

UNIVERSIDADE FEDERAL DE SÃO CARLOS
PROGRAMA DE PÓS-GRADUAÇÃO INTERDISCIPLINAR EM ESTUDOS DA
CONDIÇÃO HUMANA

ANDRÉA MARIA FIERI SILVA

**A IMPORTÂNCIA DA TRANSPARÊNCIA E INTERVENÇÃO HUMANA NAS
TECNOLOGIAS DE INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE CRÍTICA DAS
INICIATIVAS DE REGULAMENTAÇÃO ÉTICA EM ANDAMENTO**

Sorocaba

2023

UNIVERSIDADE FEDERAL DE SÃO CARLOS
PROGRAMA DE PÓS-GRADUAÇÃO INTERDISCIPLINAR EM ESTUDOS DA
CONDIÇÃO HUMANA

ANDRÉA MARIA FIERI SILVA

**A IMPORTÂNCIA DA TRANSPARÊNCIA E INTERVENÇÃO HUMANA NAS
TECNOLOGIAS DE INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE CRÍTICA DAS
INICIATIVAS DE REGULAMENTAÇÃO ÉTICA EM ANDAMENTO**

Dissertação apresentada ao Programa de Pós-Graduação em Estudos da Condição Humana da Universidade Federal de São Carlos *campus* Sorocaba, para obtenção do título de Mestre em Estudos da Condição Humana.

Orientação: Prof. Dr. Dalton Lopes Martins

Sorocaba

2023

FICHA CATALOGRÁFICA

Silva, Andréa Maria Fieri

A importância da transparência e intervenção humana nas tecnologias de Inteligência Artificial: uma análise crítica das iniciativas de regulamentação ética em andamento / Andréa Maria Fieri Silva -- 2024.
125f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus Sorocaba, Sorocaba
Orientador (a): Dalton Lopes Martins
Banca Examinadora: Márcio Antônio Gatti, Luciana Conrado Martins
Bibliografia

1. Inteligência Artificial. 2. Regulamentação ética. I. Silva, Andréa Maria Fieri. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática
(SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Maria Aparecida de Lourdes Mariano -
CRB/8 6979



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Humanas e Biológicas
Programa de Pós-Graduação em Estudos da Condição Humana

Folha de Aprovação

Defesa de Dissertação de Mestrado da candidata Andréa Maria Fieri Silva, realizada em 01/03/2024.

Comissão Julgadora:

Prof. Dr. Dalton Lopes Martins (UnB)

Prof. Dr. Marcio Antonio Gatti (UFSCar)

Profa. Dra. Luciana Conrado Martins (SESI Lab)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Estudos da Condição Humana.

DEDICATÓRIA

Dedico este trabalho de coração e alma aos meus filhos queridos, Gabriel e Mariana, cuja presença constante trouxe luz, força e inspiração a todos os dias desta jornada acadêmica. Suas risadas, paciência e amor foram minha motivação mais poderosa.

Ao meu esposo Flavio, seu apoio, paciência e compreensão foram fundamentais durante todas as horas necessárias de dedicação desta trajetória. Juntos, construímos um caminho que vai além das páginas desta dissertação.

À memória de nosso filho Felipe, que mesmo ausente fisicamente, permanece presente em cada página. Sua breve passagem por nossas vidas durante esse período de estudos, foi um aprendizado doloroso, mas sua lembrança ilumina cada passo. Esta dissertação é também dedicada a ele, uma estrela que brilha eternamente em nossos corações.

Que este trabalho seja uma homenagem ao meu pai (em memória), à minha mãe e ao meu irmão, que refletem o amor familiar e a resiliência que compartilhamos ao longo dessa trajetória, pois cada desafio superado foi fortalecido pelo vínculo familiar que nos une.

À Deus, pois sempre faz possível todas as coisas, pois como está escrito: “Tudo posso naquele que me fortalece” (Filipenses 4:13).

E por fim, dedico todo esse trabalho a mim mesma. As experiências desafiadoras que aconteceram durante esse período, fortaleceram minha resiliência e determinação, moldando-me de maneiras intensas apesar das adversidades.

AGRADECIMENTO

Quero expressar minha sincera gratidão a todas as pessoas que desempenharam um papel fundamental na conclusão desta dissertação, contribuindo significativamente para o meu percurso acadêmico e pessoal.

Em especial, meu agradecimento ao meu orientador, Professor Dalton Lopes Martins. Sua calma, paciência e expertise foram pilares para o desenvolvimento desta pesquisa. Suas contribuições foram valiosas, moldando não apenas o trabalho, mas também o meu crescimento como pesquisadora.

Agradeço ao Programa de Pós-graduação em Estudos da Condição Humana, representado pela Professora Viviane Melo de Mendonça, e a todo o corpo docente, pelo acolhimento, oportunidade e generosa partilha de conhecimento.

À Professora Mariana Faiad, minha eterna gratidão. Em um momento desafiador e doloroso, seu abraço acolhedor, sua compreensão, apoio e encorajamento foram essenciais. Sua generosidade ao compartilhar conhecimento e empatia contribuíram significativamente para a conclusão deste trabalho.

Gustavo, meu psicólogo, merece um agradecimento especial. Seu apoio emocional foi essencial para superar os desafios pessoais que surgiram durante este período. Sua abordagem compassiva, firmeza e encontros valiosos foram fundamentais para o equilíbrio entre as demandas acadêmicas.

À amiga e apoiadora, Marina Helena Pereira Vieira, agradeço do fundo do coração. Sua amizade, compreensão e palavras de estímulo foram uma fonte constante de inspiração. Seu apoio incondicional foi um pilar vital ao longo dessa jornada.

A todos que acreditaram, em especial a minha amiga Cristiane Palomar, meu agradecimento! Nossas partilhas foram um reflexo do valor da colaboração, amizade e apoio.

Sou verdadeiramente grata por ter compartilhado esta jornada com pessoas excepcionais.

Epígrafe

“.. A vida é transformação, o futuro é ancestral.”

Tukano, 2017

RESUMO

O avanço da Inteligência Artificial em diversas áreas da sociedade tem gerado preocupações em diversos âmbitos e a regulamentação ética, portanto, passou a ser um tema urgente e relevante. Esse estudo tem como objetivo examinar regulamentações éticas das tecnologias que utilizam Inteligência Artificial, em diferentes partes do mundo. Através da análise do conteúdo das legislações vigentes, busca-se oferecer uma visão abrangente do momento atual e identificar os desafios e lacunas a serem enfrentados no futuro. A regulamentação ética da Inteligência Artificial é fundamental para garantir que seu desenvolvimento e para que sua utilização seja conduzida de forma responsável, transparente e segura, sem agravar ainda mais a vulnerabilidade dos sujeitos e das instituições de Estado.

Palavras-chave: Inteligência Artificial; Ética; Legislação; Dados.

ABSTRACT

The advancement of Artificial Intelligence in various areas of society has raised concerns in various areas and ethical regulation, therefore, has become an urgent and relevant topic. This study aims to examine ethical regulations of technologies that use Artificial Intelligence in different parts of the world. By analyzing the content of current legislation, we seek to offer a comprehensive view of the current situation and identify the challenges and gaps to be faced in the future. The ethical regulation of Artificial Intelligence is essential to ensure that its development and that its use is conducted in a responsible, transparent and secure manner, without further aggravating the vulnerability of subjects and State institutions.

Keywords: Artificial Intelligence; Ethic; Legislation; Data.

LISTA DE FIGURAS

Figura 1 - Estrutura da Pesquisa	25
--	----

LISTA DE QUADROS

Quadro 1- Características da Análise Documental e Análise de Conteúdo.....	51
Quadro 2- Desenvolvimento da Análise de Conteúdo	52
Quadro 3 - Documentos Analisados	54
Quadro 4- Análise comparativa preliminar das diretrizes analisadas.....	98
Quadro 5- Análise final das diretrizes analisadas – Quadro Resumido.....	103

LISTA DE GRÁFICOS

Gráfico 1- Resultado da análise de dados referente à diretriz Regulamentação Ética.....	106
Gráfico 2 Resultado da análise de dados referente à diretriz Educação e Capacitação Profissional	107
Gráfico 3 - Resultado da análise de dados referente à diretriz Infraestrutura.....	108
Gráfico 4 - Resultado da análise de dados referente à diretriz Acordos Internacionais	109

LISTA DE ABREVIATURAS E SIGLAS

COMEST – Comissão Mundial sobre Ética do Conhecimento Científico e Tecnologia

EBIA – Estratégia Brasileira de Inteligência Artificial

EUA – Estados Unidos da América

EvEx – Evidência Express

FTC – *Federal Trade Commission*

IA – Inteligência Artificial

IoT – *Internet of Things*

LGPD – Lei Geral de Proteção de Dados

OCDE – Organização para a Cooperação e Desenvolvimento Econômico

OMB – *Office of Management and Budget*

P&D – Pesquisa e Desenvolvimento

PD&I – Projetos de Pesquisa, Desenvolvimento e Inovação

PL – Projeto de Lei

TICs – Tecnologias da Informação e Comunicação

UE – União Europeia

UFSCar – Universidade Federal de São Carlos

UnB – Universidade Federal de Brasília

UNESCO – Organização das Nações Unidas para a Educação, a Ciência e a Cultura

SUMÁRIO

1. Introdução	14
1.1. Apresentação	17
1.2. Justificativa.....	19
1.3. Objetivos.....	23
1.4. Estrutura da dissertação	24
2. Referencial Teórico.....	27
2.1 Novo contexto tecnológico – A IA na sociedade	27
2.2 Conceituação de inteligência artificial.....	36
2.2.1 Machine Learning e Deep Learning	38
2.3 Histórico e evolução	39
2.4. Ética na IA	41
2.5. A era dos dados	48
3. Metodologia	50
3.1 Recorte do estudo	53
4. Análise Documental	56
4.1 IA no mundo	56
4.1.1.Regulamentação Setorial, Regulamentação Transversal e Regulamentação.....	58
4.2 Estados Unidos	59
4.2.1 Quadro de diretrizes encontradas em documento analisado	61
4.3 União Europeia	69
4.3.1. Quadro de diretrizes encontradas em documento analisado	72
4.4 China.....	77
4.4.1. Quadro de diretrizes encontradas em documento analisado	79
4.5 Brasil.....	84
4.5.1. Quadro de diretrizes encontradas em documento analisado	87

4.6	Síntese da análise dos documentos	95
5.	Resultados.....	100
5.1.	Apresentação dos Dados Oficiais	102
5.1.1.	Quadro análise final	102
5.2.	Apresentação dos dados de resultados por referência de diretriz.....	105
5.2.1.	Diretriz - Regulamentação Ética	106
5.2.2.	Educação e capacitação profissional.....	107
5.2.3.	Infraestrutura	108
5.2.4.	Acordos Internacionais	109
5.3.	Relação com os Objetivos da Pesquisa:	110
6.	Considerações Finais	112
7.	Referências	116

1. Introdução

De maneira geral, a Inteligência Artificial (IA) refere-se a um campo de conhecimento relacionado à linguagem e à inteligência, ao raciocínio e aprendizagem voltados à resolução de problemas, a partir da interação entre o humano e a máquina (Kaufman, 2019).

Atualmente, os sistemas inteligentes permeiam quase todos os campos da ciência e a vida em sociedade. Por ser tema de pesquisa em diversas áreas do conhecimento, como Computação, Matemática, Linguística, Neurociências, dentre muitas outras, e se desdobrar em muitos subcampos de estudo, há uma dificuldade em descrever o estado da arte de Inteligência Artificial.

O termo “Inteligência Artificial” foi cunhado por John McCarthy em 1956, durante uma conferência sobre tecnologia no Dartmouth College, nos EUA. No entanto, o assunto já era objeto de pesquisa anteriormente por volta de 1950, por Alan Turing, considerado o pai da computação (Silva; Mairink, 2019).

Apesar de não se tratar de um assunto recente, as pesquisas no âmbito da Inteligência Artificial cresceram muito nos últimos anos com o substancial avanço tecnológico, que levou a evolução dos sistemas informacionais a novos patamares.

Segundo Doneda *et al.*, (2018), a disponibilidade de mais recursos computacionais e de mais informação, resultantes tanto do desenvolvimento tecnológico como do paradigma do “Big Data”, proporcionaram o surgimento de sistemas de inteligência artificial alimentados com grandes quantidades de dados¹, que subsidiam decisões e formam padrões de decisão desses sistemas.

No contexto da inteligência artificial, a utilização de grandes volumes de dados é uma prática amplamente adotada para o treinamento de modelos. Estes conjuntos de dados fornecem uma vasta gama de informações, possibilitando aos modelos aprender padrões complexos.

O processo de treinamento dos modelos de IA é iterativo, no qual o algoritmo ajusta seus parâmetros internos com base nos exemplos apresentados nos dados de treinamento. Esta

¹ Dados são registros armazenados em sistemas computacionais que geram informações processadas por meio de softwares. São fundamentais para tecnologia de informação. Controle Net. O que são dados e qual sua importância para a tomada de decisão. 2023. Disponível em <https://www.controle.net/faq/o-que-sao-dados>. Acesso em: 11 de novembro 2023.

abordagem reflete de maneira semelhante o processo de aprendizagem humana, onde a exposição a uma variedade de exemplos contribui para a melhoria da compreensão e da capacidade de reconhecimento de padrões.

Uma vez devidamente treinados, tais modelos são capazes de realizar previsões estatísticas ou tomar decisões embasadas nas informações absorvidas durante o treinamento.

Contudo, é fundamental destacar que a qualidade e a quantidade dos dados de treinamento podem impactar significativamente o desempenho e a precisão dos modelos de IA. Conjuntos de dados incompletos, tendenciosos ou mal rotulados podem resultar em previsões imprecisas ou decisões inadequadas.

Assim, a partir da segunda década do século XXI, a convergência de diversas tecnologias – Internet das Coisas (ou *Internet of Things* - IoT), *blockchain*, plataformas digitais, impressão 3D, robótica avançada, novos materiais, manipulação genética – permeadas pela IA, vêm promovendo resultados superiores a quaisquer previsões precedentes (Kaufman; Santaella, 2020).

A Inteligência Artificial, portanto, se torna cada vez mais presente em diversos setores da sociedade contemporânea, como saúde, transporte, finanças, segurança e desperta a atenção para seus potenciais benefícios e desafios, bem como as preocupações éticas envolvidas que surgem nesse novo cenário. Apesar dos benefícios da IA, sua aplicação em certos contextos pode criar riscos para usuários e desenvolvedores. Como gerenciar e preparar a estrutura regulatória para coibir tais práticas é um desafio sendo enfrentado por governos em todo o mundo.

Assim, com a crescente disseminação e utilização de sistemas de inteligência artificial e com uma legislação e regulamentação ainda em fase inicial, surgem algumas preocupações sobre o cumprimento legal e ético do uso e aplicação da inteligência artificial.

O debate sobre privacidade, viés algorítmico, a transparência, responsabilidade, são apenas algumas das questões que precisam ser consideradas ao se pensar no desenvolvimento e uso de tecnologias de IA. Segundo Rossetti e Angeluci (2021), a tomada de decisão algorítmica tem potencial de prejudicar indivíduos e grupos, aumentando a necessidade de transparência nos processos algorítmicos.

No Brasil, temos desde 2014, a Lei 12.965, conhecida como “Marco Civil da Internet”, uma legislação que estabelece direitos, garantias e responsabilidades dos usuários da internet

no país. Este marco legal descreve princípios fundamentais que regem o uso da internet no Brasil, visando assegurar a privacidade e a proteção dos usuários.

Uma das garantias oferecidas pelo Marco Civil é a inviolabilidade e sigilo do fluxo de comunicações dos usuários, bem como a inviolabilidade e sigilo de suas comunicações privadas e armazenadas, salvo por ordem judicial (TJDFT, 2014). Essas questões são essenciais para proteger a privacidade dos indivíduos e garantir que suas comunicações não sejam acessadas sem a devida autorização legal.

É importante observar que, embora o Marco Civil da Internet não profba explicitamente o uso de grandes conjuntos de dados, como é o caso dos sistemas de Inteligência Artificial, ele estabelece diretrizes que devem ser rigorosamente seguidas para garantir a conformidade com os princípios fundamentais, como a proteção da privacidade e a neutralidade na internet. Dessa forma, qualquer prática de desenvolvimento de softwares que envolva a coleta e o armazenamento de dados pessoais sem o consentimento adequado pode estar em desacordo com as leis de privacidade no Brasil, no caso o Marco Civil da Internet e a Lei Geral de Proteção de Dados².

Essas medidas visam equilibrar a inovação tecnológica e o desenvolvimento da internet com a proteção dos direitos individuais, garantindo um ambiente online seguro e respeitoso às liberdades individuais dos usuários.

A utilização de dados para alimentar os novos sistemas de inteligência artificial e a sua utilização para a tomada de decisões proporcionam uma acurácia bastante significativa para um número crescentes de aplicações. Para Doneda *et al.*, (2018) isto abre espaço, ao menos, de dois temas centrais para os debates sobre autonomia e direitos fundamentais nos próximos anos: os efeitos que a utilização desses sistemas causará para a pessoa e sua autonomia pessoal e a necessidade de qualificar a natureza desses instrumentos e sistemas de inteligência artificial.

Nesse debate, a necessidade de que sejam proporcionadas soluções que preservem os direitos fundamentais, dentro de um quadro de intenso desenvolvimento tecnológico, sugere a necessidade de recorrer à ética como instrumento capaz de encaminhar soluções que,

² Lei Geral de Proteção de Dados - Lei nº 13.709/2018, legislação brasileira que estabelece regras para o tratamento de dados pessoais por entidades públicas e privadas. Portal Câmara dos Deputados. LEI Nº 13.709, DE 14 de agosto de 2018. Disponível em: <https://www2.camara.leg.br/legin/fed/lei/2018/lei-13709-14-agosto-2018-787077-publicacaooriginal-156212-pl.html#:~:text=Art.,da%20personalidade%20da%20pessoa%20natural..> Acesso em: novembro 2023.

eventualmente possam consolidar-se em alternativas regulatórias e, posteriormente até mesmo legislativas.

Embora existam iniciativas nacionais e internacionais em andamento, a elaboração de diretrizes e regulamentações que abordem as questões éticas da inteligência artificial enfrentam grandes desafios. Dentre as dificuldades discutidas, a necessidade de estabelecer padrões globais e vinculantes para a inteligência artificial, é importante, considerando que essa tecnologia transcende fronteiras nacionais. (Wermuth; Rocha, 2022).

Nesse sentido, é necessário ampliar pesquisas sobre as iniciativas existentes e identificar lacunas que precisam ser preenchidas na regulamentação ética da Inteligência Artificial de forma global, criar normas que não se tornem obsoletas desde o seu surgimento dado o ritmo acelerado da evolução tecnológica.

1.1. Apresentação

É com muita satisfação que a partir deste texto, compartilho parte da minha trajetória acadêmica e profissional. Tenho certeza a cada dia, que partes da minha história moldaram minha evolução, sou Andréa, mãe, casada, profissional que iniciou carreira cedo, mas antes de tudo uma pessoa em formação contínua, cada vez mais entusiasta com o ser humano em sua diversidade. Características estas que moldam e enriquecem minha perspectiva de aprendizado diário.

Iniciei minha formação de nível médio, na Escola Técnica Estadual Fernando Prestes, na cidade de Sorocaba, onde cursei Processamento de Dados, mergulhando nas bases tecnológicas que pavimentariam minha carreira. Posteriormente, obtive o título de bacharel em Análise de Sistemas pela Universidade de Sorocaba.

Os primeiros passos na vida profissional foram marcados por experiências em indústrias e empresas de tecnologia, desempenhando papéis fundamentais nas áreas de qualidade de software, suporte ao cliente, e testes de sistemas. Neste momento, mesmo sem perceber, a área de humanidades já estava intrinsicamente em mim, pois para exercer essas atividades profissionais faz-se necessário uma junção de áreas, o que na formação acadêmica dentro da área de exatas faz muita falta.

Concomitantemente, iniciei a carreira de docente, a partir de uma oportunidade de um concurso público, compartilhar conhecimento como docente no Centro Paula Souza, atuando

no ensino técnico integrado ao ensino médio em tecnologia. Esta fase foi decisiva para minha percepção sobre a interconexão entre os conhecimentos de exatas e humanas, indicando um caminho promissor na confluência desses campos em benefício do desenvolvimento humano e social.

Impulsionada pela necessidade de aprimorar conhecimentos, ingressei na pós-graduação em Educação a Distância, consolidando uma compreensão mais ampla das dinâmicas educacionais contemporâneas. Anos mais tarde, decidida a explorar novos horizontes, iniciei o Mestrado em Ciências da Computação, onde cursei duas disciplinas como aluna especial. Contudo, foi no Programa Interdisciplinar em Estudos da Condição Humana que encontrei não apenas um ambiente acolhedor, mas também um espaço propício para meu desenvolvimento acadêmico e pessoal.

Durante esse período, pude perceber de forma vívida a complementaridade entre as áreas de conhecimento. A integração desses saberes revelou-se vital para compreender as complexidades da condição humana, destacando a importância de uma abordagem interdisciplinar na busca por soluções que atendam às necessidades sociais, muito mais com a Inteligência Artificial em nossos dias.

Este caminho de descobertas consolidou minha convicção de que o conhecimento não é solitário, muito menos que o conhecimento não é oposto, mas sim aliados na construção de um futuro mais justo e equitativo. Cada disciplina, em sua singularidade, contribuiu para a compreensão do humano e sua interação com o meio. Dessa forma enfatizo a importância de uma abordagem unificada, onde conhecimentos técnicos e humanísticos convergem para promover impacto positivo em nossa sociedade.

Esse breve texto resume e reflete parte da trajetória que moldou minha visão, reforçando meu compromisso com a construção de pontes entre diferentes áreas do conhecimento em prol do desenvolvimento humano integral.

Agradeço a oportunidade de partilhar essas experiências neste espaço acadêmico, consciente de que cada capítulo contribui para estruturar uma compreensão mais rica e profunda da condição humana.

1.2. Justificativa

A inteligência artificial já desempenha um papel significativo na vida humana em diversas áreas. Em plataformas de *streaming* e comércio eletrônico, como a Netflix e Amazon, os sistemas de recomendação³ guiados por essa tecnologia, interpretam os padrões comportamentais dos usuários, oferecendo sugestões personalizadas. Além disso, nos assistentes de voz, como a Siri da Apple e o *Google Assistant*, utilizam técnicas avançadas de processamento de linguagem natural⁴ capacitando-os a responder perguntas e desempenhar tarefas cotidianas de forma intuitiva. Esses exemplos tangíveis evidenciam a presença intrínseca da inteligência artificial, sublinhando sua relevância e impacto multifacetado na sociedade contemporânea.

Além disso, a IA é amplamente empregada em áreas como diagnóstico médico, previsão do tempo, condução autônoma de veículos, jogos de vídeo *game*, controle agrário, dentre outras. O crescente uso da IA demonstra seu impacto e relevância cada vez maiores na vida diária das pessoas, tornando-a uma tecnologia fundamental para o presente e o futuro.

Assim, a importância de discutir aspectos éticos na criação e uso da inteligência artificial tem sido amplamente destacada em diversos campos do conhecimento. Essas discussões são fundamentais para garantir o desenvolvimento responsável e benéfico da IA, bem como para evitar danos e consequências negativas à sociedade.

Em outras palavras, em um mundo de rápido desenvolvimento e disseminação de tecnologia, a ética desempenha um papel fundamental na análise de como essas tecnologias afetam indivíduos, empresas, grupos, sociedade e meio ambiente (Sætra; Fosch-Villaronga, 2021).

É necessário entendimento dos impactos gerados pela revolução tecnológica, especialmente no que diz respeito à interação entre ética, regulamentação e progresso tecnológico. Para abordar esse tema de maneira eficaz, é essencial compreender o conceito de

³ Sistema ou algoritmos de recomendação são aplicações que sugerem conteúdo aos seus usuários com base na probabilidade de aceitação, considerando seu histórico de acessos. Tecnoblog. Como funcionam os sistemas de recomendação?, 2022. Disponível em: <https://tecnoblog.net/responde/como-funcionam-os-sistemas-de-recomendacao/>. Acesso: novembro 2023.

⁴ Processamento de linguagem natural - mescla ciência da computação, inteligência artificial e linguística, se dedicando à geração e compreensão automática da linguagem natural em sistemas de inteligência artificial. Alura. PLN: o que é Processamento de Linguagem Natural?, 2023. Disponível em: <https://www.alura.com.br/artigos/o-que-e-pln>. Acesso em: dezembro 2023.

*Data Ethics*⁵, um mecanismo ético que desempenha um papel significativo na governança tecnológica e, portanto, contribui de maneira substancial para o debate em questão.

Ao pensarmos em *Data Ethics*, é necessário o entendimento de vertentes, como a ética de dados, onde concentra questões referentes a coleta e análise de grandes conjuntos de dados; ética dos algoritmos⁶ que traz entendimento da complexidade de soluções e autonomia de em sistemas de inteligência artificial e a ética das práticas que inclui a responsabilidade de pessoas e organizações envolvidas nas construções destes sistemas. (Bigaton; Teixeira; Bilhalva, 2022).

Muitas aplicações de IA possuem a capacidade de evoluir sem intervenção humana explícita. Isso implica, muitas vezes, que seus criadores não têm um conhecimento preciso sobre o funcionamento exato do algoritmo. A opacidade resultante é uma das preocupações centrais, contrapondo-se à transparência e intensificando a imprevisibilidade e a incontrolabilidade da IA (Cordeiro, 2022).

Para Bostrom e Yudkowsku (2011), a transparência não é a única característica desejável da IA, mas também a governança da Inteligência Artificial necessita entender a previsibilidade e a interpretabilidade.

A falta de transparência nos algoritmos de IA pode ser problemática. Em seu artigo “*The Mythos of Model Interpretability*”, Lipton (2018) ressalta que a opacidade dos modelos de IA pode dificultar a compreensão de suas decisões e impede que os usuários responsabilizem as ações tomadas por esses sistemas. A discussão incentiva a criação de métodos interpretáveis e abordagens para tornar o processo de tomada de decisão da IA mais transparente e compreensível.

Barocas e Selbst (2016) abordam os desafios de proteger a privacidade em um mundo movido a big data e os riscos de discriminação associados. Discutir ética pode ajudar a estabelecer diretrizes para a coleta, armazenamento e uso responsável desses dados, protegendo a privacidade das pessoas e minimizando riscos.

⁵ Data Ethics – A Ética de dados é um campo que estuda questões ligadas a dados, algoritmos e práticas associadas, com o objetivo de criar e administrar soluções baseadas em condutas éticas e valores apropriados. TADDEO, Mariarosaria ; FLORIDI, Luciano. Introduction: What is data ethics? *Philosophical Transactions R. Soc.*, Londres, v. 374, n. 2083, 2016. Disponível em: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2016.0360>. Acesso em nov. de 2023.

⁶ Algoritmos – conjunto de regras e operações bem definidas que aplicadas a um conjunto de dados e a um número finito de etapas conduzem a resolução de um problema. São fundamentais em processos computacionais. Baudson, Adolfo José Gonçalves Stavaux; Araújo, Francisco César Rodrigues. *Algoritmos e Programação*. Rede E-tec Brasil, 2013. Disponível em: <https://www.ifmg.edu.br/ceadop3/apostilas/algoritmos-e-programacao>. Acesso em dezembro 2023.

Anderson e Anderson (2007) já discutiam há mais de uma década sobre importância de determinar responsabilidades e propõem princípios éticos para orientar o desenvolvimento da IA. Essas discussões são fundamentais para garantir que a autonomia da IA seja acompanhada de responsabilidade e intervenção humana adequada.

Há uma preocupação de que uma eventual regulação excessiva possa inibir a inovação por parte das empresas, mas, ao mesmo tempo ignorar esse debate pode significar deixar de prevenir e mitigar danos que possam ser causados pelo uso indevido da IA. Assim, os países estão experimentando diferentes propostas que levam em conta essas dimensões, além de outras, como competitividade.

Parece claro que o tipo de transparência e a intervenção humana necessária para reorientar e preservar a autonomia dos indivíduos, exigirá legislação, supervisão e jurisprudência dedicada para reequilibrar as atuais relações de poder. O ser humano gera a informação, produto valioso para o gerenciamento da IA. (Hildebrandt, 2018)

Além disso, ao considerarmos as implicações éticas da inteligência artificial, é fundamental examinar as relações de poder subjacentes a esse cenário. Essas dinâmicas ultrapassam os interesses individuais quanto à exploração de matéria-prima. No capitalismo industrial utiliza-se a natureza para compras e vendas, emprega o trabalho para dinamizar o mercado. Já o capitalismo de vigilância⁷ explora a experiência humana gratuita como fonte de matéria-prima para previsões de comportamento a serem comercializado no mercado. (Zuboff, 2021)

Essa exploração ocorre muitas vezes sem que as pessoas saibam, inibindo assim o direito sobre a propriedade das informações coletadas, seus resultados de processamento e o futuro resultante dessa dinâmica de apropriação, única exclusivamente para atender a interesses de grandes empresas de tecnologia na movimentação da nova maneira de colonização de países a partir do controle de dados. (Silveira, 2020).

A análise dos dados coletados gera conhecimentos valiosos e precisos. Enquanto sociedade pode utilizar essas informações para aprimorar decisões e compreender processos pelos quais os indivíduos estão vivendo, atendendo a interesses coletivos, o impacto na vida das pessoas pode também ser tendencioso e excludente justamente porque os dados refletem a

⁷ Capitalismo de Vigilância - É uma especificidade do capitalismo contemporâneo, no qual as práticas de coleta massiva de dados sobre comportamentos e atividades das pessoas são fundamentais para os modelos de negócios de algumas grandes empresas. SILVEIRA, Sérgio Amadeu *et al.* Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal. 1. ed. São Paulo: Autonomia Literária, 2021.

realidade, e a IA é uma ferramenta de criação de riqueza, portanto é crucial a intervenção humana nesses processos. (Dijck, 2014).

A experiência e a condição humana se tornaram matéria-prima a ser explorada por plataformas que podem utilizar os dados não somente das camadas médias, mas também das massas pauperizadas para treinar seus algoritmos de aprendizado de máquina⁸. É impressionante que quanto mais o neoliberalismo manda reduzir custos do Estado, mais avança a extração de dados dos países empobrecidos para suas matrizes. Os dirigentes do Brasil atual fazem de tudo para assegurar as plenas condições para a extração de dados aqui na colônia. (Silveira, 2020).

Dado o rápido avanço da IA e suas implicações éticas, é provável que mais países e organizações internacionais desenvolvam e atualizem suas regulamentações para lidar com as complexidades e desafios éticos da inteligência artificial. No entanto, até o momento, não há uma regulação ética única que cubra todos os aspectos da IA em escala global.

A regulação da IA varia significativamente de país para país e até mesmo entre diferentes jurisdições dentro de um mesmo país.

Alguns países têm adotado estratégias específicas para abordar questões éticas relacionadas à IA e tecnologias associadas. Por exemplo, a União Europeia (UE) tem trabalhado em diretrizes e regulamentações para IA responsável e ética. Outros países, como os Estados Unidos e a China, têm adotado abordagens diferentes em relação à regulamentação da IA considerando princípios da autorregulação da indústria e privacidade dos dados.

