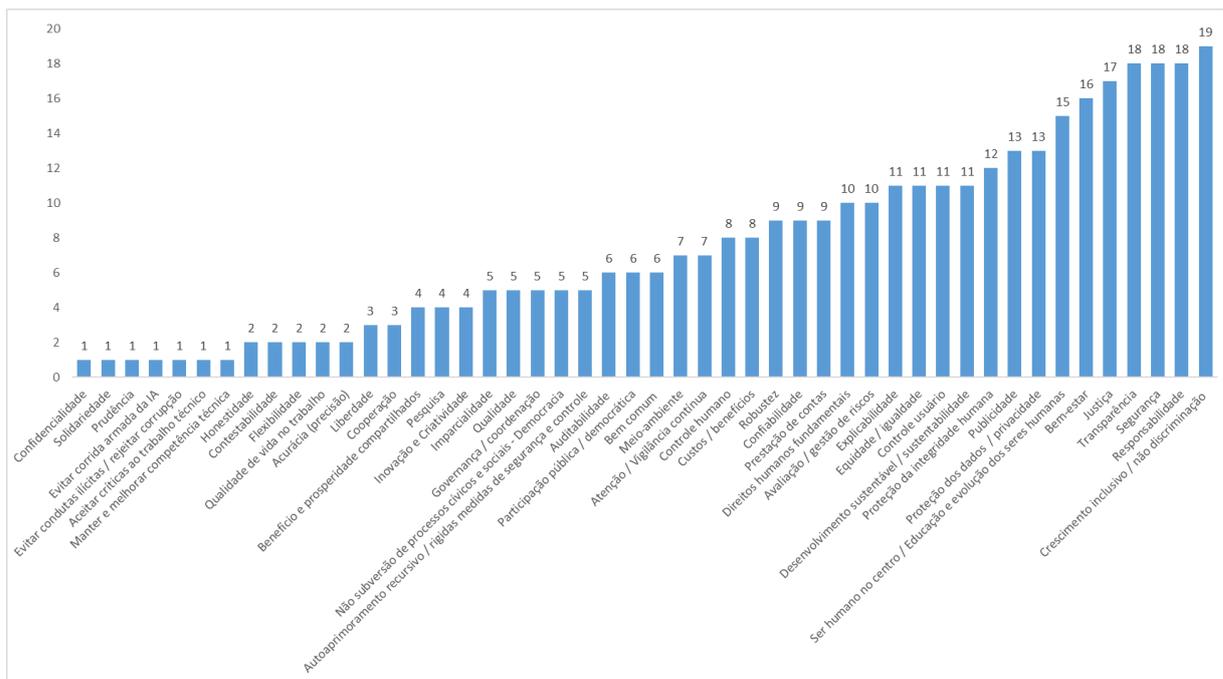
Tabela 4 - Tabela de parâmetros éticos por documento analisado - parte 04

	Avaliação /gestão de riscos	Custos / benefícios	Benefício e prosperidade compartilhados	Não subversão de processos cívicos e sociais - Democracia	Evitar corrida armada da IA	Autoaprimoramento recursivo / rígidas medidas de segurança e controle	Acurácia (precisão)	Pesquisa	Inovação e criatividade	Evitar condutas ilícitas / rejeitar corrupção	Aceitar críticas ao trabalho técnico	Manter e melhorar competência técnica
CHINA - plano			X					X				
CHINA - princípios	X		X			X			X			
REINO UNIDO												
BRASIL - EBIA												
BRASIL - RES 332 CNJ												
BRASIL - PL 21/2020				X				X	X			
BRASIL - PL 2.338/2023	X	X		X		X	X	X	X			
UNIÃO EUROPEIA - CEPEJ												
ÍNDIA												
EUA ordem executiva 13859	X	X										
EUA - DoD												
FRANÇA - CNIL												
AUSTRÁLIA												
SINGAPURA							X					
OCDE												
G20												
IEEE	X									X	X	X
ACM	X	X		X		X			X			
ASILOMAR	X	X	X	X	X	X		X				
GOOGLE	X	X										
MICROSOFT												
Public Voice Coalition	X	X										
UNIVERSIDADE DE MONTREAL												
Beijing IA Principles	X	X	X			X						
A4People	X	X		X								

Fonte: elaborada pelo autor.

Por fim, com a intenção de permitir melhor compreensão de quais foram os parâmetros éticos mais indicados e os que apresentaram menor incidência, elaboramos quadro com análise quantitativa, que segue abaixo.

Figura 1 – Incidência dos parâmetros éticos pesquisados



Fonte: elaborada pelo autor.

Vê-se, de um lado, princípios éticos com baixíssima incidência, como os da “confidencialidade”, “solidariedade”, “prudência”, “evitar corrida armada da IA”. “evitar condutas ilícitas / rejeitar corrupção”, “aceitar críticas ao trabalho técnico” e “manter e melhorar competências técnicas”, que possuem, cada um, apenas uma indicação, dentre os 25 documentos visitados.

Por outro lado, com mais de 10 inserções temos 14 princípios éticos. Dentre os mais citados, temos “crescimento inclusivo / não discriminação” (19 inclusões), responsabilidade, segurança e transparência (com 18 inclusões, cada) e justiça com 17 inserções.

Levantados os estândares éticos previstos na mais diversa literatura, impõe-nos, por fim, verificar quais desses critérios guardam relação de importância com o desenvolvimento ético da IA pelo Judiciário.

5 SURVEY – IA E ÉTICA NO JUDICIÁRIO

Levantadas hipóteses durante a entrevista narrativa quanto ao uso e desenvolvimento de IA pelo Judiciário e apresentados possíveis referenciais de parâmetros éticos para desenvolvimento da IA, convém confrontar um público mais extenso de integrantes do Poder Judiciário, quanto à pertinência desses estândares para a checagem ética dos modelos de IA em uso ou desenvolvimento na justiça.

Para tanto, valemo-nos de um *survey*.

Survey é um método de coleta de informações de uma amostra de indivíduos de uma população a ser estudada acerca de suas ideias, sentimentos, planos e opiniões.

O *survey* se presta à obtenção de dados ou informações sobre características, ações ou opiniões de determinado grupo de pessoas, indicado como representante de uma população alvo, por meio de um instrumento de pesquisa, normalmente um questionário (Tanur *apud* Pinsonneault & Kraemer, 1993).

Fink descreve como principais características do método de pesquisa *survey*: o interesse é produzir descrições quantitativas de uma população; o principal meio de coleta de informação é pelas respostas de questões estruturadas predefinidas; a informação é geralmente coletada sobre uma fração da população em estudo – a amostra; e faz uso de um instrumento predefinido.

A *survey* é apropriada como método de pesquisa quando se deseja responder a questões do tipo “o quê?”, “por quê?”, “como?” E “quanto”, ou seja, quando o foco de interesse é sobre “o que está acontecendo” ou “como e por que isso está acontecendo”. Também é adequado quando o ambiente natural é a melhor situação para estudar os fenômenos de interesse e o objeto de interesse ocorre no presente ou no passado recente.

Pinsonneault & Kraemer (1993) classificam a pesquisa *survey* quanto ao seu propósito em: a) explanatória – tem como objetivo testar uma teoria e as relações causais, estabelece a existência de relações causais, mas também questiona porque a relação existe; b) exploratória – o objetivo é familiarizar-se com o tópico ou identificar os conceitos iniciais sobre um tópico, dar ênfase na determinação de quais conceitos devem ser medidos e como devem ser medidos, buscar descobrir novas possibilidades e dimensões da população de interesse; c) descritiva – busca identificar quais situações, eventos, atitudes ou opiniões estão manifestamos em uma população, como também descreve a distribuição de algum fenômeno na população ou entre subgrupos da população ou, ainda, faz uma comparação dentre essas distribuições; neste tipo de *survey* a hipótese não é causal, mas tem proposto de verificar se a percepção dos fatos está ou não de acordo com a realidade.

5.1 SURVEY: ADEQUAÇÃO, METODOLOGIA, ORGANIZAÇÃO E CONDUÇÃO

Da entrevista narrativa realizada com membros do Poder Judiciário – que teve como perspectiva, a partir da indução, gerar percepções quanto ao uso e desenvolvimento ético de IA pelo Judiciário – e após pesquisar e visitar diversos documentos e regulamentos sobre o contorno dos parâmetros éticos pontuados pelo mercado, por universidades e por governos, que levantou 48 estândares diferentes, é que se chega à última etapa desta pesquisa, que busca, através de um *survey*, de forma mais ampla, com um público-alvo definido e maior, testar as teorias levantadas na entrevista narrativa e suas respectivas relações causais, como também

captar opiniões, costumes ou características deste público, para se familiarizar com o conceito de ética na IA sob o olhar do Judiciário.

A *survey* é adequada ao presente estudo, porque se presta a testar as hipóteses e percepções levantadas durante a entrevista narrativa² e suas respectivas relações causais (propósito explanatório), bem assim porque, em um cenário de fluidez conceitual de ética na IA, agrega a este trabalho familiarização conceitual sobre quais os estandartes éticos devem ser considerados, promovendo-se medição de sua respectiva importância.

Utilizamos de amostra não probabilística, diante da dificuldade de identificação e acesso a toda a população-alvo, que se adstringe a profissionais de tecnologia da informação de todos os tribunais e órgãos da Justiça brasileira, bem assim juízes, desembargadores, ministros e conselheiros com conhecimento em TI e/ou gestores de tecnologia da informação em seus respectivos órgãos.

A pesquisa foi disponibilizada e realizada em todos os ramos da Justiça – Federal, Estadual, Trabalho, Eleitoral e Militar –, englobando também os tribunais superiores (Supremo Tribunal Federal, Superior Tribunal de Justiça, Tribunal Superior Eleitoral, Superior Tribunal Militar e Tribunal Superior do Trabalho) e órgãos de gestão nacional (Conselho Nacional de Justiça, Conselho da Justiça Federal, Conselho da Justiça do Trabalho).

O formulário foi organizado em cinco etapas:

- 1) Primeiro buscou-se identificar o perfil do entrevistado, se magistrado (a) ou servidor (a), qual a formação e órgão da justiça que integra;
- 2) Depois buscou-se avaliar o contato e o conhecimento do entrevistado com a tecnologia da inteligência artificial, a frequência de utilização ou desenvolvimento de ferramentas de IA pelo entrevistado e pelo tribunal que ele integra;
- 3) Em sequência mediu-se o conhecimento do entrevistado quanto ao tema ética e o quanto ele se sente preparado/apto a aplicar conceitos éticos no desempenho concreto do seu trabalho;
- 4) Após, aferiu-se a percepção do entrevistado quanto ao uso da IA pela Justiça, bem como indagou-se quanto à importância de uma abordagem ética e da existência de uma lista de verificação (checklist) para conferência de conformidade ética e legal dos modelos de IA utilizados pelo Judiciário;
- 5) Por fim, foram apresentadas todas as 48 categorias identificadas no quadro 3.4, solicitando ao entrevistado que avaliasse a importância de cada um dos critérios listados com o tema ética e IA na Justiça.

Executada em um intervalo 10 dias, mediante disponibilização de formulário eletrônico, divulgado nos grupos e aplicativos de mensagens, a pesquisa obteve 102 respostas voluntárias.

² Relativas a: motivação e interesse; benefícios e riscos; governança da IA, gestão e regulamentação deste tema nos órgãos da Justiça; desafios éticos no desenvolvimento de IA pelos tribunais nacionais; fases em que a abordagem ética deve ser priorizada; definição de critérios éticos que precisam ser checados; necessidade de adoção de um framework ético objetivo e concreto.

5.2 SURVEY: RESULTADOS

Passamos, agora, à apresentação e interpretação dos dados obtidos após a execução da pesquisa.

5.2.1 Perfil do Entrevistado: Qual Órgão ou Tribunal Representa?

No que se refere à primeira etapa da pesquisa, que intenta compreender o perfil dos entrevistados, de partida é importante esclarecer que, por força de determinação da Constituição Federal (art. 92), o Poder Judiciário é composto de diversos órgãos, juízes e tribunais, que são agrupados por distribuição de competências realizada pela própria Constituição.

Os juízes e tribunais inferiores integram ramos da Justiça, podendo ser Federais, Estaduais, do Trabalho, Eleitorais e Militares. Estes tribunais inferiores possuem competências territoriais e geográficas definidas, normalmente com subdivisão correlacionada ao Estado-membro, mas, por vezes, um tribunal abrange diversos estados. Daí que há no Brasil 27 tribunais estaduais, 6 tribunais federais, 24 tribunais do trabalho, 27 tribunais eleitorais e 3 tribunais militares³.

