

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

VICTOR RIBEIRO PIRES

PRINCÍPIOS DE INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE
COMPARATIVA DE POLÍTICAS NACIONAIS E INTERNACIONAIS

RIO DE JANEIRO

2024

VICTOR RIBEIRO PIRES

PRINCÍPIOS DE INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE
COMPARATIVA DE POLÍTICAS NACIONAIS E INTERNACIONAIS

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Computação da Universidade Federal do Rio de Janeiro como parte dos requisitos para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Prof. João Carlos Pereira da Silva

Co-orientador: Prof. Claudio Miceli de Farias

RIO DE JANEIRO

2024

CIP - Catalogação na Publicação

P667p Pires, Victor Ribeiro
Princípios de Inteligência Artificial: Uma
Análise Comparativa de Políticas Nacionais e
Internacionais / Victor Ribeiro Pires. -- Rio de
Janeiro, 2024.
44 f.

Orientador: João Carlos Pereira Da Silva.
Coorientador: Claudio Miceli De Farias.
Trabalho de conclusão de curso (graduação) -
Universidade Federal do Rio de Janeiro, Instituto
de Computação, Bacharel em Ciência da Computação,
2024.

1. Inteligencia Artificial. 2. Ética. 3.
Preconceito Algorítmico. I. Da Silva, João Carlos
Pereira , orient. II. De Farias, Claudio Miceli,
coorient. III. Título.

VICTOR RIBEIRO PIRES

PRINCÍPIOS DE INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE
COMPARATIVA DE POLÍTICAS NACIONAIS E INTERNACIONAIS

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Computação da Universidade Federal do Rio de Janeiro como parte dos requisitos para obtenção do grau de Bacharel em Ciência da Computação.

Aprovado em 26 de Junho de 2024

BANCA EXAMINADORA:

Documento assinado digitalmente
 **JOAO CARLOS PEREIRA DA SILVA**
Data: 12/08/2024 18:20:03-0300
Verifique em <https://validar.iti.gov.br>

João Carlos Pereira da Silva
D.Sc (UFRJ)

Documento assinado digitalmente
 **CLAUDIO MICELI DE FARIAS**
Data: 20/08/2024 17:52:13-0300
Verifique em <https://validar.iti.gov.br>

Claudio Miceli de Farias
D.Sc (UFRJ)

Documento assinado digitalmente
 **MARIA LUIZA MACHADO CAMPOS**
Data: 12/08/2024 07:44:15-0300
Verifique em <https://validar.iti.gov.br>

Maria Luiza Machado Campos
Ph.D (UFRJ)

Documento assinado digitalmente
 **PRISCILA MACHADO VIEIRA LIMA**
Data: 20/08/2024 11:29:06-0300
Verifique em <https://validar.iti.gov.br>

Priscila Machado Vieira Lima
Ph.D (UFRJ)

Para a minha avó Nuvem, que continue me olhando do céu. Para minha família que sempre me deu apoio incondicional, mesmo para as minhas loucuras.

AGRADECIMENTOS

Primeiramente gostaria de agradecer a Deus por ter me dado minha família, que foi a base que me criou e me solidificou para que eu pudesse chegar até onde estou. Obrigado a meus pais Maria Luiza e Jorge Alonso, minhas irmãs Júlia e Jéssica e meus avós Nuvem, Olga, Siulí e Ubyrajara, que sempre puderam me guiar nos momentos mais escuros de minha mente.

Agradeço também a minha esposa Ana Luiza, que sem seu apoio e revisão ortográficas advindos de seu título como licenciada em Letras da UFRJ, não permitiu que eu escrevesse como se tivesse criptografando um código.

Agradeço aos diversos amigos do LABNET-NCE, do GRIS e da SEGTIC, em especial a Gabriel Caldas, Gabriel Martins, Marcos Araújo, Daniel La Rubia, Leonardo Schripsema, Xiao Kong, Filipe Lopes, Leonardo Gonçalves, Arthur Castro, Yago Stellet, Sidney Outeiro, Gustavo Buçard e Bruno Matos que me empurraram até o final do curso. Sem vocês a caminhada teria sido abandonada no meio.

Ao Club de Regatas Vasco da Gama que mesmo não estando em suas grandes fases, continua sendo sempre gigante.

Por fim, agradeço aos meus orientadores Claudio Miceli e João Carlos por toda paciência que tiveram comigo enquanto aluno e como orientado. Sei que não sou uma pessoa fácil e mesmo assim tive total apoio para desenvolver esse trabalho.

*“Feeding AI systems on the world’s beauty,
ugliness, and cruelty, but expecting it to
reflect only the beauty is a fantasy.”*

Ruha Benjamin

RESUMO

Neste trabalho foi realizada uma pesquisa com o objetivo de identificar os princípios e preocupações relacionadas ao desenvolvimento e uso de sistemas de inteligência artificial em documentos e legislações da União Europeia, Estados Unidos, China, UNESCO e Brasil. Por meio dessa pesquisa documental, buscou-se compreender os valores e diretrizes subjacentes aos posicionamentos dessas entidades em relação à inteligência artificial. Com base na análise desses documentos, foi elaborada uma lista de princípios que se acredita serem fundamentais para promover o desenvolvimento ético e responsável da inteligência artificial. Esses princípios visam não apenas garantir a segurança e a privacidade dos indivíduos, mas também promover a equidade, a transparência e a responsabilidade no uso dessa tecnologia. Através da identificação e sistematização desses documentos, pretende-se sintetizar os princípios para desenvolvedores, pesquisadores, legisladores e demais interessados no campo da inteligência artificial. Ao adotar tais princípios, espera-se que os desenvolvedores estejam mais conscientes dos impactos sociais, éticos e legais de suas criações, contribuindo assim para um ecossistema de inteligência artificial mais justo, inclusivo e sustentável.

Palavras-chave: Inteligencia Artificial; Ética; Preconceito Algorítmico.

ABSTRACT

In this work a research was conducted with the aim of identifying the principles and concerns regarding the development and use of artificial intelligence systems in documents and laws from the European Union, United States, China, UNESCO, and Brazil. Through this documentary research, the values and guidelines underlying the positions of these entities regarding artificial intelligence were sought to be understood. Based on the analysis of these documents, a list of principles believed to be fundamental to promote the ethical and responsible development of artificial intelligence was elaborated. These principles aim not only to ensure the safety and privacy of individuals but also to promote equity, transparency, and accountability in the use of this technology. Through the identification and systematization of these documents, it is we intend to synthesize the principles for developers, researchers, lawmakers, and other stakeholders in the field of artificial intelligence. By adopting such principles, it is expected that developers will be more aware of the social, ethical, and legal impacts of their creations, thus contributing to a fairer, more inclusive, and sustainable intelligence ecosystem.