Além disso, algumas organizações internacionais, como a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), têm trabalhado na elaboração de princípios e diretrizes para a IA ética.

Por esse motivo, essa pesquisa busca preencher uma lacuna teórica, ao levantar as principais iniciativas globais em andamento para a regulamentação ética da inteligência artificial e verificar o que elas possuem em comum.

⁸ Aprendizado de máquina ou *Machine Learning* em inglês, é um subcampo da inteligência artificial (IA) que se concentra no desenvolvimento de algoritmos e modelos que permitem aos sistemas computacionais aprenderem padrões e tomar decisões com base em dados coletados. Tem como objetivo capacitar os computadores a melhorarem seu desempenho em uma determinada tarefa à medida que são expostos a mais dados relevantes. PEREIRA, Flavia. *Machine Learning: Um guia atualizado!*, 2023. Disponível em: <https://www.dataex.com.br/machine-learning-um-guia-atualizado/>. Acesso em: dezembro 2023.

1.3. Objetivos

Esse estudo tem como objetivo geral analisar as principais iniciativas globais em andamento para a regulamentação ética da inteligência artificial, a fim de identificar as principais diretrizes, lacunas, estratégias e desafios estabelecidos por políticas públicas internacionais e nacional, com vistas a compreender posicionamentos e recomendações prioritários para promover a transparência na concepção e uso dessas tecnologias.

O recorte temporal deste estudo abrange a análise de propostas de regulamentações citadas no período de 2021 e 2023. Essa delimitação visa destacar as discussões e desenvolvimentos nesse campo dinâmico, durante o referido período, proporcionando uma visão atualizada das abordagens regulatórias em curso.

Para isso, tem como objetivos específicos:

- I. Desenvolver uma revisão bibliográfica sobre a conceituação de Inteligência Artificial;
- II. Analisar os documentos legais que tratem das diretrizes éticas para Inteligência Artificial no Brasil;
- III. Analisar os documentos que regulamentam as diretrizes éticas para Inteligência Artificial nos países: Estados Unidos, China, e a União Europeia. Esse recorte foi definido a partir da adaptação de um *benchmarking* publicado pelo Evidência Express (EvEx) da Universidade Federal de Brasília (UnB)⁹ que descreve políticas públicas internacionais, já implementadas ou em desenvolvimento, para regular o uso da Inteligência Artificial (IA) no mundo. O trabalho foi elaborado para identificar as atividades que ocorrem no âmbito da Estratégia Brasileira de Inteligência Artificial (EBIA).
- IV. Por fim, criar um quadro com os aspectos que esses documentos têm em comum, a partir de uma análise de conteúdo dos mesmos.

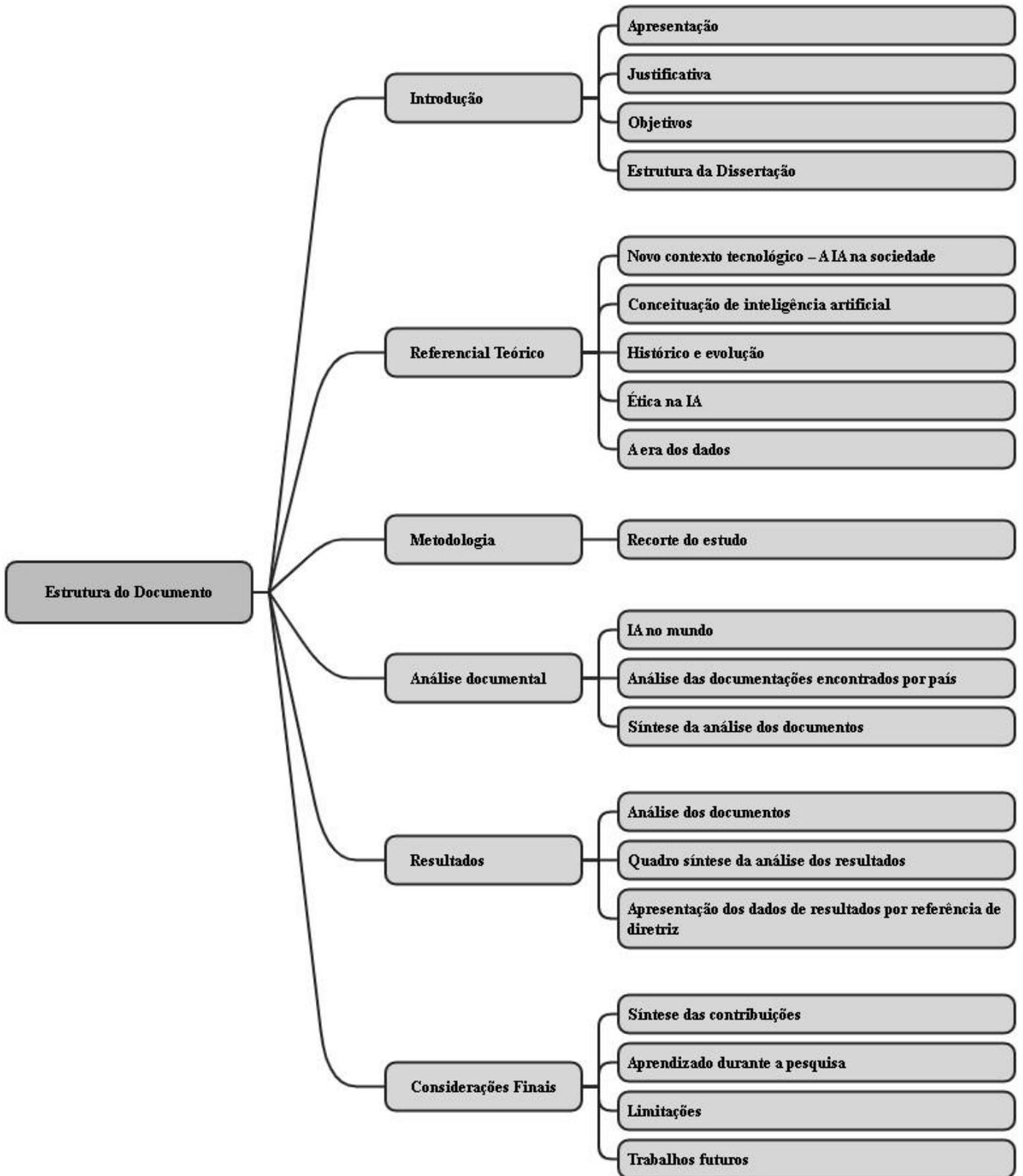
⁹ O Evidência Express (EvEx) é uma iniciativa da Diretoria de Altos Estudos da Escola Nacional de Administração Pública (Enap) em parceria com a Universidade Federal de Brasília (UnB). A missão do EvEx é melhorar a tomada de decisão do setor público. Para isso, a equipe sintetiza, produz e dissemina evidências que possam servir de base para o desenho, monitoramento e avaliação de políticas públicas.

1.4. Estrutura da dissertação

Além dessa introdução, essa dissertação apresenta um capítulo de fundamentação teórica que trata do conceito de inteligência artificial, bem como outros aspectos importantes que subsidiam a compreensão da temática como *machine learning* e *deep learning*, baseados em dados, que recebem destaque e relevância nesse novo cenário tecnológico. Ainda nesse capítulo, são apresentadas algumas discussões éticas no cenário mundial.

Na sequência, é apresentada o capítulo que descreve a metodologia empregada nesse estudo. Posteriormente, o capítulo 4 apresenta a análise documental realizada e os principais aspectos encontrados. O capítulo 5 evidencia o desenvolvimento da análise e discussões dos resultados e, por fim, são apresentadas as considerações finais, conforme Figura 1:

Figura 1 - Estrutura da Pesquisa



A partir da estrutura proposta, espera-se desenvolver as etapas necessárias para concluir a dissertação e validar sua contribuição acadêmica para a área de pesquisa.

2. Referencial Teórico

O referencial teórico é um alicerce essencial para a pesquisa, fornecendo a estrutura conceitual necessária para orientar a investigação, interpretar os resultados e contribuir para o avanço do conhecimento em uma área específica, estabelecendo uma conexão entre o trabalho realizado e o corpo de conhecimento já existente.

Nesse capítulo, serão abordados temas como o novo contexto tecnológico; o conceito de inteligência artificial; *Machine Learning* e *Deep Learning*, histórico e evolução da IA; ética na IA e a era dos dados.

2.1 Novo contexto tecnológico – A IA na sociedade

Ao longo dos séculos, as transformações sociais têm sido impulsionadas por avanços tecnológicos, os indivíduos buscam constantemente novos processos, resultando em diferentes formas de trabalhar, produzir, comprar, se relacionar e se comunicar.

Nos últimos 20 anos, as tecnologias de informação e comunicação têm crescido intensamente, promovendo mudanças profundas na maneira como as pessoas se comunicam, captam e compartilham informações, e emitem opiniões (Szwarcwald *et al.*, 2021). Essa evolução tecnológica, conforme discutido por Ramos (2005), não apenas reflete uma democratização da comunicação, mas também assume uma dimensão política importante, influenciando a organização da sociedade e a dinâmica democrática.

A reivindicação da democratização da comunicação tem diferentes conotações, muitas além das que se costuma acreditar. Compreende evidentemente o fornecimento de meios mais numerosos e variados a maior número de pessoas, mas não se pode reduzir simplesmente alguns aspectos quantitativos a um suplemento de material. Implica acesso do público aos meios de comunicação existentes, mas este acesso é apenas um dos aspectos da democratização. (Ramos, 2005).

Uma grande proliferação de meios digitais de comunicação como *sites*, *chats*, redes sociais, vem tornando o mundo uma enorme rede de pessoas conectadas, independentemente de suas localizações geográficas. Através dessa possibilidade, sites de relacionamentos, redes sociais digitais, surgiram, nesse período técnico-científico-informacional, usadas por grande parte das pessoas que se conectam à internet.

As novas tecnologias permitiram flexibilizar a comunicação entre os indivíduos a partir do surgimento de diversos meios de comunicação cada vez mais interativos. A necessidade de interação social e a disseminação das redes sociais potencializaram a tendência da comunicação digital, que por sua vez altera completamente as possibilidades de comunicação (Branco; Matsuzaki, 2009).

À medida que as novas tecnologias permitem flexibilizar a comunicação entre os indivíduos, é necessário analisar como o uso de máquinas influencia nas mudanças sociais e políticas.

De acordo com Reis (2019), as tecnologias existentes em cada época, disponíveis para utilização por determinado grupo social, transformam radicalmente as formas de organização social, a comunicação, a cultura e a própria aprendizagem.

Para Melo e Camargo (2020), os processos em que grupos e comunidades estão envolvidos, como sujeitos de sua própria Comunicação - em qualquer lugar, com quaisquer grupos sociais, idades, gêneros, com ou sem tecnologias digitais - são a base do que se pode chamar comunicação como direito humano.

Nesse contexto, a expansão das formas de comunicação, bem como o amplo acesso às suas ferramentas – da alfabetização aos meios digitais – são formas de garantir um direito social e é também o ângulo a partir do qual o desenvolvimento futuro das comunicações terá que ser considerado para ser totalmente compreendido e socialmente aplicado.

Segundo Vermelho *et al.* (2014), as novas tecnologias permitiram a criação de meios de comunicação mais interativos, liberando os indivíduos das limitações de espaço e tempo, tornando a comunicação mais flexível. Os autores afirmam que os jovens nascidos após 1995 são “nativos” da cibercultura, inseridos num modelo de comunicação com equipamentos que operam por meio da convergência de mídias, em um processo em que as trocas simbólicas que viabilizam a comunicação de qualquer natureza são mediadas por signos estritamente vinculados à evolução tecnológica, denominadas redes.

No entanto, apesar da velocidade e da aparente variedade de informações que recebemos diariamente, é essencial reconhecer que o uso dessa tecnologia traz consigo implicações comerciais muitas vezes não divulgadas claramente aos seus usuários. Enquanto consideramos o computador e seus programas como ferramentas inanimadas e mecânicas, destinadas a satisfazer nossas necessidades profissionais, acadêmicas e sociais sem qualquer exigência em

troca, é importante refletir sobre o papel da comunicação como um direito humano fundamental.

A expansão do uso das redes e a integração da Inteligência Artificial em nossas rotinas nos leva a reavaliar como concebemos e exercemos esse direito, à medida que os computadores deixam de ser apenas ferramentas de uso e se tornam parte de sistemas complexos, capazes de influenciar nossa percepção e compreensão do mundo ao nosso redor

A evolução das formas de comunicação, desde a sociedade industrial até a sociedade da informação, é explorada por Ramos (2005), que ressalta o papel da comunicação como um direito social essencial para fortalecer a democracia. Nesse contexto, a expansão das formas de comunicação e o amplo acesso às suas ferramentas tornam-se instrumentos para garantir direitos sociais, abrindo espaço para debates políticos sobre democratização. Ramos (2005) ainda observa que a sociedade da informação e da comunicação é a nova forma de organização hegemônica do capitalismo, em oposição a uma sociedade industrial em declínio, onde demonstra expectativas de avanço crescente da democracia da igualdade em todo o mundo.

Ao discutir democratização dos meios de comunicação, surge a necessidade de entender o direito à comunicação. Esse debate ganha relevância com a introdução da Inteligência Artificial nas plataformas de comunicação digital, onde são redefinidos o papel das pessoas.

A capacidade da sociedade se tornar produtora ativa de conteúdo e não somente consumidora, representa uma mudança social importante na perspectiva nos meios de comunicação de massa. Sobre esse aspecto Gomes (2001), nos apresenta a importância de debater o direito à comunicação no contexto da Inteligência Artificial, tendo em vista que a comunicação pela internet se diferencia da comunicação pelos meios de massa, pois permite a personalização de informações, mostra também preocupações inclusive de cunho político pois tem potência de influenciar a autonomia dos indivíduos com o surgimento de bolhas de informação¹⁰.

¹⁰ Bolhas de Informação - Ambiente, especialmente online, em que as pessoas são expostas apenas a informações e opiniões que confirmam aquilo em que já acreditavam. A bolha informacional é um viés construído pelos algoritmos a partir de nossos hábitos e pesquisas na internet. Educamídia. Bolha informacional/bolha de informação. Disponível em: <https://educamidia.org.br/glossario>. Acesso: dezembro 2023.

Nesse processo, as decisões autônomas e flexíveis baseadas em dados concedidos em buscas e acessos, geram ações tomadas pelos sistemas, que impulsionam a dataficação¹¹ através de temas, discussões e produtos para determinadas bolhas de consumo.

As diversas plataformas de serviços disponíveis têm como base decisões algorítmicas, por meio de dados que filtram preferências e históricos de acesso, onde esse processo de coleta de dados, age como um mecanismo de gerenciamento para oferecimento de produtos ou serviços mais rentáveis ao modelo de negócio, dentro de uma mentalidade neoliberal de empreendimentos (Dardot; Laval, 2017).

A dataficação, ao coletar e armazenar massivamente dados sobre comportamentos, preferências e interações, possibilita o desenvolvimento de produtos e serviços altamente personalizados e direcionados. A análise desses dados impulsiona estratégias de negócios, permitindo que as empresas atendam de forma mais precisa às necessidades dos consumidores, maximizando, assim, os lucros.

Os diálogos são ativos nas redes sociais e ocorrem através de acessos, onde nesses disponibilizamos milhões de dados e informações aos computadores. Nesses diálogos, programas são alimentados com a opinião e preferência dos usuários e conseguem enviar informações de forma personalizada. Essa dinâmica, contudo, ocorre de maneira invisível para os usuários, já que os sistemas utilizam algoritmos de recomendação.

Os algoritmos têm a governança que se baseia em dados e toma as decisões automatizadas nas plataformas, sem conhecimento e consentimento do usuário.

Na comunicação online, cada usuário pode atuar como emissor ou receptor de informações a qualquer momento, sem necessidade de autorização de órgãos legisladores ou grandes investimentos. Todos os participantes têm a capacidade de se transformar, permitindo qualquer pessoa se torne provedora de informação, ao distribuir ou compartilhar conteúdo produzido por outros.

Como resultado da algoritmização¹² de sistemas de recomendação, como um tipo de programação em Inteligência Artificial, culmina na criação de bolhas de informação ou bolha

¹¹ Dataficação - o processo de transformação de informações em dados digitais, permitindo a coleta, armazenamento, análise e uso desses dados para fins específicos. Aspectos da experiência humana são transformados em dados. Silveira, Sérgio Amadeu *et al.* **Colonialismo de dados:** como opera a trincheira algorítmica na guerra neoliberal. 1. ed. São Paulo: Autonomia Literária, 2021. v. 1. 172 p.

¹² Algoritmização possibilita a conexão entre demandas e ofertas em um ambiente virtual, onde os algoritmos desempenham o papel de intermediar relações de consumo. SILVEIRA, Sérgio Amadeu *et al.*

de consumo, na interação social online. Essas bolhas têm o potencial de impactar as democracias em geral, inclusive negativamente. Petit (2018) detalha que as bolhas criam um universo de informações único, para cada um de nós, alterando de forma substancial nossa maneira de encontrar ideias e informações na internet.

Esse processo embora aparentemente transparente para o usuário, acaba por minar a independência dos indivíduos, tornando-os suscetíveis a comportamentos tendenciosos, distorções de entendimento da realidade, acentuando a polaridade política a partir de linhas de personalização, sistemas de recomendação e comportamentos persuasivos nas redes afetando de forma não mensurável a comunicação e a construção social e individual.

Para Ramos (2005) a consideração da comunicação como política pública no contexto do capitalismo é essencial. Este reconhecimento se fundamenta na perspectiva da ideologia liberal das sociedades de mercado, na qual a comunicação é vista como a principal garantidora e, mesmo, alavancadora da liberdade de mercado, por meio da teoria do livre fluxo da informação. Na mencionada teoria, qualquer intervenção do Estado nos meios de comunicação é automaticamente interpretada como censura, configurando-se, portanto, como uma ameaça aos direitos e à liberdade em sua totalidade.

Gomes (2001) reforça a representação da internet como um vasto acervo de informações, que abrange diversos temas e objetos, acessível exclusivamente online. Essas informações estão distribuídas em redes globais de computadores, organizadas de maneira a serem facilmente lidas, vistas, reproduzidas e distribuídas. Enquanto se assemelha a uma gigantesca biblioteca ou enciclopédia online, a internet destaca-se como um espaço para a coleta e compartilhamento de conhecimento, ciência, cultura e arte, mais do que como um meio de comunicação tradicional.

[...] ao lugar da informação e a sua circulação veloz e eficaz como capital social determinante num século em que a crise do Estado-Nação, a transformação do capital financeiro em dados voláteis e a informação contínua e total estão na agenda do dia. (Gomes, 2001).

A característica descentralizada da internet gera debates e divide opiniões sobre a questão da regulamentação, especialmente diante da ascensão da Inteligência Artificial, pois a

censura tornou-se um tema sensível e de manipulação partidárias. O tema divide opiniões quanto aos limites de uma regulamentação dessas tecnologias.

Gomes (2001) aborda de maneira aprofundada o dilema ético inerente a conexão entre a liberdade de expressão e a necessária proteção contra informações falsas. O autor destaca a complexidade do tema particularmente nos contextos que envolvem a liberdade de expressão, ressaltando a crescente necessidade de resguardar a sociedade contra a disseminação prejudicial de notícias falsas.

Esse dilema está intrinsecamente ligado à necessidade de regulamentação. No entanto, em comparação com a regulação da mídia, a regulamentação da Inteligência Artificial no Brasil avança rapidamente, evidenciando a preocupação da sociedade diante da complexidade desse tema. Esse avanço suscita questões políticas e éticas que permeiam a sociedade contemporânea.

Ainda Gomes (2001) ressalta que a liberdade de expressão é um valor moral legítimo eticamente somente quando o seu exercício não resulta em atos que contrariem a dignidade dos outros.

A liberdade de expressão, como princípio, não é suficiente para garantir a legitimidade moral de atos específicos. Atos de fala ofensivos, difamatórios e humilhantes, mesmo quando realizados com total liberdade, podem ser considerados imorais. No estado democrático a liberdade de expressão não se limita apenas à capacidade de falar, mas também implica a obrigação, inclusive legal, de assumir a responsabilidade pelo conteúdo expresso. A presença do Estado se faz necessária para assegurar a imparcialidade ou equilíbrio de opinião pública de modo que o direito ao acesso e a dignidade humana sejam garantidos.

A regulação da mídia no Brasil ainda é um tema que envolve muitas discussões sociais permeadas por interesses de grandes instituições, mas que precisa ser feito para que os meios de comunicação estejam a serviço do público, e não o contrário. (Samir, 2015)

No contexto brasileiro, a regulamentação da Inteligência Artificial avança rapidamente em comparação com a regulação da mídia. Essa disparidade na urgência reflete, complexidade intrínseca ao tema, a sua inegável importância. A abundância de informações disponíveis, aliada a velocidade das transformações tecnológicas, gera preocupações substanciais, que demandam estudos interdisciplinares aprofundados para compreender o impacto futuro dessas tecnologias na sociedade e sua condição humana.

Para Castells (2009), as redes informacionais são estruturas complexas de comunicação construídas em torno de um conjunto de nós interconectados que garantem a unidade de propósito e a flexibilidade de execução devido sua capacidade de se adaptar à evolução do ambiente que a as circundam.

A Inteligência Artificial de hoje se configura como um agrupamento dessas tecnologias capaz de gerar outras tecnologias, novas metodologias e aplicações e, por isso mesmo, suas características são de natureza distinta da de outras inovações que chegam ao mercado. Seu impacto no crescimento da economia e na melhoria da vida social é potencialmente maior do que o impacto de outras tecnologias (Arbix, 2020).

Apesar das inúmeras vantagens da Inteligência Artificial nas sociedades contemporâneas, as preocupações éticas procuram balizar a Inteligência Artificial e garantir que seu curso esteja sempre voltado para melhorar a sociedade, e não para exacerbar seus desequilíbrios, preconceitos, desigualdades ou até mesmo corroer sua democracia (Eubanks, 2018).

Aplicações que identificam o perfil de pessoas para direcionar a publicidade (serviço oferecido on-line e também em campanhas políticas, como revelado no caso da *Cambridge Analytic*), são exemplos da capacidade da IA de capturar as preferências e características de usuários e, assim, moldar seus objetivos, influenciar seu comportamento e debilitar sua autodeterminação (Taddeo; Floridi, 2018a).

As formas tendenciosas de *designs* minimamente projetados para prender a atenção e a influência invisível da inteligência Artificial em nosso cotidiano, pode ameaçar nossa capacidade de determinar nossa própria vida a partir de nossas escolhas e identidade social (Taddeo; Floridi, 2018a).

Quando os algoritmos são direcionados para personalização individual ou de grupos específicos, pode ocorrer um distanciamento das informações, uma vez que não existem filtros editoriais nesse tipo de mídia. Isso pode resultar em uma redução na exposição do indivíduo a uma variedade diversificada de informações e em uma possível perda de perspectiva em relação a questões coletivas. Essa dinâmica ocorre devido ao fato de os programas de computador fornecerem conteúdo com base no comportamento dos usuários durante seus acessos, o que acaba limitando o acesso dos cidadãos a uma variedade ampla e não restrita de informações.

Como resultado, as pessoas são expostas a um recorte específico da informação, o que pode comprometer a compreensão de acontecimentos. A personalização permite lacunas para

as relações imaginárias dos indivíduos, representadas por determinadas ideologias e suas condições reais de existência (Althusser, 1971).

Os algoritmos muitas vezes proporcionam aos usuários uma sensação falsa de liberdade de escolha e de veracidade em relação a eventos que a tecnologia não consegue discernir como certos ou errados. Isso ocorre porque, para os processos computacionais, todos os dados são tratados como verdadeiros.

Segundo Mittelstadt *et al.* (2016), a tomada de decisão algorítmica tem potencial de prejudicar indivíduos e grupos, aumentando a necessidade de transparência nos processos algorítmicos.

Segundo Harari (2015) a capacidade de manipulação do indivíduo pelo aparato tecnológico é motivo de preocupação social. Harari (2018) discute a capacidade de manipulação dos indivíduos por meio da tecnologia sendo esta considerada um dos grandes problemas contemporâneos, algo impossível de se fazer por governos autoritários no passado, devido à inexistência de recursos tecnológicos tão avançados.

Da Costa Felipe e Mulholland (2022) destacam que, pela primeira vez na história, surge a possibilidade de um terceiro ou o próprio governo conhecer uma pessoa melhor do que ela mesma. Esse avanço implica na capacidade de manipular indivíduos e, em última instância, tomar decisões por eles, ainda que de maneira sutil. A coleta em larga escala de informações por meio das tecnologias pode ter repercussões diretas nas estratégias de campanhas políticas, impactando, sem dúvida, o conceito de participação democrática.

Em uma época de alta circulação de dados, há o impulsionamento de modelos de diversos negócios e a transformação de processos sociais em dados, na qual ocorre a mudança das relações sociais e essa permite diversas leituras de momentos e informações.

O advento das tecnologias de inteligência artificial no cotidiano e outras mudanças ocorridas a partir da expansão da tecnologia impactam o comportamento e o conceito social da humanidade. Em um passado recente, as relações eram construídas por meio do fortalecimento das comunidades e coletividade gregária, hoje relações são perpassadas pelo sistema de redes, que conectam e desconectam as pessoas entre si e independem da materialidade local (Bauman, 2001). Em outras palavras, a digitalização enfraquece os laços comunitários, uma vez que introduz um processo desencarnado. A comunicação digital é uma comunicação descorporada (Han, 2021).

De maneira geral, essa conexão gera relações superficiais, não porque não podem ser profundas, mas especialmente, porque podem ser facilmente rompidas, trazendo à tona o individualismo.

Além disso Dunker (2019) alerta quanto a possibilidade de controle e a impressão de poder manipular todo ambiente digital que a inteligência artificial nos traz, porém essa espécie de ilusão traz consequências patológicas, pois o valor do indivíduo é distorcido em relação ao tamanho do mundo real, nesses meios.

Sem avanços no campo da ética, capazes de iluminar os procedimentos dos algoritmos, a IA, em suas diferentes modalidades, poderá sofrer um processo de desgaste e corrosão da confiança das sociedades (Arbix, 2020).

A ascensão da Inteligência Artificial representa uma mudança histórica na trajetória da humanidade, marcando o início de uma nova era. No entanto, a falta de preparo adequado por parte dos países pode agravar desigualdades sociais e acelerar divisões. Estima-se que a história dos empregos em todo o mundo será significativamente alterada pela ascensão da inteligência artificial e da robótica. (Harari, 2018).

Para Harari (2018), o desenvolvimento contínuo dessas tecnologias exigirá dos profissionais do futuro uma habilidade decisiva: a capacidade de se adaptar rapidamente a novas funções de trabalho. A criação constante de novos empregos e o processo de requalificação das pessoas que irão desempenhá-los tornar-se-ão práticas recorrentes, destacando a importância da flexibilidade e do aprendizado contínuo.

Grohmann (2020) vem explicar que as lógicas algorítmicas passaram progressivamente, nos últimos trinta anos, a governar mais dimensões da vida social e comunicacional a partir dos “algoritmos de relevância pública”. Dessa forma, as relações com trânsito, alimentação, transporte, amor e trabalho passam também por mediações algorítmicas.

A partir dessa implicação, surge uma falsa simetria, que menospreza o papel e a responsabilidade do “Império da Nuvem¹³” na classificação e automatização dos processos sociais, que são tomados, por sua vez, como um imperativo não só tecnológico, mas também financeiro e político.

¹³ Império da Nuvem – se refere ao uso predominante de serviços em nuvem e tecnologias associadas, onde a nuvem é um artifício tecnológico de computação, que permite armazenamento, processamento, entre outros serviços a serem executados através da internet em servidores remotos. Silveira, Sérgio Amadeu *et al.* Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal. 1. ed. São Paulo: Autonomia Literária, 2021

Diante desse cenário complexo, é relevante abordar os principais pontos éticos relacionados à segurança de dados e à criação de uma legislação destinada a proteger os usuários de problemas cujo impacto eles, inicialmente, podem não conseguir dimensionar ou compreender.

2.2 Conceituação de inteligência artificial

A inteligência artificial (IA) teve origem como um ramo da ciência da computação, com um marco inicial na Conferência de Dartmouth realizada em 1956. Foi nesse evento que o termo "Inteligência Artificial" foi oficialmente introduzido, juntamente com definições fundamentais que propunham a possibilidade de máquinas reproduzirem aprendizado e inteligência humana. Durante essa conferência, foi definido um princípio que dizia que "cada aspecto de aprendizado ou outra forma de inteligência pode ser descrita de forma tão precisa que uma máquina pode ser criada para simular isso".

As primeiras referências à inteligência artificial são dadas ao matemático e cientista da computação Alan Turing, em seu artigo "*Computing Machinery and Intelligence*" (1950), onde ele afirma que podemos esperar que as máquinas vão competir com os seres humanos na área da inteligência. (Kaufman, 2019).

A partir desse ponto, investimentos de diferentes setores, tanto privados quanto governamentais, foram alocados para impulsionar o desenvolvimento da IA. Esses investimentos foram responsáveis pela criação de algoritmos inovadores, incluindo programas de aprendizagem e redes neurais, que permitiram às máquinas começarem a adquirir habilidades de aprendizado autônomo.

A autonomia do aprendizado de máquina é um tema amplamente debatido na atualidade. A falta de regulamentação clara nas legislações vigentes é evidente, e contribui para questões como a discriminação algorítmica ou viés algorítmico. Este fenômeno ocorre quando decisões automatizadas resultam em tratamento injusto devido ao processamento automatizado de dados (Goodman, 2016). Compreender o impacto global da Inteligência Artificial é fundamental, indo além dos benefícios para a sociedade, mas também sua interferência nas escolhas e preferências do ser humano.

os algoritmos de aprendizagem de máquina herdaram padrões sociais refletidos em seus dados de treinamento, sem qualquer intenção original de programadores para incluir tais tendências. Os cientistas da computação chamam isto de viés algorítmico (Cappra, 2021, p. 13).

Para entender a profundidade do tema, precisamos inicialmente definir conceitos importantes quanto a esse ramo da ciência da computação. A inteligência artificial possui camadas onde o foco principal delas é criação de softwares que imitam comportamentos e capacidades humanas.

Para Russell e Norvig (2013) devemos considerar o desempenho humano e a racionalidade, a Inteligência Artificial busca realizar tarefas de forma racional, e as decisões e execução de ações sejam consideradas corretas e adequadas, a partir de informações e conhecimento disponíveis previamente, imitando o desempenho humano.

O aprendizado de máquina se trata da capacidade de aprendizagem das máquinas com a experiência, onde permite a realização de tarefas com melhor performance a partir da experiência (Mitchell, 1997).

Para Mitchell (1997), as redes neurais têm inspiração do cérebro humano para processamento da aprendizagem adquirida em diversas funções na tecnologia, onde armazena o conhecimento experimental e torna disponível para uso.