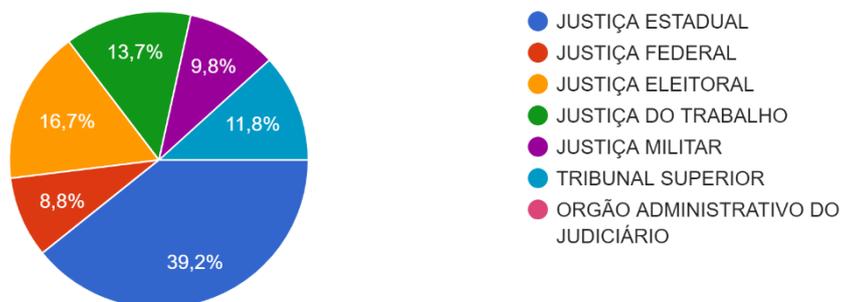
Há também os tribunais superiores (Supremo Tribunal Federal, Superior Tribunal de Justiça, Tribunal Superior Eleitoral, Superior Tribunal Militar e Tribunal Superior do Trabalho) e órgãos administrativos e de gestão nacionais (Conselho Nacional de Justiça, Conselho da Justiça Federal, Conselho da Justiça do Trabalho).

Todos estes órgãos possuem autonomia financeira, orçamentária e administrativa, de maneira que, dada a extensão do território do Brasil e as multifacetadas realidades locais, as políticas públicas adotadas em relação a qualquer tema – inclusive em tecnologia da informação e inteligência artificial – possuem certo grau de individualidade, de onde se considera que os dados da pesquisa precisam ser interpretados com atenção para as peculiaridades de cada ramo estudado.

Gráfico 1 – Qual órgão ou Instituição você integra?

QUAL ÓRGÃO OU INSTITUIÇÃO VOCÊ INTEGRA?

102 respostas



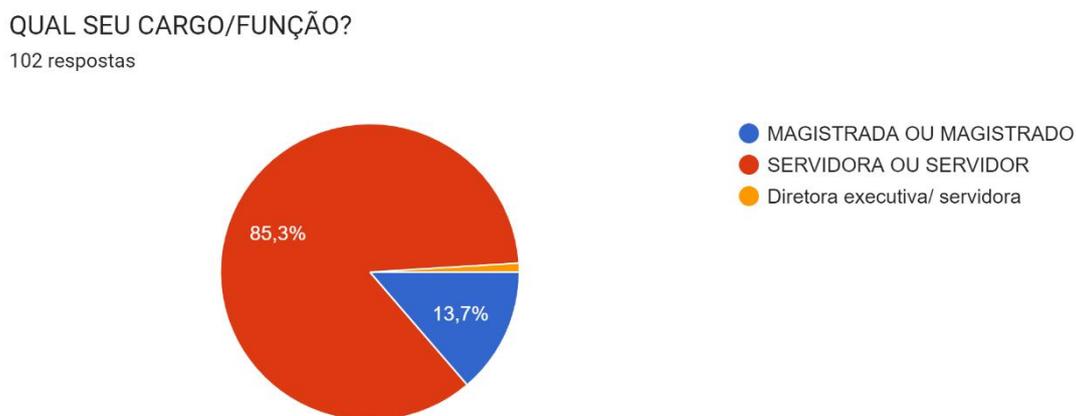
Fonte: Produzido pelo autor.

³ Para maiores informações: <https://www.cnj.jus.br/poder-judiciario/tribunais/>

Neste ponto, a pesquisa obteve satisfatória representatividade, uma vez que contou com 40 respostas da Justiça Estadual (39,2% dos entrevistados), 17 respostas da Justiça Eleitoral (16,7% dos entrevistados), 14 respostas da Justiça do Trabalho (13,7% dos entrevistados), 12 respostas de tribunais superiores (11,8% dos entrevistados), 10 respostas da Justiça Militar (9,8% dos entrevistados), 9 respostas da Justiça Federal (8,8% dos entrevistados).

5.2.2 Perfil do Entrevistado: Qual cargo/função?

Gráfico 2 – Qual seu cargo/função?



Fonte: Produzido pelo autor.

Ainda interpretando os dados relativos ao perfil dos entrevistados, vê-se que a grande maioria de respostas adveio de servidores ou servidoras públicas (88 respostas) e pequena parcela é de magistrados e magistradas (14 respondentes).

É de se ressaltar que, embora, aparentemente, a quantidade de respostas do público-alvo de magistrados e magistradas possa parecer pequena, o fato é, em verdade, o número é expressivo e significativo.

Isto porque o número de juízes e juízas (18.035) é proporcionalmente diminuto se considerada a população total no país, mas também pequena se considerada a população-alvo (pessoas que integram o Poder Judiciário). Mais que isso, são raros os profissionais do direito, especialmente juízes e juízas, que possuem algum conhecimento teórico e científico em tecnologia da informação e, especialmente, no tema ora pesquisado. Além disso, menor ainda é o número destes profissionais que exercem algum tipo de gestão na política pública de seus tribunais em tecnologia da informação.

É bem verdade que, com o advento da Res nº 370/2021 do CNJ, que determinou a instalação de comitês gestores de TI com múltipla e heterogênea composição em todos os tribunais nacionais, houve crescente participação de juízes, desembargadores, ministros e conselheiros no tratamento do tema de TI, tanto em âmbito nacional, como local, assumindo esses profissionais funções de gestão de tecnologia da informação em seus respectivos órgãos.

Entretanto, ainda que considerado que todos os tribunais nacionais tenham um magistrado ou magistrada integrante desses comitês, a amostra total seria de, no máximo, 95 pessoas, daí que, a resposta de 14 magistrados é bastante representativa e suficiente para os fins aqui pesquisados, em que não se busca amostra não probabilística.

Afora esse registro, é de se relevar que a grande parte dos serviços de TI prestados pelo Judiciário, no âmbito de sua atuação, fica a cargo de servidores e servidoras, concursados, que representam, conforme dados constantes do painel do Panorama de Tecnologia da Informação e Comunicação do Poder Judiciário, a quantidade de quase 6.822 cargos ocupados.

5.2.3 Perfil do Entrevistado: Formação Acadêmica

Gráfico 3 – Formação Acadêmica

FORMAÇÃO ACADÊMICA

102 respostas



Fonte: Produzido pelo autor.

Para finalizar a análise de perfil dos entrevistados, verificou-se que há relevante pulverização de formação acadêmica entre os entrevistados, sendo apontados mais de 24 formações diferentes no público-alvo.

Dos dados é de se extrair que imensa maioria (65,68%) possui formação em tecnologia da informação, o que representa o total de 67 entrevistados – e aqui somamos os entrevistados com formação em ciência da computação (41), análise e desenvolvimento de sistemas (10), engenharia da computação (5), sistemas da informação (5), engenharia de software (2), segurança da informação (1), estrutura de redes (1), além de dois entrevistados que reportaram múltipla formação, dentre elas em dos cursos anteriormente mencionados.

Merece destaque que um dos entrevistados não possui qualquer formação academia e outro está cursando estatística.

Digno de nota que 19,6% dos entrevistados, o que representa 20 pessoas, possuem graduação em Direito, o que se justifica pela presença de 14 magistrados e magistradas entre os entrevistados, assim como pelo costume de servidores ou servidoras públicas da área de TI que ingressam no Judiciário acabarem por cursar também Direito (há dois entrevistados com formação em TI e formação em Direito).

Interessante e, talvez, até curioso, que entre os entrevistados há pessoas com formação em áreas completamente dissociadas do ramo jurídico e até mesmo da ciência da computação, como, por exemplo, entrevistados com formação em turismo (1), psicologia (1), comunicação social (1), logística (1), engenharia civil (1), engenharia química (1) e administração de empresas (4).

Duas constatações afiguram-se importantes nesse ponto. A primeira é que, na atualidade, o interesse em temas afetos à tecnologia da informação se irradia pelos mais diversos campos do conhecimento humano e, mais que isso, também se constata a diversidade de composição das equipes de TI nos tribunais nacionais, o que, em nosso sentir, confere ainda mais variedade de visão riqueza no processo de desenvolvimento de tecnologia.

A segunda é que, se já constatado entre os profissionais do direito e profissionais de TI durante as entrevistas narrativas, que há insegurança e incertezas no campo da aplicação concreta e IA já Justiça, notadamente diante do conteúdo ético na IA, talvez os profissionais afetos a outras áreas possam ter ainda mais empecilhos e dificuldades.

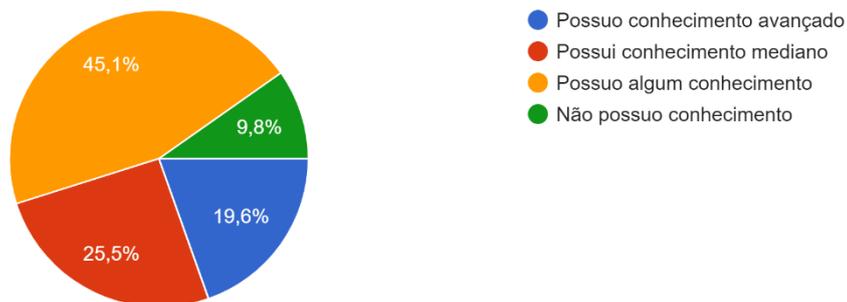
5.2.4 IA e eu: Nível de conhecimento em IA

Adentrando ao segundo tópico da pesquisa, em que se pretendeu inferir o grau de contato dos entrevistados com o tema da inteligência artificial, o que se deu pela autoavaliação de conhecimento sobre o tema, a frequência de uso da tecnologia pelo entrevistado e por seu tribunal, obtivemos resultados relevantes.

Gráfico 4 – Qual seu nível de conhecimento em IA?

QUAL O SEU NÍVEL DE CONHECIMENTO EM IA?

102 respostas



Fonte: Produzido pelo autor.

Analisando seu conhecimento pessoal sobre a inteligência artificial, 10 entrevistados responderam não possuir conhecimento sobre IA (9,8%), enquanto que a maioria do público entrevistado (46 pessoas, representando 45,1% dos entrevistados) afirmou possuir algum conhecimento sobre IA, seguidos de 26 pessoas que disseram possuir conhecimento mediano sobre o tema (25,5% dos entrevistados) e 20 pessoas asseveraram possuir conhecimento avançado (19,6% dos entrevistados).

Dáí se extrai que a maioria dos entrevistados (56 pessoas) disse não possuir ou possuir pouco conhecimento em IA, ao passo que apenas 20 pessoas afirmaram deterem conhecimentos avançados no tema.

Interessante notar que, dentre essas 56 pessoas que se autoavaliaram sem conhecimento ou com pouco conhecimento, 11 delas afirmaram fazer uso frequente ou muito frequente da tecnologia em seus tribunais

Esse dado é alarmante, quando confrontado com o contexto de amplo desenvolvimento de IA, efetiva utilização de ferramentas de IA pelos tribunais nacionais abordados no item 1. É dizer: embora a tecnologia seja muito utilizada, a maioria do pessoal não detém conhecimento mínimo.

Realizada a contextualização entre o nível de conhecimento com cargo e com formação dos entrevistados, não fora encontrado nenhuma correlação importante.

Entre os entrevistados que afirmaram não ter conhecimentos em IA, metade pertence ao a área do conhecimento de ciências da computação e a outra metade corresponde a profissionais com formação em direito, comunicação social, engenharia química ou não possuem conhecimento.

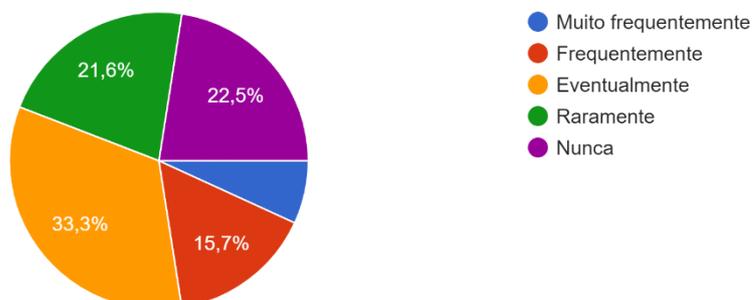
Além disso, 9 servidores afirmaram não ter conhecimento sobre IA e apenas um deles é magistrado.

Analisado o nível de conhecimento dos magistrados, sete deles (a metade) afirmaram possuir algum conhecimento, enquanto que um disse não ter conhecimento, três possuem conhecimento mediano e três possuem conhecimento avançado. Esse dado revela, surpreendentemente, que quase todos os magistrados entrevistados possuem conhecimento sobre IA, dado que se explica justamente pelo contexto de interesse nacional quanto ao tema.

5.2.5 IA e eu: Frequência do uso e desenvolvimento de ferramentas de IA

Gráfico 5 – Com qual frequência você utiliza ou desenvolve ferramentas de IA?

COM QUAL FREQUÊNCIA VOCÊ UTILIZA OU DESENVOLVE FERRAMENTAS DE IA?
102 respostas



Fonte: Produzido pelo autor.