Keywords: Artificial Intelligence; Ethics; Algorithmic Bias.

LISTA DE ILUSTRAÇÕES

Figura 1 – Pirâmide de risco (EUROPEIA, 2021)	18
---	----

LISTA DE TABELAS

Tabela 1 – Princípios Éticos em Diferentes Contextos Internacionais	33
Tabela 2 – Receios Relacionados à Inteligência Artificial em Diferentes Contextos Internacionais	34

LISTA DE ABREVIATURAS E SIGLAS

IA	Inteligência artificial
UE	União Europeia
LGPD	Lei Geral de Proteção de Dados Pessoais
<i>AI Act</i>	Artificial Intelligence Act
<i>AI Bill</i>	<i>AI Bill of Rights</i>
DoD	Departamento de Defesa dos Estados Unidos
EUA	Estados Unidos da América
PL	Projeto de Lei
UnB	Universidade de Brasília
STF	Supremo Tribunal Federal
MERCOSUL	Mercado Comum do Sul
OTAN	Organização do Tratado do Atlântico Norte

SUMÁRIO

1	INTRODUÇÃO	12
2	AS DISCUSSÕES DAS LEIS DIGITAIS E ÉTICAS ACERCA DA INTELIGÊNCIA ARTIFICIAL DE OUTROS PAÍSES E BLOCOS	15
2.1	DISCUSSÕES DA UNIÃO EUROPEIA	15
2.2	DISCUSSÕES DE INTELIGÊNCIA ARTIFICIAL NOS ESTADOS UNIDOS	20
2.3	DISCUSSÕES DE INTELIGÊNCIA ARTIFICIAL NA CHINA	23
2.4	CONCLUSÃO	26
3	DISCUSSÕES E ATOS SOBRE A INTELIGÊNCIA ARTIFICIAL NO BRASIL	28
3.1	PROJETOS DE LEI ACERCA DO USO DE INTELIGÊNCIA ARTIFICIAL NO BRASIL	28
3.2	CONCLUSÃO	31
4	CONCLUSÃO E TRABALHOS FUTUROS	33
	REFERÊNCIAS	38

1 INTRODUÇÃO

Nos últimos anos, temos observado um avanço significativo na área de Inteligência Artificial (IA) que cada vez mais é sentida no cotidiano das pessoas. Essa mesma presença do uso de IA no dia-a-dia tem criado tensões em diversos temas como empregabilidade, segurança dos dados, direitos de produções artísticas, uso de imagem, entre outras. A IA é um campo da computação voltado a desenvolver algoritmos e sistemas capazes de realizar tarefas que demandam habilidades associadas à inteligência humana (GARCIA, 2020; PEIXOTO, 2020a), sem se preocupar em recriar um cérebro humano, apenas traços das funções cognitivas. Alguns exemplos de uso de inteligência artificial que encontramos no cotidiano são o reconhecimento facial, reconhecimento de voz, *chatbots*, recomendação de produtos, automatização de processos, predição de eventos, reconhecimento de padrões, resolução de problemas, aprendizado, planejamento, compreensão da linguagem natural etc (RUSSELL; NORVIG, 2021).

Nessas discussões acerca da IA uma pergunta pode ocorrer: existe uma ética para o uso e o desenvolvimento de inteligências artificiais. Primeiramente, para termos uma possível resposta, faz-se necessário entender o conceito de "ética". A "ética" serve para conduzir as ações humanas a respeito das boas ações (virtudes) ou das não-éticas, as más (vícios) (FIGUEIREDO, 2008). Pela forma de aquisição das virtudes, Aristóteles as classifica em: intelectuais (conhecimento teórico resultante do ensino) e morais (conhecimento prático adquirido pelos hábitos). Partindo desse ponto de vista, podemos definir "ética" como "ciência dos costumes" (COTRIM; FERNANDES, 2013). Então definiremos ética como as guias das ações humanas que visam as boas ações, respeito ao próximo, respeito às escolhas individuais que não impactem o próximo e o respeito à Declaração Universal dos Direitos Humanos (BRASIL, 1998).

Se existem escolhas éticas no desenvolvimento de IAs, então podemos ter também escolhas antiéticas. Um problema de escolhas antiéticas no desenvolvimento

e do uso incorreto de IA é o preconceito algorítmico¹, que pode ocorrer tanto por uma escolha deliberada das pessoas que desenvolveram a IA, quanto por um acaso ocorrido durante o desenvolvimento da IA. Mesmo que a ética seja um conceito exclusivo aos homens (ROSSETTI; ANGELUCI, 2021), podemos definir preconceito algorítmico como o reflexo e uma interpretação de atos humanos negativos como racismo, misoginia, xenofobia, homofobia e outras atitudes consideradas antiéticas nos resultados e decisões de sistemas automatizados e/ou com inteligência artificial (PEREIRA; PERDONA, 2022). O preconceito algoritmo não necessariamente precisa ser implementado no código do sistema, ele pode ser derivado das análises dos dados usados para calibração desses sistemas. Se a máquina receber dados e informações carregados de vieses e preconceitos de raça, de gênero, de escolha sexual, de forma física ou de qualquer outro traço, ela irá não só aprendê-los como perpetuá-los(GARCIA, 2020). Por esse motivo, nem sempre é claro para os desenvolvedores de código que o sistema apresenta algum tipo de viés em suas decisões. Isso acarreta que hoje, não temos nenhum grande exemplo de IA ou sistema com IA que seja considerado uma referencia em ser não enviesado ou sem preconceitos.

Contudo, a resposta da pergunta acima é complexa. A UNESCO (Organização das Nações Unidas para a Educação, a Ciência e a Cultura) aprovou em novembro de 2021 o documento chamado "Recomendação sobre a Ética na Inteligência Artificial"(UNESCO, 2021). Este documento tem como objetivo fornecer um marco para os Estados-membros formalizarem e basearem suas políticas acerca de inteligência artificial, incentivando os debates acerca da inteligência artificial e promovendo o acesso, conhecimento e os avanços da inteligência artificial a todos, respeitando os direitos humanos e as liberdades individuais. São apresentados princípios a serem seguidos, como a proporcionalidade e não causar dano; segurança e proteção; justiça e não discriminação; sustentabilidade; direito à privacidade e proteção de dados; supervisão humana e determinação; transparência e explicabilidade; responsabilidade e prestação de contas; conscientização e alfabetização; governança, colaboração adaptáveis e com múltiplas partes interessadas.