A utilização de redes neurais, conhecidas como *Deep Learning*, revolucionou a capacidade de reconhecimento de imagens e processamento de grandes volumes de dados. Com isso, tem sido possível avançar em áreas como diagnóstico por imagem, reconhecimento facial e até mesmo na criação de obras de arte por meio de algoritmos.

2.2.1 Machine Learning e Deep Learning

No contexto das tecnologias atuais, o emprego dos termos IA (Inteligência Artificial), Aprendizado de Máquina (*Machine Learning*) e Aprendizado Profundo (*Deep Learning*) gera confusão em relação à sua aplicabilidade, enfatizando a importância de investigar de forma comparativa esses conceitos.

Diante da complexidade terminológica e da crescente integração dessas abordagens, torna-se primordial desvendar e examinar como se interconectam para impulsionar os avanços tecnológicos e moldar o panorama da inovação.

Apesar de terem suas origens nas pesquisas realizadas na década de 1960, tanto o *Machine Learning* quanto o *Deep Learning* passaram por mudanças significativas ao longo dos anos, resultando em uma distinção mais acentuada entre esses dois modelos (Russell, Norvig, 2021).

Ao analisarmos o paradigma do *Machine Learning*, entendemos ser um tipo de IA que aprimora a habilidade de um computador em aprender a partir do conhecimento acumulado e evoluir a partir dos dados assimilados. Dentro desse contexto, elementos fundamentais incluem análise estatística e preditiva, que capacita a identificação de padrões e a extração de *insights* cruciais para embasar decisões estratégicas em contextos empresariais (Russell, Norvig, 2021).

O *Deep Learning* é um ramo avançado da IA, moldada com base na maneira pela qual o cérebro humano assimila conhecimento. Dessa forma, se o cérebro humano é capaz de identificar um objeto em questão de milissegundos, o *Deep Learning* tem a capacidade de imitar esse processo com uma velocidade e precisão comparáveis. Kaufman (2019) acrescenta que o *Deep Learning* é capaz de transformar um vasto volume de dados em informação útil.

A ascensão das redes neurais profundas, tem revolucionado o reconhecimento de imagens e o processamento de grandes volumes de dados, possibilitando avanços em áreas como diagnóstico por imagem, reconhecimento facial e até mesmo na produção artística através de algoritmos. Esse considerável interesse no domínio decorre, em grande parte, dos resultados notáveis obtidos em diversas aplicações ao longo do tempo.

2.3 Histórico e evolução

Desde suas origens, a inteligência artificial passou por uma evolução notável. Nas décadas de 60 e 90, marcos significativos foram alcançados com a criação de programas de reconhecimento de linguagem e a introdução de robôs autônomos, respectivamente.

Na década de 90, ocorreram avanços tecnológicos significativos que tiveram um impacto profundo em várias áreas. Um desses marcos foi o surgimento da internet comercial, que permitiu a conexão global de pessoas e informações de uma maneira sem precedentes. Além disso, o Japão desempenhou um papel cru ao introduzir a chamada quinta geração de computadores, marcando um salto na indústria tecnológica. Com isso, houve uma modernização abrangente da indústria, impulsionada pelo avanço da inteligência artificial.

Essa combinação de tecnologias acelerou tanto os setores de microcomputadores quanto de supercomputadores. Todo o desenvolvimento tecnológico resultou na criação do protótipo da Google. Esse motor de busca revolucionário foi construído com base nos princípios da inteligência artificial e se tornou uma ferramenta fundamental para a descoberta e organização de informações online. A ascensão do Google representou um passo significativo na interconexão global e na acessibilidade ao conhecimento. Nesse mesmo período, a IA demonstrou suas capacidades ao possibilitar que um computador da IBM superasse um mestre humano no xadrez (Teixeira, 2014).

O século XXI teve uma série de avanços marcantes no campo da inteligência artificial (IA), resultando em uma revolução na interação com a tecnologia. Para Teixeira (2014), a imitação do comportamento humano torna-se cada vez mais uma camuflagem sob a qual se abriga uma máquina totalmente diferente de nós humanos.

Surgiram assistentes virtuais como Siri, Alexa e Google Assistente, redefinindo a maneira como os seres humanos interagem com dispositivos tecnológicos. Foram introduzidos robôs autônomos, notavelmente os de limpeza, que combinaram sistemas de IA com sensores de posicionamento, representando um passo significativo em direção à automação inteligente.

Nesse período iniciou também um aprofundamento das investigações acerca de veículos autônomos, com o objetivo de redefinir completamente a mobilidade através da aplicação da inteligência artificial.

Em 2012, a Google demonstrou avanços significativos ao empregar o aprendizado profundo de redes neurais para o reconhecimento de imagem em seus sistemas de IA. Essa abordagem de *deep learning* permitiu um processamento mais sofisticado de informações, inclusive o reconhecimento de emoções humanas através de rostos.

Em 2014, o *chatbot* Eugene Goostman, ao convencer jurados de sua humanidade em uma conversa escrita, destacou a evolução das capacidades de comunicação das máquinas.

Em 2016, quando o algoritmo AlphaGo, desenvolvido pela DeepMind, uma empresa da Google, derrotou um campeão mundial em um jogo complexo e estratégico. Ao aprender e aprimorar suas habilidades por meio de observação e autojogo, o AlphaGo exemplificou a capacidade da IA em abordar problemas intrincados.

Esse progresso reflete não apenas a capacidade da IA de imitar habilidades humanas, mas também sua capacidade de superá-las em domínios específicos [...] Em seguida, num futuro próximo, tentaremos ultrapassar nosso próprio cérebro, através de nossa mistura com máquinas. Nas próximas décadas não haverá uma linha divisória nítida entre robôs e humanos. Seremos nos mesmos os robôs que inventamos, o que permitirá a inteligência artificial superar a natural (Teixeira, 2014).

A IA continua a avançar, concentrando-se no processamento de linguagem natural, no reconhecimento de imagens por meio da visão computacional e na criação de assistentes virtuais cada vez mais sofisticados. Sua presença está profundamente integrada aos nossos dispositivos eletrônicos do dia a dia. No momento, vivemos a era dos algoritmos generativos conhecidos como GPT-3- *Generative Pre-Training Transformer3*, que têm a capacidade de criar ou gerar novas proposições a partir de um conjunto inicial de dados. Para Pasquinelli (2022), a linha de montagem do aprendizado de máquina é alimentada por dados, os quais alimentam os algoritmos, resultando na geração de novos modelos. Assim, partindo-se desse princípio, quanto mais dados forem disponibilizados, mais informações o algoritmo terá para aprender um contexto específico.

Para Kaufman (2019) a IA propicia a simbiose entre o humano e a máquina ao acoplar sistemas inteligentes artificiais ao corpo humano, e ao dispor interação entre o ser humano e a máquina como duas “espécies” distintas conectadas.

Para Arbix (2022) a flexibilidade humana é que permite a imitação das máquinas. A era que estamos não é pós-humana, estamos em uma dobra do tempo em que podemos construir artefatos que imitam a consciência.

Para Haraway (2000), podemos ser responsáveis pelas máquinas; elas não nos dominam ou nos ameaçam, nós somos responsáveis pelas fronteiras; nós somos essas fronteiras.

2.4. Ética na IA

A discussão ética nos processos algorítmicos ultrapassa a questão da sciência das aplicações de inteligência artificial no meio contemporâneo. O *Deep Learning*, caracterizado por aprendizados profundos em algoritmos de alto nível, que visa replicar a rede neural do cérebro humano, demonstra uma capacidade mais avançada das máquinas em realizar tarefas anteriormente exclusivas dos humanos, levantam discussões mundiais, as chamadas *Black Box*¹⁴ são praticamente inexplicáveis.

Como destacado por O'Neil (2016, p. 11), “algoritmos são opiniões embutidas no código”, é importante ter em mente que algoritmos são projetados por seres humanos e, portanto, carregam suas visões de mundo e valores implícitos. Essa compreensão é essencial ao discutir questões éticas relacionadas ao uso de algoritmos, uma vez que suas decisões podem afetar significativamente comunidades e indivíduos.

Também é fundamental considerar as possíveis consequências e impactos das decisões algorítmicas sobre aqueles que são influenciados pelo discurso, especialmente diante da construção de discursos opacos que podem estar sujeitos a vieses inerentes à natureza humana.

O uso da palavra Inteligência em seu nome, leva muitas pessoas a entendimentos equivocados sobre a verdadeira função das máquinas que utilizam tecnologia de Inteligência Artificial, a atividade atual dessas máquinas com a evolução dos computadores permitiu que desenvolvessem atividades somente executadas anteriormente pelos seres humanos.

Contudo, a inteligência está intrinsecamente vinculada ao raciocínio humano, como destaca Arbix (2022). Essas funções confundem a atividade reflexiva dos cientistas pois quando as máquinas assumem a capacidade de tomar decisões, estas são confrontadas com a realidade e a noção de intencionalidade, que está no cerne da racionalidade humana, mas não no das máquinas.

¹⁴ Black Box - é um termo usado para descrever sistemas de Inteligência Artificial, onde a entrada de dados e as atividades não são visíveis para os usuários, de difícil entendimento para seus desenvolvedores, mas também assegura que erros permaneçam nas aplicações durante muito tempo. Beworkhappy, Abrindo a Caixa Preta da Inteligência Artificial, criamos explicabilidade e transparência totais. Disponível em: <https://www.beworkhappy.com.br/pt/noblackbox>. Acesso em dezembro 2023.

Para Benasayag (2018), o termo inteligência para os processos de tecnologia de inteligência artificial é apenas uma metáfora, pois a máquina pode ultrapassar os humanos com relação à capacidade de calcular, mas ela é incapaz de atribuir qualquer significado a estes cálculos. A inteligência humana é um processo que articula afetividade e é articulada com todos os processos cerebrais e corporais, a máquina apenas processa cálculos, sendo o ser vivo o único responsável por criar significados.

O comportamento humano não é exato, mas o comportamento algorítmico é lógico, exato e segue prescrições técnicas que podem extrapolar o objetivo de seus desenvolvedores, para Savater (2014) a ética é a prática de refletir sobre o que vamos fazer e sobre os motivos pelos quais vamos fazer. Como garantiremos que os programas de aprendizagem de máquina seguirão essa premissa ética?

[...] o maior problema que a internet nos suscita é a veracidade. Dizer ou não dizer a verdade se converteu numa questão mais complicada do que no passado. Agora a partir de casa, podemos enviar mensagens falsas, errôneas e até mesmo nocivas, a conhecidos e desconhecidos. Essa possibilidade é nova e supõe um novo desafio moral, pois muitas vezes se trata de apertar ou não uma tecla; e isso ocorre em casa, sem ninguém nos vigiando, sem que possamos ser acusados e sem prestar contas a ninguém. Nosso poder aumentou e, se levarmos a vida a sério, nossa responsabilidade também aumenta (Savater, 2014, p. 26).

Na construção de sistemas de inteligência artificial, é fundamental dar mais atenção aos aspectos éticos e à influência que o meio social pode gerar durante a operação desses sistemas. Os sistemas éticos, são compostos por componentes que se não forem devidamente integrados correm o risco de gerar danos irreparáveis na sociedade. Quando pensamos em sistemas tecnológicos temos como pressuposto que a responsabilidade da ação desses programas nas organizações sociais é de seus criadores.

Existem diferentes maneiras de entender ética, pois o termo possui muitos significados. Para Chauí (2012) o termo é usado para se referir a costumes considerados valores e obrigações para a conduta dos membros de um grupo ou de uma sociedade.

Freitas (2014), defende a ideia de que a ética não deve ser encarada apenas como uma prescrição, mas sim como um processo construído, negociado e (re)significado por diversas vozes, não significa negligenciar a ética em pesquisas, mas sim focar na competência ética de todos os envolvidos no estabelecimento de valores e normas de conduta. Essa perspectiva ética não implica renunciar ao rigor científico da pesquisa.

Valls (1994) situa a ética como um conceito universalmente reconhecido, embora sua definição prove ser desafiadora quando solicitada uma explicação precisa. Discutir ética significa falar sobre liberdade, mesmo que remeta a discussão sobre normas e responsabilidade. Por sua vez a moral, muitas vezes considerada sinônimo de ética, e destaca o aspecto de interiorização dessas normas. A moral representa a racionalidade do sujeito.

Em relação a discussão sobre os a inteligência artificial, precisamos observar que os algoritmos desenvolvem procedimentos até o momento, exclusivamente humanos, as aplicações de Inteligência Artificial podem cometer erros por não conter um código moral e ético, e justamente por não fazer assimilações, essas aplicações podem reproduzir resultados imorais ou antiéticos sem a percepção e identificação clara da sociedade.

Nos ambientes virtuais, os indivíduos são imersos em um universo amplamente lúdico, que lhes confere a liberdade de fazer escolhas variadas. Contudo, essas decisões muitas vezes desencadeiam uma série de ações com impactos sociais significativos, frequentemente resultantes em sofrimento. Por exemplo, as inúmeras discussões em redes sociais, muitas vezes fundamentadas em uma interpretação distorcida da liberdade de expressão, levam ofensas e disseminação de notícias falsas, culminando em casos de suicídio. De acordo com Galvão (2023), a falta de transparência gera danos à sociedade, muitas vezes de maneira irreversível.

No digital, temos uma realidade que pode ser alterada apenas por desejos e intenções, quando os algoritmos tomam decisões sem visibilidade dos resultados, esses são desconhecidos e imagináveis. Para Homem (2022) os indivíduos desenvolvem estratégias para estabelecer sua presença, garantir proteção, facilitar a comunicação, intermediar interações e adaptar-se ao ambiente, sendo a tecnologia a técnica contemporânea empregada para tal. Porém a dinâmica do vínculo social e o domínio do imaginário sofrem alterações nesse contexto tecnológico. Afirma ainda que a interação com a tecnologia transita desde as esferas mais íntimas até as mais públicas, visando atribuir significado à experiência vivenciada.

Na busca de atribuir significados, corremos o risco de confiar nas máquinas por entender como aplicações inteligentes, porém essas muitas vezes nos dão respostas por vezes equivocadas. Os sistemas operam através de armazenamento de informações e estatística, porém, com previsibilidade mínima. Enquanto os algoritmos podem analisar e processar dados de forma eficiente, a compreensão verdadeira e empática das nuances humanas permanece exclusiva dos seres humanos.

Somente através dessa compreensão profunda da condição humana é possível obter *insights* significativos em relação aos desafios apresentados. No entanto, é importante considerar que confiar unicamente em algoritmos para tomar decisões éticas pode ser problemático, especialmente quando esses algoritmos estão nas mãos de entidades capitalistas. A busca por lucro pode levar a uma contradição ética, o que destaca a necessidade de uma criticidade das implicações da inteligência artificial em nossa sociedade.

Na realidade, sistemas autônomos operam eticamente com reduzida previsibilidade, seja porque não foram concebidos ou não são adequados para envolver representações de raciocínio moral, seja porque os valores desses sistemas não foram devidamente sintonizados com os padrões éticos que regem as respectivas sociedades. Por isso, mostra-se inconsistente a visão de vários pesquisadores de que o aumento da autonomia dos sistemas isentaria os designers da sua responsabilidade. O movimento, de fato, ocorre no sentido inverso, pois, quanto maior a autonomia dos algoritmos, maior será a responsabilidade dos seus criadores. É o que toda sociedade espera para consolidar sua confiança nessas tecnologias (Arbix, 2020).

Para mitigar esses equívocos na sociedade, é imperativo aumentar a transparência das decisões algorítmicas nas aplicações de inteligência artificial. Nesse sentido, os processos devem ser tornados explicáveis e interpretáveis, sendo as chamadas *White box*¹⁵, apresentadas como uma ferramenta construída para diminuir a opacidade das decisões algorítmicas. Entretanto é essencial destacar a necessidade de verificação e validação humana se faz necessária para monitorar e verificar resultados. Isso reforça a urgência de estabelecer um código ético específico para essa transformação tecnológica em curso na sociedade.

A inteligência artificial (IA) não é apenas uma nova tecnologia que requer regulamentação. É uma força poderosa que está remodelando as práticas diárias, as interações pessoais e profissionais e os ambientes. Para o bem-estar da humanidade é crucial que este poder seja usado como uma força do bem. A ética desempenha um papel fundamental nesse processo, garantindo que as regulamentações de IA aproveitem seu potencial e mitiguem seus riscos (Taddeo; Floridi, 2018b).

A necessidade de transparência nos algoritmos é importante para garantir que as decisões tomadas sejam compreensíveis e justificáveis, pois a responsabilidade deve ser assumida pelas organizações de quaisquer decisões que seus sistemas tomem e garantir que haja

¹⁵ White box – são modelos de machine learning que fornecem resultados compreensíveis e explicáveis por especialistas da área. NUNES, Dierle José Coelho; ANDRADE, Otávio Morato. O uso da inteligência artificial explicável enquanto ferramenta para compreender decisões automatizadas: possível caminho para aumentar a legitimidade e confiabilidade dos modelos algorítmicos? Disponível em: <https://periodicos.ufsm.br/revistadireito/article/download/69329/61096/379227>. Acesso em: dezembro 2023.

uma maneira de responsabilizar as pessoas por quaisquer danos causados por decisões inadequadas.

Dada a vasta amplitude da ética como um tema, surge na sociedade informacional a necessidade de definir estruturas éticas específicas que permeiam esse contexto, de dados. A ética de dados, também conhecida como Data Ethics conforme denominada por Floridi e Taddeo (2016), fundamenta-se nos princípios éticos estabelecidos pela ética no uso da informática e da informação. Ela refina e aprimora a abordagem endossada até então neste campo de investigação, promovendo uma mudança no nível de abstração das investigações éticas, passando de uma ênfase centrada na informação para uma perspectiva centrada em dados.

Para Bynum (2015) O termo *Data Ethics* é utilizado para aplicar teorias éticas ocidentais tradicionais, é empregado para descrever uma ética profissional na qual os profissionais de tecnologia aplicam códigos éticos e normas de boas práticas em sua profissão. Além disso, expressões como "ciberética", "ética da computação" e "ética na Internet" referem-se a questões éticas na área de informática relacionadas à internet.

A ética de dados aborda questões relacionadas à coleta e análise de grandes conjuntos de dados. Definida como o ramo da ética que estuda problemas morais relacionados a dados, algoritmos e práticas, buscando soluções.

Os desafios éticos são mapeados nos eixos da ética dos dados, ética dos algoritmos e ética das práticas, destacando a importância de abordar questões éticas antes de referir-se à informação, além da responsabilização moral de designers e cientistas de dados quanto a consequências indesejadas ou oportunidades perdidas. Mecanismos como esse, podem atuar como construtor de uma nova visão do direito, permitindo uma coesão sistemática na interdisciplinaridade estrutural social. (Bigaton, Teixeira, Bilhalva, 2022).

Para Galvão (2023), os ambientes digitais são meros meios de expressão de opiniões, muitas vezes distorcidas, funcionando como um megafone que amplifica e reflete as tendências da sociedade. Por trás de todos os desafios do mundo digital, há seres humanos que demonstram falta de maturidade, humanismo e valores, não evoluindo ao mesmo ritmo da tecnologia que eles próprios geraram. (Galvão, 2023).

Savater (2014) destaca a educação como o único meio de revolução pacífica disponível. Priorizar a educação digital torna-se crucial, pois somente por meio do conhecimento a sociedade pode proteger seus indivíduos de serem totalmente controlados por direcionamentos

externos à sua própria realidade. O entendimento dos conceitos de realidade e virtualidade¹⁶ é essencial para que os meios existentes não controlem as escolhas individuais.

A evolução da Inteligência Artificial destaca a necessidade de uma distribuição equitativa de educação digital é um processo de atualização contínua, que transcende a questão de renda ou classe social. No cenário brasileiro, exige o desenvolvimento de políticas públicas efetivas voltadas para a inclusão de acesso à internet, uma vez que a acessibilidade ainda é um desafio. A prevenção de desigualdades emerge como urgentes, e a tecnologia de inteligência artificial apresenta um papel significativo nesse processo. Seu potencial de identificar áreas de disparidades sociais e econômicas permite a adoção de medidas preventivas eficientes por parte dos governos.

Para Eubanks (2018), é importante entender que a Inteligência Artificial pode ser uma ferramenta que protege a sociedade de distanciamento de classes, mas também que pode perpetuar as desigualdades existentes.

Para reduzir o distanciamento entre classes sociais ocasionado pela tecnologia, é imprescindível que ocorra a democratização dessa ferramenta, tornando-a mais acessível e de fácil utilização, de modo a permitir que pessoas de diferentes origens e habilidades tenham acesso às mesmas oportunidades. Nesse contexto, é fundamental reduzir a disparidade existente entre indivíduos com habilidades técnicas e aqueles desprovidos delas.

Para Garcia (2020), a necessária conscientização da sociedade para essa nova realidade em que vivemos, é de suma importância, uma vez os dados não são neutros, e registram decisões humanas tomadas em algum contexto.

A tecnologia deveria ser capaz de detectar e corrigir potenciais desigualdades, independentemente do viés em que possam surgir. No entanto entendemos que alcançar a neutralidade é uma meta desafiadora na área da tecnologia, visto que as ferramentas desenvolvidas refletem a visão de mundo de seus desenvolvedores. A busca pela neutralidade é importante para evitar qualquer forma de discriminação, sendo este um tema de discussão urgente.

faz-se indispensável uma abordagem multidisciplinar, que inclua especialistas em Ética, cientistas sociais e especialistas que entendam as nuances de cada área de aplicação de Inteligência Artificial. (Garcia, 2020).

¹⁶ Virtualidade são novas formas de relação mediada pelas tecnologias digitais que eliminam a necessidade de proximidade física entre os envolvidos. LEVY, Pierre. O que é o virtual?. São Paulo: Ed. 34, 1996, p. 17-18.

Loren Spinola (2023) ressalta que os principais organismos internacionais reconhecem a importância da multidisciplinaridade para o desenvolvimento de IA responsável e emitem recomendações nesse sentido. No entanto, no Brasil apesar dos esforços colaborativos em andamento nas iniciativas, parece haver um desalinhamento com outros países. Nota-se que a proposta de regulação da IA foi elaborada por um grupo composto exclusivamente por juristas, que, possuem uma visão "limitada" da tecnologia devido à ausência de conhecimento técnico sobre seu funcionamento e desenvolvimento.

A regulamentação e o uso da inteligência artificial devem respeitar e beneficiar os indivíduos e a sociedade como um todo, incluindo a diversidade de perspectivas, origens e experiências, sendo esse um desafio mundial pois os princípios éticos variam de acordo com o contexto social da localidade do indivíduo. O monitoramento se faz necessário para garantir que a tecnologia seja usada de maneira ética e responsável para garantir a justiça e a imparcialidade nos processos.

Um ambiente sem regulação, onde se pode fazer toda espécie de coisas que são negativas para o bom funcionamento da sociedade que, não obstante, por leniência política, e para não ofender os piratas, ficam totalmente impunes. Pois a impunidade é corrupção (Savater, 2014).

A sociedade informacional surge a partir de processos de inovação que formam complexas relações e necessitam do acompanhamento e amadurecimento do processo de formação regulatória, sem desconsiderar diferenças culturais, sociais e econômicas, inerentes a este novo modelo social (Jorge *et al.*, 2020).

Para Loren Spinola (2023) embora seja compreensível que exista um período para o aprimoramento do processo de elaboração de regulamentações, sendo esta uma prática padrão, existe uma preocupação significativa em relação à possibilidade de que, devido à rápida evolução tecnológica, essas normas possam se tornar obsoletas logo após sua implementação.

Com o avanço tecnológico e o desenvolvimento acelerado da inteligência artificial, é essencial que a sociedade esteja ciente da necessidade de uma intervenção humana adequada no uso dessa tecnologia. A sociedade apresenta ambientes altamente complexos que exige uma formação regulatória capaz de garantir a ética e a responsabilidade no uso da inteligência artificial.

Isso significa que é fundamental que os reguladores considerem as diferenças culturais, sociais e econômicas inerentes a esse novo modelo social, juntamente com a participação da sociedade, a fim de garantir que as intervenções sejam justas e adequadas para todos os indivíduos e comunidades envolvidos. Além disso, é importante que essas intervenções não restrinjam desnecessariamente a inovação e a evolução da tecnologia.

A intervenção humana apropriada no uso da Inteligência Artificial é essencial para garantir que essa tecnologia seja usada de maneira responsável e benéfica para a sociedade em geral. A inteligência artificial pode trazer inúmeros benefícios em diferentes áreas, desde a saúde e a educação até a indústria e a economia. No entanto, esses benefícios só serão alcançados se houver uma intervenção humana consciente e responsável no desenvolvimento de seu uso.

2.5. A era dos dados

Diversos países discutem políticas de regulação para Inteligência Artificial, porém o consenso tem sido difícil de alcançar devido ao grande interesse em se tornar uma superpotência nesse campo. Esse interesse é impulsionado por uma série de fatores, incluindo motivações geopolíticas, benefícios econômicos, desenvolvimento tecnológico, preocupações com segurança nacional e o desejo de prestígio e influência internacional. Esses elementos combinados criam uma dinâmica complexa que torna desafiador chegar a acordos globais sobre a regulação da IA.

Os dados gerados por acessos a plataformas digitais são considerados a nova riqueza mundial, atualmente chamado de economia de dados, esse é um modelo de negócios baseado na oferta de interfaces gratuitas, onde a coleta e tratamento dos dados gerados, é considerado um novo produto. Surge a necessidade de entender e transformar esses dados em produto, dando origem assim a uma nova ciência: a ciência de dados, que se dedica ao estudo da extração de informações significativas para negócios. (Silveira, Avelino, 2021).

Essa tendência aqueceu ainda mais a revolução tecnológica, construindo paradigmas de uma economia digital e movimentando o mercado de dados, onde o destaque é para esse “novo petróleo mundial”.

O mercado de dados se tornou um dos principais mercados do capitalismo contemporâneo. É altamente lucrativo e tem gerado plataformas gigantescas que não param de coletar dados e colonizar o planeta, o que Couldry e Mejias qualificaram como a conversão dos fluxos de vida em dados (Silveira, 2021, p. 42).

No aprendizado de máquina os algoritmos conseguem prever o comportamento humano, podendo assim estar presente em muitas tarefas e processos sociais, agilizando e adiantando prazos, avaliando credibilidade, programas de bonificação de governos entre outras iniciativas e implantações, porém essas demandas ainda esbarram em análises extremamente problemáticas nas questões de liberdade civil.

Por vezes, a interação da Inteligência Artificial presente nos acessos da população mundial, em plataformas digitais e serviços gratuitos disponíveis na tecnologia, ocorre através de sobreposição aos conceitos éticos e legislativos vigentes do local em que aquele programa irá atuar, contudo não se pode negar os avanços em diversos setores sociais em inúmeros países.

Os aplicativos desenvolvidos em *Machine Learning*, um dos principais aspectos da Inteligência Artificial, impulsionam a economia de dados e são processados com base nas escolhas feitas por seres humanos, portanto, falíveis. Algumas dessas escolhas mesmo sendo feitas com as melhores intenções, eventualmente codificam preconceitos, mal-entendidos e geram grandes discussões sociotécnicas, já que os sistemas de software cada vez mais administravam nossas vidas (O'Neil, 2016).

Tendo em vista a interação de programas de Inteligência Artificial com os seres humanos em várias escalas e segmento sociais, é importante o reconhecimento dos limites dessa autonomia, refletindo sobre os efeitos desses sistemas computacionais inteligentes no cotidiano.

Sem o conhecimento e conscientização da população sobre as estruturas decisórias que os algoritmos provêm aos usuários, a autonomia social se põe em risco, pois comportamentos persuasivos são dificilmente identificados por quem usa os sistemas e se identifica com a rede. Vale destacar que questões de vulnerabilidade social, acesso à informação, renda, formação, dentre outros aspectos socioeconômicos e culturais, podem se acentuar diante das possibilidades virtuais e serem ainda mais críticos quando contextualizados em segmentos sociais com maior dificuldade de interpretação e maior influência das redes.

3. Metodologia

Essa pesquisa tem caráter exploratório que, segundo Gil (2019) têm como propósito proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito, considerando os mais variados aspectos relativos ao fato ou fenômeno estudado. As pesquisas exploratórias têm por objetivo familiarizar-se com o contexto real ou obter uma nova percepção dele e descobrir novas ideias (Cervo; Bervian; Da Silva, 2010).

A abordagem é prioritariamente qualitativa, que utiliza dados descritivos, como entrevistas, observações participantes, análise de documentos e registros audiovisuais, com o objetivo principal de aprofundar a compreensão dos fenômenos estudados, permitindo que os pesquisadores obtenham *insights* ricos e detalhados sobre a complexidade do comportamento humano, das interações sociais e dos contextos culturais. A abordagem qualitativa oferece três diferentes possibilidades de se realizar pesquisa: a pesquisa documental, o estudo de caso e a etnografia (Godoy, 1995).

Neste estudo, será realizada uma pesquisa documental, que é uma abordagem de coleta de dados em pesquisa científica que se concentra na análise e interpretação de documentos. O processo envolve a identificação e seleção de fontes relevantes, seguida de uma análise minuciosa dos conteúdos. O pesquisador examina os documentos em busca de informações, padrões, tendências e elementos que possam contribuir para a compreensão do tema em estudo.

Uma vez que as fontes relevantes foram selecionadas, o pesquisador procede com a leitura crítica e análise do conteúdo. Essa etapa inclui a identificação de ideias-chave, conceitos importantes, métodos de pesquisa utilizados, resultados obtidos e conclusões dos autores.

Segundo Bardin (2011), uma análise de conteúdo não deixa de ser uma análise de significados que se ocupa de uma descrição objetiva, sistemática e qualitativa do conteúdo extraído das comunicações e sua respectiva interpretação.

A autora relaciona análise do conteúdo e análise documental, pois, segundo ela, algumas técnicas e procedimentos da análise de conteúdo fazem menção à análise documental como forma de condensação das informações, para consulta e armazenamento. O Quadro 1 ilustra algumas de suas características.

Quadro 1- Características da Análise Documental e Análise de Conteúdo

Análise de Conteúdo	Análise Documental
Foca nas mensagens (comunicações); Geralmente é categorial-temática; Objetivo: manipulação de mensagens para confirmar os indicadores que permitam inferir sobre outra realidade que não a da mensagem.	Foca-se em documentos; Classificação – Indexação; Agrupamento de dados; Objetivo: representação condensada da informação para consulta e armazenagem.

Fonte: Adaptado de Bardin (2011)

A Análise de Conteúdo é classificada como uma opção válida quando o objetivo for analisar os dados provenientes das comunicações, buscando compreender os significados e os sentidos das mensagens, que vão além de uma leitura comum (Cardoso; Oliveira; Ghelli, 2021).

Como método de investigação, Moraes (1999) afirma que a Análise de Conteúdo pode ser considerada como um único instrumento, mas marcado por uma grande variedade de formas, justamente porque seu campo de aplicação, a comunicação, é muito vasto, sendo adaptável a muitos tipos de pesquisa.