Indagados sobre a frequência pessoal de utilização ou desenvolvimento de ferramentas de IA, 6,9% dos entrevistados afirmaram fazer uso muito frequente (7 pessoas), 15,7% disseram

usar frequentemente (16 pessoas), 33,3% relataram usar eventualmente (34 pessoas), 21,6% raramente (22 pessoas) e 22,5% nunca (23 pessoas).

Aqui ficou evidenciado que, embora em franco desenvolvimento da IA no Judiciário nacional, o fato é que a imensa maioria dos entrevistados (77,4%) disse fazer uso eventual, raro ou nunca ter feito uso da tecnologia, ao tempo em que apenas 16 entrevistados reportaram uso frequente e apenas 7 pessoas indicaram uso muito frequente.

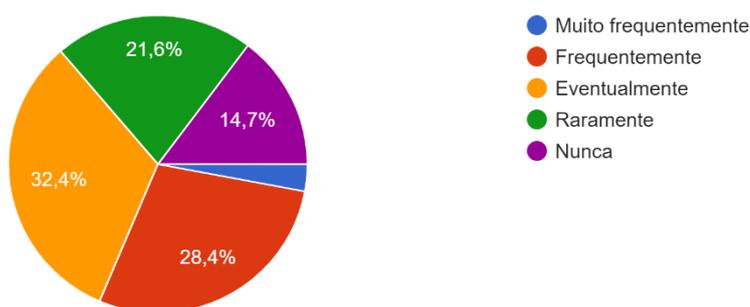
Dentre os que afirmaram fazer uso “muito frequente”, majoritariamente são os entrevistados com formação nos cursos de computação e estatística (71% dos respondentes), tendência que se repetiu, em maior escala (81% dos casos) quando feito o recorte dos que responderam “frequentemente”. Feita a leitura de modo diverso, apenas 18% dos entrevistados que responderam “frequentemente” e 29% dos que responderam “muito frequentemente” possuem graduação em direito. Este dado revela que o uso intensivo da tecnologia de IA ainda está adstrita aos profissionais formados em tecnologia da informação.

No público de magistrados, apenas 5 deles, o que representa 35% dos entrevistados, afirmaram ter utilização frequente ou muito frequentemente ferramentas de IA. A maioria deles reportou utilização eventual (5), raro (2) ou nunca (2). Embora em franco desenvolvimento e utilização da IA pelos tribunais nacionais, esse dado revela que os magistrados ainda usam pouco dos benefícios desta tecnologia.

5.2.6 IA e eu: Frequência do uso e desenvolvimento de ferramentas de IA pelo órgão ou Tribunal

Gráfico 6 – Seu Tribunal/órgão usou ou desenvolveu alguma ferramenta que utilize IA?

SEU TRIBUNAL/ÓRGÃO USOU OU DESENVOLVEU ALGUMA FERRAMENTA QUE UTILIZE IA?
102 respostas



Fonte: Produzido pelo autor.

Ainda quanto à frequência do uso e desenvolvimento de ferramentas de IA, desta vez indagados na perspectiva do tribunal ou órgão que integra o entrevistado, 2,9% afirmaram utilização ou desenvolvimento muito frequente de IA pelo tribunal (3 pessoas), ao passo em que 28,4% disseram ser frequente (29 pessoas), 32,4% eventualmente (33 pessoas), 21,6% raramente (22 pessoas) e 14,7% nunca (15 pessoas).

De igual forma, aqui constatamos que a maioria dos tribunais (quase 70%) eventualmente, raramente ou nunca utilizam ou desenvolvem sistemas utilizando de tecnologia da IA.

Feito o recorte no perfil de magistrados, é de se ver que nenhum magistrado reportou utilização muito frequente, ao passo que 6 juízes afirmaram fazer utilização frequente, mas a maioria (8), o que representa 60% dos entrevistados, reportaram uso eventual ou raro. Nenhum magistrado reportou que seu tribunal nunca utilizou ferramenta de IA.

Realizado o recorte por ramo da Justiça, verificamos que todos os magistrados federais disseram ter uso frequente (2) e o único magistrado do trabalho reportou uso raro. Não houve relevante achado quando isolados os magistrados estaduais, que seguira a regra geral, disposta no parágrafo acima.

Entre os entrevistados que responderam “nunca”, pudemos observar que todos são servidores, a grande maioria com formação em tecnologia da informação (80%). Outro achado de pesquisa nesse ponto foi verificar que os que reportaram nunca ter envolvimento do tribunal/órgão com IA estão concentrados na Justiça Militar (40%), Justiça do Trabalho (27%), Justiça Federal (13%) e Tribunais Superiores (13%). Não houve qualquer entrevistado que tenha respondido “nunca” que integre tribunal estadual.

Dentre os que responderam “muito frequentemente” ou “frequentemente” podemos observar que se concentram na Justiça Estadual (40%), Justiça Eleitoral (22%), Tribunais Superiores (18%) e Justiça Federal (12%).

Esse dado reflete e até corrobora pesquisa do CNJ, de onde se extrai que a IA está em maior desenvolvimento e uso nos tribunais estaduais, eleitorais e federais, além dos tribunais superiores.

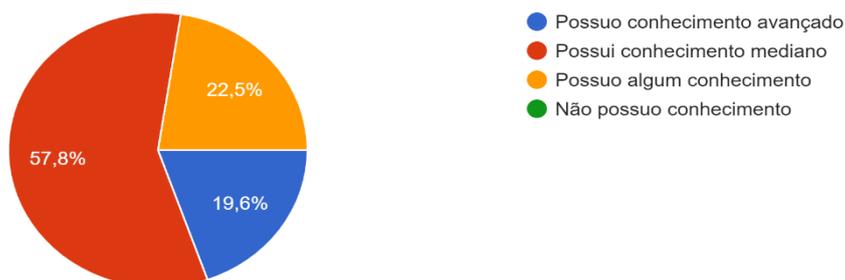
5.2.7 Ética e eu: Nível de conhecimento

No terceiro capítulo da pesquisa, os entrevistados foram indagados quanto ao seu preparo e conhecimento sobre ética e, mais que isso, o quanto o entrevistado se sente preparado para aplicar conceitos éticos no desempenho concreto de suas atribuições.

Gráfico 7 – Qual seu nível de conhecimento sobre ética?

QUAL O SEU NÍVEL DE CONHECIMENTO SOBRE ÉTICA?

102 respostas



Fonte: Produzido pelo autor.

Na primeira pergunta, 19,6% dos entrevistados (20 pessoas) afirmaram possuir conhecimento avançado sobre ética, ao tempo em que a maioria (57,8%, o que representa 59 pessoas) disse deter conhecimento mediano no tema e 22,5% concluíram possuir algum conhecimento (23 pessoas). Registre-se, por oportuno, que nenhum entrevistado afirmou não possuir conhecimentos sobre ética.

Visto globalmente, 65% dos entrevistados que responderam possuir conhecimento avançado sobre ética têm formação em direito. Dentre os que responderam conhecimento mediano, apenas 12% são da área do direito, ao passo que imensa maioria tem formação em tecnologia da informação (78%), o que se repetiu quando realizado o corte em relação aos que responderam “possuir algum conhecimento”, que revelou que 8% possuem formação em direito e 87% dos respondentes são formados nas ciências exatas.

Analisado o público de magistrados, todos afirmaram ter conhecimento, destacando-se que 85% desses entrevistados relataram conhecimento avançado (8) ou mediano (4). Não há nenhuma correlação quando isolados os magistrados nos ramos de Justiça.

Esses dados que relacionam a formação dos entrevistados ao conhecimento sobre ética advêm, certamente, da própria estrutura curricular dos cursos de ciência exatas e ciências sociais e humanas. É de se dizer que o curso de direito possui na grade curricular filosofia e sociologia gerais e do direito, além de que, há mais de 10 anos, nos concursos da magistratura, são conteúdos obrigatórios as disciplinas de cunho humanístico, dentre elas filosofia jurídica, sociologia jurídica e ética.

Também merece registro que os entrevistados com conhecimento mediano ou algum conhecimento estão majoritariamente concentrados entre os profissionais de tecnologia da informação, de onde se confirma a hipótese levantada na entrevista semiestruturada, quando se pontuou a situação de incerteza e insegurança quanto à atuação ética no uso e desenvolvimento de IA entre os programadores e profissionais da TI.

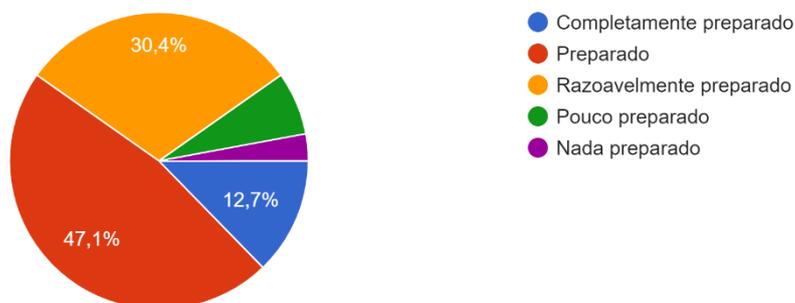
Nesse, também se faz relevante mencionar que esses dados são confirmados pela literatura, especialmente pela constatação de Ximenes e Ramalho, que referenciam estudos que concluem pela falta de preparo dos desenvolvedores e profissionais de *machine learning* para discutir e lidar com questões éticas. Afirmam que “apesar das crescentes questões éticas em aplicativos de computador, a ética é pouco abordada nos currículos universitários de Ciência da Computação (CS) em todo o mundo”.

5.2.8 Ética e eu: Aptidão e preparo para aplicar conceitos éticos no desempenho concreto do seu trabalho

Gráfico 8 – Você se sente apto/preparado a aplicar conceitos éticos no desempenho concreto do seu trabalho?

VOCÊ SE SENTE APTO/PREPARADO A APLICAR CONCEITOS ÉTICOS NO DESEMPENHO CONCRETO DO SEU TRABALHO?

102 respostas



Fonte: Produzido pelo autor.

A partir da medição do conhecimento de cada entrevistado com o tema ética, a pesquisa seguiu examinando o quanto os entrevistados se sentiam preparados a aplicar conceitos éticos no desempenho concreto do seu trabalho.

Isto porque, a despeito do nível de conhecimento quanto aos conceitos e contornos éticos, o fato é que todos os profissionais, em qualquer área do conhecimento, têm necessidade de atuar eticamente.

Os resultados revelaram que 12,7% dos entrevistados se sentem completamente preparados, 47,1% preparados, 30,4% razoavelmente preparados, 6,9% pouco preparados e 3% nada preparados.

Dentre os completamente preparados, verificamos majoritariamente a formação em direito (61,5%). Já entre os nada ou pouco preparados, identificamos maioria com formação em tecnologia da informação (80%).

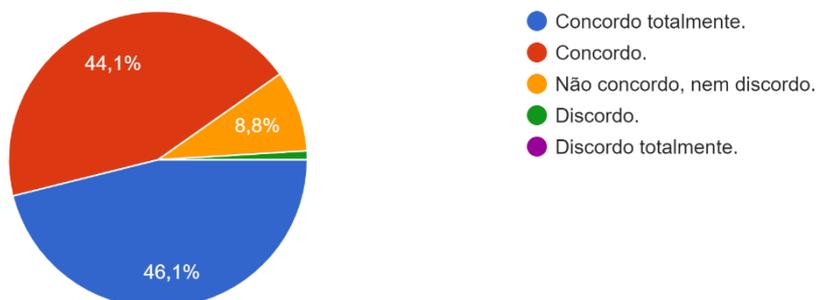
Analisando a segmentação de perfil, especificamente quanto aos magistrados, somente um magistrado federal afirmou se sentir pouco preparado para aplicar conceitos éticos no desempenho concreto do seu trabalho, ao passo em que a imensa maioria (quase 93%) reportou sentir-se razoavelmente preparado (2), preparado (5) ou completamente preparado (8).

5.2.9 IA e o Judiciário: Aplicação, utilidade e adequação para enfrentamento dos desafios atuais

Gráfico 9 –Na sua opinião, a IA é tecnologia aplicável, útil e adequada para resolução dos desafios enfrentados pelo Poder Judiciário?

NA SUA OPINIÃO, A INTELIGÊNCIA ARTIFICIAL É TECNOLOGIA APLICÁVEL, ÚTIL E ADEQUADA PARA RESOLUÇÃO DOS DESAFIOS ENFRENTADOS PELO PODER JUDICIÁRIO?

102 respostas



Fonte: Produzido pelo autor.

A penúltima etapa da pesquisa, buscou coletar as percepções dos entrevistados quanto ao uso da inteligência artificial pelo Poder Judiciário.