¹ Adotamos o uso do termo "preconceito algorítmico" por ser o mais usado dentro dos documentos encontrados em português. Entende-se que poderia ser utilizado o termo "viés algorítmico" mas os autores preferem usar a palavra preconceito, devido ao sentimento de algo negativo que as pessoas tendem a associar ao uso da palavra.

Embora a UNESCO possua 95 países membros e oito membros associados, os países não tem nenhuma obrigação de adotar essas medidas em lei, sendo assim, o documento é visto como simbólico ou uma referência e não como um tratado internacional. O documento da UNESCO mostra que hoje em dia existe uma base de discussão para qualquer cientista ou leigo. Também dá direcionamento aos países que ainda não apresentam uma discussão madura sobre o tema. Para os cientistas da computação, é possível ver a amplitude que as questões acerca da IA permeiam a sociedade. Os impactos que anteriormente eram imaginários apenas nas discussões acadêmicas agora são uma realidade tangível e a certo ponto, popular.

Diante deste cenário, este trabalho nasce da ideia de se perguntar como legisladores têm se questionado e abordado o tema da IA, numa época onde o tema nunca foi tão debatido, buscando levantar como a regulamentação do uso e do desenvolvimento da IA tem sido abordada em diferentes países. Neste trabalho fazemos uma análise das principais propostas de regulação do desenvolvimento e uso de IA no mundo, em particular União Europeia, China, Estados Unidos e Brasil. Buscamos apresentar seus problemas e limitações, bem como seus pontos fortes. Este trabalho está organizado da seguinte forma: no segundo capítulo, apresentamos e discutimos as leis, propostas e documentos de IA que estão sendo implementados ou discutidos na União Europeia, Estados Unidos e China; no terceiro capítulo, é apresentada a proposta e a discussão do Projeto de Lei em tramitação no Senado Federal do Brasil e o quarto capítulo apresenta a conclusão do trabalho e um conjunto de princípios para o desenvolvimento e uso de IA.

2 AS DISCUSSÕES DAS LEIS DIGITAIS E ÉTICAS ACERCA DA INTELIGÊNCIA ARTIFICIAL DE OUTROS PAÍSES E BLOCOS

Neste capítulo, iremos apresentar as discussões sobre inteligência artificial na União Europeia, Estados Unidos e China. A decisão por esses países é devido à importância geopolítica dos mesmos, visto que são as maiores economias do mundo. As discussões da Europa e dos Estados Unidos, regiões essas que possuem maior maturidade nas legislações que abordam os temas de tecnologia e possuem também maior controle da produção das mesmas, tendem a influenciar outros países no resto do mundo. As discussões chinesas sobre inteligência artificial tem como base a alta liderança que esse país tem exercido na criação de ferramentas de inteligência artificial (HINE; FLORIDI, 2022). Este capítulo tem como missão apresentar as discussões dos documentos mais importantes dessas regiões e seus debates que fundamentaram as mesmas.

2.1 DISCUSSÕES DA UNIÃO EUROPEIA

A aprovação do *AI Act* (MADIEGA, 2021), documento que demonstrou a maturidade da União Europeia na discussão sobre a importância e os perigos da inteligência artificial em seus países, se deu no dia 21 de julho de 2023 (LOPES, 2023). Ao longo dos anos, a UE já havia preparado uma série de documentos sobre uso e regulação de casos que contenham uso de inteligência artificial. Esses mesmos documentos serviram não somente como base para a criação do *AI Act*, como para identificar os riscos sobre o uso de inteligência artificial (MADIEGA, 2021; SOVRANO et al., 2022). Entre os documentos que já existiam na UE sobre o uso de inteligência artificial, podemos destacar:

- Princípios éticos para o desenvolvimento e uso de inteligência artificial (EUROPEU, 2020a; EUROPEU, 2021b);
- Responsabilidade Civil para operadores de inteligência artificial de alto risco (EUROPEU, 2020b);

- Documentos para inteligência artificial em propriedade intelectual, militar, educação, cultura e audiovisual (EUROPEU, 2020c; EUROPEU, 2021b; EUROPEU, 2021a; EUROPEU, 2021c);
- Documentos de lei criminal para inteligência artificial (EUROPEU, 2021a; EUROPEU, 2021d).

2.1.1 O *AI Act*

O *AI Act* foi construído pensando não somente nos desafios e riscos que a população europeia tem mapeado no desenvolvimento da inteligência artificial, como no medo de potências rivais ao bloco econômico se desenvolverem mais na criação e aprimoramento, enquanto a UE teria barreiras legais para o uso. Para evitar esse problema, a *AI Act* atinge somente o setor civil enquanto o ramo militar não está sujeito às mesmas regras citadas no documento. (MADIEGA, 2021).

Para a criação do *AI Act*, inicialmente foram identificados 6 problemas principais que nortearam as discussões. São eles:

1. Risco aos cidadãos e a segurança: diz respeito a aplicação ou desenvolvimento de IA que crie riscos a segurança física dos habitantes da UE, salvo em aplicações militares pois não estão cobertas pelo *AI Act*;
2. Risco aos direitos fundamentais e aos valores da UE: são riscos que ofereçam ofensas aos direitos humanos fundamentais estabelecidos pela UNESCO, bem como aos valores éticos dos países membros da UE;
3. Autoridades sem poder para regular e monitorar o desenvolvimento das inteligências artificiais: diz respeito a gerência e controle das ferramentas de desenvolvidas e usadas na UE;
4. Incerteza legal e complexidade sobre implementar inteligências artificiais: diz respeito a tratativa de unificar dentro da UE as leis que regem IAs, para que evite problemas legais na medida do possível, visto que não existe como saber quais os problemas futuros que a IA podem ser criados;

5. Atrapalhar o desenvolvimento de inteligência artificial na Europa: é a preocupação de que o *AI Act* não atrapalhasse o desenvolvimento de IA em relação ao resto do mundo;
6. Problemas no cenário além das fronteiras: diz respeito as interações de IAs de outros países na EU e das IAs da EU em outros países.