Há várias maneiras para analisar conteúdos de materiais de pesquisa. Conforme Bardin (1977), as principais técnicas de Análise de Conteúdo são: Análise Categorial, Análise do Discurso, Análise de Avaliação, Análise de Enunciação, Análise de Expressão, Análise das Relações (Co-ocorrências e Estrutural).

Contudo, será a Análise Categorial a servir de base para descrever as principais fases da Análise de Conteúdo dos documentos pesquisados, já que no conjunto das técnicas da análise de conteúdo, Bardin (1977) afirma ser esta, a mais antiga e, na prática, a mais utilizada.

As diferentes fases da Análise de Conteúdo, segundo Bardin (1977), organizam-se em torno de três polos cronológicos: i) a pré-análise; ii) a análise do material; iii) o tratamento dos resultados, a inferência e a interpretação.

Nessa pesquisa, esses polos serão desenvolvidos de forma adaptada, conforme apresentado no Quadro 2.

Quadro 2- Desenvolvimento da Análise de Conteúdo

Fase proposta por Bardin	Objetivos	Prática dessa pesquisa
Pré-análise	a) a escolha dos documentos a serem submetidos à análise (<i>corpus</i>);	Foram considerados documentos representativos, homogêneos e pertinentes em sites oficiais dos governos dos países definidos no recorte
	b) a formulação das hipóteses e dos objetivos;	Objetivo: encontrar similaridades no conteúdo dos documentos
	c) a referência dos índices e a elaboração de indicadores	Escolher e definir as principais categorias a serem observadas nos documentos
Análise do material	a) unidade de registro;	Cada documento / regulamento será uma unidade de registro para análise
	b) unidade de contexto;	Será analisado o conteúdo de cada documento selecionado
	c) regras de enumeração;	Serão consideradas a presença ou ausência de determinados elementos em cada documento, para registrar nas categorias
	d) categorização	Será realizada a classificação e agrupamento de dados considerando a parte comum existente entre eles
Tratamento dos resultados obtidos e interpretação	a) sintetização	Na análise dos resultados, será produzido um texto síntese para cada uma das categorias, de modo a expressar o conjunto de significados presentes nos diversos documentos analisados

Fonte: Elaborado pela autora, 2023

Ressalta-se a importância de que o pesquisador adote uma abordagem crítica na análise dos documentos, considerando o contexto histórico e social em que foram produzidos, bem como possíveis vieses e perspectivas subjacentes. Além disso, a triangulação de fontes, ou seja, o uso de várias fontes diferentes, pode fortalecer a validade e a confiabilidade dos resultados obtidos.

Para a coleta de dados foram realizadas buscas em publicações de organizações mundiais cujo objetivo se dá através de cooperação e desenvolvimento social entre as nações sobre a criação de um código de ética que regulamente o uso da inteligência artificial.

Nesta pesquisa, foram examinados documentos divulgados pelas principais potências mundiais, incluindo países renomados por seu poder econômico e tecnológico, como: Estados Unidos, União Europeia e a China além do Brasil, país de relevância para o contexto da pesquisa, apesar de não ser considerado uma potência global.

3.1 Recorte do estudo

Na impossibilidade de esgotar o assunto e de avaliar a evolução da legislação em todas as partes do mundo, é proposta uma análise documental das regulamentações em andamento nos seguintes países:

Estados Unidos: **Lei de iniciativa de Inteligência Artificial de 2020** / A abordagem dos EUA para a regulamentação da IA tem sido mais setorial e dependente das agências reguladoras existentes. A *Federal Trade Commission* (FTC), por exemplo, tem autoridade para regular práticas de comércio desleais ou enganosas, que podem incluir o uso de IA. Ainda assim, não existe um marco regulatório de IA abrangente em nível nacional nos EUA.

União Europeia (UE): Na União Europeia, se encontra em processo final de implementação e propõe a regulação de uma ampla gama de aplicações de IA por meio de uma abordagem baseada em risco para a segurança do usuário, e sanções em caso de infringimento de regras. Em 2021, a Comissão Europeia propôs novas regras abrangentes sobre IA, conhecidas como o **Regulamento de Inteligência Artificial**. Essas regras visam criar um ambiente de confiança e segurança em torno da utilização da IA na UE, focando na transparência, no controle humano sobre a IA, na proteção da privacidade e dos dados pessoais, e na responsabilidade das partes envolvidas. Elas abrangem desde a proibição de sistemas de IA considerados perigosos até requisitos rigorosos para sistemas de IA de alto risco, passando por regras de transparência para sistemas de IA que interagem com as pessoas.

China: A China também tem se movimentado em direção à regulamentação da IA, principalmente através de suas políticas de segurança cibernética e diretrizes sobre a ética da IA. Em 2020, o Ministério da Ciência e Tecnologia da China divulgou princípios para a “nova geração” da IA incluindo princípios de governança e chamadas para a promoção da pesquisa em ética da IA, por meio do **Plano de Desenvolvimento de Inteligência Artificial de Nova Geração**.

Brasil: Sendo discutido o **Projeto de Lei (PL) 21/20 “Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil “**, **Projeto de Lei (PL) 2338/2023 “Estabelece normas gerais de caráter nacional para o**

desenvolvimento, implementação e uso responsável de sistemas de inteligência artificial (IA) no Brasil” e “Estratégia Brasileira de Inteligência Artificial”, que estabelecem normas gerais para o desenvolvimento, implementação e uso de sistemas de inteligência artificial (IA) no Brasil.

Os documentos têm como objetivo 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. Os projetos de Lei mostram as frentes de trabalho sendo a garantia de direitos para proteção de usuários impactados pela IA e disposição de ferramentas de governança e de arranjo institucional de fiscalização e supervisão para o desenvolvimento tecnológico. Sendo que o projeto de 2338/2023 vem com atualizações quanto ao uso da IA Generativa¹⁷, pois o documento de 2020 não possuía ainda referências sobre essa evolução tecnológica.

No documento Estratégia Brasileira de Inteligência Artificial disponível são encontradas ações desenvolvidas em prol do desenvolvimento, as vertentes das indicações para o uso da Inteligência Artificial e possíveis estratégias de crescimento para o País.

Considerando a relevância dessas discussões nas regulamentações sobre o futuro da IA, o Quadro 3 apresenta os documentos analisados:

Quadro 3 - Documentos Analisados

País	Título do documento	Site oficial (<i>link</i>)
Estados Unidos (EUA)	Lei de iniciativa de Inteligência Artificial de 2020	https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210
União Europeia (UE)	Regulamento de Inteligência Artificial	https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN#:~:text=A%20proposta%20apoia%20o%20objetivo,especificamente%20pelo%20Parlamento%20Europeu%204%20

¹⁷ IA Generativa - se refere ao uso de IA para criar novos conteúdos, como texto, imagens, música, áudio e vídeos. A IA generativa usa um modelo de machine learning para aprender os padrões e as relações em um conjunto de dados de conteúdo criado por humanos. Em seguida, ele usa os padrões aprendidos para gerar novo conteúdo. Google Cloud. Gere texto, imagens, códigos e muito mais com a IA do Google Cloud. Disponível em: <https://cloud.google.com/use-cases/generative-ai?hl=pt-BR>. Acesso em dezembro 2023.

China	Plano de Desenvolvimento de Inteligência Artificial de Nova Geração	https://www.gov.cn/zhengce/zhengceku/2017-07/20/content_5211996.htm
Brasil	Projeto de Lei (PL) 21/20 “Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil “	https://www25.senado.leg.br/web/atividade/materias/-/materia/151547
	Projeto de Lei (PL) 2338/2023 “Dispõe sobre o uso da Inteligência Artificial.”	https://www25.senado.leg.br/web/atividade/materias/-/materia/157233
	Estratégia Brasileira de Inteligência Artificial	https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-documento_referencia_4-979_2021.pdf

Fonte: Elaborado pela autora, 2023

4. Análise Documental

Esse capítulo apresenta a análise das documentações selecionadas metodologicamente. O preâmbulo faz um apanhado sobre o contexto da IA no mundo. Na sequência, apresenta como cada país do recorte vem tratando as discussões em IA. Por fim, apresenta um quadro com a síntese dos principais achados dessas análises.

4.1 IA no mundo

As práticas sobre utilização e aplicações de tecnologias de inteligência artificial são cada vez mais discutidas em diversos países ao redor do mundo. Considerando a criticidade do poder decisório dos usuários *versus* máquina, é necessário garantir que o processo de digitalização não impacte forma negativa a sociedade, sendo assim a discussão sobre a evolução da inteligência artificial permeia grandes poderes mundiais.

Em 2017 o presidente da Rússia, Vladimir Putin, expôs claramente o interesse e a importância estratégica da Inteligência Artificial discursando que A IA é o futuro, não apenas para a Rússia, mas para toda a humanidade e que quem se tornar líder nessa área, se tornará o dono do mundo.

O governo chinês reconheceu a criticidade da inteligência artificial para o futuro do país, e a importância da autonomia do desenvolvimento dessa tecnologia para que não dependa de empresas estrangeiras. Em 2017, o governo chinês anunciou ao mundo que até 2030 alcançaria os Estados Unidos e lideraria o mundo em várias questões e entre elas a tecnologia de Inteligência artificial, se coloca como de extrema importância.

Em 2021, na Conferência Geral da Unesco, onde a Comissão Mundial sobre Ética do Conhecimento Científico e Tecnologia (COMEST), preparou um documento intitulado de “Estudo preliminar sobre um possível instrumento de definição de padrões de ética para Inteligência Artificial”.

No cenário atual, os Estados Unidos ocupam uma posição dominante no desenvolvimento de Inteligência Artificial, liderando o desenvolvimento dessa tecnologia. No entanto, a China apresenta um notável crescimento e está se aproximando rapidamente da liderança. Ambos os países se destacam em importantes aspectos da IA incluindo pesquisa,

coleta de dados e investimento. Esse investimento robusto impulsiona suas respectivas indústrias e instituições de pesquisa, permitindo que avancem em ritmo acelerado na aplicação e desenvolvimento de soluções inovadoras em IA (Castro, Mclaughlin, Chivot, 2019).

Vale ressaltar que a China lidera no que diz respeito ao desenvolvimento de tecnologias e estratégias do governo. Além disso, detém uma posição de destaque na abundância excedente de dados, que é a principal matéria-prima para impulsionar a IA superando os Estados Unidos em volume absoluto como o principal país produtor de informações (Lee, 2020).

Paralelamente, a União Europeia destaca-se pela quantidade significativa de pesquisadores dedicados à área, o que possibilita avanços relevantes em pesquisas (Castro, Mclaughlin, Chivot, 2019).

Dessa forma, o domínio em IA é uma dinâmica em constante evolução, com diferentes regiões apresentando vantagens distintas. Atualmente, diversos elementos envolvendo a ética na aplicação das tecnologias de inteligência artificial estão em destaque globalmente, abrangendo áreas como regulamentação, privacidade de dados, transparência algorítmica, responsabilidade e governança.

Dentre as recomendações das organizações, duas importantes publicam recomendações de uso da Inteligência artificial:

A Organização das Nações Unidas para a Educação, a Ciência e a Cultura - UNESCO em sua Conferência Geral no ano de 2021, apresentou através da Comissão Mundial sobre Ética do Conhecimento Científico e Tecnologia (COMEST), um documento de “Estudo preliminar sobre um possível instrumento de definição de padrões a ética da inteligência artificial” (UNESCO, 2021).

A Organização para a Cooperação e Desenvolvimento Econômico – OCDE, é uma organização internacional que reúne as economias mais avançadas do mundo, bem como alguns países emergentes, que trabalha para construir uma vida melhor. Estabelecem padrões internacionais baseada em evidências para encontrar solução para os desafios da sociedade, por sua vez elaborou o documento "Princípios éticos para a inteligência artificial", onde apresenta cinco princípios fundamentais que devem ser seguidos na implementação de tecnologias de IA: transparência, diversidade, justiça, responsabilidade e privacidade (OCDE, 2019).

Analisando as publicações recentes de vários países, não há como negar a necessidade de uma regulamentação, porém é um debate extenso, necessário, demorado e desafiador para

que isso se transforme em regulamentação mundial. Qualquer tentativa de regulamentação exige obrigatoriamente um diálogo geopolítico nas diferentes vertentes e abordagens sociais.

4.1.1. Regulamentação Setorial, Regulamentação Transversal e Regulamentação

A abordagem para a regulamentação da Inteligência Artificial assume diversas formas, muitas vezes transitórias, em consonância com o amadurecimento das normas estabelecidas em países ao redor do mundo, destaca-se as complexidades e evoluções observadas nesse cenário dinâmico.

A forma como cada governo ou entidade reguladora desenvolve seu plano está alinhada com o enfoque adotado, o qual reflete a diversidade de estratégias para lidar com os desafios éticos, legais e sociais associados ao desenvolvimento e uso da Inteligência Artificial.

A Regulamentação setorial se concentra em setores ou indústrias específicas que utilizam a tecnologia de IA. Esse enfoque busca adaptar as regulamentações de acordo com características e necessidades particularidade de cada setor e seus contextos específicos de suas áreas. (Barbosa, Pinheiro, 2023).

A Regulamentação Transversal utiliza uma abordagem mais abrangente e horizontal, com diretrizes e normas aplicáveis a várias indústrias e setores. Essa abordagem estabelece princípios gerais que podem ser aplicados de maneira uniforme na sociedade, sendo mais genérica para diversas esferas das áreas existentes. (Barbosa, Pinheiro, 2023).

A Regulamentação Mista combina elementos da regulamentação setorial e da regulamentação transversal, reconhecendo a necessidade de regras específicas para determinados setores sociais, industriais e empresariais, ao mesmo tempo em que estabelece princípios gerais que beneficiam o país como um todo. Esse enfoque busca um equilíbrio entre a adaptabilidade às necessidades de setores específicos e a consistência ética em um contexto mais amplo. (Barbosa, Pinheiro, 2023).

4.2 Estados Unidos

Em 2019, o então presidente dos Estados Unidos, Donald J. Trump, em seu discurso sobre o Estado da União, publicado no site da Casa Branca, enfatizou claramente a importância do investimento do país em tecnologias de inteligência artificial e a liderança nacional nesse campo: “A liderança americana contínua em Inteligência Artificial é de suma importância para manter a segurança econômica e nacional dos Estados Unidos”. Na mesma ocasião, ele assinava uma Ordem Executiva de iniciativas, em que o governo federal concentrava recursos no desenvolvimento da IA, acelerando assim, a liderança nacional (White House, 2021).

De acordo com informações disponíveis no site do Governo dos Estados Unidos, foi realizada uma análise do documento intitulado “**Lei de iniciativa de Inteligência Artificial**” de 2020, que contém diretrizes éticas para sistemas de inteligência artificial no país. Este documento encontra-se disponível no site do governo americano.

A Casa Branca lançou várias iniciativas, com objetivo de garantir a liderança dos Estados Unidos nas pesquisas e desenvolvimento da IA, para a integração de sistemas de inteligência artificial em todos os setores da economia e da sociedade, garantindo sua governança ética e preparar a força de trabalho atual e futura dos EUA.

O documento analisado é resultado de iniciativas de diversas agências governamentais no País, o qual estabeleceu no país o comitê *National Artificial Intelligence Initiative* (NAII), para tratar de temas relacionados a IA (Roberts, 2021).

O documento contém 2.214 páginas, e trata-se de um relatório de conferência da lei de autorização de defesa nacional para o ano fiscal de 2021 americano. O estudo concentrou o assunto, no início do documento até a página 150, e a partir da página 1150, onde estão as tratativas sobre a inteligência artificial. Contém várias tratativas de diretrizes do uso da inteligência artificial, no início do documento encontramos sessões referentes a padrão do uso de Inteligência artificial para uso do Departamento de defesa e órgãos governamentais, onde mostra preocupação na coleta de dados e estabelecer linhas de comunicação eficazes para a ampla adoção e disseminação da inteligência artificial.

Contém uma descritiva das modificações nas atividades conjuntas de pesquisa, desenvolvimento e transição de inteligência artificial. Nessa descrição, são abordadas preocupações relacionadas ao avanço no desenvolvimento de tecnologias de inteligência

artificial, bem como à aplicação de soluções de inteligência artificial e aprendizado de máquina em problemas operacionais. (GPO, 2020).

Foram observadas informações referentes a atribuição de responsabilidades às empresas que desenvolvem tecnologias de inteligência artificial, envolvendo áreas como a pesquisa, desenvolvimento, prototipagem, testes, aquisição de requisitos e uso operacional. Além disso, estabelece um conselho para assuntos relacionados a questões éticas no desenvolvimento e uso de inteligência artificial (GPO, 2020).

Além disto, encontramos descritiva sobre desenvolvimento ético e responsável das tecnologias de inteligência artificial, onde as aquisições e desenvolvimento dessas tecnologias devem obedecer a padrões para assegurar a segurança do país pelo departamento de defesa. (GPO, 2020).

O documento descreve a iniciativa de estabelecer institutos nacionais de pesquisa em Inteligência Artificial. Nesse contexto, define um programa destinado a conceder assistência financeira para investimento no planejamento e apoio à criação de um ecossistema inovador, bem como aprimoramento de ecossistemas existente. O documento expressa suporte a atividades de educação interdisciplinar e experiências de pesquisa, promove o desenvolvimento da força de trabalho em disciplinas relacionadas à inteligência artificial no país, além de enfatizar a divulgação e o envolvimento para ampliar a participação na investigação em inteligência artificial e na força de trabalho correspondente, evidenciado no item 4.2.1.

De maneira geral, os Estados Unidos evidenciam diretrizes éticas para sistemas de inteligência artificial através das publicações no site da Casa Branca, onde o governo americano reconhece a importância estratégica da IA e promove medidas para seu desenvolvimento responsável e impacto positivo para manter a liderança americana em desenvolvimento de tecnologias de inteligência artificial.

A abordagem dos EUA para a regulamentação da IA tem sido mais setorial e dependente das agências reguladoras existentes. A *Federal Trade Commission* (FTC), por exemplo, tem autoridade para regular práticas de comércio desleais ou enganosas, que podem incluir o uso de IA. Ainda assim, não existe um marco regulatório de IA abrangente em nível nacional nos EUA (GPO, 2020).

Essas estratégias buscam fortalecer a liderança dos Estados Unidos na área da IA, ao mesmo tempo em que assegura a proteção da privacidade, a segurança cibernética e a igualdade de oportunidades.

Os Estados Unidos por meio de sua iniciativa AI.gov, mostra as diretrizes éticas para sistemas de inteligência artificial, que busca promover o uso responsável e ético dessas tecnologias, onde os pilares estratégicos são inovação, avanço da IA confiável, educação e treinamento, infraestrutura, formulários e cooperação internacional.

4.2.1 Quadro de diretrizes encontradas em documento analisado

Os trechos abaixo foram encontrados e retirados integralmente do documento oficial, traduzido do original e dispostos após análise, para comprovação das diretrizes pesquisadas e confecção do Quadro 4 e Quadro5 de análise comparativa.

Regulamentação Ética

IA Confiável

SEG. 22A. PADRÕES DE INTELIGÊNCIA ARTIFICIAL. - O Instituto deverá: " (1) desenvolver estruturas colaborativas, padrões, diretrizes e métodos e técnicas associados para inteligência artificial; "(2) apoiar o desenvolvimento de um quadro de mitigação de riscos para a implantação de sistemas de inteligência artificial; "(3) apoiar o desenvolvimento de padrões e diretrizes técnicas que promovam sistemas de inteligência artificial confiáveis; e "(1) apoiar a pesquisa de medição e o desenvolvimento de melhores práticas e padrões voluntários para sistemas de inteligência artificial confiáveis, que podem incluir— (I) todas as outras áreas consideradas pelo Diretor como críticas para o desenvolvimento e implantação de inteligência artificial confiável;. (GPO, 2020. p 1178).

A estrutura deve: "(1) identificar e fornecer padrões, diretrizes, melhores práticas para metodologias, procedimentos e processos para A) desenvolver sistemas de inteligência artificial confiáveis; "(B) avaliar a confiabilidade dos sistemas de inteligência artificial; e "(C) mitigação dos riscos dos sistemas de inteligência artificial; "(2) estabelecer definições e caracterizações comuns para aspectos de confiabilidade, incluindo explicabilidade, transparência, segurança, privacidade, proteção, robustez, justiça, preconceito, ética, validação, verificação, interpretabilidade e outras propriedades relacionadas à inteligência artificial sistemas que são comuns a todos os setores; (GPO, 2020. p 1179).

DIVISÃO E – INTELIGÊNCIA ARTIFICIAL NACIONAL - Os conferencistas acreditam ainda que tais sistemas prejudiciais de inteligência artificial podem incluir sistemas de alto risco que carecem de robustez suficiente para prevenir ataques adversários; sistemas de alto risco que prejudicam a privacidade ou a segurança dos utilizadores ou do público em geral; sistemas artificiais de inteligência geral que se tornam autoconscientes ou incontroláveis; e sistemas de inteligência artificial que discriminam ilegalmente classes protegidas de

pessoas, inclusive com base em sexo, raça, idade, deficiência, cor, credo, origem nacional ou religião. Finalmente, os conferencistas acreditam que os Estados Unidos devem adotar uma abordagem governamental global para a liderança em inteligência artificial confiável, inclusive através da coordenação entre o Departamento de Defesa, a Comunidade de Inteligência e as agências civis. (GPO, 2020. p 2136).

Privacidade

Os conferencistas acreditam ainda que tais sistemas prejudiciais de inteligência artificial podem incluir sistemas de alto risco que carecem de robustez suficiente para prevenir ataques adversários; sistemas de alto risco que prejudicam a privacidade ou a segurança dos utilizadores ou do público em geral; sistemas artificiais de inteligência geral que se tornam autoconscientes ou incontroláveis; e sistemas de inteligência artificial que discriminam ilegalmente classes protegidas de pessoas, inclusive com base em sexo, raça, idade, deficiência, cor, credo, origem nacional ou religião. Finalmente, os conferencistas acreditam que os Estados Unidos devem adotar uma abordagem governamental global para a liderança em inteligência artificial confiável, inclusive através da coordenação entre o Departamento de Defesa, a Comunidade de Inteligência e as agências civis. (GPO, 2020. p637).

Os conferencistas acreditam ainda que tais sistemas prejudiciais de inteligência artificial podem incluir sistemas de alto risco que carecem de robustez suficiente para prevenir ataques adversários; sistemas de alto risco que prejudicam a privacidade ou a segurança dos utilizadores ou do público em geral; sistemas artificiais de inteligência geral que se tornam autoconscientes ou incontroláveis; e sistemas de inteligência artificial que discriminam ilegalmente classes protegidas de pessoas, inclusive com base em sexo, raça, idade, deficiência, cor, credo, origem nacional ou religião. Finalmente, os conferencistas acreditam que os Estados Unidos devem adotar uma abordagem governamental global para a liderança em inteligência artificial confiável, inclusive através da coordenação entre o Departamento de Defesa, a Comunidade de Inteligência e as agências civis. (GPO, 2020. p 1168).

(D) As normas jurídicas, incluindo as concebidas para garantir a utilização de sistemas de inteligência artificial, são consistentes com os direitos de privacidade, os direitos civis e as liberdades civis, bem como com as questões de direitos das pessoas com deficiência levantadas pela utilização destas tecnologias. (GPO, 2020. p 1171).

(A) privacidade e segurança, inclusive para conjuntos de dados usados para treinar ou testar sistemas de inteligência artificial e software e hardware usados em sistemas de inteligência artificial; (GPO, 2020. p 1178).

Os conferencistas acreditam que os sistemas de inteligência artificial têm o potencial de transformar todos os sectores da economia dos Estados Unidos, aumentando a produtividade, melhorando a investigação científica e aumentando a competitividade dos EUA e que o governo dos Estados Unidos deve utilizar esta iniciativa para permitir os benefícios da confiança. inteligência artificial digna, evitando ao mesmo tempo a criação e

o uso de sistemas de inteligência artificial que se comportam de maneiras que causam danos. Os conferencistas acreditam ainda que tais sistemas prejudiciais de inteligência artificial podem incluir sistemas de alto risco que carecem de robustez suficiente para prevenir ataques adversários; sistemas de alto risco que prejudicam a privacidade ou a segurança dos utilizadores ou do público em geral; sistemas artificiais de inteligência geral que se tornam autoconscientes ou incontrolláveis; e sistemas de inteligência artificial que discriminam ilegalmente classes protegidas de pessoas, inclusive com base em sexo, raça, idade, deficiência, cor, credo, origem nacional ou religião. Finalmente, os conferencistas acreditam que os Estados Unidos devem adotar uma abordagem governamental global para a liderança em inteligência artificial confiável, inclusive através da coordenação entre o Departamento de Defesa, a Comunidade de Inteligência e as agências civis. (GPO, 2020. p 2136).

Participação da sociedade civil na elaboração da regulamentação

Alguns princípios são norteadores para que agências formulem abordagens regulatórias e não regulatórias para o projeto, desenvolvimento, implantação e operação de aplicativos de IA, tanto gerais quanto específicos do setor: Confiança pública na IA; Participação Pública; Integridade Científica e Qualidade da Informação; Avaliação e Gestão de Riscos; Benefícios e Custos; Flexibilidade; Justiça e Não Discriminação; Divulgação e Transparência, Segurança e Proteção; e, Coordenação Interinstitucional. (ENAP, 2022).

(a) EM GERAL. - O Secretário de Comércio deverá, em consulta com o Diretor do Escritório de Política Científica e Tecnológica, o Secretário de Defesa, o Secretário de Energia, o Secretário de Estado, o Procurador-Geral e o Diretor de Inteligência Nacional estabelecerem um comitê consultivo que será conhecido como Comitê Consultivo Nacional de Inteligência Artificial. (b) QUALIFICAÇÕES. - O Comitê Consultivo será composto por membros, nomeados pelo Secretário de Comércio, que representem conhecimentos e perspectivas amplas e interdisciplinares, incluindo instituições acadêmicas, empresas de diversos setores, entidades sem fins lucrativos e da sociedade civil, incluindo organizações de direitos civis e de direitos das pessoas com deficiência, e laboratórios federais, que representam a diversidade geográfica e que estão qualificados para fornecer aconselhamento e informações sobre pesquisa científica e tecnológica, desenvolvimento, ética, padrões, educação, transferência de tecnologia, aplicação comercial, segurança, e competitividade econômica relacionada com a inteligência artificial. (GPO, 2020. p 1169).

Garantia de Direitos

Representem conhecimentos e perspectivas amplas e interdisciplinares, incluindo instituições acadêmicas, empresas de diversos setores, entidades sem fins lucrativos e da sociedade civil, incluindo organizações de direitos civis e de direitos das pessoas com deficiência, e laboratórios federais, que representam a diversidade geográfica e que estão qualificados para fornecer aconselhamento e informações sobre pesquisa científica e tecnológica, desenvolvimento, ética, padrões, educação, transferência de tecnologia,

aplicação comercial, segurança, e competitividade econômica relacionada com a inteligência artificial. (GPO, 2020. p 1169).

(2) Uma descrição dos tipos de falsificações de conteúdo digital, incluindo aquelas usadas para cometer fraude, causar danos, assediar, coagir ou silenciar grupos ou indivíduos vulneráveis, ou violar direitos civis reconhecidos. (GPO, 2020. p 1419).

(12) responsabilidade e direitos legais, incluindo questões relacionadas com a supervisão de sistemas de inteligência artificial utilizando abordagens regulamentares e não regulamentares, a responsabilidade por quaisquer violações das leis existentes por parte de um sistema de inteligência artificial, e formas de equilibrar o avanço da inovação e ao mesmo tempo proteger os direitos individuais ; e (13) como a inteligência artificial pode aumentar as oportunidades para diversas regiões geográficas dos Estados Unidos, incluindo comunidades urbanas, tribais e rurais. (GPO, 2020. p 1170).

(D) As normas jurídicas, incluindo as concebidas para garantir a utilização de sistemas de inteligência artificial, são consistentes com os direitos de privacidade, os direitos civis e as liberdades civis, bem como com as questões de direitos das pessoas com deficiência levantadas pela utilização destas tecnologias. (GPO, 2020. p 1171).

Educação e Capacitação profissional

Capacitação técnica / programas de Treinamento

É detalhado um item para Educação e Desenvolvimento da Força de Trabalho, onde aparecem criação de programas de proficiência em linguagem de codificação e inteligência artificial, além de incentivos salariais para membros das forças armadas e civis que tenham desempenho bem-sucedido em testes específicos de proficiência em codificação de linguagem e de competência em inteligência artificial. O documento mostra diretrizes para contratação de professores universitários e estudantes de ciências de defesa e empresa de tecnologia, onde conduzirão pesquisas na área de inteligência artificial e máquinas (GPO, 2020).

(3) Apoio a programas interdisciplinares de investigação, educação e formação de mão de obra para estudantes e investigadores que promovam a aprendizagem nos métodos e sistemas utilizados na inteligência artificial e promovam perspectivas e colaborações interdisciplinares. (7) Apoio a uma rede de institutos interdisciplinares de pesquisa em inteligência artificial, conforme descrito na seção 5201(b). (GPO, 2020. p 1166).

(2) Apoio a programas de ensino fundamental e médio e de ensino póssecundário, incluindo treinamento de força de trabalho e programas de educação profissional e técnica, e programas de educação informal para preparar a força de trabalho americana e o público em geral para serem capazes de criar, usar e interagir com sistemas de inteligência artificial. (GPO, 2020. p 1166).

(1) fornecer a coordenação interagências de atividades federais de pesquisa, desenvolvimento e demonstração de inteligência artificial e atividades e programas de educação e treinamento de força de trabalho de departamentos e agências federais realizados de acordo com a Iniciativa; (2) o mais tardar 2 anos após a data de promulgação

desta Lei, desenvolver um plano estratégico para inteligência artificial (a ser atualizado pelo menos a cada 3 anos) que estabeleça metas, prioridades e métricas para orientar e avaliar como o agências que executam a Iniciativa irão. A) determinar e priorizar áreas de pesquisa, desenvolvimento e demonstração de inteligência artificial que exijam Liderança e investimento do Governo Federal; (B) apoiar o financiamento de longo prazo para pesquisa, desenvolvimento, demonstração e educação interdisciplinares em inteligência artificial; (F) apoiar e coordenar atividades federais de educação e treinamento de mão de obra relacionadas à inteligência artificial; e (G) apoiar e coordenar a rede de institutos de pesquisa em inteligência artificial descritos na seção 5201(b). (GPO, 2020. p 1168).

(1) RECURSO E RECURSO NACIONAL DE PESQUISA DE INTELIGÊNCIA ARTIFICIAL. - Os termos "Recurso Nacional de Pesquisa de Inteligência Artificial" e "Recurso" significam um sistema que fornece a pesquisadores e estudantes de todos os campos e disciplinas científicas acesso a recursos de computação, colocados

com conjuntos de dados governamentais e não governamentais, disponíveis ao público e preparados para a inteligência artificial, e um ambiente de investigação com ferramentas educativas adequadas e apoio ao utilizador. (GPO, 2020. p 1175).

(A) melhorar o ensino e a aprendizagem de tópicos relacionados a sistemas de inteligência artificial no ensino fundamental e médio e no ensino pós-secundário programas, incluindo treinamento de força de trabalho e programas de educação profissional e técnica, programas de educação de graduação e pós-graduação e em ambientes formais; e (B) aumentar a participação em campos relacionados à inteligência artificial, inclusive por indivíduos identificados (GPO, 2020. p 1182).