Para tanto, foram indagados quanto à sua concordância com a seguinte frase: “A inteligência artificial é tecnologia aplicável, útil e adequada para resolução dos desafios enfrentados pelo Poder Judiciário?”.

Os resultados indicaram que 46,1% dos entrevistados (47 pessoas) concordam totalmente e 44,1% (45 pessoas) concordam; 8,8% dos entrevistados nem concordaram nem discordaram (9 pessoas), 1% (1 pessoa) discorda e nenhum discorda totalmente.

Relevante destacar nesse ponto que esmagadora maioria dos entrevistados (mais de 90%) concorda ou concorda totalmente com a afirmação da adequação da IA para resolução dos problemas enfrentados pelos seus tribunais, sendo que apenas 9 pessoas responderam com neutralidade. Mais que isso, somente uma pessoa discordou e nenhuma discordou totalmente. Este dado de que só um entrevistado negou a hipótese revela que há forte indicativo, por parte dos entrevistados, de que a IA seja uma tecnologia capaz de auxiliar o trabalho desempenhado pelo Judiciário.

Todos os magistrados entrevistados afirmaram concordar (7) ou concordar totalmente (7) que a IA é tecnologia aplicável, útil e adequada para resolução dos desafios atuais do Judiciário.

Apenas um servidor, de tribunal superior, com formação em administração é que registrou discordância quanto à aderência da IA para resolução dos desafios do Poder Judiciário.

Essas respostas corroboraram hipótese levantada na entrevista narrativa, em que se verificou que todos os entrevistados naquela oportunidade elegeram a IA como tecnologia aderente aos interesses do Judiciário e melhoria da prestação dos serviços à população.

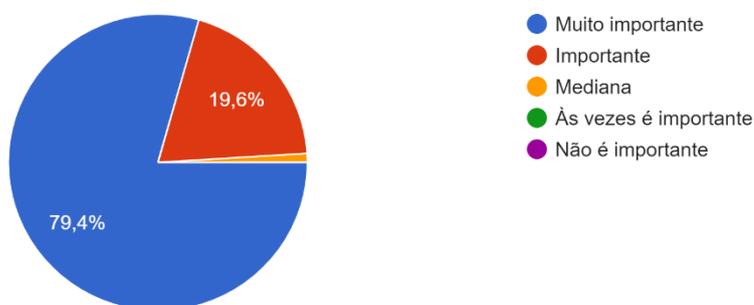
Além disso, esta conclusão também justifica a razão pela qual há crescente atenção do Judiciário com o tema, além de que explica o porquê de tanto esforço na governança e regulamentação da IA pelo CNJ.

5.2.10 IA e o Judiciário: Importância da abordagem ética para uso e desenvolvimento de ferramentas de IA pelo Judiciário

Gráfico 10 –A abordagem ética é importante para uso e desenvolvimento de ferramentas de IA pelo Poder Judiciário?

A ABORDAGEM ÉTICA É IMPORTANTE PARA USO E DESENVOLVIMENTO DE FERRAMENTAS DE IA PELO JUDICIÁRIO?

102 respostas



Fonte: Produzido pelo autor.

Em sequência, os entrevistados foram instados a avaliar, em grau de importância, o quanto é relevante uma abordagem ética para uso e desenvolvimento de ferramentas de IA pelo Judiciário.

Nenhum entrevistado avaliou com não importante, bem como nenhum disse que às vezes é importante. Apenas 1% dos entrevistados (uma pessoa) reportou importância mediana, ao tempo em que 19,6% (20 pessoas) relataram ser importante e 79,4% dos entrevistados (81 pessoas) asseveraram muita importância no enfoque ético para desenvolvimento e uso de IA pelos tribunais.

Entre os magistrados, todos reafirmaram a máxima relevância da abordagem ética, vez que um dos entrevistados classificou como importante (7%) e todos os demais como muito importante (93%).

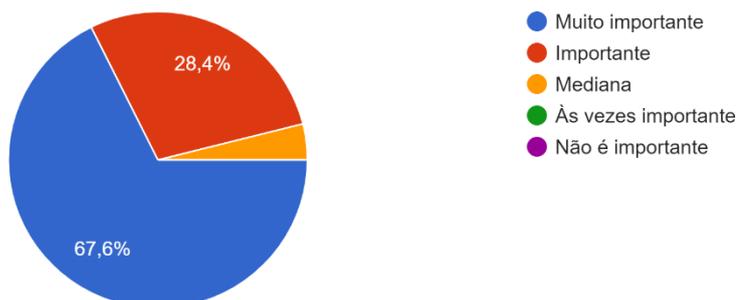
Neste ponto, mais uma vez, ficou claro que a hipótese levantada unanimemente durante a entrevista narrativa, da máxima relevância de abordagem ética no desempenho de IA, foi confirmada pela totalidade dos entrevistados, isso se considerado que 99% dos entrevistados (101 pessoas) registraram importância ou muita importância e um pontuou mediana importância.

5.2.11 IA e o Judiciário: Importância de uma lista de verificação (checklist) para conferência de conformidade ética e legal nos modelos de IA pelo Judiciário

Gráfico 11 –Avalie a importância de uma lista de verificação (*checklist*) para conferência de conformidade ética e legal dos modelos de IA utilizados pelo Judiciário

Avalie a importância de uma lista de verificação (checklist) para conferência de conformidade ética e legal dos modelos de IA utilizados pelo Judiciário.

102 respostas



Fonte: Produzido pelo autor.

Concluindo essa etapa e, partindo da premissa levantada nas entrevistas narrativas, foi solicitado aos entrevistados que avaliassem a importância de uma lista de verificação (checklist) para conferência de conformidade ética e legal dos modelos de IA utilizados pelo Judiciário.

Nenhum entrevistado classificou como não importante, assim como ninguém disse que seria importante às vezes. Todos os entrevistados registraram importância, da seguinte maneira: 4% dos entrevistados (4 pessoas) indicaram ter mediana importância, 28,4% afirmaram ser importante (29 pessoas) e 67,6% asseveraram ser muito importante (69 pessoas).

A unanimidade dos entrevistados pontuou a importância de um checklist para checagem ética no uso e desenvolvimento de IA pelo Judiciário.

Realizado o recorte de perfil de entrevistados, em relação aos magistrados, viu-se que dois (15%) deles classificaram como importante prévia lista de verificação ético e legal no desenvolvimento de IA, enquanto que todos os demais registraram como “muito importante” esse checklist (85%), de forma que para essa classe há extrema relevância na definição de um framework ético para desenvolvimento de IA na Justiça.

Isto confirma o que foi registrado nas entrevistas narrativas, quando se revelou que a ausência de parâmetros de análise gera insegurança, incerteza e até medo no uso da tecnologia entre os servidores e desenvolvedores, assim como pelos magistrados.

A ausência de um parâmetro, em última análise também retarda, dificulta ou até impede o desenvolvimento de ferramentas de IA pelo Judiciário.

As três últimas constatações, em última análise, corroboram a necessidade e utilidade desta pesquisa, no sentido de demonstrar lançar luzes sobre a o desenvolvimento ético de IA no

Judiciário. Daí que, se uma abordagem ética é importante, o “como” abordar e o contorno desse enfoque é medida urgente.

5.2.12 Critério e Parâmetros Éticos

A última etapa da entrevista avaliou a importância, sob a ótica dos membros do Judiciário, dos diversos critérios mapeados no capítulo anterior. Convém relembrar que naquela oportunidade foi traçado um panorama atual do cenário internacional, identificando, especificamente, quantos e quais padrões éticos estão sendo considerados pelos governos, pelo mercado/indústria e no âmbito acadêmico.

No ponto atual da pesquisa, a intenção foi apresentar aos entrevistados os critérios éticos globalmente considerados quando se fala em abordagem ética da IA, em um total de 48, indagando-os especificamente qual o grau de importância de cada um destes parâmetros para o desenvolvimento ético da IA no Judiciário, utilizando-se, metodologicamente, da escala de importância de Likert, que escalona a relevância em “muito importante”, “importante”, “mediana”, “às vezes é importante” e “não é importante”.

Para evitar introdução de viés do próprio pesquisador, optamos por não realizar, neste momento da pesquisa, nenhum tipo de edição ou interpretação, de forma que não foi realizado agrupamento, fusão, diluição, reagrupamento. A intenção da pesquisa foi a de permitir o choque cru entre o que a literatura mapeia como critério ético e avaliação dos profissionais da Justiça quanto à pertinência e sua área de atuação. Trabalhos futuros poderão investir na definição de uma lista mais concisa e coesa.

Considerando que alguns desses conceitos poderiam estar fora da zona de conhecimento do entrevistado, notadamente pelo perfil múltiplo dos envolvidos e dada a tecnicidade do tema, para evitar que fosse dada resposta enviesada, que comprometeria a validade dos resultados, optamos por inserir, dentre as opções de resposta, a alternativa “não sei opinar”.

Interessante registrar, de logo, que este confronto entre critérios éticos possíveis e existentes e critérios adequados e úteis na ótica do Judiciário, se torna extremamente relevante, inclusive sob a perspectiva acadêmica, na medida em que, em um cenário de ausência de parâmetros claros definidos para desenvolvimento e uso éticos de IA no Judiciário – o que restou evidenciado durante as entrevistas narrativas – a opinião e a percepção de magistrados e servidores envolvidos com a temática das ciências computacionais, pode apontar para um caminho, uma trilha a ser perseguida na adoção de contorno ético concreto.

Da análise das respostas, uma primeira constatação é a de que nenhum dos 48 critérios éticos apresentados foi majoritariamente avaliado com “não é importante” pelos entrevistados, de onde se extrai que, na avaliação dos profissionais da justiça brasileira, todos os padrões levantados importam.

Para melhor interpretar os dados obtidos, tabulamos as respostas conforme o seguinte critério: primeiramente, atribuímos nota 4 às respostas “muito importante”, 3 às respostas “importante”, 2 às respostas “mediana”, 1 aos retornos “às vezes é importante” e zero para “não é importante”; após, apuramos a soma de todas as respostas, por padrões, dividindo-se pelas quantidade respostas válidas, ou seja, excluindo-se as respostas “não sei opinar”, de modo que as médias foram calculadas só entre os respondentes que possuíam conhecimento sobre o

parâmetro indagado; obtida a média de pontuação conferida a cada estandarte ético, elaboramos um quadro comparativo, com as médias de todos os 48 indicadores éticos.

Ressalte-se, de logo, que houve baixa incidência de respostas “não sei opinar”, o que se verificou em todos os critérios, de maneira que, de modo geral, não houve dificuldade em responder aos estâdareos éticos elencados.

Registre-se que os três critérios em que os entrevistados tiveram mais dificuldade foram “Não subversão de processos cívicos e sociais”, que registrou 7 respostas “não sei opinar” (6,8% dos entrevistados), “benefício e prosperidade compartilhados”, com 6 respondentes que não opinaram (5,8% dos entrevistados) e “solidariedade”, com 5 (4,9% dos entrevistados). Como se vê todos os critérios contabilizaram expressiva quantidade de respostas válidas (mais de 93%).

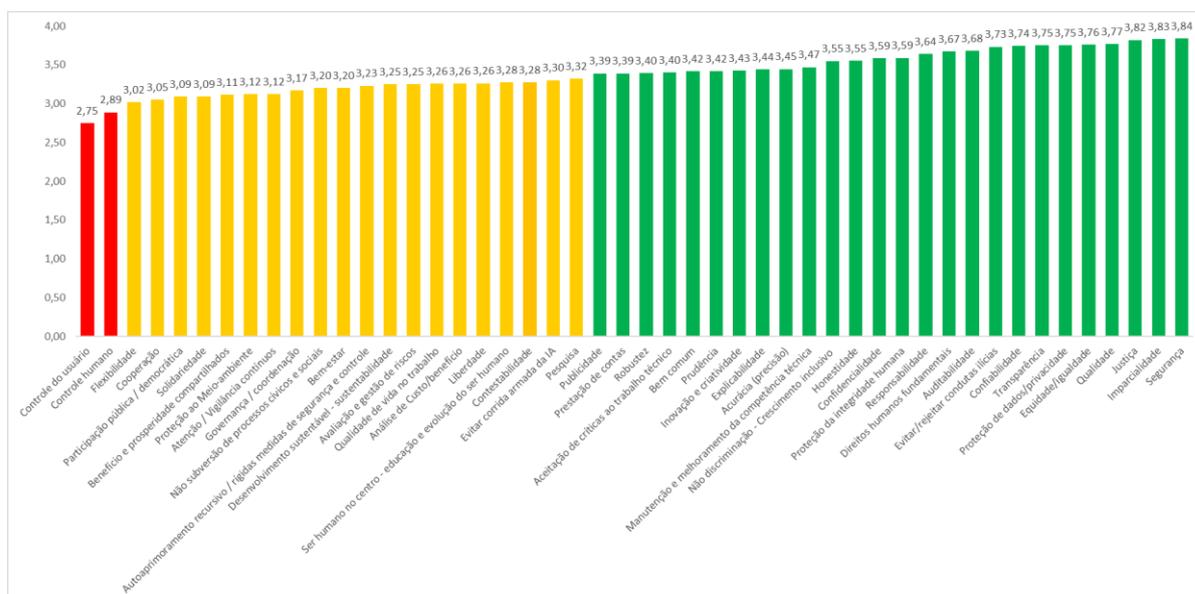
Dentre todos os critérios, somente um deles teve 100% de respostas, que foi o critério da “segurança”, que contabilizou com 102 respostas, ou seja, nenhum entrevistado registrou “não sei opinar”.