Com os principais pontos de preocupação mapeados, a UE decidiu por tomar como caminho uma abordagem de risco no *AI Act*. Dessa forma, em vez de decidir quais providências tomar para todos os usos de inteligência artificial, foram criados 4 níveis de risco e descrições de quais aplicações se encaixam nesses níveis e quais atitudes tomar. Tais níveis, e suas abordagens, são aplicados para quaisquer empresas na UE ou empresas fora da UE que tenham como alvos cidadãos na Europa, tornando assim grandes empresas de tecnologia como o X, Meta, Alpha, Apple e Microsoft sujeitos à aplicação da lei. A Meta foi uma das grandes motivadoras das discussões na UE sobre direito digital, devido ao caso Facebook–Cambridge Analytica, onde a coleta de dados de redes sociais e sua posterior análise de dados possibilitou a manipulação do referendo do Brexit e de diversos processos eleitorais ao redor do mundo (FORNASIER; BECK, 2020). O *AI Act* também não será aplicado sobre governos, ferramentas militares ou organizações internacionais, uma vez que existem tratados internacionais que devem ser respeitados, criados antes ou podendo ser criados posteriormente à lei. Também havia a preocupação de países que fazem parte da Organização do Tratado do Atlântico Norte (OTAN) tenham que se desenvolver mais rapidamente no campo da IA para fins bélicos e dessa forma, o *AI Act* poderia ser um empecilho. Com a guerra da Ucrânia, o uso de IA no campo de batalha se mostrou cada vez mais necessário como o uso nos setores de defesa de mísseis antiaéreos, desenvolvimento de softwares que apoiam o movimento e a logística, gestão de conhecimento e administração e pequenos sistemas aéreos não tripulados (CHACÓN, 2019; TOSCANO, 2023). A IA poderá ser a chave da desigualdade bélica entre países.

O *AI Act* usa uma pirâmide de risco para elucidar as práticas de inteligência artificial e relacioná-las com seu risco em potencial. A pirâmide apresenta 4 níveis, sendo o nível superior (topo) tratado como a maior ameaça e o nível inferior

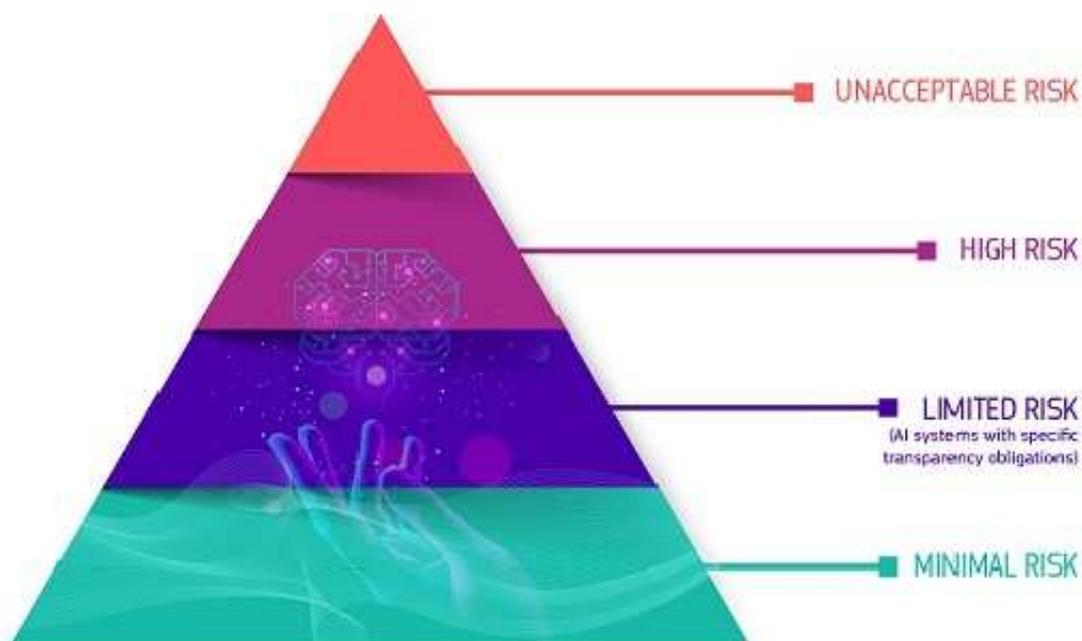


Figura 1 – Pirâmide de risco (EUROPEIA, 2021)

(base) a menor. Para cada nível, existe uma ação recomendada atrelada a uma classificação.

Serão classificadas como Risco Inaceitável as inteligências artificiais que deverão ser proibidas. Para ser classificada como Risco Inaceitável, a inteligência artificial deve estar enquadrada em uma das características abaixo:

- Ser capazes de criar controles subliminares a determinados grupos de usuários;
- Ser capazes de explorar grupos vulneráveis específicos, como pessoas com deficiências físicas e mentais;
- Ser usadas por governos (ou empresas a mando dos mesmos ou em benefícios dos mesmos) para classificar pessoas;
- Ser usada para à exploração de biometria para rastrear pessoas em espaços públicos, exceto em casos especiais.

Serão classificadas como Alto Risco as inteligências artificiais que deverão ser permitidas ante avaliação de conformidade e sujeitas aos requisitos de inteligência artificial. São considerados inteligências artificiais de Alto Risco componentes de sistemas de segurança de um produto ou item que sejam classificados na legislação de saúde e segurança da UE, como brinquedos, aviões, automóveis, elevadores, aparelhos médicos entre outros igualmente sensíveis à população. Também serão como Alto Risco inteligências artificiais que atuem em 8 categorias especiais determinadas pela *AI Act*:

- Biometria e categorização de pessoas;
- Gerenciamento e operação de sistemas críticos;
- Educação e treinamento vocacional;
- Empregos, gerenciamento de pessoas e acesso ao trabalho;
- Acesso a serviços privados essenciais, benefícios oferecidos pelos governos ou serviços públicos em geral;
- Aplicação da lei;
- Migração, asilo e controle de fronteiras;
- Administração da justiça e processos democráticos.

Serão classificadas como Risco Limitado as inteligências artificiais que estão permitidas, mas estão sujeitas à prestação de informações e relatórios de transparência para órgãos e à população. Nesta categoria encontram-se as IAs de sistemas com os quais atualmente as pessoas têm um maior convívio no seu dia a dia. São elas:

- Sistemas que interagem com humanos como por exemplo *chatbots*;
- Sistemas de reconhecimento e categorização de emoções e sentimentos;
- Sistemas que geram ou manipulam áudio, vídeo ou imagens.

Por fim, serão classificadas como Baixo Risco ou Mínimo inteligências artificiais que não se encaixam nas outras categorias e serão permitidas sem nenhuma ação necessária.

Os documentos utilizados como base para o *AI Act* bem como o número de reuniões e discussões sobre a proposta criaram uma maturidade para a UE sobre o assunto. As decisões sobre não embarcar fins militares no *AI Act* acabou por se mostrar importante aos países do bloco, visto que entre as discussões (2018) e a aprovação (2023), ocorreu a Guerra da Ucrânia contra a Rússia (2022), que está sendo um campo de batalha travado no mundo real e digital, sendo o primeiro grande conflito Europeu com o uso de IAs capazes de gerar desinformação em massa, técnicas sofisticadas de batalha e uso de armas não pilotadas. A Polônia foi um dos fornecedores de armas para o lado ucraniano, sendo ela membro da UE (MCCLUSKEY, 2023). Outro ponto que acabou por ser bastante discutido é a não discriminação e o preconceito algoritmo, assuntos delicados na UE, devido ao passado histórico da Segunda Guerra e dos atuais problemas causados pelas novas culturas e povos que imigraram, legalmente ou não, e que agora fazem um número expressivo de participantes da população da UE. Os textos gerados por inteligência artificial, bastante populares atualmente, não são uma preocupação pois diferente dos manipuladores de vídeo, áudio e imagem, não causam no momento problemas relevantes. Os outros, podem ser usados para gerar provas falsas, acusações falsas e crimes, sendo uma preocupação maior para os legisladores da UE.