EM GERAL.—O Diretor da National Science Foundation concederá subsídios para pesquisa, desenvolvimento e atividades relacionadas em educação em inteligência artificial para apoiar programas e atividades de ensino fundamental e médio e pós-secundário, incluindo treinamento de força de trabalho e programas e atividades de educação profissional e técnica, educação de graduação, pós-graduação e pós-doutorado, e programas e atividades de educação informal que: (A) apoiem o desenvolvimento de um fluxo diversificado de força de trabalho para ciência e tecnologia com relação a sistemas de inteligência artificial (GPO, 2020. p 1183).

Oportunidades de carreira em IA

consultar as orientações aplicáveis do Departamento de Defesa e, quando for benéfico, emitir novas orientações aos secretários dos departamentos militares e aos chefes dos componentes de defesa sobre o melhor uso dos processos de contratação direta para profissionais de inteligência artificial e outro pessoal de ciência de dados e desenvolvimento de software.

(b) O objetivo da orientação emitida na subseção (a) será garantir que os líderes organizacionais assumam maior responsabilidade pelos resultados da contratação civil de profissionais de inteligência artificial e outro pessoal de ciência de dados e desenvolvimento de software.

4) incentivar o uso de bônus de indicação para recrutamento e contratação de profissionais de inteligência artificial altamente qualificados e outro pessoal de ciência de dados e desenvolvimento de software, de acordo com o volume 451 da Instrução 1400.25 do Departamento de Defesa.

(B) O efeito da orientação sobre o processo de contratação de profissionais de inteligência artificial e outro pessoal de ciência de dados e desenvolvimento de software. (GPO, 2020. p 764).

(2) BOLSAS DE FACULDADE DE INTELIGÊNCIA ARTIFICIAL. - (A) BOLSAS DE RECRUTAMENTO DE FACULDADE. - (i) EM GERAL. - O Diretor da National Science Foundation estabelecerá um programa para conceder bolsas a instituições elegíveis de ensino superior. cação para recrutar e reter professores efetivos ou efetivos em inteligência artificial e áreas afins. (GPO, 2020. p 1183).

(ii) USO DE FUNDOS. - Uma instituição de ensino superior deverá usar os fundos da subvenção fornecidos sob a cláusula (i) para fins de - (I) recrutar novos membros efetivos ou docentes efetivos que conduzam pesquisa e ensino em inteligência artificial e campos e áreas de pesquisa relacionados, incluindo ética tecnológica (GPO, 2020. p 1184

Programa para aumentar a participação dos contratantes em atividades de ciência, tecnologia, engenharia e matemática." Esta seção também orientaria o Secretário de Defesa a executar um programa sob o qual o Secretário procurará entrar em parcerias com empreiteiros do Departamento de Defesa para realizar atividades de serviços comunitários para promover o interesse em carreiras nas disciplinas de ciência, tecnologia, engenharia e matemática, e permitir que essas atividades sejam consideradas como custos permitidos em um contrato governamental. (GPO, 2020. p 1551).

Estudos de Impactos Profissionais

(4) questões relacionadas à inteligência artificial e aos Estados Unidos. Força de trabalho dos Estados, incluindo questões relacionadas com o potencial de utilização da inteligência artificial para a formação da força de trabalho, as possíveis consequências da deslocação tecnológica e o apoio a oportunidades de formação de mão de obra para ocupações que conduzam à autossuficiência económica para indivíduos com barreiras de acesso (GPO, 2020. p 1169).

1) impactos na força de trabalho em todos os setores causados pela crescente adoção de inteligência artificial, automação e outras tendências relacionadas; (2) necessidades de mão de obra e oportunidades de emprego geradas pela crescente adoção da inteligência artificial em todos os setores; (3) lacunas de pesquisa e dados necessários para melhor compreender e acompanhar os parágrafos (1) e (2); (GPO, 2020. p 1171).

(B) CENTROS DE EXCELÊNCIA DE INTELIGÊNCIA ARTIFICIAL. – O Diretor da Fundação Nacional de Ciência estabelecerá centros nacionais de educação científica e técnica para promover a educação e o desenvolvimento da força de trabalho em áreas relacionadas à inteligência artificial, As atividades de tais centros podem incluir: (i) o desenvolvimento, disseminação e avaliação de currículos e outras ferramentas e métodos educacionais em campos relacionados à inteligência artificial e áreas de pesquisa, incluindo ética tecnológica; (ii) o desenvolvimento e avaliação de certificações relacionadas com inteligência artificial para programas de 2 anos; e (iii) investigação interdisciplinar em ciências e engenharia na aprendizagem de adultos baseada no emprego e na reconversão profissional relacionada com os domínios da inteligência artificial. (GPO, 2020. p 1185).

Infraestrutura

Computação de alto desempenho

Consta importância de análise de dados, tecnologia autônoma, materiais avançados, software, alta computação de desempenho de inteligência artificial deve ser identificadas pelo comitê (GPO, 2020).

5) Auxiliar o Secretário e o Diretor no desenvolvimento de orientações de nível estratégico sobre aquisição de hardware relacionado à inteligência artificial, questões da cadeia de suprimentos e outros assuntos técnicos relacionados à inteligência artificial. (GPO, 2020. p 98).

(E) fornecer ou facilitar as instalações necessárias de computação, redes e dados para pesquisa e desenvolvimento de inteligência artificial; (GPO, 2020. p 1168).

B) chips e hardware de computador avançados projetados para sistemas de inteligência artificial. (GPO, 2020. p 1178).

(5) o desenvolvimento de hardware de computação avançado e de arquitetura de computadores adaptada a sistemas de inteligência artificial, incluindo a concepção codificada de redes e hardware computacional (GPO, 2020. p 1186).

(1) disponibilizar infraestrutura de computação de alto desempenho cultura em laboratórios nacionais; 2) efetuar todas as atualizações necessárias para melhorar a utilização dos recursos informáticos existentes para sistemas de inteligência artificial, incluindo atualizações de hardware; (3) estabelecer novas capacidades computacionais necessárias para gerir dados e realizar computação de alto desempenho que permita a utilização de sistemas de inteligência artificial; e (4) manter e melhorar, conforme necessário, a infraestrutura de rede, os mecanismos de entrada e saída de dados e as capacidades de análise, armazenamento e serviços de dados. (GPO, 2020. p 1187).

Recurso de Dados

i) identificar casos de uso tático e operacional para armas artificiais inteligência; "(ii) melhorar a coleta de dados relativos à inteligência artificial. (GPO, 2020. p 95).

(I) estabelecer acordos de uso de dados e modelos e acordos de parceria colaborativa para desenvolvimento de produtos de inteligência artificial com cada organização e elemento do Departamento, incluindo cada uma das Armadas. (GPO, 2020. p 97).

(A) privacidade e segurança, inclusive para conjuntos de dados usados para treinar ou testar sistemas de inteligência artificial e software e hardware usados em sistemas de inteligência artificial. (GPO, 2020. p 1178).

C) gerenciamento de dados e técnicas para aumentar a usabilidade dos dados, incluindo estratégias para limpar, rotular e padronizar sistematicamente os dados em formatos úteis

para o treinamento de sistemas de inteligência artificial e o uso de licenças abertas e comuns. (GPO, 2020. p 1178).

3) a análise de conjuntos de dados existentes em grande escala provenientes de experiências e simulações científicas e de engenharia, incluindo simulações energéticas e outras prioridades no Departamento, conforme determinado pelo Secretário, utilizando ferramentas e técnicas de inteligência artificial; (4) o desenvolvimento de sistemas de operação e controle que aprimorar capacidades de tomada de decisão automatizadas e inteligentes. (GPO, 2020. p 1186).

6) o desenvolvimento de conjuntos de dados padronizados para campos e aplicações emergentes de pesquisa em inteligência artificial, incluindo métodos para lidar com a escassez de dados. (GPO, 2020. p 1186).

Tecnologia em Nuvem

(2) outras capacidades tecnológicas potencialmente inovadoras no que diz respeito a essas infraestruturas, incluindo bases de dados baseadas na nuvem, tecnologias de inteligência artificial, automação e capacidades melhoradas de modelização e simulação. (GPO, 2020. p 1437).

Acordos Internacionais

Diálogo Internacional

(8) Apoiar oportunidades de cooperação internacional com aliados estratégicos, conforme apropriado, em matéria de investigação e desenvolvimento, avaliação e recursos para sistemas de inteligência artificial fiáveis. (GPO, 2020. p 1166).

O Comitê Consultivo aconselhará o Presidente e o Escritório da Iniciativa sobre assuntos relacionados à Iniciativa, incluindo recomendações relacionadas a: (1) o estado atual da competitividade e liderança dos Estados Unidos em inteligência artificial, incluindo o escopo e a escala dos investimentos dos Estados Unidos em investigação e desenvolvimento de inteligência artificial no contexto internacional; (GPO, 2020. p 1169).

(8) avaliar oportunidades de colaboração internacional com aliados estratégicos na investigação e desenvolvimento de inteligência artificial. (GPO, 2020. p 1182).

Parcerias Internacionais

11) oportunidades de cooperação internacional com aliados estratégicos em atividades de pesquisa em inteligência artificial. (GPO, 2020. p 1170).

(4) alinhar-se com os padrões internacionais, conforme apropriado. (GPO, 2020. p 1179).

A disposição instruiria o Secretário da Defesa a garantir que a autoridade pública privada de intercâmbio de talentos seja usada para trocar pessoal com experiência no setor privado trabalhando em aplicações de inteligência artificial. (GPO, 2020. p 1766).

4.3 União Europeia

Segundo dados disponíveis no site da União Europeia, foram analisados os documentos - Regulamento de Inteligência Artificial publicado em 2021 (Comissão Europeia, 2021).

A União Europeia em abril de 2021, publicou sua proposta de Regulamento de Inteligência Artificial, que aborda questões de transparência, responsabilidade e privacidade em sistemas de inteligência artificial. Também houve a publicação do Regulamento Geral de Proteção de Dados, uma legislação da União Europeia que estabelece regras rigorosas para o uso de dados pessoais, incluindo o uso de tecnologias de inteligência artificial. Ele entrou em vigor em maio de 2018 e se aplica a todas as empresas que coletam, armazenam ou processam dados pessoais de cidadãos.

As diretrizes éticas para Inteligência Artificial da União Europeia foram publicadas em 2019 e estabelecem princípios para o uso responsável de tecnologias de inteligência artificial. As diretrizes incluem princípios como transparência, equidade, segurança e privacidade. Elas foram desenvolvidas em consulta com especialistas em inteligência artificial, empresas, organizações da sociedade civil e autoridades públicas.

Em 2021, a Comissão Europeia propôs novas regras abrangentes sobre IA, conhecidas como o “Regulamento de Inteligência Artificial” disponível em site oficial da União Europeia. Essas regras visam criar um ambiente de confiança e segurança em torno da utilização da IA na UE, focando na transparência, no controle humano sobre a IA, na proteção da privacidade e dos dados pessoais, e na responsabilidade das partes envolvidas.

O documento contém 67 páginas e inicia num contexto em que são delineadas as bases de uma proposta legislativa, voltada para uma abordagem transversal, coordenada às implicações humanas e éticas da inteligência artificial. Nele são devidamente tratados os benefícios e os riscos associados à IA. A proposta alinha-se ao objetivo da União de liderar mundialmente o desenvolvimento de uma inteligência artificial segura, ética e confiável, conforme mencionado pelo Conselho Europeu.

Na proposta apresentada pela Comissão, busca-se regulamentar a inteligência artificial com o objetivo de assegurar que os sistemas de IA colocados no mercado da União e utilizados sejam seguros e estejam em conformidade com a legislação vigente em matéria de direitos fundamentais e valores da União; ·garantir a segurança jurídica para facilitar os investimentos e a inovação no domínio da IA; ·melhorar a governança e a aplicação efetiva da legislação em vigor em matéria de direitos fundamentais e dos requisitos de segurança aplicáveis aos sistemas

de IA; · facilitar o desenvolvimento de um mercado único para as aplicações de IA legítimas, seguras e de confiança e evitar a fragmentação do mercado (Comissão Europeia, 2021, p.2).

No segundo capítulo do documento são explorados assuntos de base jurídica, subsidiariedade e proporcionalidade, onde é exposto o tratado sobre o Funcionamento da União Europeia. Na terceira parte são demonstradas avaliação de impacto, onde de uma forma geral, as partes interessadas definiram os termos risco, risco elevado, risco baixo e prejuízo/dano, abrangem desde a proibição de sistemas de IA considerados perigosos até requisitos rigorosos para sistemas de IA de alto risco, passando por regras de transparência para sistemas de IA que interagem com as pessoas. (Comissão Europeia, 2021, p.5).

O exposto no documento evidencia a preocupação do Parlamento em assegurar que os sistemas de inteligência artificial empregados na União Europeia sejam seguros, transparentes, passíveis de rastreamento, livres de discriminação e ecologicamente conscientes, tendo como descrito os direitos fundamentais assegurados como direito à dignidade, respeito a vida privada e familiar e a proteção dos dados pessoais, direito como a não discriminação, igualdade entre homens e mulheres.

A proposta pretende evitar um efeito inibidor nos direitos à liberdade de expressão, garantir a proteção do direito à ação e a um tribunal imparcial e dos direitos de presunção de inocência e de defesa, bem como do direito a uma boa administração. Além disso são descritos nessa proposta os direitos dos trabalhadores a condições de trabalho justas e equitativas, o direito a um elevado nível de defesa dos consumidores, os direitos das crianças e o direito de integração das pessoas com deficiência. Itens como a preocupação referente a gestão de riscos e à supervisão humana também são citados no documento (Comissão Europeia, 2021).

O documento demonstra ser fundamental que a supervisão humana prevaleça sobre a automatização na governança dos sistemas de IA, a fim de prevenir eventuais consequências negativas. Para isso os riscos foram separados em níveis, assegurando assim os direitos fundamentais e a segurança dos usuários. Para a supervisão humana o regulamento proposto deixa claro que os sistemas de IA de alto risco devem permitir a supervisão eficaz por indivíduos durante o uso, evitando assim riscos para os seres humanos e/ou a sociedade que podem ocorrer durante seu processamento.

Todos os sistemas de IA considerados uma clara ameaça à segurança, meios de subsistência e direitos das pessoas serão banidos, desde pontuação social pelos governos até

brinquedos que usam assistência de voz que incentivam comportamentos perigosos (Comissão Europeia, 2021).

São citadas práticas de inteligência artificial proibidas, sistemas de inteligência artificial de risco elevado, obrigações de transparência aplicáveis, medidas de apoio à inovação, inclusive código de conduta à sociedade.

A Comissão Europeia (2021) prevê medidas para a implementação da supervisão humana que integralizem pelo fornecedor essa condição antes de ser disponibilizado no mercado ou usado, e identifiquem essas medidas adequando assim para implantação pelo usuário antes do sistema de IA de alto risco ser disponibilizada.

Essas medidas viabilizam uma compreensão abrangente das capacidades e limitações do sistema permitindo o controle e detecção mais rápidos de problemas. Isso possibilita que os usuários estejam mais conscientes do risco na confiança do resultado desses sistemas de IA para tomada de decisões, ao mesmo tempo possibilita uma interpretação precisa dos resultados gerados pelo sistema. Essas medidas permitem também a intervenção no funcionamento podendo interromper seu processamento a qualquer momento.

O documento indica a necessidade da verificação de pelo menos duas pessoas nesses processos, para assegurar a confiabilidade dos resultados.

Essas análises estão em fase avançada de implementação e visam regular uma vasta gama de aplicações de IA através de uma abordagem que prioriza a segurança do usuário, com a imposição de sanções em caso de violação das regras, inclusive tem descrito obrigações de transparência aplicáveis nos sistemas de inteligência artificial.

4.3.1. Quadro de diretrizes encontradas em documento analisado

Os trechos abaixo foram encontrados e retirados integralmente do documento oficial, traduzido do original e dispostos após análise, para comprovação das diretrizes pesquisadas e confecção do Quadro 4 e Quadro5 de análise comparativa.

Regulamentação ética

IA confiável

A presente proposta visa dar execução ao segundo objetivo, desenvolvendo um ecossistema de confiança mediante a proposta de um quadro jurídico para uma IA de confiança. A proposta tem como base os valores e os direitos fundamentais da UE e pretende dar às pessoas e a outros utilizadores a confiança necessária para adotarem soluções baseadas em IA, ao mesmo tempo que incentiva as empresas para que as desenvolvam. (Comissão Europeia, 2021, p.01).

A inteligência artificial deve ser uma ferramenta ao serviço das pessoas e uma força positiva para a sociedade com o objetivo final de aumentar o bem-estar dos seres humanos. As regras aplicáveis às tecnologias de inteligência artificial disponibilizadas no mercado da União ou que afetam as pessoas da União devem, por isso, centrar-se no ser humano, de modo que as pessoas possam confiar que a tecnologia é utilizada de uma forma segura e em cumprimento da lei, incluindo em matéria de respeito dos direitos fundamentais. (Comissão Europeia, 2021, p.01).

artigo 15.º - Os sistemas de IA de risco elevado devem ser concebidos e desenvolvidos de maneira que alcancem, tendo em conta a finalidade prevista, um nível apropriado de exatidão, solidez e cibersegurança e apresentem um desempenho coerente em relação a tais aspetos durante o ciclo de vida.

O nível de exatidão, solidez e cibersegurança a que se refere o artigo 15.º relativamente ao qual o sistema de IA de risco elevado foi testado e validado e que pode ser esperado, bem como quaisquer circunstâncias conhecidas e previsíveis que possam ter um impacto nesse nível esperado de exatidão, solidez e cibersegurança. (Comissão Europeia, 2021, p.27).

As instruções de utilização que acompanham os sistemas de IA de risco elevado devem declarar os níveis de exatidão e a métrica de exatidão aplicável. (Comissão Europeia, 2021, p.28).

Artigo 54º - Existem mecanismos de controlo eficazes para identificar quaisquer riscos elevados para os direitos fundamentais dos titulares dos dados que possam surgir durante a experimentação no ambiente de testagem, bem como um mecanismo de resposta para atenuar prontamente esses riscos e, se necessário, interromper o tratamento;

- d) Todos os dados pessoais a tratar no contexto do ambiente de testagem se encontram num ambiente de tratamento de dados funcionalmente separado, isolado e protegido sob o controlo dos participantes, sendo apenas acessíveis a pessoas autorizadas;
- e) Nenhum dados pessoais tratados são transmitidos, transferidos ou acedidos, de outro modo, por terceiros;
- f) Nenhum tratamento de dados pessoais no contexto do ambiente de testagem dá origem a medidas ou decisões que afetem os titulares dos dados;
- g) Todos os dados pessoais tratados no contexto do ambiente de testagem são apagados assim que a participação no ambiente de testagem terminar ou que os dados pessoais atingirem o fim do respetivo período de conservação;
- h) Os registos do tratamento de dados pessoais no contexto do ambiente de testagem são mantidos durante a participação no ambiente de testagem e pelo período de um ano após o respetivo termo, apenas enquanto forem necessários para efeitos exclusivos de cumprimento de obrigações em matéria de responsabilidade e documentação previstas no presente artigo ou em outra legislação da União ou dos Estados-Membros aplicável. (Comissão Europeia, 2021, p.39).

Privacidade

Além das inúmeras utilizações benéficas da inteligência artificial, essa tecnologia pode ser utilizada indevidamente e conceder instrumentos novos e poderosos para práticas manipuladoras, exploratórias e de controlo social. Essas práticas são particularmente prejudiciais e devem ser proibidas, pois desrespeitam valores da União, como a dignidade do ser humano, a liberdade, a igualdade, a democracia e o Estado de direito, bem como direitos fundamentais da União, incluindo o direito à não discriminação, à proteção de dados pessoais e à privacidade, e os direitos das crianças. (Comissão Europeia, 2021, p 11).

5. Na medida do estritamente necessário para assegurar o controlo, a deteção e a correção de enviesamentos em relação a sistemas de IA de risco elevado, os fornecedores desses sistemas podem tratar categorias especiais de dados pessoais a que se refere o artigo 9.º, n.º 1, do Regulamento (UE) 2016/679, o artigo 10.º da Diretiva (UE)2016/680 e o artigo 10.º, n.º 1, do Regulamento (UE) 2018/1725, assegurando salvaguardas adequadas dos direitos fundamentais e liberdades das pessoas singulares, incluindo impor limitações técnicas à reutilização e utilizar medidas de segurança e preservação da privacidade de última geração, tais como a pseudonimização ou a cifragem nos casos em que a anonimização possa afetar significativamente a finalidade preconizada. (Comissão Europeia, 2021, p.26).

Participação da sociedade civil na elaboração da regulamentação

No total, foram recebidos 1 215 contributos, dos quais 352 de empresas ou organizações/associações comerciais, 406 de cidadãos (92 % eram cidadãos da UE), 152 em nome de instituições académicas/de investigação e 73 de autoridades públicas. As opiniões da sociedade civil foram representadas por 160 respondentes (9 dos quais eram organizações de consumidores, 129 eram organizações não governamentais e 22 eram sindicatos), sendo que 72 respondentes contribuíram identificando-se como Outros. Das 352 empresas e representantes da indústria, 222 eram empresas e representantes comerciais, sendo que 41,5 % eram micro, pequenas e médias empresas. As restantes eram associação empresariais. De um modo geral, 84 % das respostas das empresas e da indústria eram provenientes da UE-27. Dependendo da pergunta, entre 81 e 598 dos respondentes utilizaram a opção de texto livre para inserir observações. Foram apresentadas mais de 450 posições escritas através do

portal EU Survey, quer como complemento das respostas aos inquéritos (mais de 400), quer como contributos autónomos (mais de 50). (Comissão Europeia, 2021, p 04).

Garantia de Direitos

Garantir que os sistemas de IA colocados no mercado da União e utilizados sejam seguros e respeitem a legislação em vigor em matéria de direitos fundamentais e valores da União. p2

Dadas as suas características específicas (por exemplo, a opacidade, a complexidade, a dependência dos dados, o comportamento autónomo), a utilização da inteligência artificial pode afetar negativamente um conjunto de direitos fundamentais consagrados na Carta dos Direitos Fundamentais da UE (a seguir designada por «Carta»). A presente proposta procura assegurar um nível elevado de proteção desses direitos fundamentais e visa fazer face aos vários riscos mediante uma abordagem baseada no risco claramente definida. Graças a um conjunto de requisitos relativos a uma IA de confiança e obrigações proporcionadas para todos os participantes da cadeia de valor, a proposta melhorará e promoverá a proteção dos direitos consagrados na Carta: o direito à dignidade do ser humano (artigo 1.º), o respeito pela vida privada e familiar e a proteção de dados pessoais (artigos 7.º e 8.º), a não discriminação (artigo 21.º) e a igualdade entre homens e mulheres (artigo 23.º). A proposta pretende evitar um efeito inibidor nos direitos à liberdade de expressão (artigo 11.º) e à liberdade de reunião (artigo 12.º), garantir a proteção do direito à ação e a um tribunal imparcial e dos direitos de presunção de inocência e de defesa (artigos 47.º e 48.º), bem como do direito a uma boa administração. Além disso, conforme aplicável em determinados domínios, a proposta afetará de forma positiva os direitos de um conjunto de grupos especiais, como os direitos dos trabalhadores a condições de trabalho justas e equitativas (artigo 31.º), o direito a um elevado nível de defesa dos consumidores (artigo 28.º), os direitos das crianças (artigo 24.º) e o direito de integração das pessoas com deficiência (artigo 26.º). O direito a um elevado nível de proteção do ambiente e melhoria da sua qualidade (artigo 37.º) também é relevante, incluindo em relação à saúde e à segurança dos cidadãos. As obrigações relativas à testagem ex ante, à gestão de riscos e à supervisão humana também facilitarão o respeito de outros direitos fundamentais, graças à minimização do risco de decisões assistidas por IA erradas ou enviesadas em domínios críticos como a educação e a formação, o emprego, serviços essenciais, a manutenção da ordem pública e o sistema judicial. Caso continuem a ocorrer violações dos direitos fundamentais, as pessoas afetadas têm acesso a vias eficazes de recurso graças à garantia da transparência e da rastreabilidade dos sistemas de IA, associadas a fortes controlos ex post. (Comissão Europeia, 2021, p.06).

Educação e capacitação profissional

Capacitação técnica / Programas de treinamento

Não consta

Oportunidades de carreira em IA

Não consta

Estudo de impactos profissionais

Os sistemas de IA utilizados nos domínios do emprego, da gestão de trabalhadores e do acesso ao emprego por conta própria, nomeadamente para efeitos de recrutamento e seleção, de tomada de decisões sobre promoções e despedimentos, de repartição de tarefas e de controlo ou avaliação de pessoas no âmbito de relações contratuais de trabalho também devem ser classificados como de risco elevado, uma vez que podem ter um impacto significativo nas perspectivas de carreira e na subsistência dessas pessoas. O conceito de «relações contratuais relacionadas com o trabalho» deve abranger os funcionários e as pessoas que prestam serviços por intermédio de plataformas, conforme mencionado no programa de trabalho da Comissão para 2021. Em princípio, essas pessoas não devem ser consideradas utilizadores na acepção do presente regulamento. Ao longo do processo de recrutamento e na avaliação, promoção ou retenção de pessoas em relações contratuais relacionadas com o trabalho, esses sistemas podem perpetuar padrões históricos de discriminação, por exemplo, contra as mulheres, certos grupos etários, pessoas com deficiência ou pessoas de uma determinada origem racial ou étnica ou orientação sexual. Os sistemas de IA utilizados para controlar o desempenho e o comportamento destas pessoas podem ter ainda um impacto nos seus direitos à proteção de dados pessoais e à privacidade. (Comissão Europeia, 2021, p.14).

Infraestrutura

Computação de alto desempenho

Não consta

Recurso de dados

Dados de treino, os dados usados para treinar um sistema de IA mediante o ajustamento dos seus parâmetros passíveis de serem aprendidos, incluindo os pesos de uma rede neuronal;

30) Dados de validação, os dados utilizados para realizar uma avaliação do sistema de IA treinado e para ajustar os seus parâmetros não passíveis de serem aprendidos e o seu processo de aprendizagem, a fim de, entre outros objetivos, evitar um sobre ajustamento; sendo que o conjunto de dados de validação pode ser um conjunto de dados separado ou parte de um conjunto de dados de treino, quer como divisão fixa ou variável;

31) Dados de teste, os dados utilizados para realizar uma avaliação independente do sistema de IA treinado e validado, a fim de confirmar o desempenho esperado desse sistema antes de ser colocado no mercado ou em serviço;

32) Dados de entrada, os dados fornecidos a um sistema de IA, ou por ele obtidos diretamente, com base nos quais o sistema produz um resultado;

33) Dados biométricos, dados pessoais resultantes de um tratamento técnico específico das características físicas, fisiológicas ou comportamentais de uma pessoa singular, os quais permitem obter ou confirmar a identificação única dessa pessoa singular, nomeadamente imagens faciais ou dados dactiloscópicos; (Comissão Europeia, 2021, p.21).

Sistemas e procedimentos de gestão de dados, incluindo recolha de dados, análise de dados, rotulagem de dados, armazenamento de dados, filtragem de dados, prospecção de dados, agregação de dados, conservação de dados e qualquer outra operação relativa aos dados que seja realizada antes e para efeitos da colocação no mercado ou colocação em serviço de sistemas de IA de risco elevado; (Comissão Europeia, 2021, p.29).

Tecnologias em nuvem

Não consta

Acordos Internacionais

Diálogo internacional

Face à natureza digital dos sistemas de IA, determinados sistemas devem ser abrangidos pelo âmbito do presente regulamento, mesmo quando não são colocados no mercado ou em serviço, nem são utilizados na União. Esta situação verifica-se, por exemplo, quando um operador estabelecido na União contrata determinados serviços a um operador estabelecido fora da União relativamente a uma atividade a realizar por um sistema de IA que seria considerado de risco elevado e cujos efeitos afetam pessoas singulares localizadas na União. Nessas circunstâncias, o operador fora da União poderia utilizar o seu sistema de IA para tratar dados recolhidos e transferidos licitamente da União e fornecer ao operador contratante na União o resultado desse sistema de IA decorrente desse tratamento, sem que o sistema de IA em causa fosse colocado no mercado ou em serviço ou utilizado na União. Para evitar que o presente regulamento seja contornado e para assegurar uma proteção eficaz das pessoas singulares localizadas na União, o presente regulamento deve ser igualmente aplicável a fornecedores e utilizadores de sistemas de IA estabelecidos num país terceiro nos casos em que o resultado desses sistemas seja utilizado na União. No entanto, para ter em conta os mecanismos existentes e as necessidades especiais de cooperação com os parceiros estrangeiros com quem são trocadas informações e dados, o presente regulamento não deve ser aplicável às autoridades públicas de um país terceiro e às organizações internacionais quando estas atuam no âmbito de acordos internacionais celebrados a nível nacional ou europeu para efeitos de cooperação policial e judiciária com a União ou com os seus Estados-Membros. Tais acordos têm sido celebrados bilateralmente entre Estados-Membros e países terceiros ou entre a União Europeia, a Europol e outras agências da UE e países terceiros e organizações internacionais. (Comissão Europeia, 2021, p.11).

Parcerias internacionais

Não consta

4.4 China

O governo chinês tem tomado medidas significativas com o propósito de regulamentar o desenvolvimento e a aplicação de tecnologias relacionadas à Inteligência Artificial. Uma das iniciativas governamentais foi a publicação oficial do “**Plano de Desenvolvimento de Nova Geração de Inteligência Artificial**”, em 2017, disponível no site do governo. Esse plano estabelece metas específicas para o progresso da IA na China até 2030 e estabelece diretrizes para sua implementação em setores chave como manufatura, saúde e transporte. Inclui medidas para garantir a segurança, privacidade e proteção de dados na aplicação dessas tecnologias.

O plano intitulado como “Aviso do Conselho Estadual de Impressão e Distribuição do Plano de Desenvolvimento de uma Nova Geração de Inteligência Artificial”, pelo conselho estadual da China, é um documento com aproximadamente 15 páginas que articula o papel estratégico da inteligência artificial destacando seu papel na aceleração da construção de uma nação inovadora e ascensão como uma potência mundial em ciência e tecnologia” (China, 2017).

No documento encontramos descrições da situação estratégica do País, como foco de competição internacional, e abordagem de regulamentação mista, com detalhamentos para os setores, mas também com uma abordagem geral e ampla para a sociedade do país. Mostra-se itens importantes como transformação industrial através da IA, produtividade social, e tratativa de desafios do país como envelhecimento da população e restrição de recursos ambientais, neste documento é descrito como a inteligência artificial pode ser amplamente utilizada para melhorar a qualidade de vida da população.

A tecnologia de inteligência artificial pode perceber, prever e alertar com precisão sobre as principais tendências nas operações de infraestrutura e seguridade social, compreender oportunamente a cognição do grupo e as mudanças psicológicas e tomar a iniciativa de tomar decisões e responder, o que melhorará significativamente a capacidade e o nível de governança social, e é indispensável para a efetiva manutenção da estabilidade social. (China, 2017. p.3).