As cinco menores médias registradas foram as dos critérios: “Controle do usuário”, “Controle humano”, “flexibilidade”, “cooperação” e “solidariedade” e “Participação pública/democrática”, com, respectivamente, 2,75, 2,89, 3,02, 3,04 e 3,09 pontos (os dois últimos empatados).

As cinco maiores médias foram conferidas aos seguintes critérios éticos: “segurança”, com 3,84 pontos; “imparcialidade”, com 3,83 pontos; “Justiça”, com 3,82 pontos; “qualidade”, com 3,77 pontos e; “equidade/igualdade”, com 3,76 pontos.

Encontramos que a média de pontuação geral foi a de 3,38. Daí que optamos por identificar, no quadro abaixo, de verde os critérios éticos que conseguiram pontuação superior à média geral, de amarelo os que obtiveram pontuação média menor do que a média geral e de vermelho os critérios que obtiveram pontuação média inferior a 3.

Gráfico 12 – Critérios e parâmetros éticos.



Fonte: Produzido pelo autor.

Para que um critério ético fosse considerado validado, ou seja, pudesse ser considerado importante sob a ótica dos membros do Poder Judiciário em uma checagem ética, considerou-se que a pontuação média de Likert deveria ser superior a 3, pois de acordo com a escala proposta, um estândar ético deveria receber no mínimo 3 de 4 para fosse considerado pertinente pelo público-alvo.

Somente dois critérios éticos podem ser vistos como refutados, a saber, “Controle do usuário”, que obteve média 2,75 e “Controle humano”, com média de 2,89.

Como se vê, ao que parece, os entrevistados refutaram, como critério ético necessário para desenvolvimento de IA pelo Judiciário a necessidade de controle humano nas ferramentas de IA, assim como o controle do próprio usuário sobre os modelos de IA.

Além desses dois critérios, que obtiveram performance abaixo de 3, outros quatro critérios pontuaram bem próximo a 3, de forma que merecem registro. São os critérios éticos da “flexibilidade”, que contou com nota 3,02, “cooperação”, com nota 3,04 e “solidariedade” e “Participação pública/democrática”, que empataram em 3,09 pontos.

Esses estândares éticos são os que implicam diretamente na necessidade de conectar, comunicar e explicar as descobertas ao cidadão, descrevendo as informações usadas, resultados possíveis e fornecendo um meio para os usuários fornecerem respostas e percepções sobre as decisões do modelo.

Interpretados esses dados de forma conjunta, é de se notar que a má performance esteve ligada à interferência do usuário-cidadão no modelo de IA ou no resultado das pesquisas de IA dentro do Poder Judiciário, bem como à maleabilidade, colaboração e participação públicas, o que talvez seja explicado por, historicamente, o Poder Judiciário, único cujo acesso não se dá por eleição popular, estar mais distante da sociedade, mais blindado às interferências sociais no gerir administrativo diário.

Em sentido oposto, os estândares éticos melhor pontuados, com notas superiores a 3,7 são: “Evitar/rejeitar condutas ilícitas”, “Confiabilidade”, “Transparência”, “Proteção de dados/privacidade”, “Equidade/igualdade”, “Qualidade”, “Justiça”, “Imparcialidade” e “Segurança”.

Como se pode notar, orbitam todos em torno de um mesmo sentimento de justiça, equidade, imparcialidade, segurança, transparência e privacidade. São valores, direitos fundamentais e princípios basilares de um estado de Direito, previstos na própria Constituição Federal brasileira. É o mínimo que os entrevistados esperam: que as ferramentas em uso e desenvolvimento de IA pelo Judiciário estejam em conformidade com ordem legal vigente.

Encaminhando-se para o fim, realizada a comparação entre os sentidos e significados que foram percebidos durante as entrevistas narrativas, contrapondo-as ao que for apurado no survey, não obtivemos achados relevante. Isto porque, conforme já registrado, nas entrevistas narrativas não houve um linear mínimo na indicação de critérios éticos entre os entrevistados. O máximo que podemos inferir é que o primeiro e o segundo entrevistados indicaram critérios de igualdade e imparcialidade, o que foi corroborado dentre os critérios melhor pontuados no survey.

Dito isso, por derradeiro, cumpre-nos avaliar, também, a incidência dos parâmetros éticos constantes do levantamento bibliográfico (quadro-resumo item 4.4), que cotou com análise de 25 documentos, em contraposição aos dados obtidos pelo *survey*.

No levantamento de parâmetros éticos observou-se que os estandartes que obtiveram maior incidência foram Inclusão/não-discriminação(19), Segurança (18), Transparência (18), Responsabilidade (18), Justiça (17), Bem-estar (16).

É de se ver parcial coincidência entre os critérios que obtiveram maior pontuação média no survey com os parâmetros mais ocorrentes nos documentos pesquisados.

Interessante notar que embora o “controle do usuário” conte com 11 inserções nos documentos analisados e o “controle humano” esteja em 8 deles, estes dois critérios foram justamente os que foram refutados no survey, com nota média inferior a 3. Vemos aí uma diferente avaliação de critérios pelos órgãos elaboradores dos documentos em contraposição com o público da Justiça.

Corroboram os dados do survey, entretanto, a pouca ocorrência dos critérios “solidariedade”, que contou com uma inserção (Universidade de Montreal), flexibilidade (com duas ocorrências, no PL 2.338/2023 e na ordem executiva 13859 dos EUA) e cooperação, que teve apenas 3 inserções nos documentos (China, Asilomar e Google).

Merece registro que parâmetros que obtiveram baixa ocorrência nos documentos pesquisados, como “Evitar corrida armada da IA” (com 1 inserção, em Asilomar), “qualidade de vida no trabalho” (1 inserção em ACM) e “evitar condutas ilícitas / rejeitar corrupção” (com uma indicação, no IEEE), tiveram validação expressiva no *survey*, com médias de 3.30, 3,26 e 3,73, respectivamente.

6 CONCLUSÕES

A primeira conclusão que se pode extrair deste trabalho – que pode parecer tautológica, mas não é – consiste na identificação de que a inteligência artificial é vista pelo Poder Judiciário como uma tecnologia capaz de impactar profundamente a prestação dos serviços jurisdicionais, seja no apoio à tomada de decisões, na melhoria dos processos de trabalho, predição de eventos futuros e tendências. As hipóteses levantadas nas entrevistas narrativas foram contundentemente confirmadas no *survey*, quando a esmagadora maioria dos entrevistados (mais de 90%) concorda ou concorda totalmente com a afirmação da adequação da IA para resolução dos problemas enfrentados pelos seus tribunais.

Conclui-se também que embora seja uma tecnologia com desenvolvimento e expansão nítidos, a utilização ainda se dá em “ilhas”, ou seja, poucos órgãos da Justiça se utilizam efetivamente de IA e a imensa maioria dos entrevistados (77,4%) disse fazer uso eventual, raro ou nunca ter feito uso da tecnologia. Tal constatação se justifica pela autonomia administrativa de cada órgão/tribunal - que pode optar ou não pela adoção da IA, a despeito da coordenação central exercida pelo CNJ - assim como pela falta de capacidade de arremeter pessoal especializado na matéria.

Quanto à abordagem ética da IA no Judiciário, ficou claro que a hipótese levantada unanimemente durante a entrevista narrativa relativamente à máxima relevância de abordagem ética no desempenho de IA, foi confirmada pela totalidade dos entrevistados do *survey*.

Outra conclusão possível é a de que, inobstante exista regulamentação específica para pesquisa, uso e desenvolvimento ético da IA no Judiciário, as questões éticas não estão na pauta do desenvolvimento dos produtos e serviços da Justiça do hoje, hipótese levantada nas entrevistas narrativas, confirmada no *survey*, quando se verificou público quase unânime pleiteando definição de um framework para desempenho ético da IA na Justiça.

Não há um procedimento bem definido ou especificado de checagem ética, com requisitos objetivos e lógica científica, fator que causa insegurança na pesquisa, no uso e no desenvolvimento de IA pelo Judiciário.

Também se conclui que a falta de definição concreta de um delinear ético mínimo para checagem e validação é um dos principais desafios para implementação de como uma IA ética no Judiciário. Aliado a esse vazio, associa-se o fato de que a maioria de quem está desenvolvendo IA no Judiciário tem pouco ou não tem conhecimento sobre IA, além do baixo conhecimento das equipes técnicas, de programadores sobre ética, o que dificulta, ainda mais, o uso e desenvolvimento da tecnologia pela Justiça brasileira.

É, portanto, premente a necessidade de definição de um checklist ético para pesquisa, uso e desenvolvimento ético da IA no Judiciário, hipótese validada por 96% dos entrevistados no *survey*, sendo registrado nas entrevistas narrativas que a ausência de parâmetros de análise gera insegurança, incerteza e até medo no uso da tecnologia entre os servidores, desenvolvedores e magistrados. A ausência de um parâmetro, em última análise, também retarda, dificulta ou até impede o desenvolvimento de ferramentas de IA pelo Judiciário.

Estas constatações corroboram a utilidade desta pesquisa, no sentido de apresentar o fenômeno do uso da IA pelo Judiciário e lançar luzes sobre a pesquisa, uso e desenvolvimento ético de IA na Justiça.

Especificamente quanto ao conteúdo ético essencial para desenvolvimento de IA pelo Judiciário, houve dissenso quanto ao núcleo ético nas entrevistas narrativas, que indicaram pulverizada e aleatoriamente alguns parâmetros. Por outro lado, no *survey*, depois de levantados 48 estândares, percebeu-se que somente 2 deles foram afastados e 26 obtiveram média superior à média geral, de modo, embora haja forte indicativo de caminho a se percorrer, ainda não há muito claramente um contorno bem delineado e preciso do que seria eticamente exigível para uma IA ética na Justiça.

Além disso, seria relevante dissecar, explicar como cada um desses critérios podem ser considerados atendidos, concretamente, com um mínimo de objetividade necessária para se alcançar segurança no ambiente de desenvolvimento.

Dáí que, como possível evolução desta pesquisa e, em trabalhos futuros, poder-se-ia investigar com mais profundidade o conteúdo de cada um deles, inclusive as razões de afastamento de dois parâmetros e definir “como” realizar checagem concreta. A definição desse contorno de enfoque ético é medida urgente.

6.1 FRAQUEZAS À VALIDADE, CONFIABILIDADE E FIDEDIGNIDADE

Podemos listar as seguintes fraquezas à validade, confiabilidade e fidedignidade listados abaixo, com a respectiva justificativa ou medida de mitigação adotada na pesquisa.

O primeiro ponto foi o eventual desconhecimento dos entrevistados quanto aos conceitos apresentados como parâmetros éticos. Embora não tenhamos indicado aos entrevistados previamente o contorno de cada um dos parâmetros – mesmo porque isto importaria necessário viés ou encaminhamento não queridos – o fato é que o público-alvo da pesquisa é extremamente qualificado tecnicamente, composto de servidores de área de TI e magistrados que atuam nesta área, de modo que possuem algum conhecimento sobre critérios comuns. Além disso, tentou-se minimizar o risco oferecendo a opção “não sei opinar”.

A quantidade de parâmetros também pode ser visto como dificultador, no entanto, para objetivo de pesquisa aqui definido, de perceber quais valores e critérios mapeados na literatura possuem conexão com o serviço judiciário, qualquer clusterização, agrupamento ou simplificação importaria, necessariamente, inclusão de viés interpretativo do pesquisador, mormente se considerado que cada documento analisado estabelece a adota conceitos díspares, com agrupamentos próprios.

Extensão e quantidade de estândares, o que poderia gerar confusão: tentou-se reduzir, ao máximo, as demais questões, para que o *survey* ficasse mais fluido, além de que, foi inserida a opção “não sei opinar”.

A nomenclatura não uniforme merece registro, porque torna complexa a compilação dos dados e a busca, podendo gerar dubiedade na interpretação quanto ao significado e conteúdo dos diferentes documentos: buscou-se utilizar um mesmo conceito para cada parâmetro, coletando a essência do parâmetro indicado, em detrimento do nome ou título empregado.