2.2 DISCUSSÕES DE INTELIGÊNCIA ARTIFICIAL NOS ESTADOS UNIDOS

Nos Estados Unidos (EUA), as discussões sobre inteligência artificial ocorrem há muitos anos. É possível ver nos filmes, séries e livros, discussões sobre utopias e distopias regidas por inteligência artificial e manipulação de pessoas por inteligências artificiais são vistas desde o filme *Jogos de Guerra* de 1983, em que o adolescente David conecta acidentalmente seu computador ao do Departamento de Defesa Americano com Inteligência Artificial capaz de desencadear a Terceira Guerra Mundial. O filme gerou tanto impacto que a *Computer Fraud and Abuse Act*(CFAA) nos Estados Unidos, que criminaliza acessos não autorizados a computadores e atividades relacionadas à computação (CONSTANT, 2013). Entretanto,

mesmo com diversas representações midiáticas do medo da inteligência artificial e seus abusos, somente após o governo Trump, os EUA começaram a formalizar suas discussões em projetos de lei para a regulação dessas inteligências. Embora não apresente ainda nenhuma lei que regula a inteligência artificial em seu território, os Estados Unidos apresentaram em 2022 o *Blueprint for an AI Bill of Rights* (HOUSE, 2022), que é descrito como um *whitepaper* e um *framework* para guiar os usos e desenvolvimentos de inteligência artificial, mesmo que sem uma lei como o caso da UE. Além do *Blueprint*, os Estados Unidos apresentam um código de ética do Departamento de Defesa, que tem como objetivo apresentar as bases éticas para o desenvolvimento e uso de inteligência artificial para fins bélicos (BOARD, 2019).

2.2.1 *Blueprint for an AI Bill of Rights*

O *Blueprint for an AI Bill of Rights (AI Bill)* foi criado para ser um guia para o desenvolvimento de inteligências artificiais no território americano, que tem como objetivo impactar os direitos civis dos Estados Unidos, as liberdades civis, a privacidade, o acesso a oportunidades e a recursos críticos ou serviços. O *AI Bill* utiliza a mesma abordagem orientada ao produto que foi utilizada na *AI Act* da UE, embora não utilize o mesmo mapeamento de risco ou foque nos mesmos problemas que a UE. O *AI Bill* utiliza 5 princípios como guias para a sua formalização:

- Sistemas Seguros e Eficazes (*Safe and Effective Systems*) - Os sistemas de inteligência artificial devem ser feitos com a consulta de diversas comunidades, especialistas e *stakeholders* para identificar problemas, riscos e impactos.
- Proteções contra discriminação algorítmica (*Algorithmic Discrimination Protections*) - Os sistemas não devem ser implantados com discriminação algorítmica e devem ser usados para a equidade.
- Privacidade de Dados (*Data Privacy*) - Os sistemas devem ser protegidos de práticas abusivas com proteções pré-estabelecidas e devem ter o controle de como os dados são usados.

- Alerta e Explicação (*Notice and Explanation*) - Os usuários devem entender como que um sistema com inteligência artificial é usado e como ele contribui para impactar o usuário e os outros no seu entorno.
- Alternativas humanas (*Human Alternatives, Considerations and Fallback*) - Deve-se permitir que o usuário possa opinar ou não, quando apropriado, e dar acesso a suporte humano de forma rápida.

O maior problema do *AI Bill* é que ele não prevê nenhuma sanção a sistemas que violam seus princípios. Na realidade, ele é mais um motivador para incentivar a população, empresas e pesquisadores a entender os impactos que a inteligência artificial pode criar na sociedade e fazer os mesmos a contribuir com as futuras discussões e projetos de lei de inteligência artificial que devem ser debatidas nos próximos anos.

2.2.2 *AI Act* do Departamento de Defesa

Embora o *AI Bill* seja o maior guia nos Estados Unidos acerca dos princípios para o uso de inteligência artificial, outro documento se destacou. O *AI Act* (DEFENSE, 2020) criado pelo Departamento de Defesa dos Estados Unidos (DoD). Neste documento, o DoD estabelece 5 princípios para o desenvolvimento e uso de inteligência artificial nos seus domínios. Para a elaboração do documento, foram utilizados como base a Constituição dos EUA, a Lei da Guerra dos EUA (DEFENSE, 2023), os tratados internacionais em vigor nos Estados Unidos e Título 10 do Código dos Estados Unidos (*Title 10 of the United States Code*) (CONGRESS, 1956), sendo este último a base legal para as funções, missões e organização de cada um dos serviços públicos dos EUA, bem como do DoD. Segue abaixo os princípios listados pelo DoD:

- Responsável (*Responsible*): O DoD exercerá níveis apropriados de julgamento e cuidado, enquanto permanecerá responsável pelo desenvolvimento, implantação e uso de capacidades de inteligência artificial.
- Equitativo (*Equitable*): O DoD tomará medidas para minimizar preconceitos não intencionais nas capacidades de IA.

- Rastreável (*Traceable*): As capacidades de inteligência artificial do DoD serão desenvolvidas e implantadas de forma que o pessoal relevante possua uma compreensão apropriada da tecnologia, processos de desenvolvimento e métodos operacionais aplicáveis às capacidades de inteligência artificial, incluindo metodologias transparentes e auditáveis, fontes de dados e documentação de procedimentos de design.
- Confiável (*Reliable*): As capacidades de inteligência artificial do DoD terão usos explícitos e bem definidos, e a segurança, a proteção e a eficácia de tais capacidades estarão sujeitas a testes e garantia dentro desses usos definidos em todo o seu ciclo de vida.
- Governável (*Governable*): O DoD projetará e construirá capacidades de inteligência artificial para cumprir suas funções pretendidas enquanto possui a capacidade de detectar e evitar consequências não intencionais, e a capacidade de desativar sistemas implantados que demonstrem comportamento não intencional.