O documento descreve preocupações quanto aos empregos, mudança de carreiras já existentes, impacto quanto a ética social, privacidade entre outros pontos. Propõe medidas no aspecto de pesquisa científica e tecnológica, e demandas tecnológicas necessárias para essa evolução. Nesta publicação destaca a importância e implementação de pesquisa e desenvolvimento (P&D), também a criação de ambiente e formulação ética e regulamentos, o governo defende também o conceito de partilha de código e promove a criação de partilha de

vários temas como inovação, investigação nos meios industriais e meios acadêmicos. (China, 2017. p.4).

Na descrição documental, observa-se preocupação em liderar atualização das tecnologias de IA como inteligência de big data, inteligência híbrida humano-computador, onde concentra-se em resolver problemas com aprendizagem não supervisionada¹⁸.

O governo descreve como pretende acelerar treinamento e formação de equipes de talentos em desenvolvimento de IA, como melhorar o sistema educacional do país desenvolvendo equipes inovadoras de inteligência artificial, apoio a pesquisa aplicada e incentivo a faculdades e universidades a expandir conteúdo de educação profissional de inteligência artificial, além do fortalecimento da cooperação indústria-universidade-pesquisa na construção de disciplinas que tratem o desenvolvimento de Inteligência Artificial.

Inovação em *hardware* também é uma preocupação do governo chinês, citando o fortalecimento de indústrias de hardware e software para aplicações de IA, onde tem propósito de forte incentivo a criação de robôs inteligentes, realidade aumentada, dispositivos de IOT, carros autônomos e sistemas de trânsito ferroviário e integração tecnológica, com desenvolvimento de chips de sensores inteligentes de alta sensibilidade e alta confiabilidade que suportem a nova geração de Internet das Coisas.

Em menos de três anos a China possui já empresas bilionárias em tecnologia de Inteligência Artificial, um progresso rápido e inevitável. A China tem amplo apoio em infraestrutura do governo, com mais usuários de tecnologia do que qualquer lugar do mundo, produz dados enviados para nuvem em uma velocidade gigantescas para a implementação de novas tecnologias e a aprendizagem de máquina se torna muito mais rápida, quanto mais dados a IA funciona melhor e a China lidera mundialmente em quantidade de dados. (Castro, 2019).

[..] a vantagem dos dados da China se estende da quantidade à qualidade. O grande número de usuários de internet do país – maior que os Estados Unidos e a Europa juntos – cria a quantidade de dados, mas é o que esses usuários fazem on-line que garante a qualidade. A natureza do universo alternativo de aplicativos da China significa que os dados coletados também serão muito mais úteis na construção de empresas orientadas para IA (Lee, 2019, p. 74).

¹⁸ Aprendizagem não supervisionada - Usa algoritmos de machine learning para analisar e agrupar conjuntos de dados não rotulados. Esses algoritmos descobrem padrões ocultos ou agrupamentos de dados sem a necessidade de intervenção humana. IBM. O que é Aprendizado não Supervisionado? Disponível em: <https://ibm.com/br-pt/topics/unsupervised-learning>. Acesso em dezembro 2023.

O plano estabelece metas para alcançar avanços significativos até 2025 em pesquisa fundamental, tecnologias e aplicações de IA considerando-os como fatores impulsionadores para a melhoria e transformação produtiva. Para 2030, o objetivo é alcançar níveis que permitam competir pela liderança global em termos de inovação, alcançando resultados notáveis na construção de uma nova sociedade e economia, visando sustentar uma economia inovadora e foco em ser uma potência econômica (Hiratuka, Diegues, 2021).

Além disso, a China estabeleceu uma série de regulamentações e padrões para garantir a qualidade e segurança das aplicações de IA. Também tem se movimentado em direção à regulamentação da IA, principalmente através de suas políticas de segurança cibernética e diretrizes sobre a ética da IA.

4.4.1. Quadro de diretrizes encontradas em documento analisado

Os trechos abaixo foram encontrados e retirados integralmente do documento oficial, traduzido do original e dispostos após análise, para comprovação das diretrizes pesquisadas e confecção do Quadro 4 e Quadro5 de análise comparativa.

Regulamentação Ética

IA Confiável

A incerteza no desenvolvimento da inteligência artificial traz novos desafios.

(2) Princípios básicos. A inteligência artificial é uma tecnologia disruptiva com uma ampla gama de impactos. Pode trazer problemas como mudar a estrutura do emprego, impactar as leis e a ética social, infringir a privacidade pessoal e desafiar as normas das relações internacionais. Terá um impacto profundo sobre gestão governamental, segurança econômica, estabilidade social e até mesmo governança global. Enquanto desenvolvemos vigorosamente a inteligência artificial, devemos atribuir grande importância a possíveis riscos e desafios de segurança, fortalecer a prevenção voltada para o futuro e a orientação de restrição, minimizar os riscos e garantir o desenvolvimento seguro, confiável e controlável da inteligência artificial. (China, 2017. p3).

Privacidade

A inteligência artificial é uma tecnologia disruptiva com uma ampla gama de impactos. Pode trazer problemas como mudar a estrutura do emprego, impactar as leis e a ética social,

infringir a privacidade pessoal e desafiar as normas das relações internacionais. Terá um impacto profundo sobre gestão governamental, segurança econômica, estabilidade social e até mesmo governança global. Enquanto desenvolvemos vigorosamente a inteligência artificial, devemos atribuir grande importância a possíveis riscos e desafios de segurança, fortalecer a prevenção voltada para o futuro e a orientação de restrição, minimizar os riscos e garantir o desenvolvimento seguro, confiável e controlável da inteligência artificial. (China, 2017, p3).

Teoria avançada de aprendizado de máquina. Pesquise teorias básicas de aprendizagem estatística, raciocínio de incerteza e tomada de decisão, aprendizagem distribuída e interação, oculta Teorias de aprendizagem e modelos eficientes, como aprendizagem com preservação de privacidade, aprendizagem com pequenas amostras, aprendizagem por reforço profundo, aprendizagem não supervisionada, aprendizagem semissupervisionada e aprendizagem ativa. (China, 2017, p6).

Construir um sistema de infraestrutura inteligente onipresente, seguro e eficiente. Promover vigorosamente a construção de infraestrutura de informação inteligentes, melhorar o nível de inteligência das infraestruturas tradicionais e formar um sistema de infraestrutura que satisfaça as necessidades da economia inteligente, da sociedade inteligente e da construção da defesa nacional. Acelerar a transformação da infraestrutura de informação digital e em rede com a transmissão de informações como núcleo para uma infraestrutura de informação inteligente que integre detecção, transmissão, armazenamento, computação e processamento. Otimizar e atualizar a infraestrutura de rede, desenvolver e implantar o sistema de comunicações móveis (5G) de quinta geração, melhorar a infraestrutura da Internet das Coisas, acelerar a construção de redes de informação integradas espaço-solo e melhorar as capacidades de transmissão de baixa latência e alto rendimento. Coordenar o uso da infraestrutura de big data, fortalecer a segurança dos dados e a proteção da privacidade e fornecer suporte massivo de dados para pesquisa e desenvolvimento de inteligência artificial e aplicação generalizada. Construir infraestruturas de computação de alta eficiência e melhorar as capacidades de suporte de serviços dos centros de supercomputação para aplicações de inteligência artificial. Construir uma Internet energética distribuída e eficiente para formar um apoio. (China, 2017, p11).

Estabelecer padrões de tecnologia de inteligência artificial e sistemas de propriedade intelectual. Reforçar a investigação sobre o sistema-quadro padrão de inteligência artificial. Aderir aos princípios de segurança, usabilidade, interoperabilidade e rastreabilidade e estabelecer e melhorar gradualmente padrões técnicos para uniformidade básica de inteligência artificial, interconexão, aplicações industriais, segurança de rede e proteção de privacidade. (China, 2017, p13).

Participação da sociedade civil na elaboração da regulamentação

Implementar profundamente a estratégia de desenvolvimento de integração militar-civil e promover a formação de um padrão de integração militar-civil de inteligência artificial multifatorial, multicampo e de alta eficiência. Implantar a investigação e o desenvolvimento de teorias básicas e tecnologias comuns chave de uma nova geração de inteligência artificial com a orientação da partilha militar e civil, e estabelecer um mecanismo normalizado de

comunicação e coordenação entre institutos de investigação científica, universidades, empresas e unidades industriais militares. Promover a transformação militar-civil bidirecional da tecnologia de inteligência artificial, fortalecer o forte apoio da tecnologia de inteligência artificial de nova geração para tomada de decisões de comando, exercícios militares, equipamentos de defesa nacional, e orientar a transformação e aplicação de inteligência artificial científica e conquistas tecnológicas no campo da defesa nacional para o campo civil. Incentivar as forças superiores de investigação científica civil a participarem nas principais tarefas de inovação científica e tecnológica de inteligência artificial no domínio da defesa nacional e promover a rápida integração de várias tecnologias de inteligência artificial no domínio da inovação da defesa nacional. (China, 2017, p11).

Garantia de Direitos

Formular leis, regulamentos e normas éticas para promover o desenvolvimento da inteligência artificial. Reforçar a investigação sobre questões jurídicas, éticas e sociais relacionadas com a inteligência artificial e estabelecer um quadro jurídico, regulamentar e ético para garantir o desenvolvimento saudável da inteligência artificial. Realizar pesquisas sobre questões jurídicas, como confirmação de responsabilidade civil e criminal, proteção de privacidade e direitos de propriedade e utilização de segurança da informação relacionada a aplicações de inteligência artificial, estabelecer um sistema de rastreabilidade e responsabilização e esclarecer as questões jurídicas da inteligência artificial e direitos e obrigações relacionados e responsabilidades. (China, 2017, p13).

Educação e capacitação profissional

Capacitação técnica / Programas de treinamento

Formar uma série de bases mundiais de inovação tecnológica e treinamento de talentos em inteligência artificial e construir uma lei de inteligência artificial mais completa leis, regulamentos, ética e sistemas políticos. (China, 2017, p4).

Oportunidades de carreira em IA

Basear, fortalecer a equipe de talentos de inteligência artificial de ponta, promover a interação colaborativa entre entidades inovadoras e formar capacidades contínuas de inovação em inteligência artificial. (China, 2017, p4).

Estudo de impactos profissionais

A inteligência artificial é uma tecnologia disruptiva com uma ampla gama de impactos. Pode trazer problemas como mudar a estrutura do emprego, impactar as leis e a ética social,

infringir a privacidade pessoal e desafiar as normas das relações internacionais. (China, 2017, p3).

Infraestrutura

Computação de alto desempenho

A teoria da computação quântica inteligente concentra-se em avanços nos métodos de aprendizado de máquina acelerado quântico, estabelecendo um modelo híbrido de computação de alto desempenho e algoritmos quânticos e formando uma arquitetura de sistema de inteligência artificial quântica eficiente, precisa e autônoma(China, 2017, p4).

Recurso de dados

Processamento de informações chinês, monitoramento inteligente, identificação biométrica, robôs industriais, serviços robôs, a condução não tripulada está gradualmente a entrar em aplicações práticas, a inovação da inteligência artificial e o empreendedorismo estão a tornar-se cada vez mais ativos, e uma série de empresas líderes aceleraram o seu crescimento e ganharam ampla atenção e reconhecimento internacional. A acumulação acelerada de capacidades técnicas, a combinação orgânica de enormes recursos de dados. (China, 2017, p3).

Tecnologias em nuvem

Infraestrutura computacional de alto desempenho. Continuar a fortalecer a construção de infraestruturas de supercomputação, infraestruturas de computação distribuída e centros de computação em nuvem. (China, 2017, p12).

Acordos Internacionais

Diálogo internacional

Formular ética e códigos de conduta para desenvolvedores e designers de produtos de inteligência artificial, fortalecer a avaliação dos potenciais danos e benefícios da inteligência artificial e construir soluções para emergências em cenários complexos de inteligência artificial. Participar ativamente na governança global da inteligência artificial, fortalecer a pesquisa sobre as principais questões internacionais comuns em inteligência artificial, como alienação de robôs e supervisão de segurança, aprofundar a cooperação internacional em leis e regulamentos de inteligência artificial, regras internacionais, e responder conjuntamente aos desafios globais. (China, 2017, p13).

Parcerias internacionais.

Não consta.

4.5 Brasil

De acordo com informações disponíveis no site do Governo Brasileiro, foram analisados três documentos principais que abordam do uso de tecnologias de Inteligência Artificial no país: Projeto de Lei (PL) 21/2020 “**Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil**”, Projeto de Lei 2338/2023 “**Dispõe sobre o uso da Inteligência Artificial**” de 2023 e documento “**Estratégia Brasileira de Inteligência Artificial**” (EBIA) de 2021. (Brasil, 2021).

O documento do Projeto de Lei 21/2020 traz o texto em 11 páginas onde é apresentado uma proposta de lei que estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da Inteligência Artificial no Brasil. Este projeto foi aprovado na Câmara dos Deputados em regime de urgência, e traz um modelo de regulamentação transversal. A aceleração recebeu críticas devido à ausência de tempo hábil e debate junto a setores da sociedade brasileira. (Nogueira, 2021).

Devido a essa situação, uma comissão de juristas foi instaurada no Senado Federal em 2022, que ficou responsável pela elaboração de um projeto substitutivo para a regulamentação da Inteligência Artificial no Brasil.

Destacam-se pontos como a definição de sistemas de IA, os objetivos da aplicação da IA no país, os fundamentos para seu desenvolvimento, e princípios orientadores, como a busca por resultados benéficos para a humanidade, a centralidade do ser humano, a não discriminação, a transparência, entre outros.

A lei propõe diretrizes para o poder público, enfatiza a gestão baseada em risco, e a participação social na elaboração de normas relacionadas à IA. Também inclui responsabilidades, regulamentações setoriais e a promoção da cooperação internacional.

Em 2021, o governo brasileiro publicou oficialmente o documento “**Estratégia Brasileira de Inteligência Artificial**” (EBIA), com 52 páginas, no qual encontramos descrições sobre importantes diretrizes em relação a estratégia do país no desenvolvimento de tecnologia de inteligência artificial.

São disponibilizadas ações estratégicas para as legislações, regulação e uso ético, governança de IA, aspectos internacionais, qualificação para um futuro digital, força de trabalho e capacitação, pesquisa, desenvolvimento inovação e empreendedorismo, aplicações nos

setores produtivos, aplicações no poder público e segurança pública (BRASIL, 2021). A abordagem dos documentos aparece de forma transversal, porém já existem projetos utilizando a IA de forma setorial.

O Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), definiu como prioridade a área de Inteligência Artificial, no que se refere a projetos de pesquisa, de desenvolvimento de tecnologias e inovações, para o período 2020 a 2023, onde foi elaborado o documento EBIA. O documento descreve ações que estimulam pesquisa, inovação e desenvolvimento dessas tecnologias, bem como o uso consciente e ético protegendo a sociedade para um futuro melhor, e tem como objetivo potencializar o desenvolvimento e a utilização da tecnologia a promover o avanço científico no país.

Ao analisar o documento, encontra-se preocupações do país com a competitividade e a produtividade bem como a melhoria da qualidade de vida e a redução da desigualdade social.

A estratégia foi construída a partir de contratação de uma consultoria especializada em IA, com referências nacionais e internacionais e processo de consulta pública entre os anos 2019 e 2020. A partir do projeto de cooperação técnica internacional junto a UNESCO, o estudo do documento tem objetivo entender os potenciais impactos sociais e econômicos das tecnologias de Inteligência artificial, bem como seus efeitos negativos e a maximização dos efeitos positivos na sociedade.

São descritas ações para atrair empresas que desenvolvam projetos de pesquisa, desenvolvimento e inovação (PD&I), parcerias internacionais em IA e uso de IA em segmentos da sociedade brasileira, inclusive focada em pesquisas acadêmicas.

O documento segue diretrizes da OCDE e descreve a citação do secretário geral Angel Gurría “A Inteligência Artificial está a revolucionar a forma como vivemos e trabalhamos, oferecendo benefícios extraordinários para as nossas sociedades e economias” (Brasil, 2021).

São encontrados objetivos claros descritos no documento da estratégia, como contribuir para a elaboração de princípios éticos para o desenvolvimento e uso de IA responsáveis; promover investimentos sustentados em pesquisa e desenvolvimento em IA; remover barreiras à inovação em IA; capacitar e formar profissionais para o ecossistema da IA; estimular a inovação e o desenvolvimento da IA brasileira em ambiente internacional; promover ambiente de cooperação entre os entes públicos e privados, a indústria e os centros de pesquisas para o desenvolvimento da Inteligência Artificial (Brasil, 2021).

Além das estratégias, o documento dispõe de *ranking* nacionais e internacionais de uso de inteligência artificial, onde segue o desenvolvimento de políticas e estratégias internacionais, como pesquisa e desenvolvimento, competências profissionais e o futuro do trabalho, industrialização, padrões éticos para IA, governança de dados, infraestrutura digital e bem-estar social.

No referido documento, destacam-se as principais iniciativas do governo brasileiro relacionadas diretamente ao uso de tecnologias de inteligência artificial. Entre essas iniciativas encontram-se a Lei Geral de Proteção de Dados, a política de dados abertos¹⁹ do Poder Executivo Federal e a portaria referente ao uso de software público brasileiro. O documento aborda também questões como vieses algorítmicos, incentivando a implementação de sistemas de IA responsáveis, enfatizando a importância de princípios para uma IA centrada nos humanos e ressaltando a necessidade de interação humana nos processos.

O documento evidencia preocupações primordiais, como proteger os direitos fundamentais e garantir a implementação de sistemas seguros e confiáveis, visando o benefício da pessoa humana, a preservação do regime democrático e o estímulo ao desenvolvimento científico e tecnológico. O projeto enfatiza a necessidade de garantir direitos para proteção usuário impactado pela Inteligência Artificial, além de prever ferramentas de governança e um arranjo institucional de fiscalização e supervisão para o desenvolvimento tecnológico.

Em 2023, ocorreu a publicação do Projeto de Lei 2338/2023, sendo esta uma incorporação de projetos anteriores que tratam da regulamentação do desenvolvimento, implementação e uso responsável de sistemas de inteligência artificial (IA) no Brasil, com a participação de diversos setores da sociedade e a Comissão de Juristas. Esse projeto segue a linha da proposta da União Europeia, onde cita detalhes de obrigações e responsabilidades de fornecedores de tecnologia de Inteligência Artificial.

O documento intitulado "Marco Legal da IA no Brasil" é composto por 31 páginas, adota uma abordagem mista de regulamentação. Ele estabelece princípios, direitos e responsabilidades para agentes de IA, como fornecedores e operadores, e define categorias de riscos associados a esses sistemas. (Müller, 2023)

¹⁹ Política de dados abertos - define regras para promover a abertura de dados governamentais no âmbito dos órgãos e entidades federais, como ministérios, autarquias, agências reguladoras e fundações públicas. Tendo como pilar as disposições da Lei de Acesso à Informação. BRASIL, Política de Dados Abertos. Disponível em: <https://dados.gov.br/dados/conteudo/politica-de-dados-abertos>. Acesso em dezembro de 2023.

O texto trata também detalhes sobre direitos das pessoas afetadas por sistemas de inteligência artificial, incluindo o direito à informação prévia, explicação das decisões, contestação, participação humana, não-discriminação e privacidade. Além de descrever questões quanto a responsabilidade civil por partes dos fornecedores e operadores de tecnologia de IA.

No documento são encontradas descrições sobre fiscalização e coordenação de órgãos reguladores específicos. É abordado a criação e um ambiente regulatório experimental para incentivar inovação tecnológica no país.

No projeto de lei aparece enfatizado a valorização do trabalho e a dignidade humana como ponto central que deve andar ao lado das inovações tecnológicas.

4.5.1. Quadro de diretrizes encontradas em documento analisado

Os trechos abaixo foram encontrados e retirados integralmente dos documentos oficiais, dispostos após análise, para comprovação das diretrizes pesquisadas e confecção do Quadro 4 e Quadro5 de análise comparativa.

Regulamentação ética

IA confiável

Contudo, a ideia de *accountability* (aqui traduzida como responsabilidade e a prestação de contas) impõe que, a depender da aplicação de IA e dos riscos a ela associados, sejam estabelecidas estruturas de governança de IA, que possam assegurar a adoção de princípios para IA confiável e implementar mecanismos para sua observância. Tais mecanismos podem incluir: (i) a designação de indivíduos ou de grupos específicos dentro da organização para promover a conformidade com os princípios; (ii) a adoção de medidas para aumentar a conscientização interna sobre a necessidade dessa conformidade, inclusive por meio de orientações e treinamentos em toda a empresa; e (iii) a implementação de um processo de escalação por meio do qual os funcionários possam levantar preocupações de conformidade e resolver essas preocupações. Podem, ainda, envolver a criação de selos, certificações e códigos de conduta corporativos ou governamentais. (Brasil, 2021, p24).

A Recomendação do Conselho da OCDE sobre IA (2019) sugere que Estados nacionais devam promover e incentivar investimentos públicos e privados em P&D, contemplando esforços interdisciplinares para promover inovação em IA confiável, de modo a focar não apenas em desafios técnicos, mas também em implicações sociais, jurídicas e éticas

associadas a IA. Além disso, o documento recomenda que governos viabilizem investimentos em bases de dados abertas, que sejam representativas e respeitem direitos à privacidade e à proteção de dados pessoais, de modo a: (i) promover um ambiente para pesquisa e desenvolvimento em IA que seja livre de viés; e (ii) aperfeiçoar a interoperabilidade e o uso de padrões comuns. Para além desses aspectos, a Recomendação indica que governos devem promover um ambiente de políticas públicas que apoie uma transição ágil da fase de P&D para a fase de desenvolvimento e operação de sistemas de IA. (Brasil, 2021, p35).

A Pesquisa e Desenvolvimento da Inteligência Artificial deve adotar abordagens éticas de *design* 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, prover explicação e previsibilidade e tornar o sistema mais rastreável, auditável e responsável. (Brasil, 2021, p35).

Privacidade

VIII – a privacidade, a proteção de dados e a autodeterminação informativa; (Lei nº 2338/2023, p2).

VI – direito à privacidade e à proteção de dados pessoais, nos termos da legislação pertinente. (Lei nº 2338/2023, p4).

IV – legitimação do tratamento de dados conforme a legislação de proteção de dados, inclusive por meio da adoção de medidas de privacidade desde a concepção e por padrão e da adoção de técnicas que minimizem o uso de dados pessoais; (Lei nº 2338/2023, p13).

VIII – a segurança, a privacidade e a proteção de dados pessoais; (Lei nº 21/20, p2).

II – centralidade do ser humano: respeito à dignidade humana, à privacidade, à proteção de dados pessoais e aos direitos fundamentais, quando o sistema tratar de questões relacionadas ao ser humano; (Lei nº 21/20, p4).

A Inteligência Artificial oferece uma imensa oportunidade de beneficiar pessoas em todo o mundo. Mas, à medida em que a tecnologia da IA impacta cada vez mais a sociedade, é preciso também assegurar a utilização e desenvolvimento responsável de IA, assegurando que as pessoas sejam tratadas de maneira justa, com proteção de sua segurança e respeito a sua privacidade. (BRASIL, 2021, p17).

Participação da sociedade civil na elaboração da regulamentação

A presente 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 EBIA foi concebida durante os anos de 2019 e de 2020 a partir de visões diversas e da convergência de estudos, reflexões, pesquisas e de consulta aos especialistas, empresas, pesquisadores e órgãos públicos, tendo como propósito o planejamento tecnológico de longo prazo para o país. (BRASIL, 2021, p5).

A referida Comissão realizou uma série de audiências públicas, além de seminário internacional, ouvindo mais de setenta especialistas sobre a matéria, representantes de diversos segmentos: sociedade civil organizada, governo, academia e setor privado. Abriu ainda oportunidade para a participação de quaisquer interessados, por meio de contribuições escritas, tendo recebido 102 manifestações, individualmente analisadas e organizadas de acordo com suas propostas. Finalmente, a Comissão demandou à Consultoria Legislativa do Senado Federal estudo sobre a regulamentação da inteligência artificial em mais de trinta países integrantes da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), o que permitiu analisar o panorama normativo mundial da matéria. (Lei nº 2338/2023, p28).

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; (Lei nº 21/20, p9).

Garantia de Direitos

Art. 1º Esta Lei estabelece normas gerais de caráter nacional para o desenvolvimento, implementação e uso responsável de sistemas de inteligência artificial (IA) no Brasil, com o objetivo de 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. (Lei nº 2338/2023, p1).

II – o respeito aos direitos humanos e aos valores democráticos; (Lei nº 2338/2023, p1).

V – a igualdade, a não discriminação, a pluralidade e o respeito aos direitos trabalhistas; (Lei nº 2338/2023, p1).

VI – discriminação: qualquer distinção, exclusão, restrição ou preferência, em qualquer área da vida pública ou privada, cujo propósito ou efeito seja anular ou restringir o reconhecimento, gozo ou exercício, em condições de igualdade, de um ou mais direitos ou liberdades previstos no ordenamento jurídico, em razão de características pessoais como origem geográfica, raça, cor ou etnia, gênero, orientação sexual, classe socioeconômica, idade, deficiência, religião ou opiniões políticas; (Lei nº 2338/2023, p3).

VII – discriminação indireta: discriminação que ocorre quando normativa, prática ou critério aparentemente neutro tem a capacidade de acarretar desvantagem para pessoas pertencentes a grupo específico, ou as coloquem em desvantagem, a menos que essa normativa, prática ou critério tenha algum objetivo ou justificativa razoável e legítima à luz do direito à igualdade e dos demais direitos fundamentais; (Lei nº 2338/2023, p3).

CAPÍTULO II DOS DIREITOS

Seção I Disposições Gerais

Art. 5º Pessoas afetadas por sistemas de inteligência artificial têm os seguintes direitos, a serem exercidos na forma e nas condições descritas neste Capítulo:

I – direito à informação prévia quanto às suas interações com sistemas de inteligência artificial;

II – direito à explicação sobre a decisão, recomendação ou previsão tomada por sistemas de inteligência artificial;

III – direito de contestar decisões ou previsões de sistemas de inteligência artificial que produzam efeitos jurídicos ou que impactem de maneira significativa os interesses do afetado;

IV – direito à determinação e à participação humana em decisões de sistemas de inteligência artificial, levando-se em conta o contexto e o estado da arte do desenvolvimento tecnológico;

V – direito à não-discriminação e à correção de vieses discriminatórios diretos, indiretos, ilegais ou abusivos; e

VI – direito à privacidade e à proteção de dados pessoais, nos termos da legislação pertinente.

Parágrafo único. Os agentes de inteligência artificial informarão, de forma clara e facilmente acessível, os procedimentos necessários para o exercício dos direitos descritos no *caput*. (Lei nº 2338/2023, p4)

Art. 6º A defesa dos interesses e dos direitos previstos nesta Lei poderá ser exercida perante os órgãos administrativos competentes, bem como em juízo, individual ou coletivamente, na forma do disposto na legislação pertinente acerca dos instrumentos de tutela individual, coletiva e difusa.

Seção II

Dos direitos associados a informação e compreensão das decisões tomadas por sistemas de inteligência artificial

Art. 7º Pessoas afetadas por sistemas de inteligência artificial têm o direito de receber, previamente à contratação ou utilização do sistema, informações claras. (Lei nº 2338/2023, p5).

§ 3º Os sistemas de inteligência artificial que se destinem a grupos vulneráveis, tais como crianças, adolescentes, idosos e pessoas com deficiência, serão desenvolvidos de tal modo que essas pessoas consigam entender seu funcionamento e seus direitos em face dos agentes de inteligência artificial. (Lei nº 2338/2023, p6).

Seção III

Do direito de contestar decisões e de solicitar intervenção humana **Art. 9º** A pessoa afetada por sistema de inteligência artificial terá o direito de contestar e de solicitar a revisão de decisões, recomendações ou previsões geradas por tal sistema que produzam efeitos jurídicos relevantes ou que impactem de maneira significativa seus interesses.

§ 1º Fica assegurado o direito de correção de dados incompletos, inexatos ou desatualizados utilizados por sistemas de inteligência artificial, assim como o direito de solicitar a anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade com a legislação.

§ 2º O direito à contestação previsto no *caput* deste artigo abrange também decisões, recomendações ou previsões amparadas em inferências discriminatórias, irrazoáveis ou que atentem contra a boa-fé objetiva, assim compreendidas as inferências que:

I – sejam fundadas em dados inadequados ou abusivos para as finalidades do tratamento;

II – sejam baseadas em métodos imprecisos ou estatisticamente não confiáveis; ou

III – não considerem de forma adequada a individualidade e as características pessoais dos indivíduos. (Lei nº 2338/2023, p6).

Os sistemas de IA devem ser projetados de maneira a respeitar o Estado de Direito, os direitos humanos, os valores democráticos e a diversidade, e devem incluir salvaguardas apropriadas - possibilitando a intervenção humana sempre que necessário - para garantir uma sociedade justa. (BRASIL, 2021, p6).

Governança de Dados e Infraestrutura Digital: financiamento para parcerias que envolvam a utilização de dados abertos, plataformas de desenvolvimento compartilhado de *software* de IA e conjuntos de dados, bem como o compromisso para criar ambientes de testes de proteção aos direitos do cidadão. (Brasil, 2021, p15).

O desenvolvimento tecnológico da Inteligência Artificial tem sido acompanhado de intensas discussões acerca da necessidade de desenvolvimento de parâmetros jurídicos, regulatórios e éticos para orientar o desenvolvimento e aplicação da tecnologia. No centro de tais debates encontra-se a preocupação em estabelecer um ponto de equilíbrio entre: (i) a proteção e a salvaguarda de direitos, inclusive aqueles associados à proteção de dados pessoais e à prevenção de discriminação e viés algorítmico; (ii) a preservação de estruturas adequadas de incentivo ao desenvolvimento de uma tecnologia cujas potencialidades ainda não foram plenamente compreendidas; e (iii) o estabelecimento de parâmetros legais que confirmem segurança jurídica quanto à responsabilidade dos diferentes atores que participam da cadeia de valor de sistemas autônomos. (Brasil, 2021, p17).

Um dos principais atributos da IA é a capacidade de tomar decisões sem intervenção humana. Um dos pontos de discussão mais importantes, nesse sentido, é o de avaliar quais tipos de decisão podem ser delegadas a uma máquina e quais exigiriam, necessariamente, a intervenção humana. A LGPD endereça tal questão dispondo sobre o direito de indivíduos solicitarem a revisão das decisões tomadas apenas com base no processamento automatizado de dados pessoais, quando estas afetam seus interesses. Isso inclui decisões que visam definir o perfil pessoal, profissional, de consumidor e crédito, bem como quaisquer aspectos da personalidade dos titulares de dados.

Outro aspecto importante de tal discussão diz respeito à ideia de que os indivíduos devem ter ciência de suas interações com sistemas de IA. De fato, a informação aos indivíduos quanto à existência de processos de tomada de decisão baseados em IA caracteriza-se como pressuposto para o exercício do direito de revisão de decisões automatizadas previsto na LGPD. (Brasil, 2021, p19).