A fidedignidade depende da motivação, honestidade, memória e habilidade para responder as perguntas por parte dos participantes; tentou-se minimizar através de uma

abordagem individualizada e próxima dos colegas magistrados e servidores, conscientizando-os da importância da pesquisa.

REFERÊNCIAS

- ACM. **ACM Code of Ethics and Professional Conduct**. Disponível em <<https://www.acm.org/code-of-ethics>>. Acesso em 07 de out. de 2022.
- AI4PEOPLE. **AI4People's 7 AI Global Frameworks**, p. 174-175. Disponível em <<https://2020.ai4people.eu/wp-content/pdf/AI4People7AIGlobalFrameworks.pdf>>. P. 174-175. Acesso em 20 de jan. de 2020.
- AI4PEOPLE. **AI4People's Ethical Framework for a good AI Society: Opportunities, Risks, Principles, and Recommendations**. Disponível em <https://www.eismd.eu/wp-content/uploads/2019/11/AI4People%E2%80%99s-Ethical-Framework-for-a-Good-AI-Society_compressed.pdf>. Acesso em 20 de jan. de 2023.
- AIVERIFY. **Generative AI: Implications for Trust and Governance**. Disponível em <https://aiverifyfoundation.sg/downloads/Discussion_Paper.pdf>. Acesso em 15 de fev. de 2023.
- AUSTRALIA. **Australia's AI Ethics Principles**. Disponível em <<https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles>>. Acesso em 10 de out. de 2022.
- BAUER, Martin W.; GASKELL, George. **Pesquisa qualitativa com texto, imagem e som: um manual prático**. Editora Vozes Limitada, 2017.
- BÉLGICA. Universal Guidelines for Artificial Intelligence. 2018. **The Public Voice Coalition**. Disponível em <<https://thepublicvoice.org/ai-universal-guidelines/>>. Acesso em 10 de out. de 2022.
- BENCH-CAPON, Trevor; PRAKKEN, Henry; SARTOR, Giovanni. **Argumentation in legal reasoning**. Springer US, 2009.
- BLIGHT, K. J. (2018). **Artificial Intelligence, AI biases and risks, and the need for AI-regulation and AI ethics: some examples**. Available at: <http://www.researchgate.net/publication/326377798_Artificial_Intelligence_AI_biases_and_risks_and_the_need_for_AI-regulation_and_AI_ethics_some_examples_17_Nov_2018>. Acesso em 11 de out. de 2019.
- BODDINGTON, P. **Towards a Code of Ethics for Artificial Intelligence**. Springer, 2017.
- BODEN, Margaret A. (Ed.). **Artificial intelligence**. Elsevier, 1996.
- BOONE JR, Harry N.; BOONE, Deborah A. **Analyzing likert data**. The Journal of extension, v. 50, n. 2, p. 48, 2012.

BOSTROM, Nick; YUDKOWSKY, Eliezer. The ethics of artificial intelligence. In: **Artificial intelligence safety and security**. Chapman and Hall/CRC, 2018. p. 57-69.

BOSTROM, N.; YUDKOWSKY, E. **The Cambridge Handbook of Artificial Intelligence** Cambridge Univ. Press (eds Frankish, K. & Ramsey, W. M.), 2014. P 316–334. Disponível em <<https://doi.org/10.1017/CBO9781139046855.020>>. Acesso em 11 de out. de 2019.

BRASIL. [Constituição (1988)]. Constituição da República Federativa do Brasil. Brasília, DF, 1988.

BRASIL. Congresso Nacional. Projeto de Lei Nº 21 de 2020. Disponível em <<https://www.camara.leg.br/propostas-legislativas/2236340>>. Acesso em 21 de nov. de 2022.

BRASIL. Congresso Nacional. Projeto de Lei Nº 2338 de 2023. Disponível em <<https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>>. Acesso em 10 de mai. de 2023.

BRASIL. Congresso Nacional. Projeto de Lei Nº 5051 de 2021. Disponível em <<https://www25.senado.leg.br/web/atividade/materias/-/materia/138790>>. Acesso em 10 de mai. de 2023.

BRASIL. Congresso Nacional. Projeto de Lei Nº 872 de 2021. Disponível em <<https://www25.senado.leg.br/web/atividade/materias/-/materia/147434>>. Acesso em 10 de mai. de 2023.

BRASIL. Conselho Nacional de Justiça. Inteligência artificial na Justiça. Coordenação: José Antônio Dias Toffoli; Bráulio Gabriel Gusmão. Brasília: CNJ, 2019.

BRASIL. Conselho Nacional de Justiça. Panorama de Tecnologia da Informação e Comunicação do Poder Judiciário – 2016 a 2021. Disponível em <<https://bit.ly/3Oxhh3E>>. Acesso em 20 de abr. de 2023.

BRASIL. Conselho Nacional de Justiça. Portaria Nº 25/2019, de 19/02/2019. Diário de Justiça Eletrônico/CNJ, nº 274, p. 4-8, de 25/08/2020.

BRASIL. Conselho Nacional de Justiça. Portaria Nº 271/2020, de 04/12/2020. Diário de Justiça Eletrônico/CNJ, nº 393, p. 2-4, de 14/12/2020.

BRASIL. Conselho Nacional de Justiça. Resolução Nº 332/2020, de 21/08/2020. Diário de Justiça Eletrônico/CNJ, nº 35, p. 4-7, de 22/02/2019.

BRASIL. Conselho Nacional de Justiça. Justiça em Números 2022. Disponível em <<https://www.cnj.jus.br/wp-content/uploads/2022/09/justica-em-numeros-2022-1.pdf>>. Acesso em 20 de abr. 2023.

BRASIL. Conselho Nacional de Justiça. Panorama de Tecnologia da Informação e Comunicação do Poder Judiciário – 2016 a 2021. Disponível em <<https://bit.ly/42UqOpC>>. Acesso em 20 de abr. 2023.

BRASIL. Conselho Nacional de Justiça. Portaria MCTI nº 4.617, de 6 de abril de 2021.

BRASIL. Conselho Nacional de Justiça. Portaria Nº 25/2019, de 19/02/2019. Diário de Justiça Eletrônico/CNJ, nº 274, p. 4-8, de 25/08/2020.

BRASIL. Conselho Nacional de Justiça. Portaria Nº 271/2020, de 04/12/2020. Diário de Justiça Eletrônico/CNJ, nº 393, p. 2-4, de 14/12/2020.

BRASIL. Conselho Nacional de Justiça. Resolução Nº 332/2020, de 21/08/2020. Diário de Justiça Eletrônico/CNJ, nº 35, p. 4-7, de 22/02/2019.

BRASIL. Conselho Nacional de Justiça. Panorama de Tecnologia da Informação e Comunicação do Poder Judiciário – 2016 a 2021. Disponível em <<https://bit.ly/3Oxhh3E>>. Acesso em 20 de abr. de 2023.

BRASIL. Congresso Nacional. Projeto de Lei Nº 21 de 2020. Disponível em <<https://www.camara.leg.br/propostas-legislativas/2236340>>. Acesso em 21 de nov. de 2022.

BRASIL. Congresso Nacional. Projeto de Lei Nº 2338 de 2023. Disponível em <<https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>>. Acesso em 10 de mai. de 2023.

BRASIL. Justiça 4.0: Inteligência Artificial está presente na maioria dos tribunais brasileiros. Conselho Nacional de Justiça, 2022. Disponível em <<https://www.cnj.jus.br/justica-4-0-inteligencia-artificial-esta-presente-na-maioria-dos-tribunais-brasileiros/>>. Acesso em 10 de out. de 2022.

BRASIL. Ministério da Ciência, Tecnologia, Inovações e Comunicações. Portaria MCTI nº 4.617, de 6 de abril de 2021.

BRASIL. Ministério das Relações Exteriores. Declaração de Osaka dos Líderes do G20. Nota à Imprensa nº 170/2019. 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 10 de out. de 2022

CANADÁ. Declaração de Montreal pelo Desenvolvimento Responsável da Inteligência Artificial. 2018. Disponível em <https://www.sbmac.org.br/wp-content/uploads/2021/02/Portugue%CC%82s-UdeM_Decl-IA-Resp_LA-Declaration_vf.pdf>. Acesso em 20 de jan. de 2023.

CHINA. A Next Generation Artificial Intelligence Development Plan. **State Council Notice on the Issuance of the Next Generation Artificial Intelligence Development Plan**. Completed: July 8, 2017. Released: July 20, 2017. Disponível em <<https://flia.org/wp->

content/uploads/2017/07/A-New-Generation-of-Artificial-Intelligence-Development-Plan-1.pdf> Acesso em 10 de out. de 2022

CORTELLA, M. S. **Qual é a tua obra?** Inquietações, propositivas sobre gestão, liderança e ética. Petrópolis: Vozes, 2007

CORTELLA, Mario Sergio; DE BARROS FILHO, Clóvis. **Ética e vergonha na cara!**. Papirus Editora, 2014.

CUEVA, Ricardo Villas Bôas. Inteligência artificial no Judiciário. NUNES, D.; LUCON, PHS; WOLKART, EN. **Inteligência Artificial e Direito Processual: Os Impactos da Virada Tecnológica no Direito Processual**, v. 2, 2021.

DE GÓMEZ, Maria Nelida Gonzalez. Luciano Floridi e os problemas filosóficos da informação: da representação à modelização. **InCID: Revista de Ciência da Informação e Documentação**, v. 4, n. 1, p. 3-25, 2013.

EASTERBROOK, Steve *et al.* Selecting empirical methods for software engineering research. **Guide to advanced empirical software engineering**, p. 285-311, 2008.

ESTADOS UNIDOS DA AMÉRICA. Executive Order 13.859 of February 11, 2019. Maintaining American Leadership in Artificial Intelligence. Disponível em <<https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>>. Acesso em 10 de out. de 2022

ESTADOS UNIDOS DA AMÉRICA. Memorandum for the Heads of Executive Departments and Agencies. Disponível em <<https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf>>. Acesso em 10 de out. de 2022

ESTADOS UNIDOS DA AMÉRICA. The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update. Disponível em <<https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf>>. Acesso em 10 de out. de 2022

ETZIONI, Amitai; ETZIONI, Oren. AI assisted ethics. **Ethics and Information Technology**, v. 18, p. 149-156, 2016.

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From data mining to knowledge discovery in databases. **AI magazine**, v. 17, n. 3, p. 37-37, 1996.

FINK, Arlene. **The survey handbook**. Thousand Oaks, Sage, 1995a.

FLORIDI, Luciano. **Ética de la información: su naturaleza y alcance**, 2006.

FLORIDI, Luciano. Soft ethics and the governance of the digital. **Philosophy & Technology**, v. 31, p. 1-8, 2018.

FLORIDI, Luciano *et al.* AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. **Minds and machines**, v. 28, p. 689-707, 2018.

FLORIDI, Luciano. **The ethics of information**. Oxford: Oxford University Press, 2013.

FRANÇA. Comissão Nacional de Computação e Liberdades da França (CNIL - *Commission Nationale de l'Informatique et des Libertés*). **The ethical matters raised by algorithms and artificial intelligence**. Dezembro de 2017. Disponível em <https://www.cnil.fr/sites/cnil/files/atoms/files/cnil_rapport_ai_gb_web.pdf>. Acesso em 10 de mai. de 2023.

GOOGLE. **IA no Google: nossos princípios**. 08 de junho, 2018. Disponível em <<https://brasil.googleblog.com/2018/06/ia-no-google-nossos-principios.html>>. Acesso em 18 de out. de 2022.

HAN, J.; KAMBER, M. **Data Mining Concepts and Techniques**. Ed. Elsevier. 2006. Second Edition.

HARTMANN, H. **Modos de Existência, Aprofundamento e Revelação**. 2019. 156 f. Tese (Doutorado em Filosofia) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2019.

HARTMANN, F.; ZUMBLICK, R. **Inteligência Artificial e Direito**. Vol. 1. Curitiba: Editora Alteridade, 2019, pág. 27/28.

IEEE. **IEEE Code of Ethics**. Disponível em <<https://www.ieee.org/about/corporate/governance/p7-8.html>>. Acesso em 9 de fev de 2022.

ÍNDIA. NITI Aayog. **Data Empowerment and Protection Architecture (DEPA)**, 2021. Disponível em <<https://www.niti.gov.in/sites/default/files/2020-09/DEPA-Book.pdf>>. Acesso em 10 de abr. de 2023.

ÍNDIA. PRS Legislative Research. **The Personal Data Protection Bill**, 2019. Disponível em <https://prsindia.org/files/bills_acts/bills_parliament/2019/Legislative%20Brief%20-%20Personal%20Data%20Protection%20Bill%202019.pdf>. Acesso em 10 de abr. de 2023.