O DoD com o *AI Act* sinaliza que vê a IA como uma fronteira bélica que precisa ser explorada e que não perca de vista os valores patrióticos dos EUA. Os EUA também são um país pertencente a OTAN e ao Conselho de Segurança da Organização das Nações Unidas. Desta forma, assim como o *AI Act* da UE, qualquer lei que vier a ser debatida pelos EUA provavelmente não irá ambarcar fins militares, visto que essa foi uma das discussões que também aconteceram na UE. Como os EUA são fornecedores de materiais bélicos para a Ucrânia e outros membros da OTAN, é provável que seus avanços em IA e os princípios descritos acima já estejam sendo aplicados em campo de batalha.

2.3 DISCUSSÕES DE INTELIGÊNCIA ARTIFICIAL NA CHINA

Enquanto a UE aprova a primeira lei de inteligência artificial com o *AI Act*, que servirá de base para o resto do ocidente para a elaboração de suas leis sobre o tema, a China já apresenta uma série de leis e regulamentos para o uso de inteligência artificial. Esses esforços para regulamentar e legislar foram impulsionados em 2016, quando o AlphaGo, programa criado pelo Google, venceu o duelo contra o

campeão mundial de Go, Lee Sedol, 18 vezes campeão mundial (HINE; FLORIDI, 2022). Diferente dos Estados Unidos e da Europa, os regulamentos e leis chinesas sobre inteligência artificial tem como foco principal manter a estabilidade social e combater os efeitos do envelhecimento da população. A estabilidade social é um ponto tão importante para o Partido Comunista Chinês (PCC) que eles deixam especificamente descrito em todos seus regulamentos que esta é a prioridade de todas as inteligências artificiais que forem desenvolvidas no país. Entende-se que a definição de humanidade nesses documentos é especificamente voltada para a ideia de que "só se é cidadão aquele que apoia o Partido Comunista" (HINE; FLORIDI, 2022). A China tem um modelo de autoritarismo fragmentado para reger suas determinações sobre inteligência artificial. As metas determinadas são executadas pelo Estado e pelas empresas privadas chinesas. O autoritarismo fragmentado é uma visão do poder de organização do PCC, caracterizado por ser descentralizado e com competição interna. Isso resulta em autonomia regional, com competição intra-partidária e fragmentação burocrática, afetando a implementação de políticas e a economia politizada. Isso é um reflexo não somente do modelo político da China, como pelo Confucionismo, que serve como base filosófica para a China e que tem como princípio a preservação do equilíbrio (IP, 2009).

2.3.1 *New Generation Artificial Intelligence Code of Ethics*

New Generation Artificial Intelligence Code of Ethics (SCIENCE; CHINA, 2021) é o principal documento que visa integrar a ética em todo o ciclo de vida da inteligência artificial e fornecer uma base para o envolvimento em atividades relacionadas à inteligência artificial. Ele fornece a orientação ética para pessoas físicas, jurídicas e outras instituições relevantes no país. No documento elaborado, os pilares éticos são montados para definir e/ou proibir como um software que embarca uma inteligência artificial deve ser criado, monitorado, auditado e usado. Este documento estabelece que qualquer IA ou uso de IA que tenha a intenção de abalar a estrutura social do país, cometer crimes ou ir contra o governo deve ser proibido. As aplicações devem ser desenvolvidas com cautela, visando o bem estar social, promover a igualdade, a não-discriminação e garantir a segurança dos dados dos usuários. O documento é composto por 24 artigos, sendo 8 de Considerações

Iniciais, 4 de Especificações de Pesquisa e Desenvolvimento, 4 sobre Especificações de Fornecimento das Tecnologias, 5 artigos sobre Especificações de Uso e 3 sobre Organização e Implementação. Podemos extrair do documento, as seguintes sínteses de cada grupos de artigos:

- Considerações Iniciais - Neste conjunto de artigos, estabelece-se que uma visão ética deve integrar todo o ciclo de vida de desenvolvimento de sistemas de inteligência artificial. Além disso, o desenvolvimento de sistema inteligentes deve promover equidade, justiça, harmonia, segurança e evitar problemas como preconceito, discriminação, invasão de privacidade e vazamento de informações. É determinado também quem são os agentes e atividades sujeitos a aplicação do documento.
- Especificações de Pesquisa e Desenvolvimento - Artigos que promovem a conscientização da governança, segurança e riscos da inteligência artificial
- Especificações de Fornecimento das Tecnologias - Artigos que reforçam as questões de maturidade das bases da pesquisa científica e como a coleta e processamento de dados em inteligencia artificial deve ser madura e auditável para evitar problemas e preconceitos gerados pelos algoritmos utilizados.
- Especificações de Uso - Estimula a maturidade da governança dos processos que envolvam inteligencias artificiais, define os limites baseados nas considerações iniciais do documento e reforça os ideais de estabilidade e avanço social que devem ser promovidos.
- Organização e Implementação - Esse bloco de encerra citando quem emitiu o documento e quais instituições podem e devem formular complementos e/ou normas de inteligência artificial, usando o *New Generation Artificial Intelligence Code of Ethics* como base.

A própria arquitetura política descentralizada da China acaba por ser um problema na unificação das estratégias e legislações de IA. Por isso o *New Generation Artificial Intelligence Code of Ethics* se destaca, sendo um documento bem estruturado, claro e direto sobre as ambições e missões da IA para a China nos próximos anos. O PCC mostra que IA será a base da nova visão de tecnologia que

a China vislumbra e uma ferramenta crucial para solidificar cada vez mais o PCC. A China não abordar preconceito racial ou religioso no documento (embora fale de preconceito algorítmico), devido ao fato de ser um país praticamente homogêneo racialmente e ter regras rígidas sobre a demonstração e pregação de religiões na China.

2.4 CONCLUSÃO

Nesse capítulo, verificamos as principais documentações a respeito de guias éticos e legislações da Europa, EUA e China em relação a IA. Com esses documentos, foi possível traçar um paralelo entre a abordagem de risco da UE como uma possível inspiração no *AI Bill*. Também foi possível perceber um isolamento da China na forma de determinar seus tratados de inteligência artificial, se preocupando mais com as questões nacionais do que com suas relações internacionais. Também é possível perceber uma naturalidade da China com a adoção de inteligência artificial no futuro cotidiano chinês, tentando minimizar ao máximo o impacto dessas tecnologias na estrutura social chinesa.

Pode se verificar que a UE apresenta uma abordagem de risco bem definida e estruturada, com o bloqueio de determinadas aplicações de IA. Essas restrições não são aplicadas para fins militares na UE, visto que seus países tem estruturas próprias para tratar do assunto e alguns deles serem membros da Organização do Tratado do Atlântico Norte (OTAN). Com a criação e aprovação da *AI Act*, a UE mostra que está madura em suas discussões acerca da IA, visto que além da mesma, possui uma série de outros documentos sobre IA em diversos cenários de aplicação (como por exemplo os documentos para uso de IA sobre o regime de responsabilidade civil aplicável à inteligência artificial (EUROPEU, 2020b)).