III – o respeito à ética, aos direitos humanos e aos valores democráticos;

IV – a livre. (Lei nº 21/20, p1).

V – a não discriminação, a pluralidade, o respeito às diversidades regionais, a inclusão e o respeito aos direitos e garantias fundamentais do cidadão; (Lei nº 21/20, p2)

II – centralidade do ser humano: respeito à dignidade humana, à privacidade, à proteção de dados pessoais e aos direitos fundamentais, quando o sistema tratar de questões relacionadas ao ser humano; (Lei nº 21/20, p2).

V – transparência: direito das pessoas de serem informadas de maneira clara, acessível e precisa sobre a utilização das soluções de inteligência artificial, salvo disposição legal em sentido contrário e observados os segredos comercial e industrial. (Lei nº 21/20, p4).

VIII – disponibilidade de dados: não violação do direito de autor pelo uso de dados, de banco de dados e de textos por ele protegidos, para fins de treinamento de sistemas de inteligência artificial, desde que não seja impactada a exploração normal da obra por seu titular. (Lei nº 21/20, p5).

Educação e capacitação profissional

Capacitação técnica / Programas de treinamento

Entre tais iniciativas, os principais eixos identificados são destacados a seguir:

- Pesquisa e Desenvolvimento: atrair, reter e treinar talentos em IA do próprio país ou estrangeiros, com financiamentos acadêmicos, programas de bolsas de estudo e a criação de programas específicos de mestrado e doutorado em IA; criar novos centros ou programas em pesquisa básica e aplicada, específicos para a IA.
- Competências profissionais e o futuro do trabalho: iniciativas para capacitar a força de trabalho, em geral, que desenvolvam habilidades para o futuro do trabalho, como investimento em educação ao longo da vida e habilidades digitais. (Brasil, 2021, p14).

Estabelecer parcerias com o setor privado e com a academia para definir políticas públicas que incentivem a formação e a capacitação de profissionais, considerando as novas realidades de mercado de trabalho.

- Estimular que as empresas e os órgãos públicos implementem programa de treinamento contínuo da sua força de trabalho voltado às novas tecnologias.
- Criar campanhas de conscientização sobre a importância de se preparar para o desenvolvimento e uso ético da IA.
- Estimular a retenção de talentos especializados em TIC no Brasil.
- Estimular a composição diversificada de equipes de desenvolvimento em IA, quanto ao gênero, raça, orientação sexual e outros aspectos socioculturais. Reforçar políticas voltadas à educação continuada e ao lifelong learning, promovendo maior interação entre o setor privado e as instituições de ensino (universidades, institutos de pesquisa e de capacitação profissional e técnica). (BRASIL, 2021, p33).

VIII – disponibilidade de dados: não violação do direito de autor pelo uso de dados, de banco de dados e de textos por ele protegidos, para fins de treinamento de sistemas de inteligência artificial, desde que não seja impactada a exploração normal da obra por seu titular. (Lei nº 21/20, p5).

g) treinamento e ações de conscientização dos riscos associados ao sistema de inteligência artificial; (Lei nº 2338/2023, p17).

Oportunidades de carreira em IA

Diante deste cenário, pode-se ressaltar que novos empregos, assim como muitos empregos existentes, exigirão novas habilidades. A capacitação, e a sua continuidade ao longo do tempo, tem função bastante importante de preparar a força de trabalho, não somente em vista

da criação de novas funções e carreiras, como também para melhor adequar os profissionais atuais para as mudanças exigidas em razão do uso da tecnologia. (BRASIL, 2021, p32).

Estudo de impactos profissionais

Ainda no que diz respeito ao mercado de trabalho, a Fast Company listou sete tipos de carreiras que possuem baixa chance de serem substituídas pela automatização: empregos que requerem criatividade, como artistas e músicos; trabalhos que envolvam resolução de problemas; cabeleireiros; psicólogos e profissões que demandem um trabalho social; professores; trabalhadores da área da saúde e cuidadores. (BRASIL, 2021, p14).

Estruturar e implementar programas de aprendizagem de adultos e de requalificação profissional para lidar com o excesso de mão de obra que resulta da automação de atividades nos campos público e privado, também fazem parte do rol de preocupações no campo das políticas de trabalho e qualificação profissional relacionadas à Inteligência Artificial. (BRASIL, 2021, p32).

Infraestrutura

Computação de alto desempenho

Não consta.

Resumo de dados

V – adoção de parâmetros adequados de separação e organização dos dados para treinamento, teste e validação dos resultados do sistema; (Lei nº 2338/2023, p13).

Tecnologias em nuvem

Não consta.

Acordos Internacionais

Diálogo internacional

O Brasil deve adotar uma postura proativa e propositiva no âmbito internacional, fomentando discussões, iniciativas e parcerias sobre Inteligência Artificial em organismos e foros internacionais, bem como em discussões e negociações entre países e blocos. É relevante, também, promover a construção de plataformas de cooperação para trocas de informação sobre Inteligência Artificial e reforçar a participação em oportunidades de cooperação advindas da intensificação das relações entre os países do BRICS (Brasil, Rússia, Índia, China e África do Sul), IBAS (Índia, Brasil e África do Sul) e Mercosul. Também é importante pontuar o trabalho sobre IA desenvolvido pela ONU e na Comissão Econômica para a América Latina e o Caribe (CEPAL). (BRASIL, 2021, p26).

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. (Lei nº 21/20, p9).

Parcerias internacionais

Estimular a inovação e o desenvolvimento da IA brasileira em ambiente internacional. (BRASIL, 2021, p7).

V – Promover ações de cooperação com autoridades de proteção e de fomento ao desenvolvimento e à utilização dos sistemas de inteligência artificial de outros países, de natureza internacional ou transnacional; (Lei nº 2338/2023, p21).

4.6 Síntese da análise dos documentos

Os documentos foram selecionados através de semelhança de conteúdo, a facilidade na disponibilidade do documento no site dos governos, e suas citações em documentos científicos.

Na análise de conteúdo realizada, foram identificadas diretrizes em comum nos documentos coletados que refletem as preocupações manifestadas pelos países, como a Regulamentação Ética, questões relacionadas a Educação e capacitação profissional, Infraestrutura e os Acordos Internacionais, de forma detalhada

- regulamentação ética e preocupações quanto a confiabilidade, privacidade, liberdade civil e garantia de direitos existentes.
- programas de educação digital, capacitação profissional e preocupações quanto a existência de programas de capacitação técnica, alfabetização digital, estudo de impactos profissionais e oportunidades de carreira em tecnologia de inteligência artificial.
- infraestrutura e preocupações quanto a tecnologia em nuvem, recurso de dados e computação de alto desempenho
- acordos internacionais e preocupações quanto a intenção de abertura de mercado em IA, parcerias internacionais, garantias de direitos humanos, oportunidades para todos.

No decorrer deste procedimento, foram reunidas e compreendidas as preocupações presentes nos documentos, as quais estavam relacionadas aos elementos citados abaixo. Essa abordagem foi implementada em conformidade com as diretrizes preexistentes, de forma a estabelecer uma ligação íntegra e coesa entre as preocupações identificadas. A partir dessas diretrizes, foi realizada a categorização dos aspectos abordados em cada documento.

4.6.1 Categorização das preocupações existentes

A **Inteligência artificial confiável** é aquela que demonstra consistência, previsão e previsibilidade em seus resultados e ações.

A **privacidade no uso de Inteligência artificial** se constitui de informação, consentimento, coleta de dados responsável, proteção de dados, onde os direitos da sociedade são garantidos.

A **participação da sociedade civil na elaboração da regulamentação** é fundamental para garantir equidade social e legitimar decisões, onde somente com a participação da sociedade é possível garantir a inclusão diversificada, mitigar preconceitos e promover a igualdade. A participação educa e conscientiza a comunidade sobre desafios e também oportunidades relacionadas à IA, além de fortalecer a aceitação pública das regulamentações, tornar legítimas e eficazes.

No caso da **garantia de direitos**, se faz necessário assegurar que a IA seja utilizada de maneira ética, justa e responsável, garantindo direito individual e coletivo das pessoas, este ainda é um desafio complexo, porém essencial para que o uso de tecnologia de IA venha a beneficiar a sociedade.

A preocupação com **capacitação técnica/ programas de treinamentos e oportunidades de carreira em IA**, bem como **estudo de impactos profissionais**, diz respeito à preparação dos indivíduos de uma sociedade quanto à constante evolução da inteligência artificial. A preocupação na educação continuada, inclusão digital e carreiras em IA são assuntos que devem estar em pauta na sociedade pois devem ser acessíveis e trazer impacto profissional, promovem o desenvolvimento e bem-estar econômico para os indivíduos.

A **computação de alto desempenho e recurso de dados**, refere-se ao uso de sistemas computacionais com grande poder de capacidade e processamento de dados, muito utilizados em sistemas de inteligência artificial. Esses recursos são essenciais para alimentar os algoritmos para aprendizado de máquina.

O uso de **tecnologia em nuvem** se dá pela escalabilidade e flexibilidade no uso dos recursos processados pela IA, uma vez que fornece infraestrutura para implantação de modelos de IA, armazenamento e processamento de grandes volumes de dados.

As **parcerias internacionais** envolvem a colaboração entre países, instituições de pesquisa, empresas e organizações no que se refere ao uso e futuro da IA mundial, sendo esse assunto de interesse global. O **diálogo internacional** é importante para estabelecer padrões, regulamentação e diretrizes comuns que tragam garantia de ética, segurança social e responsabilidade no uso de tecnologias de IA. Envolve troca de conhecimento, pesquisa e discussões sobre questões no cenário global.

No Quadro 4 é possível visualizar como análise preliminar, a existência destas diretrizes /preocupação dos governos publicada em pelo menos um dos documentos oficiais de cada país.

A legenda **NC** (não citado) demonstra a inexistência de preocupação/ diretrizes do governo referente ao assunto, **CP** (citado parcialmente), demonstra preocupação parcial do governo referente ao assunto e pôr fim a legenda **CD** (citado) demonstra existência de citações referente a preocupações da referida diretriz analisada.

Quadro 4- Análise comparativa preliminar das diretrizes analisadas

País	Documentos	Ano	Local de publicação	Diretrizes Analisadas											
				Regulamentação Ética		Educação e capacitação profissional			Infraestrutura			Acordos Internacionais			
				Preocupações existentes nos documentos internacionais pesquisados											
				IA Confiável	Privacidade	Garantia de direitos	Capacitação técnica/ programas de treinamentos	Oportunidades de carreira em IA	Estudo de Impactos profissionais	Computação de alto desempenho	Recurso de dados	Tecnologia em nuvem	Diálogo internacional	Parcerias internacionais	
Estados Unidos	Lei de iniciativa de Inteligência Artificial de 2020	2020	Site do governo Americano	CD	CD	CD	CD	CD	CD	CD	CD	CD	CD	CD	
China	Plano de Desenvolvimento de Inteligência Artificial de Nova Geração	2017	Site do governo Chines	CD	CD	NC	CD	CD	CD	CD	CD	NC	CD	CP	

União Europeia	Regulamento de Inteligência Artificial / Proposta de resolução do parlamento Europeu	2020 / 2021	Site Parlamento Europeu	CD	CD	CD	CD	CD	CP	CP	CD	CP	CP	NC
Brasil	Projeto de Lei (PL) 21/20 / Estratégia Brasileira de Inteligência Artificial / Projeto de Lei (PL) 2338/2023	2020 / 2021/2023	Site do governo brasileiro	CD	CD	CP	CD	CD	CP	CP	NC	NC	CP	CP

Fonte: Elaborado pela autora, 2023

As legendas utilizadas para demonstração dos resultados preliminares foram:

- **NC** (não citado), onde demonstra a inexistência de preocupação/ diretrizes do governo referente ao assunto.
- **CP** (citado parcialmente), demonstra preocupação parcial do governo referente ao assunto.
- **CD** (citado) demonstra existência de citações referente a preocupações da referida diretriz analisada.

5. Resultados

Este capítulo apresenta os resultados da análise concentrada em documentos oficiais de quatro nações-chave. Dada a impossibilidade de esgotar integralmente o tema e acompanhar a evolução legislativa global sobre as iniciativas de regulamentação da Inteligência Artificial, optamos por focar em países específicos.

Os Estados Unidos foram incluídos na pesquisa, devido ao reconhecimento em literaturas de seu papel como líder mundial em inovação tecnológica e no desenvolvimento de Inteligência Artificial.

A União Europeia é destacada por ser pioneira na adoção de legislação específica para a Inteligência Artificial, aprovada pelo Parlamento Europeu em junho de 2023.

A China devido ao impacto significativo de sua revolução tecnológica no desenvolvimento de Inteligência Artificial, e sua influência internacional, inclusive no Brasil, também é objeto de análise. Por fim, o Brasil foi incluído, uma vez que é essencial compreender como a regulamentação está sendo desenvolvida em nosso país, considerando seu papel emergente no contexto global de tecnologias de Inteligência Artificial.

Para a análise dos documentos, priorizamos a escolha de fontes similares nos diferentes países, todos obtidos de sites oficiais, e foram criteriosamente selecionados com referências em documentos acadêmico-científicos. Esta análise abrange minuciosamente o conteúdo proveniente dos quatro países, resultando na síntese de seis documentos, classificados de acordo com as categorias preestabelecidas.

Ao abordar as categorias da análise, é essencial considerar a complexidade do desenvolvimento desses sistemas de Inteligência Artificial. Dentre as categorias destacadas na pesquisa, é importante ter em mente a multidisciplinaridade em todo esse processo. Como Loren Spinola (2023) nos traz, a multidisciplinaridade é fundamental para desenvolver sistemas de Inteligência Artificial mais justos e inclusivos.

A participação de especialistas de diversas áreas e suas perspectivas contribui para evitar a perpetuação de discriminações favorece a promoção da transparência na tomada de decisões automatizadas, permitindo que desenvolvedores e usuários compreendam o processo decisório e avaliem as consequências sociais dessas decisões. (Spinola, 2023)

As categorias utilizadas foram: Inteligência artificial confiável, privacidade no uso de Inteligência artificial, garantia de direitos, capacitação técnica/ programas de treinamentos, oportunidades de carreira em IA, estudo de impactos profissionais, computação de alto desempenho, recurso de dados, tecnologia em nuvem, parcerias internacionais, diálogo internacional, por último, foi incluída na análise a categoria “participação da sociedade civil na elaboração da regulamentação”.

Nesse contexto, a participação da sociedade civil na elaboração da regulamentação surge como um elemento chave. Além das áreas técnicas e especializadas, a inclusão de diferentes perspectivas contribui para minimizar discriminações e promover a transparência nas decisões automatizadas. (Fico, Alencar, Sousa, 2021). Assim, a análise dessas categorias não apenas ressalta a complexidade do campo, mas também destaca a necessidade de uma abordagem integrada, que valorize a diversidade de conhecimentos na construção de sistemas de Inteligência Artificial mais éticos e equitativos.

A inovação é impulsionada pela diversidade, diferentes perspectivas trazem novas visões. A diversidade social, combinada com variações na formação acadêmica e experiência profissional, desempenha um papel fundamental na busca de alternativas para desafios persistentes. Afinal, é impraticável para um indivíduo possuir conhecimento abrangente em todos os domínios.

Dado o envolvimento da tecnologia em múltiplas esferas do conhecimento, a Inteligência Artificial se destaca como um campo interdisciplinar de grande relevância. As diversas áreas do saber têm o potencial de desempenhar um papel decisivo no desenvolvimento de algoritmos imparciais e equitativos, essenciais para a redução de vieses na IA. À medida que a diversidade de conhecimento se amplia, surgem oportunidades para não apenas a tecnologia, mas também a sociedade como um todo. (Fico, Alencar, Sousa, 2021).

Com o avanço tecnológico e novas descobertas científicas são realizadas, emergem novas oportunidades e desafios. Essa dinâmica requer que a sociedade permaneça aberta a inovações. O contínuo estado de abertura é essencial para o aprendizado. Apesar das muitas mudanças positivas proporcionadas pela tecnologia, há sempre territórios inexplorados e descobertas a serem feitas, sugerindo um futuro dinâmico e repleto de potencial.

5.1. Apresentação dos Dados Oficiais

5.1.1. Quadro análise final

O Quadro 5 mostra de forma resumida a análise que inclui a categoria “Participação da sociedade civil na elaboração da regulamentação” nas diretrizes. Essa inclusão se baseia na necessidade de existência identificada a partir da preocupação dos governos, conforme detalhamento nos documentos oficiais.

Após a conclusão da análise preliminar, foi iniciada outra etapa da pesquisa, que consistiu na extração de partes dos documentos oficiais. Isso permitiu uma investigação mais aprofundada, resultando em uma pequena divergência nos resultados, corrigida na apresentação do Quadro 5.

Foram mantidas as legendas utilizadas no quadro 4 de análises preliminares, pela facilidade de entendimento. A legenda **NC** (não citado) demonstra a inexistência de preocupação/ diretrizes do governo referente ao assunto, **CP** (citado parcialmente), demonstra preocupação parcial do governo referente ao assunto e pôr fim a legenda **CD** (citado) demonstra existência de citações referente a preocupações da referida diretriz.

Quadro 5- Análise final das diretrizes analisadas – Quadro Resumido

País	Documentos	Ano	Local de publicação	Diretrizes Analisadas											
				Regulamentação Ética				Educação e capacitação profissional			Infraestrutura			Acordos Internacionais	
				Preocupações existentes nos documentos internacionais pesquisados											
				IA Confiável	Privacidade	Participação da sociedade civil na elaboração da regulamentação	Garantia de direitos	Capacitação técnica/ programas de treinamentos	Oportunidades de carreira em IA	Estudo de Impactos profissionais	Computação de alto desempenho	Recurso de dados	Tecnologia em nuvem	Diálogo internacional	Parcerias internacionais
Estados Unidos	Lei de iniciativa de Inteligência Artificial de 2020	2020	Site do governo Americano	CD	CD	CD	CD	CD	CD	CD	CD	CD	CD	CD	CD
China	Plano de Desenvolvimento de Inteligência Artificial de Nova Geração	2017	Site do governo Chines	CD	CD	CD	CD	CD	CD	CD	CD	CD	CP	CD	NC

União Europeia	Regulamento de Inteligência Artificial / Proposta de resolução do parlamento Europeu	2020 / 2021	Site Parlamento Europeu	CD	CD	CD	CD	NC	NC	CP	NC	CD	NC	CP	NC
Brasil	Projeto de Lei (PL) 21/20 / Estratégia Brasileira de Inteligência Artificial / Projeto de Lei (PL) 2338/2023	2020 / 2021/2023	Site do governo brasileiro	CD	CD	CD	CD	CD	CP	CD	NC	CD	NC	CD	CD

Fonte: Elaborado pela autora, 2024

As legendas utilizadas para demonstração dos resultados se mantiveram, sendo:

- **NC** (não citado), onde demonstra a inexistência de preocupação/ diretrizes do governo referente ao assunto.
- **CP** (citado parcialmente), demonstra preocupação parcial do governo referente ao assunto.
- **CD** (citado) demonstra existência de citações referente a preocupações da referida diretriz analisada.

5.2. Apresentação dos dados de resultados por referência de diretriz

Para apresentar os resultados da análise realizada, foi necessário compilar os dados finais da pesquisa, de forma a demonstrar a análise sistemática desenvolvida a partir dos documentos oficiais selecionados.

Esses dados coletados foram transformados em legendas conforme consta nos Quadros 4 e 5, apresentados anteriormente. Em seguida, foram desenvolvidas planilhas para extrair os resultados a partir das coletas, possibilitando gerar gráficos para uma compreensão e visualização mais eficazes.

A elaboração dos gráficos foi conduzida de acordo com as diretrizes estabelecidas, utilizando os dados coletados e compilados por meio das planilhas eletrônicas.

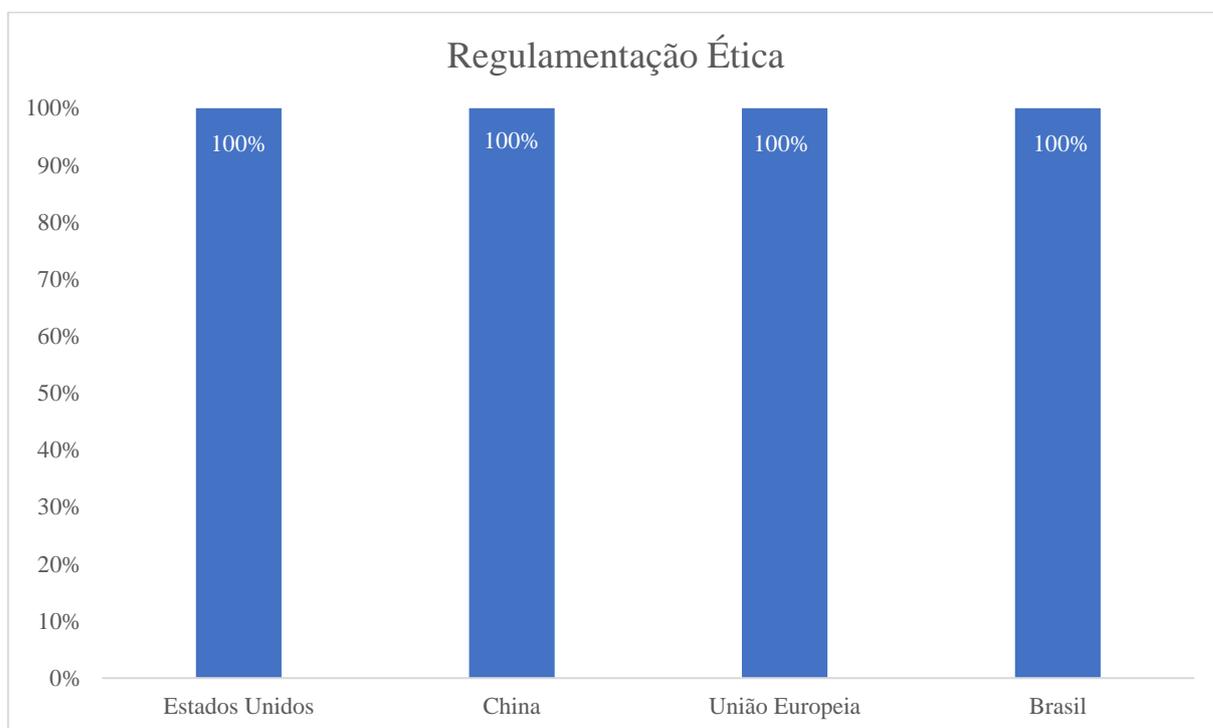
Esse processo possibilitou a apresentação dos resultados finais de forma clara e precisa.

5.2.1. Diretriz - Regulamentação Ética

O Gráfico 1 demonstra o percentual de atendimento às diretrizes pesquisadas nos documentos oficiais. A partir da análise, verifica-se que todos os países pesquisados atendem completamente às diretrizes relacionadas à Regulamentação Ética.

Nas documentações oficiais analisadas, constam preocupações referentes à IA Confiável, Privacidade, Participação da sociedade civil na elaboração da regulamentação e Garantia de Direitos, o que reflete a importância do tema.

Gráfico 1- Resultado da análise de dados referente à diretriz Regulamentação Ética



Fonte: Elaborado pela autora, 2024.

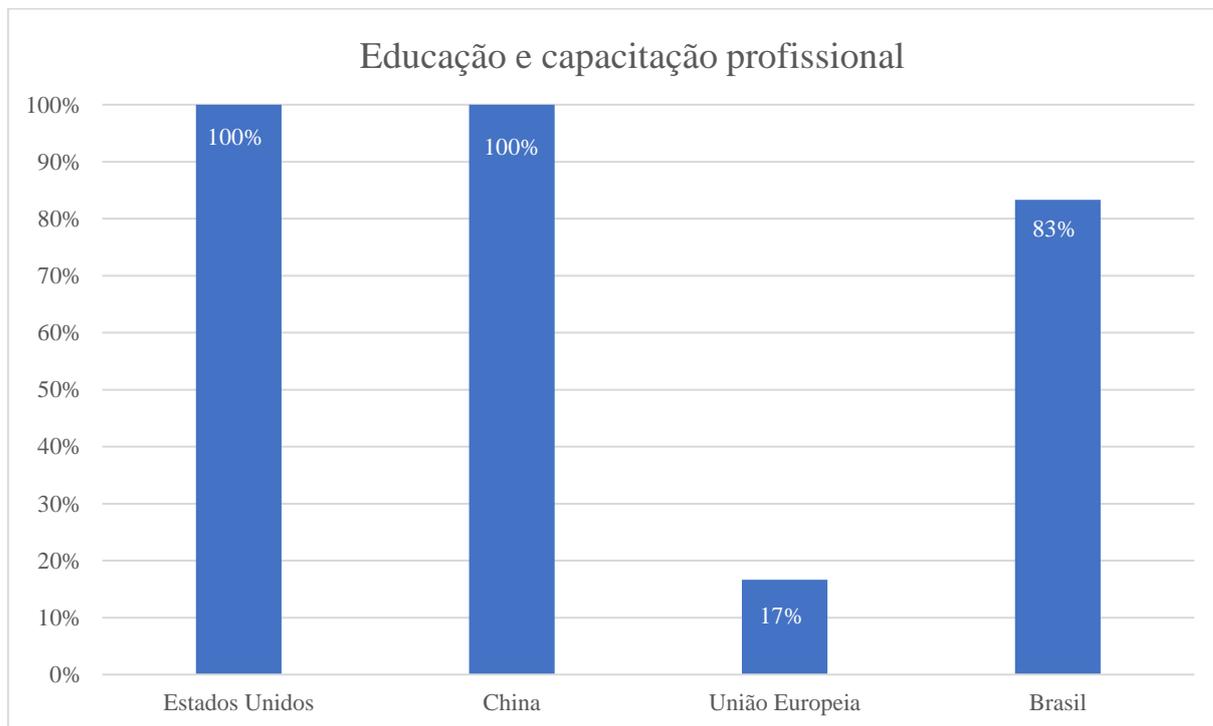
5.2.2. Educação e capacitação profissional

Nas diretrizes referente a projetos de Educação e Capacitação Profissional foram analisadas a existência de planos de capacitação técnica/programas de treinamentos, oportunidades de carreira em IA e estudo dos impactos profissionais nos países analisados.

A partir desta análise os Estados Unidos e a China possuem as diretrizes claras em seus documentos oficiais como demonstra o Gráfico 2.

No caso da União Europeia, as recomendações para esta diretriz não são diretas, pois em algumas áreas, como é o caso da Educação, cada país membro tem sua autonomia política em seus sistemas educacionais e formativos, sendo que o papel principal da União Europeia é apenas a coordenação e apoio a diversas políticas. O conteúdo e programas em geral são de responsabilidade dos estados. (Neves, 2018).

Gráfico 2 Resultado da análise de dados referente à diretriz Educação e Capacitação Profissional



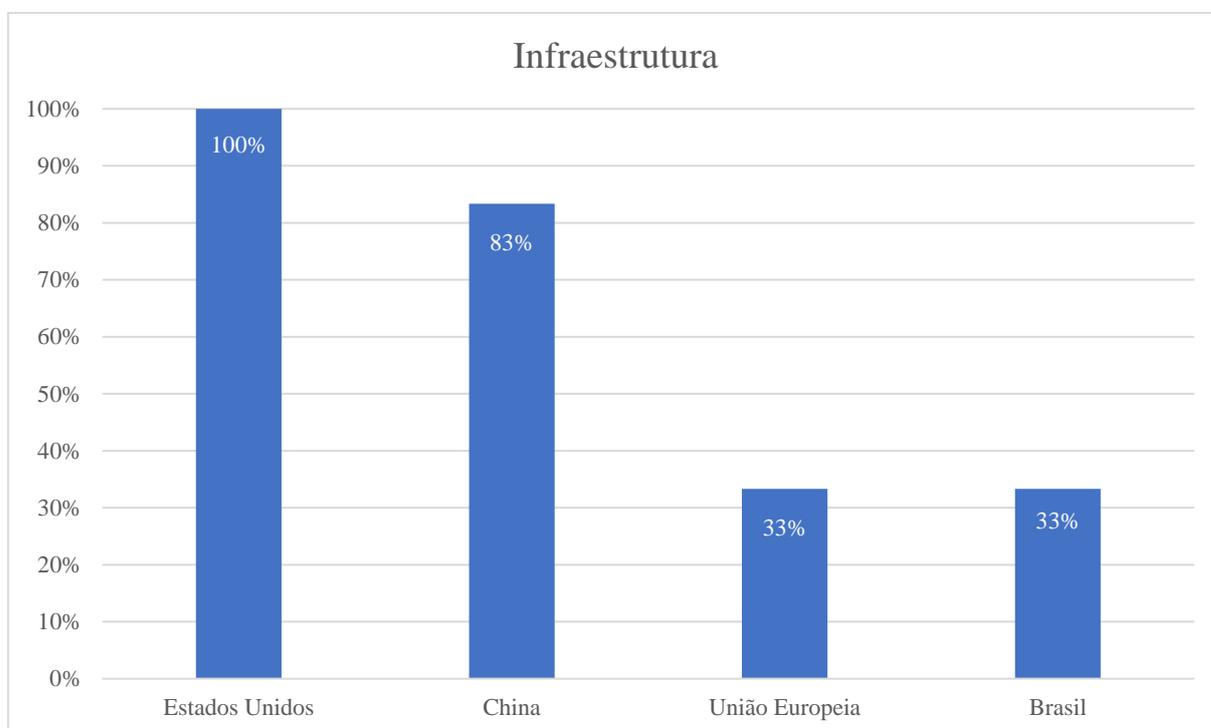
Fonte: Elaborado pela autora, 2024.

5.2.3. Infraestrutura

A análise abordou a existência e/ou preocupação em relação à infraestrutura, incluindo computação de alto desempenho na execução de tarefas intensivas em dados, como simulações de IA e análises de big data. Países com infraestrutura robusta nesse aspecto estão mais preparados para lidar com os desafios da IA em escala. Além disso, recursos de dados adequados é fundamental para o desenvolvimento e treinamento de sistemas de IA e existência e/ou planejamento de estrutura para tecnologia em nuvem, que desempenha um papel importante na facilitação da colaboração e na viabilização de soluções de IA baseadas em nuvem

A partir desta análise somente os Estados Unidos contempla completamente as diretrizes conforme demonstra o Gráfico 3.