ÍNDIA. Securities and Exchange Board of India – SEBI. **Reporting for Artificial Intelligence and Machine Learning Applications and Systems offered and used by Market Infrastructure**. Disponível em <https://www.sebi.gov.in/legal/circulars/jan-2019/reporting-for-artificial-intelligence-ai-and-machine-learning-ml-applications-and-systems-offered-and-used-by-market-infrastructure-institutions-miis-_41927.html>. Acesso em 10 de abr. de 2023.

JOBIN, Anna; IENCA, Marcello; VAYENA, Effy. The global landscape of AI ethics guidelines. **Nature Machine Intelligence**, v. 1, n. 9, p. 389-399, 2019.

JOSHI, Ankur et al. **Likert scale: Explored and explained**. British journal of applied science & technology, v. 7, n. 4, p. 396, 2015.

JOVCHELOVICH, S.; BAUER, M. W. **Entrevista Narrativa**. In: Bauer M. W., Gaskell G. (Org.). Pesquisa qualitativa com texto, imagem e som: um manual prático. Petrópolis: Editora Vozes, 2002, p. 90-113

JOVCHELOVICH, S.; BAUER, M. W. **Entrevista Narrativa**. In: Bauer, Martin W.; Gaskell, George. Pesquisa qualitativa com texto, imagem e som: um manual prático. Petrópolis: Editora Vozes Limitada, 2017.

JUSTIÇA 4.0: Inteligência Artificial está presente na maioria dos tribunais brasileiros. **Conselho Nacional de Justiça**, 2022. Disponível em < <https://www.cnj.jus.br/justica-4-0-inteligencia-artificial-esta-presente-na-maioria-dos-tribunais-brasileiros/>>. Acesso em 14 de out. de 2022.

KADAM, Tanmay. China's AI-Enabled 'Smart Courts' To Recommend Laws & Draft Legal Docs; Judges To Take Consult AI Before Verdict. **The EurAsian Times**, 2022. Disponível em <<https://eurasianimes.com/chinas-ai-enabled-smart-court-to-recommend-laws-judges/>>. Acesso em 14 de out. de 2022.

KANT, I. **Fundamentação da Metafísica dos Costumes**. Tradução: Paulo Quintela. Lisboa: Edições 70, 1996.

KAPLAN, Andreas; HAENLEIN, Michael. Rulers of the world, unite! The challenges and opportunities of artificial intelligence. **Business Horizons**, v. 63, n. 1, p. 37-50, 2020.

KIRKPATRICK, Keith. It's not the algorithm, it's the data. **Communications of the ACM**, v. 60, n. 2, p. 21-23, 2017.

KONG, Eun Bae; DIETTERICH, Thomas G. Error-correcting output coding corrects bias and variance. In: **Machine learning proceedings** 1995. Morgan Kaufmann, 1995. p. 313-321.

LAVILLE, Christian; DIONE, Jean. **A construção do saber: manual de metodologia da pesquisa em ciências humanas**. Editora Artmed, 1999.

LUHMANN, Niklas. **La ciencia de la sociedad**. 1 ed. ed. México: Universidad Iberoamericana, 1996.

MCCULLAGH, Karen. In the Shadows of the Smart City. **SSRN Scholarly Paper**. ID 3423556. Rochester, NY: Social Science Research Network, 2019.

MEHRABI, Ninareh et al. A survey on bias and fairness in machine learning. **ACM computing surveys (CSUR)**, v. 54, n. 6, p. 1-35, 2021.

MICROSOFT. **Princípios do Microsoft Responsible AI na prática**. Disponível em <<https://www.microsoft.com/pt-br/ai/responsible-ai?activetab=pivot1%3aprimariy6>>. Acesso em 10 de out. de 2022.

MICHALSKI, Ryszard Stanislaw; CARBONELL, Jaime Guillermo; MITCHELL, Tom M. **Machine learning an artificial intelligence approach**. by Springer-Verlag Berlin Heidelberg New York Tokyo in 1983, 1984.

MILAN, Stefania. Social Media and the City: Rethinking Urban Sociality. **Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries**. Oslo, Norway, 2010.

MORO, Suelen Teodoro *et al.* Exploring the Challenges of Ethical AI Design: Insights from User Perceptions. In: **Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems**. ACM, 2020. p. 1-15.

NORVIG, P. Russel; INTELLIGENCE, S. Artificial. A modern approach. Prentice Hall Upper Saddle River, NJ, USA: Rani, M., Nayak, R., & Vyas, OP (2015). An ontology-based adaptive personalized e-learning system, assisted by software agents on cloud storage. **Knowledge-Based Systems**, v. 90, p. 33-48, 2002.

NÓS. **A Carta de Princípios para a Inteligência Artificial**. Disponível em <<https://www.nos.pt/particulares/sobre-nos/grupo/manifesto-inteligencia-artificial/Pages/carta-principios-IA.aspx>>. Acesso em 18 de out. de 2022

OECD. **AI Principles Overview**. Disponível em <<https://oecd.ai/en/ai-principles>>. Acesso em 06 de set. de 2022.

OECD. **OECD Principles on Artificial Intelligence**. Disponível em <<http://www.oecd.org/going-digital/ai/principles/>>. Acesso em 12 de out. de 2022.

PENTLAND, Alex Sandy. **Social Physics: How Good Ideas Spread - The Lessons from a New Science**. Penguin Books, 2014.

PINSONNEAULT, Alain; KRAEMER, Kenneth. Survey research methodology in management information systems: an assessment. **Journal of management information systems**, v. 10, n. 2, p. 75-105, 1993.

PREFEITURA DO RECIFE. Marco Legal de Inovação do Recife. Disponível em <<http://legis.cidaderecife.org/jspui/bitstream/123456789/1460/1/marco-legal-inovacao.pdf>>. Acesso em 16 de nov. de 2022.

REESE, Simon; LEVY, Yotam; MILTON, Adam. **Introducing the Trust in AI Framework**. Disponível em <<https://trustinaiframework.org/>>. Acesso em 25 de agosto de 2022.

REINO UNIDO. Câmara dos Lordes. Carta Ética Sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu Ambiente. Publicada em 16/04/2018.

RICHMOND, Sheldon. Superintelligence: Paths, Dangers, Strategies. By Nick Bostrom. Oxford University Press, Oxford, 2014. **Philosophy**, v. 91, n. 1, p. 125-130, 2016.

ROSENBERG, Scott. “Why AI is still waiting for its ethics transplant.”. **WIRED onlajn**, v. 11, jan. 2018.

RUNESON, Per; HÖST, Martin. Guidelines for conducting and reporting case study research in software engineering. **Empirical software engineering**, v. 14, p. 131-164, 2009.

RUSSELL, Stuart. **Q&A: The Future of Artificial Intelligence**. University of Berkeley, 2016. Disponível em: <<http://people.eecs.berkeley.edu/~russell/temp/q-and-a.html>>. Acesso em: 22 jul. 2019

RUSSELL, Stuart; NORVIG, Peter. **Inteligência Artificial**. Tradução de Regina Célia Simile de Macedo. Elsevier. 8ª tiragem. Ano 2013.

RUSSELL, Stuart; DEWEY, Daniel; TEGMARK, Max. **Reserch Priorities for Robust and Beneficial Artificial Intelligence**. Cornell University. AI Magazine, 36, n.4(2015). Disponível em: <<http://arxiv.org/pdf/162.03506.pdf>>. Acesso em: 19 jul. 2019

SANFELICE, Viviane. Inovação no Poder Judiciário. RIC-JUN. **Revista Interdisciplinar de Cultura e Justiça**, v. 4, n. 9, p. 51-60, 2019.

SANTOS, Maria Helena de Castro. **Da Previsibilidade do Direito**. 1997. 372f. Tese (Doutorado em Filosofia do Direito) - Faculdade de Direito, Universidade de São Paulo, São Paulo, 1997.

SCHUTZE, F. **Die Technik des Narrativen Interviews in Interaktionsfeldstudien - Dargestellt an Einem Projekt zur Erforschung von Kommunalen Machtstrukturen**. Unpublished manuscript, University of Bielefeld, Department of Sociology. 1977.

SCHÜTZE, F. Pesquisa biográfica e entrevista narrativa. In: WELLER, W.; PFAFF, N. (Org.). **Metodologias da pesquisa qualitativa em educação: teoria e prática**. 2. ed. Petrópolis: Vozes, 2011. p. 210-222.

SHAFIQUE, Umair; QAISER, Haseeb. **A Comparative Study of Data Mining Process Models**. International Journal of Innovation and Scientific Research ISSN 2351-8014 Vol. 12 Nº. 1 Nov. 2014, pp. 217-222© 2014. Disponível em: <https://www.researchgate.net/publication/268770881_A_Comparative_Study_of_Data_Mining_Process_Models_KDD_CRISP-DM_and_SEMMA>. Acesso em: 19 jul. 2019.

SHEARER, C. “**The CRISP-DM Model: The New Blueprint for Data Mining**”. J Data Warehousing, v5, n.4, p. 13-22, 2000.

SHIH, Munique. Tribunais na China permitem que IAs tomem o lugar de juizes. **Canal Tech**, 2022. Disponível em < <https://canaltech.com.br/seguranca/tribunais-na-china-permitem-que-ias-tomem-o-lugar-de-juizes-220922/>>. Acesso em 19 de out. de 2022.

SILVA, Cristiano Marçal Vieira da; GODOY, Aruana Brianezi Danelon; SANTOS, Paulo Alves dos. **Inteligência Artificial no Direito: Impactos, Desafios e Perspectivas**. Revista de Informação Legislativa, Brasília, v. 54, n. 215, p. 107-127, jul./set. 2017.

SINGAPURA. Singapore's Approach to AI Governance. Advisory Council on the Ethical Use of AI and Data. Disponível em < <https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework>>. Acesso em 19 de out. de 2022.

SPARKS, Hannah. China's AI attorney claims to prosecute crimes 'with 97% accuracy', **New York Post**, 2021. Disponível em <<https://nypost.com/2021/12/27/chinas-ai-attorney-prosecutes-crimes-with-97-accuracy/>>. Acesso em 15 de out. de 2022.

SPIELKAMP, Matthias. Inspecting algorithms for bias. **Technology Review**, v. 120, n. 4, p. 96-98, 2017.

The Beijing Artificial Intelligence Principles. **Wired**, 2019. Disponível em <<https://www.wired.com/beyond-the-beyond/2019/06/beijing-artificial-intelligence-principles/>>. Acesso em 10 de out. de 2022.

The FLI Team. A Principled AI Discussion in Asilomar. **Future of Life Institute**. Disponível em <<https://futureoflife.org/2017/01/17/principled-ai-discussion-asilomar/>>. Acesso em 06 de out. de 2022.

TURING, Alan M. Computing machinery and intelligence (1950). **The Essential Turing: the Ideas That Gave Birth to the Computer Age**, p. 433-464, 2012.

UNIÃO EUROPEIA. **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 de maio de 2023.

UNIÃO EUROPEIA. **Estudo aprofundado sobre a utilização da IA nos sistemas judiciais, nomeadamente as aplicações de IA que processam decisões e dados judiciais**. Disponível em <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0#_Toc530141218>. Acesso em 10 de maio de 2023.

UNIÃO EUROPEIA. High-Level Expert Group On Artificial Intelligence Set Up By The European Commission Ethics Guideline. **Ethics Guidelines for trustworthy AI**. Disponível em < <https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html>>. Acesso em 10 de out. de 2022.

UNIÃO EUROPEIA. **Quais as utilizações da IA nos sistemas judiciais europeus?**. Disponível em < https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0#_Toc530141230>. Acesso em 10 de maio de 2023.

UNITED KINGDOM. AI in the UK: ready, willing and able? A report by the House of Lords Select Committee on Artificial Intelligence. London, UK: The Stationery Office Limited, 2018.

VALE, Luís Manoel Bordes do; PEREIRA, João Sérgio dos Santos Soares. **Teoria do Processo Tecnológico**. São Paulo: Thomson Reuters Brasil, 2023.

VALLS, Álvaro L. M. **O que é ética**. 9ª ed. São Paulo: Brasiliense, 1994.

VAKKURI, V.; ABRAHAMSSON, P. The Key Concepts of Ethics of Artificial Intelligence. In: **2018 IEEE International Conference of Engineering Technology and Innovation ICE/ITMC**. IEEE. Disponível em <<http://dx.doi.org/10.1109/ICE.2018.8436265>>. Acesso em: 15 de out. de 2019.

VILLAS BÔAS CUEVA, R. Inteligência Artificial no Judiciário In: **Inteligência Artificial e Direito Processual: Os Impactos da Virada Tecnológica no Direito Processual**. Coord. Dierle Nunes, Paulo Henrique dos Santos Lucon, Erik Navarro Wolkart. Salvador: Editora JusPodivm, 2021.