Nos EUA, os documentos pesquisados mostram que o país tem a visão de que os fins militares devem ser separados dos civis no desenvolvimento e sobre as leis que abordam IAs. Os EUA não eram membros da UNESCO entre 2018 e 2023, quando a administração do Governo Trump os retirou como signatário (ZANCHETTA, 2023), fazendo que os documentos sobre IAs gerados nesse período não tivessem nenhuma obrigação de consultar as recomendações da UNESCO. Foi nesse período que foram publicados os *AI Act* do DoD e *AI Bill*. O *AI Act* do DoD não tem

nenhuma obrigação de adotar pontos sobre as questões bélicas recomendadas pela UNESCO, visto que os EUA pertencem a OTAN, sendo assim desinteressante embarcar IAs em equipamentos bélicos. A UNESCO não tem força política ou policial para forçar uma nação a acatar ou consultar as recomendações, mesmo que ela tenha assinado. Os EUA tem como base o capitalismo e liberalismo nos seus documentos e decisões, demonstrando preocupação em exercer liderança global em IA, mesmo que não esteja no momento nessa posição, ocupada pela China. O país tem como ideal o desenvolvimento rápido e disruptivo de IA (HINE; FLORIDI, 2022).

A China, diferente dos EUA, não o enxerga como um rival a ser superado, ou pelo menos não transparece isso em seus documentos. Sua preocupação maior se encontra no seu próprio desenvolvimento de forma que não crie um mal estar social ou hierárquico no país. Manter a estrutura das camadas sociais, uma unidade no avanço das tecnologias de IA, usando seus princípios éticos bem definidos, com prosperidade ambiental e ecossistêmica sendo uma preocupação, garantindo diversidade e inclusão é o foco dos documentos chineses.

A China tem uma visão de preconceito algorítmico diferente da UE e EUA, vista na comparação entre seus documentos. Isso é devido às brigas raciais serem uma discussão mais presente dentro da UE e dos EUA. Os EUA tem enfrentado diversos conflitos raciais nos últimos anos que esquentaram e evidenciaram o debate, enquanto a UE além de também ter essas questões, tem como catalisador a questão religiosa, trazida principalmente pelas relações atuais com o islamismo e a herança da Segunda Guerra.

A leitura dos documentos também evidencia que os princípios para IA escolhidos tem como base os direitos humanos e a soberania nacional, mesmo que esses países sejam diferentes culturalmente. Com isso, podemos definir esse como o norte para decisões que serão usadas e criadas em outros países no mundo pelos próximos anos. Isso não significa que essa seja a melhor decisão para todos os países e nem para a humanidade, mas que considerando o cenário global atual, essa será a tendência pelos governos.

3 DISCUSSÕES E ATOS SOBRE A INTELIGÊNCIA ARTIFICIAL NO BRASIL

Neste capítulo, iremos apresentar as discussões de inteligência artificial que moldam os pensamentos dos poderes legislativo e judiciário do Brasil, através de resoluções e projetos vigentes de inteligência artificial, já que o Brasil no momento não possui nenhuma lei ou recomendação ética sobre o tema inteligência artificial na Constituição. Mesmo sem apresentar tais documentos, o Brasil apresenta usos públicos de IA, como o Projeto Victor que auxilia na análise dos recursos recebidos pelo Supremo Tribunal Federal (STF) para reconhecer padrões de texto dentro dos recursos e classificá-los dentro de temas de repercussão geral já definidos pelo tribunal (MORAIS, 2022; FILHO; JUNQUILHO, 2018) e que é uma ferramenta que surgiu da necessidade de reduzir o tempo e esforço necessários para uma etapa preliminar de preparação de ações de classificação e uma forma de tratamento preliminar da massa jurídica de documentos para execução, na tarefa-fim, da classificação em temas de repercussão geral. De acordo com (PEIXOTO, 2020b), esse processo custava 1/3 da força de trabalho do STF, uma vez que eram recebidos cerca de 400 novos processos por dia.

3.1 PROJETOS DE LEI ACERCA DO USO DE INTELIGÊNCIA ARTIFICIAL NO BRASIL

Em 3 de fevereiro de 2022 havia três projetos de lei simultâneos tramitando no Senado Federal: Projeto de Lei (PL) nº 5.051, de 2019, de autoria do Senador Styvenson Valentim, que estabelece os princípios para o uso da Inteligência Artificial no Brasil (VALENTIM, 2019); o PL nº 21, de 2020, do Deputado Federal Eduardo Bismarck, que estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil (BISMARCK, 2020) e o PL nº 872, de 2021, do Senador Veneziano Vital do Rêgo, que dispõe sobre o uso da Inteligência Artificial (RÊGO, 2021). Devido ao número de PL que tramitavam simultaneamente no Senado, o Presidente do Senado na ocasião, o Senador Rodrigo Pacheco, por meio do Ato do Presidente do Senado Federal nº 4, de 2022,

instituiu uma comissão para a elaboração de uma minuta que substituiria todos os PL apresentados até o momento e elaboraria um PL que tivesse participação de membros da academia, sociedade civil, especialistas, governo e setor privado. O resultado da comissão foi o PL N° 2338, de 2023 (PACHECO, 2023).

O PL n° 2338 tem parágrafos que apresentam ideias do *AI Act* da UE, uma vez que este já estava aprovado enquanto o PL era desenvolvido. Ele também apresenta uma abordagem de risco, em 3 camadas onde a maior é a de Risco Excessivo, seguida pela camada de Alto Risco e a última camada não apresenta um nome específico no texto, mas é composta de inteligências artificiais que não se encaixam em nenhuma das duas outras camadas. O texto é composto por 45 Artigos, dividido por 9 capítulos. O documento foi resumido de maneira explicativa a seguir:

1. Disposições Preliminares - O objetivo deste capítulo é formalizar as normas gerais do uso da inteligência artificial, respeitando os direitos já estabelecidos no Brasil, definindo seus fundamentos e outras definições jurídicas sobre o PL.
2. Dos Direitos - Estabelece as diretrizes para o desenvolvimento e uso de sistemas de inteligência artificial no Brasil. Define que estes sistemas devem ser desenvolvidos e utilizados de forma a garantir a proteção dos direitos fundamentais e a promoção do bem-estar humano, descritas no capítulo. Define também que o usuário sempre pode pedir revisão acerca do processo de decisão e que um humano deve tomar a decisão final quando o sistema tomar decisões, previsões ou recomendações que tenham um impacto irreversível, de difícil reversão ou envolvam decisões que possam gerar riscos à vida ou à integridade física de indivíduos. É nesse capítulo que afirma que não pode haver preconceito algorítmico.
3. Da Categorização dos Riscos - Estabelece que os mesmos devem ser desenvolvidos e utilizados de forma a garantir a transparência e a prestação de contas, com a divulgação de informações relevantes sobre o sistema e seus resultados, sendo eles classificados pelo seu grau de risco assim como na Europa.