Gráfico 3 - Resultado da análise de dados referente à diretriz Infraestrutura



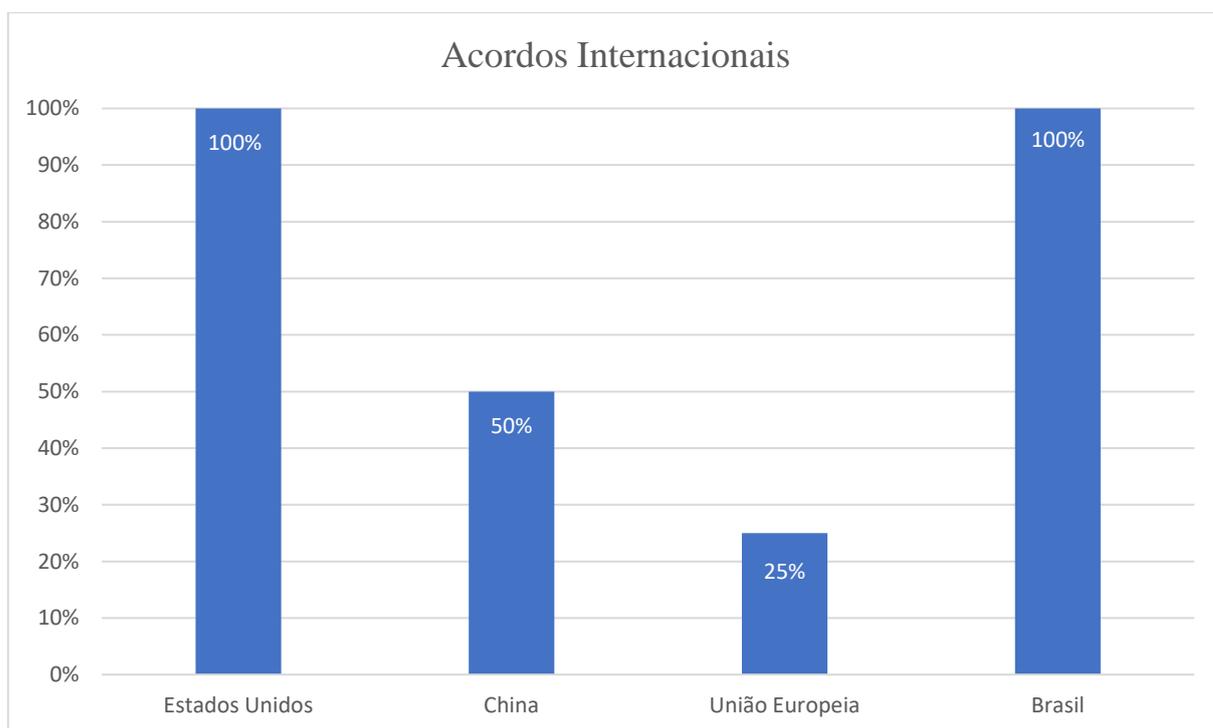
Fonte: Elaborado pela autora, 2024.

5.2.4. Acordos Internacionais

Referente aos acordos internacionais, foram analisadas questões sobre a existência/preocupação de diálogos internacionais e parcerias internacionais em sua documentação.

A partir desta análise somente os Estados Unidos e Brasil contemplam completamente as diretrizes conforme demonstra o Gráfico 4.

Gráfico 4 - Resultado da análise de dados referente à diretriz Acordos Internacionais



Fonte: Elaborado pela autora, 2024.

5.3. Relação com os Objetivos da Pesquisa:

Nesse estudo foram identificadas as principais diretrizes e desafios estabelecidos por políticas públicas nacionais e internacionais com os aspectos que esses documentos têm em comum, a partir de uma análise de conteúdo dos mesmos, onde no recorte pré-estabelecido destaca-se as principais discussões nesse campo de pesquisa, permitindo assim de uma maneira macro adquirir uma visão atualizada das abordagens regulatórias em curso.

Para isso foi necessário entendimento de conceitos sobre análise dos documentos legais que tratem das diretrizes éticas para Inteligência Artificial no Brasil, nos Estados Unidos, China, e a União Europeia.

A análise aprofundada das diretrizes de regulamentação da inteligência artificial nos Estados Unidos revela que o país abraçou de maneira abrangente as orientações contidas em seu documento oficial. A regulamentação nos Estados Unidos reflete uma compreensão profunda das nuances éticas envolvidas, incorporando princípios que buscam equilibrar o avanço tecnológico com a proteção dos valores fundamentais da sociedade. Conhecido por sua liderança em inovação e tecnologia, tem adotado uma abordagem criteriosa para abordar as complexidades éticas associadas à inteligência artificial a partir dos princípios fundamentais propostos nessa pesquisa, mesmo considerando a União Europeia como a precursora na formalização dessa regulamentação.

A coerência entre as diretrizes éticas adotadas por diferentes nações, demonstra a dimensão global do diálogo ético em torno da inteligência artificial. Isso ressalta a importância essencial de uma abordagem colaborativa entre diferentes países, visando garantir uma regulamentação eficaz que promova transparência e intervenção humana nas tecnologias de inteligência artificial.

O reconhecimento mútuo de diretrizes éticas entre nações contribui para a construção de uma base sólida de entendimento global sobre as melhores práticas na governança da inteligência artificial, conforme explorado nesta dissertação de mestrado de natureza interdisciplinar.

Portanto, esta pesquisa evidencia práticas identificadas no desenvolvimento de regulamentações éticas uniformes e eficazes. Ao reconhecer a consonância entre as diretrizes

éticas dos países, os resultados oferecem uma base sólida para a harmonização de abordagens regulatórias.

6. Considerações Finais

Dada a complexidade e o impacto significativo da Inteligência Artificial na sociedade contemporânea, este estudo procurou consolidar uma reflexão abrangente sobre os desafios e as perspectivas éticas associados a essa tecnologia em constante evolução.

A presença da IA em setores essenciais para a sociedade, evidencia não apenas seus benefícios, mas também os desafios emergentes. Ao aprofundarmos os estudos referentes ao debate ético, nos deparamos com preocupações cada vez mais urgentes. A tomada de decisão algorítmica, o viés algorítmico e a transparência nos processos algorítmicos surgem como desafios principais que exigem atenção imediata. O conflito entre a inovação tecnológica e a proteção dos direitos fundamentais da sociedade destaca a urgência de uma abordagem ética de regulamentação.

É fundamental buscar soluções que possam orientar alternativas regulatórias e legislativas. No entanto, a elaboração de diretrizes e regulamentações se mostra complexa mundialmente, especialmente na necessidade de estabelecer padrões globais vinculantes, dada a natureza transcultural da IA.

Portanto, o desenvolvimento de uma estrutura ética e regulatória global para a IA é uma jornada complexa e multifacetada. É importante a expansão de pesquisas existentes, a identificação de lacunas nas regulamentações e a criação de normas que resistam ao teste do tempo diante da acelerada evolução tecnológica. Essa iniciativa exige colaboração internacional e esforços contínuos para garantir que a IA beneficie a humanidade de maneira ética e responsável.

Para garantir que as decisões relacionadas à IA sejam tomadas de forma responsável, é essencial envolver ativamente a sociedade nesse processo. Isso pode ser alcançado por meio de campanhas de conscientização, programas educacionais e iniciativas que ajudem as pessoas compreenderem melhor as complexidades e impactos da IA e a tomar decisões informadas sobre seu uso e regulamentação. Ao promover uma cultura de transparência e participação pública, podemos garantir que a evolução da IA seja guiada por valores que respeitem e promovam a dignidade humana.

Durante o desenvolvimento deste estudo, foi possível aprofundar a compreensão sobre o funcionamento de tecnologia de IA, as questões éticas e sociais que envolvem a integração de programas de IA na sociedade. Identificaram-se percepções significativas sobre os desafios

e as oportunidades associados ao desenvolvimento e uso ético das tecnologias de IA. Ficou evidente a necessidade crítica de transparência e intervenção humana. A transparência emerge como um elemento essencial para promover a confiança e a responsabilidade na utilização da IA, enquanto a intervenção humana se mostra indispensável para garantir que as decisões tomadas por sistemas de IA sejam éticas e alinhadas com os valores humanos fundamentais.

Além disso, esta pesquisa proporcionou uma imersão na metodologia de análise de conteúdo, revelando sua eficácia como ferramenta para extrair significativos a partir de documentos e dados relacionados. A análise de conteúdo permitiu uma compreensão aprofundada das diretrizes existentes nas iniciativas de regulamentação em curso, destacando a importância de abordagens mais inclusivas e abrangentes em seu desenvolvimento.

Este estudo reforça a importância de uma abordagem reflexiva no desenvolvimento e uso das tecnologias. À medida que avançamos em um mundo cada vez mais permeado por essas tecnologias, é essencial que consideremos não apenas os avanços técnicos, mas também as implicações éticas e sociais dessas inovações. A análise crítica é fundamental para garantir que a evolução da IA seja guiada por princípios éticos e respeite a dignidade e os valores humanos.

Para o desenvolvimento desta pesquisa, foi necessário desenvolver uma revisão bibliográfica sobre a conceituação de Inteligência Artificial; analisar os documentos legais que tratam das diretrizes éticas para Inteligência Artificial no Brasil; analisar os documentos que caminham para regulamentar as diretrizes éticas do uso de tecnologias de Inteligência Artificial nos países: Estados Unidos, China, e a União Europeia.

Definimos um recorte temporal para identificar as atividades que ocorrem no âmbito da Estratégia Brasileira de Inteligência Artificial e os Projetos de Lei existentes referente ao uso e a regulamentação da IA. A pesquisa investigou propostas de regulamentações e realizou coletas sistematizadas de dados publicados em documentos oficiais de cada país. Os detalhes sobre essas coletas podem ser encontrados na sessão 4.2 deste documento, onde são abordadas partes específicas dos documentos oficiais de cada país.

Com base nessa coleta, realizou-se uma análise de conteúdo e documental, onde foram identificadas diretrizes, estratégias, e algumas lacunas estabelecidas nos países em questão. Isso permitiu compreender os posicionamentos e recomendações tanto internacionais quanto nacionais.

Foi elaborado um quadro destacando os aspectos comuns dos documentos analisados, os quais foram posteriormente submetidos a uma análise de conteúdo, como evidenciado no Quadro 5 deste documento.

Por fim, foram desenvolvidos gráficos com base na coleta das diretrizes dos documentos, categorizando a quantidade das legendas encontradas. Esse processo resultou em uma melhor visualização dos resultados em percentual, de cada diretriz estudada nos documentos dos países em questão. Esses gráficos se encontram no item 5 desta pesquisa.

No entanto, é importante reconhecer que este estudo enfrentou limitações, entre elas, a análise se baseou principalmente em documentos oficiais encontrados em publicações acadêmicas, o que pode ter deixado de fora perspectivas importantes existentes em documentos complementares. Além disso, devido às restrições de tempo e escopo, não foi possível realizar uma análise aprofundada de todas as implicações éticas abordadas e lacunas encontradas. Portanto, recomendamos que estudos futuros considerem uma abordagem mais abrangente e inclusiva.

À medida que a IA continua a moldar e transformar nossa sociedade, é importante refletirmos sobre como essas mudanças afetam diretamente a experiência humana. Desde questões relacionadas à privacidade e autonomia até preocupações mais amplas sobre o significado da humanidade em um mundo cada vez mais digitalizado, este estudo destaca a importância de uma abordagem que leve em consideração não apenas os aspectos técnicos, mas também os impactos sociais e culturais da IA.

A regulamentação ética da IA deve ser dinâmica, capaz de se adaptar ao ritmo acelerado da evolução tecnológica, promovendo assim um ambiente global seguro e ético para o desenvolvimento e aplicação dessa tecnologia impactante.

Para estudos futuros, foram identificadas lacunas que necessitam de reflexões mais aprofundadas através de continuidade de pesquisa, como os impactos das tecnologias de inteligência artificial na autonomia pessoal e nos direitos fundamentais da sociedade, sobre a utilização maciça de dados para alimentar sistemas de IA.

Este estudo destaca a importância de uma abordagem ética e regulatória abrangente para lidar com os desafios e oportunidades apresentados pela Inteligência Artificial. À medida que a IA continua a moldar e transformar nosso mundo, é importante que a sociedade avance, porém com responsabilidade, garantindo que os benefícios do uso das tecnologias de Inteligência Artificial sejam maximizados e que seus riscos sejam mitigados de maneira ética e equitativa,

preservando sempre os valores e a dignidade da condição humana. A colaboração entre diferentes setores da sociedade e a contínua pesquisa e reflexão são essenciais para alcançar esse objetivo comum.

Em conclusão, este estudo destaca a necessidade de um diálogo contínuo entre a comunidade acadêmica, os legisladores, os desenvolvedores de tecnologia, e a sociedade civil. A ética surge como o alicerce essencial para guiar práticas inovadoras e assegurar que a evolução da IA seja conduzida com responsabilidade, respeitando a condição humana.

7. Referências

- ALTHUSSER, Louis, **Ideologia e aparelhos ideológicos do Estado**, Editora Presença.
- ANDERSON, Michael; ANDERSON, Susan Leigh. **Machine ethics: Creating an ethical intelligent agent**. AI Magazine, v. 28, n. 4, p. 15-15, 2007.
- ARBIX, Glauco. **A transparência no centro da construção de uma IA ética**. Novos Estudos CEBRAP, v. 39, p. 395-413, 2020.
- ARBIX, Glauco. **Seriam as máquinas dotadas de alma e da capacidade de pensar?** Jornal da USP. Disponível em <https://jornal.usp.br/radio-usp/seriam-as-maquinas-dotadas-de-alma-e-da-capacidade-de-pensar/> Acesso em 21 jun. 2022.
- BARDIN, Laurence. **Análise de Conteúdo**. Lisboa, 1977.
- BARDIN, Laurence. **Análise de Conteúdo**. São Paulo: Edições 70, 2011, 229 p.
- BAROCAS, Solon; SELBST, Andrew D. **Big data's disparate impact**. California Law Review, p. 671-732, 2016.
- BAUMAN, Zygmunt. **Modernidade líquida**. Rio de Janeiro, Jorge Zahar, 2001.
- BENASAYAG, Miguel. **São os humanos, não as máquinas, que criam significado**. Unesco, 2018. Disponível em: <https://pt.unesco.org/courier/2018-3/sao-os-humanos-nao-maquinas-que-criam-significado>. Acesso em 19 mar. 2023.
- BIGATON, Tainá Rafaela; TEIXEIRA, Marcelo Markus; BILHALVA, João Vítor Massaro **A ética de dados como uma possível resposta para alguns desafios do direito internacional frente a inteligência artificial**. Revista Eletrônica do Curso de Direito da UFSM, v. 17, n. 1, p. e65051-e65051, 2022.
- BOSTROM, Nick; YUDKOWSKY, Eliezer. **A ética da inteligência artificial**. Rev. de Pesquisa em Filosofia, v. 1, n. 3, 2011.
- BRANCO, Camila.; MATSUZAKI, Luciano. **Olhares da rede**. São Paulo: Momento, 2009.
- BRASIL. Ministério da Ciência, Tecnologia e Inovações. **Estratégia brasileira de inteligência artificial: EBIA**. Brasília: MCTI, 2021. Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-documento_referencia_4-979_2021.pdf. Acesso em janeiro 2023.

BYNUM, T. Stanford Encyclopedia of Philosophy Archive **Ética Computer and Information Encyclopedia of Philosophy**, Edward N. Zalta (ed.), 2015. Disponível em: <http://plato.stanford.edu/archives/win2015/entries/ethics-computer/>. Acesso em nov. 2023.

CAPPRA, Ricardo. **Ética na IA: como desenvolver um sistema ético para Inteligência Artificial**. Cappra Institute for Data Science. 2021. Disponível em: <https://www.cappra.institute/estudos>. Acesso em 20 jan. 2022.

CARDOSO, Márcia Regina Gonçalves; OLIVEIRA, Guilherme Saramago; GHELLI, Kelma Gomes Mendonça. **Análise de conteúdo: uma metodologia de pesquisa qualitativa**. Cadernos da FUCAMP, v. 20, n. 43, 2021.

CASTRO, Daniel; MCLAUGHLIN, Michael; CHIVOT, Eline. **Who Is Winning the AI Race**., the EU or the United States? Center For Data Innovation, 2019. Disponível em: <https://www2.datainnovation.org/2019-china-eu-us-ai.pdf>. Acesso em: julho 2023.

CASTELLS, Manuel. **Communicattion power**. New York: Oxford University Press, 2009. 590 p.

CERVO, Amado L.; BERVIAN, Pedro A.; SILVA, Roberto da. **Metodologia científica**. 5. reimp. São Paulo, 2010.

CHAUÍ, Marilena. **Convite à filosofia**. 14. ed. São Paulo: Ática, 2012.

CHINA. **Conselho Estadual de impressão e distribuição Aviso do Plano de Desenvolvimento de Inteligência Artificial de Nova Geração**. 2017. Disponível em: https://www.gov.cn/zhengce/zhengceku/2017-07/20/content_5211996.htm. Acesso em: jun. 2023.

COMISSÃO EUROPEIA, **Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence**. 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>. Acesso em maio 2022.

COMISSÃO EUROPEIA, **Shaping Europe's digital future**. 2021. Disponível em: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>. Acesso em ago. 2023.

CORDEIRO, Tiago. **Por que é importante regular a inteligência artificial no Brasil?** 2022. Disponível em: <https://www.insper.edu.br/noticias/por-que-e-importante-regular-a-inteligencia-artificial-no-brasil/#A%20IA>. Acesso em 10 jul. 2023.

DA COSTA FELIPE, Bruno Farage; MULHOLLAND, Caitlin. **Filtro bolha e Big Nudging: a decomocracia participativa na era dos algoritmos**. Revista Direitos Fundamentais & Democracia, v. 27, n. 3, p. 06-18, 2022.

DARDOT, Pierre; LAVAL, Christian. **A nova razão do mundo**. Boitempo, 2017.

DIJCK, José Van. **Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology**. Surveillance & society, v. 12, n. 2, p. 197-208, 2014.

DONEDA, Danilo Cesar Maganhoto *et al.* **Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal**. Pensar-Revista de Ciências Jurídicas, v. 23, n. 4, p. 1-17, 2018.

DUNKER, Christian. **Sufrimento Narcísico Digital: quando as redes sociais mostram nosso espelho**. Blog do Dunker, 13 set. 2019. Disponível em: <https://blogdodunker.blogosfera.uol.com.br/2019/09/13/sofrimento-narcisico-digital-quando-as-redes-sociais-mostram-nosso-espelho/>. Acesso em abr. 2022.

EUBANKS, Virginia. **Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor**. Nova York: St. Martin's Press, 2018.

FICO, Bernardo; ALENCAR, Ana Catarina; SOUSA, Beatriz. **Discriminação de gênero na era do Big Data e da Inteligência Artificial**, 2021. Disponível em: <https://opiceblum.com.br/wp-content/uploads/2021/02/Vieses-inconscientes-de-genero-na-Inteligencia-Artificial.pdf>. Acesso em jan. 2024.

GALVÃO, Lúcia Helena. **A evolução do homem**. Nova Acrópole Brasil, 2023. Disponível em: <https://youtu.be/8ccY7hCP7I4>. Acesso em dez. 2023.

GALVÃO, Lúcia Helena. **A triste realidade da sociedade atual**. *Mente em Evolução*, 2023. Disponível em: <https://www.youtube.com/watch?v=IW5RunIOqMA>. Acesso em dezembro 2023.

GARCIA, Ana Cristina. **Ética e inteligência artificial**. *Computação Brasil*, n. 43, p. 14-22, 2020.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 6.ed. São Paulo: Atlas, 2019.

GODOY, Arilda Schmidt. **Pesquisa qualitativa: tipos fundamentais**. *Revista de Administração de Empresas*, v. 35, p. 20-29, 1995.

GOMES, Wilson. **Opinião política na internet: uma abordagem ética das questões relativas a censura e liberdade de expressão na comunicação em rede**. Apresentado na 10ª Reunião Anual da COMPÓS, GT Comunicação e Política. Brasília: UnB, 2001.

GOODMAN, Bryce W. **A step towards accountable algorithms?: Algorithmic discrimination and the European Union General Data Protection**. In: 29th Conference on Neural Information Processing Systems (NIPS), 2016. Disponível em: <http://www.mlandthelaw.org/papers/goodman1.pdf>. Acesso em 20 abr. 2023.

GPO, (U.S. GOVERNMENT PUBLISHING OFFICE). **National Artificial Intelligence Initiative Act of 2020**. Washington, DC, 2020. Disponível em: <https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf>. Acesso em fev. 2023.

GROHMANN, Rafael. **Plataformização do trabalho: entre dataficação, financeirização e racionalidade neoliberal**. Revista Eletrônica Internacional de Economia Política da Informação, da Comunicação e da Cultura, v. 22, n. 1, p. 106-122, 2020.

HAN, Byung-Chul. **O desaparecimento dos rituais: Uma topologia do presente**. Tradução de Gabriel Salvi Philipson. Petrópolis, RJ: Vozes, 2021.

HARARI, Yuval Noah. **21 lições para o século 21**. Editora Companhia das Letras, 2018.

HARARI, Yuval Noah. **A tecnologia permitirá ‘hackear’ seres humanos**. 2018. Disponível em: https://brasil.elpais.com/brasil/2018/08/20/eps/1534781175_639404.html. Acesso em: nov. 2023.

HARARI, Yuval Noah. **Sapiens - Uma Breve História da Humanidade**. Porto Alegre: L&pm, 2015. 464 p.

HARAWAY, Donna. **Manifesto ciborgue**. In: **Antropologia do Ciborgue: As vertigens do pós-humano**. Belo Horizonte: Autêntica, 2000, p. 33-118.

HILDEBRANDT, Mireille. **Law as computation in the era of artificial legal intelligence: Speaking law to the power of statistics**. University of Toronto Law Journal, v. 68, n. 1, p. 12-35, 2018.

HIRATUKA, Celio; DIEGUES, Antônio Carlos. **Inteligência artificial na estratégia de desenvolvimento da China contemporânea**. Texto para discussão. Instituto de Economia Unicamp, 2021. Disponível em: https://www.economia.unicamp.br/images/arquivos/artigos/TD/TD422_1.pdf. Acesso em ago. 2023.

HOMEM, Maria; DUNKER, Christian. **A tecnologia e todas as relações com o nosso tempo.** Fronteiras do Pensamento, 2022. Disponível em:
<https://www.fronteiras.com/leia/exibir/a-tecnologia-e-todas-as-relacoes-com-o-nosso-tempo>. Acesso em: jan de 2023.

JORGE, Carlos Francisco Bitencourt *et al.* **Proteção de dados pessoais e Covid-19: entre a inteligência epidemiológica no controle da pandemia e a vigilância digital.** Liinc em Revista, v. 16, n. 2, p. e5251-e5251, 2020.

KAUFMAN, Dora. **A inteligência artificial irá suplantar a inteligência humana?** Estação das Letras e Cores, 2019.

KAUFMAN, Dora; SANTAELLA, Lucia. **O papel dos algoritmos de inteligência artificial nas redes sociais.** Revista Famecos, v. 27, n. 1, p. e34074-e34074, 2020.

LEE, Kai-Fu. **AI superpowers: China, Silicon Valley, and the new world order.** New York: Harper Business, 2020.

LIPTON, Zachary C. The mythos of model interpretability: **In machine learning, the concept of interpretability is both important and slippery.** Queue, v. 16, n. 3, p. 31-57, 2018.

MELO, Teresa Mary Pires de Castro; CAMARGO, André Luiz Rodrigues. **Educomunicação e autoetnografia: diálogos e aproximações.** Revista Mídia e Cotidiano, v. 14, n. 3, p. 283-302, 2020.

MIRANDA, Luana Alonso Xavier de. **Liderança global em Big Data e inteligência artificial: objetivos da China e estratégia de desenvolvimento.** 2021. Disponível em:
<https://www.lume.ufrgs.br/handle/10183/242062>. Acesso em jun. 2023.

MITCHELL, T. **Machine Learning.** Boston: McGraw-Hill, 1997.

MITTELSTADT, Brent Daniel, *et al.* **"The ethics of algorithms: Mapping the debate"** .Big Data & Society v.3 n.2 2053951716679679, (2016). Disponível em:
<https://journals.sagepub.com/doi/10.1177/2053951716679679>. Acesso em 11 de mar. 2022.

MORAES, Roque. **Análise de conteúdo.** Revista Educação, Porto Alegre, v. 22, n. 37, p. 7-32, 1999.

MÜLLER, William; SILVEIRAS, Raphael. **A Evolução e a Regulamentação da Inteligência Artificial no Brasil**. Revista InterCiência-IMES Catanduva, v. 1, n. 11, p. 2-2, 2023.

NEVES, Cláudia. **A regulação das políticas de educação na União Europeia e os desafios para a Educação a Distância no ensino superior: uma perspectiva crítica e uma proposta de investigação**. 2018. Disponível em: <http://educa.fcc.org.br/pdf/rbpae/v34n1/1678-166X-rbpae-34-1-0035.pdf>. Acesso em jan. de 2024.

NOGUEIRA, Fabio. **Projeto de marco legal da IA no Brasil é pouco consistente e pode ser inútil, dizem especialistas**. 2021. Jornal da Unesp. Disponível em: <https://jornal.unesp.br/2021/07/29/projeto-de-marco-legal-da-ia-no-brasil-e-pouco-consistente-e-pode-ser-inutil-dizem-especialistas/>. Acesso em out. de 2023.

NORVIG, Peter; RUSSELL, Stuart J. **Inteligência artificial**. Tradução Regina Célia Simille, Rio de Janeiro: Elsevier, 2013.

O'NEIL, Cathy. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. **Hardcover, College & Research Libraries**, v. 78, n. 3, 2016.

OECD. **Recommendation of the Council on Artificial Intelligence**. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em mar. 2023.

PASQUINELLI, Matteo et al. **Cómo una máquina aprende y falla—Una gramática del error para la Inteligencia Artificial**. Hipertextos, v. 10, n. 17, p. 13-29, 2022.

PETIT, Martí. **Por una crítica de la razón algorítmica. Estado de la cuestión sobre la inteligencia artificial, su influencia en la política y su regulación**. Quaderns del Cac, [s. l], n. 44, p. 5-15, 2018. Quadrimestral.

RAMOS, Murilo César. **Comunicação, direitos sociais e políticas públicas**. MARQUES DE MELO, J.; SATHLER, L. Direitos à Comunicação na Sociedade da Informação. São Bernardo do Campo, SP: Umesp, p. 245-53, 2005.

REIS, Vera Lúcia Geiss dos. **Percepções sobre o uso da Plataforma Khan Academy nas aulas de Matemática com alunos do 9º ano de uma escola municipal**. Dissertação de Mestrado. Universidade Federal de Pelotas, 2019.

ROBERTS, H., COWLS, J., HINE, E. *et al.* **Achieving a ‘Good AI Society’: Comparing the Aims and Progress of the EU and the US**. Science and engineering ethics, v.27 (p. 1-25, 2021).

ROSSETTI, Regina; ANGELUCI, Alan. **Ética Algorítmica**: questões e desafios éticos do avanço tecnológico da sociedade da informação. São Paulo: Galáxia, 2021.

SÆTRA, Henrik Skaug; FOSCH-VILLARONGA, Eduard. **Healthcare digitalisation and the changing nature of work and society**. Editora MDPI, 2021.

SAVATER, Fernando. **Ética urgente**. Edições Sesc, 2016.

SILVA, Jennifer Amanda Sobral da; MAIRINK, Carlos Henrique Passos. Inteligência artificial. **LIBERTAS: Revista de Ciências Sociais Aplicadas**, v. 9, n. 2, p. 64-85, 2019.

SILVEIRA, Sérgio Amadeu *et al.* **Colonialismo de dados**: como opera a trincheira algorítmica na guerra neoliberal. 1. ed. São Paulo: Autonomia Literária, 2021. v. 1. 212p.

SILVEIRA, Sérgio Amadeu. **Brasil, colônia digital**. Instituto Humanitas Unisinos, 2020. Disponível em <https://www.ihu.unisinos.br/categorias/600360-brasil-colonia-digital-artigo-de-sergio-amadeu>. Acesso em 10 de nov. de 2023.

SPINK, Mary Jane; BRIGAGÃO, Jaqueline; NASCIMENTO, Vanda Lúcia Vitoriano; CORDEIRO, Mariana Prioli (Orgs.). **A produção de informação na pesquisa social: compartilhando ferramentas**. Rio de Janeiro: Centro Edelstein de Pesquisas Sociais, 2014.

SPINOLA, Loren. **Proposta de regulação de IA no Brasil: o risco de o tiro sair pela culatra**. ABES, 2023. Disponível em: <https://abes.com.br/proposta-de-regulacao-de-ia-no-brasil-o-risco-de-o-tiro-sair-pela-culatra/>. Acesso em: jun 2023.

SZWARCWALD, Celia Landmann *et al.* **ConVid-Pesquisa de Comportamentos pela Internet durante a pandemia de COVID-19 no Brasil: concepção e metodologia de aplicação**. *Cadernos de Saúde Pública*, v. 37, p. e00268320, 2021.

TADDEO, Mariarosaria; FLORIDI, Luciano “**How AI Can Be a Force for Good**”. *Science*, v. 361, n. 6.404, ago. 2018(b). DOI: 10.1126/science.aat5991. Acesso em jun. 2022.

TADDEO, Mariarosaria; FLORIDI, Luciano. “**Regulate Artificial Intelligence to Avert Cyber Arms Race**”. *Nature*, v. 556, 2018(a). Disponível em: DOI: 10.1038/d41586-018-04602-6.» <https://doi.org/10.1038/d41586-018-04602-6>. Acesso em jun. 2022.

TADDEO, Mariarosaria ; FLORIDI, Luciano. **Introduction: What is data ethics?** *Philosophical Transactions R. Soc., Londres*, v. 374, n. 2083, 2016. Disponível em: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2016.0360>. Acesso em nov. 2023.

TEIXEIRA, João. **Inteligência artificial** São Paulo, Paulus, 2014.

TJDFT (Tribunal de Justiça do Distrito Federal e dos Territórios). **Lei nº 12.965, de 23 de abril de 2014:** Marco civil da internet. Disponível em:

<https://www.tjdft.jus.br/institucional/imprensa/campanhas-e-produtos/direito-facil/edicao-semanal/marco-civil-da-internet>. Acesso em: nov. 2023.

TUKANO, Álvaro. **O mundo Tukano antes dos brancos**. 1. ed. [s. l.]: Um Mestre Tukano, 2017.

TURING, Alan M. Computing machinery and intelligence (1950). *In: The Essential Turing: the Ideas That Gave Birth to the Computer Age*. Oxford: Clarendon Press, 2012, p. 433-464.

UNESCO. **Report of the Social and Human Sciences Commission (SHS)**, UNESCO. General Conference, 41st, 2021 [881]. Code. 41 C/73.

VALLS, Álvaro LM. **O que é ética**. Brasiliense, 2017.

VERMELHO, Sônia Cristina *et al.* **Refletindo sobre as redes sociais digitais**. Educação & Sociedade, v. 35, p. 179-196, 2014. Acesso em jun. 2022.

WERMUTH, Maiquel; ROCHA, Leonel (Org.). **XXIX Congresso Nacional do CONPEDI Balneário Camboriú-SC**. Direito, governança e. novas tecnologias, Florianópolis: CONPEDI, 2022.

WHITE HOUSE. **Accelerating America's Leadership in Artificial Intelligence**, 11/02/2019. Disponível em: President Donald J. Trump Is Accelerating America's Leadership in Artificial Intelligence – The White House (archives.gov). Acesso em: mai. 2023.

ZUBOFF, Shoshana **A era do capitalismo de vigilância**. Rio de Janeiro, Intrínseca, 2021.