VON NEUMANN, John; MORGENSTERN, Oskar. **Theory of games and economic behavior**. Princeton University Press: 1944.

XIMENES, Bianca Helena; RAMALHO, Geber L. Concrete ethical guidelines and best practices in machine learning development. In: **2021 IEEE International Symposium on Technology and Society (ISTAS)**. IEEE, 2021. p. 1-8.

WHITTAKER, Meredith et al. **AI now report 2018**. New York: AI Now Institute at New York University, 2018.

APÊNDICE A – ENTREVISTA NARRATIVA

SEÇÃO I – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Aceito participar da pesquisa sobre **INTELIGÊNCIA ARTIFICIAL E ÉTICA NO PODER JUDICIÁRIO**, desenvolvida pelo pesquisador e aluno **José Faustino Macêdo de Souza Ferreira**, pesquisador na Pós-Graduação (Mestrado) do Centro de Informática – UFPE.

Declaro que fui informado de que a pesquisa, nesta fase, pretende apreender elementos sobre a utilização de inteligência artificial no âmbito do Poder Judiciário, especificamente no que se refere aos parâmetros éticos relacionados ao desenvolvimento, testes e uso de ferramentas.

Fui informado(a) pelo pesquisador de que permanecerão sigilosas quaisquer informações que permitam minha identificação.

Estou ciente de que a participação na pesquisa é voluntária e não remunerada e de que os resultados serão publicados e estarão à disposição para consulta, quando concluída.

Fui informado de que o pesquisador poderá ser contatado para esclarecimentos por intermédio do e-mail: jfmsf@cin.ufpe.br.

De posse de tais esclarecimentos, AUTORIZO a utilização das informações prestadas para os propósitos da pesquisa acima descritos.

- 1) O senhor teve prévio conhecimento quanto ao Termo de Consentimento Livre e Esclarecido (TCLE)?
- 2) Autoriza a utilização desta entrevista no projeto de pesquisa que desenvolvo na UFPE?

SEÇÃO II – QUALIFICAÇÃO DO ENTREVISTADO

- 3) Nome completo do entrevistado ?
- 4) Tribunal a que pertence?
- 5) Qual seu cargo no tribunal?
- 6) Qual sua formação acadêmica?
- 7) Como você se define em relação às ciências computacionais?
- 8) Como você define o seu nível de conhecimento em relação à inteligência artificial?
- 9) Como você define o seu nível de conhecimento em relação à ética e IA?

SEÇÃO III – IA, ÉTICA E JUDICIÁRIO

- 10) Seu tribunal já desenvolveu ou usou alguma ferramenta que utilize IA?
- 11) Se sim:
 - a) Relaciona algum fato específico, condição ou circunstância que tenha motivado o interesse na implementação da solução?

- b) O que sabe sobre a ideação, o desenvolvimento e a operação do modelo de IA?
- c) A sua experiência com a operação/utilização do modelo de IA?
- d) Os benefícios do uso do modelo?
- e) Vislumbra riscos ou ameaças no uso do modelo?
- f) Questões éticas da IA estiveram na pauta de desenvolvimento do modelo?
- g) Se sim, com qual contorno? O que especificamente foi considerado?
- h) Quais critérios/standards éticos foram considerados/verificados?
- i) Conhecem a Res 332 CNJ? Ela foi utilizada?
- j) Quando o modelo foi idealizado e desenvolvido, houve preocupação com os riscos e ameaças gerados pelo seu uso?
- k) Quando o modelo foi idealizado e desenvolvido, foram catalogados (uma memória visual e histórica) os riscos e ameaças gerados pelo seu uso?
- l) Você pode citar riscos e ameaças que foram considerados?

SEÇÃO IV - PERCEPÇÕES FINAIS

- 12) O entrevistado mudaria algo na fase de ideação, desenvolvimento ou governança do modelo
- 13) Como o entrevistado vê uma metodologia de governança dos modelos de IA utilizados pelo Poder Judiciário?
- 14) Como o entrevistado avalia a proposição de uma lista de verificação (checklist) ordenado para conferência de tarefas e de etapas para conformação ética e legal dos modelos de IA produzidos pelo Judiciário?
- 15) Se positivo, que valores esse checklist deveria perseguir?
- 16) O entrevistado sugere algo ou gostaria de contribuir com alguma ideia?

APÊNDICE B – ENTREVISTA NARRATIVA

SEÇÃO I – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Aceito participar da pesquisa sobre **INTELIGÊNCIA ARTIFICIAL E ÉTICA NO PODER JUDICIÁRIO**, desenvolvida pelo pesquisador e aluno **José Faustino Macêdo de Souza Ferreira**, pesquisador na Pós-Graduação (Mestrado) do Centro de Informática – UFPE.

Declaro que fui informado de que a pesquisa, nesta fase, pretende apreender elementos sobre a utilização de inteligência artificial no âmbito do Poder Judiciário, especificamente no que se refere aos parâmetros éticos relacionados ao desenvolvimento, testes e uso de ferramentas.

Fui informado(a) pelo pesquisador de que permanecerão sigilosas quaisquer informações que permitam minha identificação.

Estou ciente de que a participação na pesquisa é voluntária e não remunerada e de que os resultados serão publicados e estarão à disposição para consulta, quando concluída.

Fui informado de que o pesquisador poderá ser contatado para esclarecimentos por intermédio do e-mail: jfmsf@cin.ufpe.br.

De posse de tais esclarecimentos, AUTORIZO a utilização das informações prestadas para os propósitos da pesquisa acima descritos.

SEÇÃO II – PERFIL DO ENTREVISTADO

A fim de possibilitar um recorte preciso, conte-nos um pouco sobre você.

Não se preocupe que as respostas aos próximos quesitos não possibilitam identificação dos entrevistados.

- 1) QUAL SEU CARGO/FUNÇÃO?
 - a) MAGISTRADA OU MAGISTRADO
 - b) SERVIDOR OU SERVIDORA
 - c) OUTROS:
- 2) QUAL ÓRGÃO OU INSTITUIÇÃO VOCÊ INTEGRA?
 - a) JUSTIÇA ESTADUAL
 - b) JUSTIÇA FEDERAL
 - c) JUSTIÇA ELEITORAL
 - d) JUSTIÇA DO TRABALHO
 - e) JUSTIÇA MILITAR
 - f) TRIBUNAL SUPERIOR
 - g) ÓRGÃO ADMINISTRATIVO DO JUDICIÁRIO
- 3) FORMAÇÃO ACADÊMICA
 - a) DIREITO
 - b) ADMINISTRAÇÃO

- c) CIÊNCIA DA COMPUTAÇÃO
- d) ENGENHARIA DA COMPUTAÇÃO
- e) ENGENHARIA DE SOFTWARE
- f) ANÁLISE E DESENVOLVIMENTO DE SISTEMAS
- g) SISTEMAS DE INFORMAÇÃO
- h) OUTROS:

SEÇÃO III – INTELIGÊNCIA ARTIFICIAL E EU

Conte-nos um pouco sobre seu contato com o tema inteligência artificial. Queremos saber qual seu conhecimento sobre o tema e também qual seu envolvimento prático com a tecnologia da IA.

- 1) QUAL O SEU NÍVEL DE CONHECIMENTO EM IA?
 - a) Posso conhecimento avançado
 - b) Possui conhecimento mediano
 - c) Posso algum conhecimento
 - d) Não posso conhecimento
- 2) COM QUAL FREQUÊNCIA VOCÊ UTILIZA OU DESENVOLVE FERRAMENTAS DE IA?
 - a) Muito frequentemente
 - b) Frequentemente
 - c) Eventualmente
 - d) Raramente
 - e) Nunca
- 3) SEU TRIBUNAL/ÓRGÃO USOU OU DESENVOLVEU ALGUMA FERRAMENTA QUE UTILIZE IA?
 - a) Muito frequentemente
 - b) Frequentemente
 - c) Eventualmente
 - d) Raramente
 - e) Nunca

SEÇÃO IV – ÉTICA E EU

Nesta seção, precisamos de uma autoavaliação quanto ao seu conhecimento com o tema ética, que pode advir de estudos em qualquer fase da sua formação acadêmica ou mesmo extracurricular.

- 1) QUAL O SEU NÍVEL DE CONHECIMENTO SOBRE ÉTICA?
 - a) Posso conhecimento avançado
 - b) Possui conhecimento mediano
 - c) Posso algum conhecimento
 - d) Não posso conhecimento

- 2) VOCÊ SE SENTE APTO/PREPARADO A APLICAR CONCEITOS ÉTICOS NO DESEMPENHO CONCRETO DO SEU TRABALHO?
- a) Completamente preparado
 - b) Preparado
 - c) Razoavelmente preparado
 - d) Pouco preparado
 - e) Nada preparado

SEÇÃO V – INTELIGÊNCIA ARTIFICIAL E O PODER JUDICIÁRIO

Aqui queremos compreender suas percepções sobre o fenômeno de utilização de IA pelos órgãos da Justiça.

- 1) NA SUA OPINIÃO, A INTELIGÊNCIA ARTIFICIAL É TECNOLOGIA APLICÁVEL, ÚTIL E ADEQUADA PARA RESOLUÇÃO DOS DESAFIOS ENFRENTADOS PELO PODER JUDICIÁRIO?
- a) Concordo totalmente.
 - b) Concordo.
 - c) Não concordo, nem discordo.
 - d) Discordo.
 - e) Discordo totalmente
- 2) A ABORDAGEM ÉTICA É IMPORTANTE PARA USO E DESENVOLVIMENTO DE FERRAMENTAS DE IA PELO JUDICIÁRIO?
- a) Muito importante
 - b) Importante
 - c) Mediana
 - d) Às vezes é importante
 - e) Não é importante
- 3) AVALIE A IMPORTÂNCIA DE UMA LISTA DE VERIFICAÇÃO (CHECKLIST) PARA CONFERÊNCIA DE CONFORMIDADE ÉTICA E LEGAL DOS MODELOS DE IA UTILIZADOS PELO JUDICIÁRIO.
- a) Muito importante
 - b) Importante
 - c) Mediana
 - d) Às vezes é importante
 - e) Não é importante

SEÇÃO VI – CRITÉRIOS E PARÂMETROS ÉTICOS

Finalizando a pesquisa, com apenas um clique em cada categoria/parâmetro, avalie a importância de cada um dos critérios listados com o tema ética e IA na Justiça.

Caso não conheça o critério, marque a opção "Não sei opinar".

AVALIE CADA UM DOS CRITÉRIOS ABAIXO DE ACORDO COM A IMPORTÂNCIA PARA O DESENVOLVIMENTO ÉTICO DA IA NO JUDICIÁRIO

1. Bem-estar
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

2. Proteção ao meio-ambiente
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

3. Auditabilidade
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

4. Explicabilidade
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

5. Transparência
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

6. Publicidade
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante

- e. Não é importante
 - f. Não sei opinar
7. Confidencialidade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
8. Justiça
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
9. Imparcialidade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
10. Equidade/igualdade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
11. Solidariedade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
12. Liberdade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante

- f. Não sei opinar
13. Honestidade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
14. Controle humano
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
15. Controle do usuário
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
16. Contestabilidade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
17. Qualidade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
18. Robustez
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

19. Segurança
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

20. Confiabilidade
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

21. Prudência
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

22. Responsabilidade
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

23. Flexibilidade
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

24. Cooperação
 - a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar

25. Prestação de contas
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
26. Proteção de dados/privacidade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
27. Participação pública / democrática
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
28. Ser humano no centro - educação e evolução do ser humano
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
29. Proteção da integridade humana
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
30. Desenvolvimento sustentável – sustentabilidade
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
31. Bem comum

- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
32. Não discriminação – crescimento inclusivo
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
33. Direitos humanos fundamentais
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
34. Qualidade de vida no trabalho
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
35. Governança / coordenação
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
36. Atenção / Vigilância contínuos
- a. Muito importante
 - b. Importante
 - c. Mediana
 - d. Às vezes é importante
 - e. Não é importante
 - f. Não sei opinar
37. Avaliação e gestão de riscos
- a. Muito importante

- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

38. Análise de Custo/benefício

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

39. Benefício e prosperidade compartilhados

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

40. Não subversão de processos cívicos e sociais

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

41. Evitar corrida armada da IA

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

42. Autoaprimoramento recursivo / rígidas medidas de segurança e controle

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

43. Acurácia (precisão)

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

44. Pesquisa

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

45. Inovação e criatividade

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

46. Evitar/rejeitar condutas ilícitas

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

47. Aceitação de críticas ao trabalho técnico

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar

48. Manutenção e melhoramento da competência técnica

- a. Muito importante
- b. Importante
- c. Mediana
- d. Às vezes é importante
- e. Não é importante
- f. Não sei opinar