4. Da Governança dos Sistemas de Inteligência Artificial - Estabelece a robustez da governança, as avaliações de impacto do uso e as formas de apresentação desses dados para órgãos governamentais e usuários. Os resultados deverão ser publicados, assegurando os segredos industriais e comerciais dos sistemas contratados, respeitando a Lei Geral de Proteção de Dados.
5. Da Responsabilidade Civil - Aborda quais são as responsabilidades civis dos fornecedores e operadores de sistemas de inteligência artificial. Os agentes de inteligência artificial são responsáveis pelos danos causados por seus sistemas, devendo reparar integralmente o dano causado, independentemente do grau de autonomia do sistema, além de estar descrito neste capítulo que o sistema está sujeito a outras leis do Brasil.
6. Códigos de Boas Práticas e de Governança - Estabelece as condições para que os agentes de inteligência artificial possam formular códigos de boas práticas e de governança. Estabelece que os agentes podem desenvolver códigos de boas práticas e de governança para garantir a segurança, a transparência e a responsabilidade dos sistemas de inteligência artificial e que os mesmos devem comunicar à autoridade competente (que ainda será criada ou apontada entre as existentes) a ocorrência de graves incidentes de segurança, e define os tipos de incidentes que devem ser comunicados.
7. Da Comunicação de Incidentes Graves - Designa aos agentes de inteligência artificial a função de comunicar à autoridade competente a ocorrência de graves incidentes de segurança, incluindo quando houver risco à vida e integridade física de pessoas, a interrupção de funcionamento de operações críticas de infraestrutura, graves danos à propriedade ou ao meio ambiente, bem como graves violações aos direitos fundamentais, nos termos do regulamento.
8. Da Supervisão e Fiscalização - Estabelece as condições para que a autoridade competente possa supervisionar e fiscalizar a implementação da lei. Estabelece a missão da autoridade competente determinada pelo Poder Executivo que deverá zelar para sistemas com inteligência artificial (provavelmente, assim como na Lei Geral de Proteção de Dados Pessoais (LGPD))

(REPÚBLICA, 2018), será criado um órgão para esse fim). A autoridade competente e os órgãos reguladores devem coordenar suas atividades para garantir o cumprimento da lei; manterá um fórum permanente de comunicação com os órgãos reguladores para facilitar suas competências regulatórias, fiscalizatórias e sancionatórias; e que nos ambientes regulatórios experimentais (*sandbox* regulatórios) que são ambientes controlados para a testagem de IA, a autoridade competente será notificada e poderá se manifestar quanto ao cumprimento das finalidades e princípios da Lei. Determina que os participantes nos *sandbox* regulatórios da IA são responsáveis por quaisquer danos causados a terceiros durante a experimentação. Por fim, estabelece que cabe à autoridade competente a criação e manutenção de base de dados de inteligência artificial de alto risco, acessível ao público, que contenha os documentos públicos das avaliações de impacto, respeitados os segredos comercial e industrial, nos termos do regulamento.

3.2 CONCLUSÃO

Pode ser observada a influência do *AI Act* no PL n° 2338, fortemente evidenciado na categorização dos riscos, que define a quais grupos é vedado o uso de IA, sendo estes grupos os mesmos do *AI Act*. Também podemos ver que a abordagem de risco seguiu os mesmos padrões do *AI Act*, quando definiu os grupos de risco excessivo e alto risco, utilizando as mesmas definições do documento da UE. A União Europeia (UE) tem sido uma referência para o Brasil como foi evidenciado mais fortemente durante a preparação e discussões da LGPD no Brasil (MASSENO; MARTINS; JÚNIOR, 2020). Na ocasião, o legislativo utilizou uma série de ideias e de resoluções adotadas no Regulamento Geral sobre a Proteção de Dados (EUROPEU, 2016) para formalizar os debates e o próprio documento da Lei Geral da Proteção de Dados. A inspiração do modelo europeu foi tão forte, que acelerou o debate e a implementação da LGPD no Brasil. Então, é compreensível a influência do *AI Act* no PL n° 2338.

Mesmo com a influência da UE no texto brasileiro, podemos verificar que diferente da Europa, o Brasil se preocupou também na criação da lei em conjunto com a criação do órgão que irá supervisionar a mesma. O mesmo movimento já

havia sido praticado na criação da LGPD (REPÚBLICA, 2018). Podemos verificar também uma preocupação com a testagem dos sistemas que utilizaram modelos de IA em fins considerados sensíveis, uma preocupação que não ocorre na China e nos Estados Unidos. Existe também uma possível futura fonte de renda para o Governo descrita, quando se mostra a questão dos ambientes regulatórios experimentais (*sandbox* regulatórios), uma vez que o texto deixa em aberto se os mesmos ambientes serão alugados para empresas, pessoas e instituições de ensino e pesquisa, ou se serão usados apenas em casos determinados.

O texto brasileiro não abrange de forma direta todos os problemas debatidos na atualidade, como manipulação de imagens e vídeos. Este problema já ocorreu com conteúdos eróticos gerados por manipulação de imagens por IA, como o incidente em um colégio particular do Rio de Janeiro, em que estudantes teriam manipulado imagens de outras estudantes (POVO, 2023). Não há também menção a controle do comportamento dos jovens na rede ou o redirecionamento de conteúdo aos mesmos, embora o texto fale sobre exploração de grupos específicos baseados na idade. É um problema do Brasil a sexualização de menores sendo alimentada por redes sociais (RICARDO, 2022).

Enquanto os textos dos Estados Unidos, China e UE tem como maior foco questões como classificação de população e preconceito algorítmico, o texto brasileiro trabalha mais a questão de segurança pública. A segurança é um dos maiores problemas enfrentados no Brasil. Mesmo sem uma regulamentação, o Brasil apresenta um mercado de vigilância munida de IA, que já se apresenta de forma comum nos grandes centros urbanos, que conseguem alertar e reportar crimes e ajudando as autoridades a reconhecer e deter criminosos. Hoje, não existe nenhuma forma de comprovar que tais sistemas usam ou não algum tipo de viés para gerar esses alertas e reportes.

A preocupação com questões de identidade visual, direitos de mídia, imagem e artística não são citados na lei. Embora exista documentos na EU que tratam essas questões como a "Resolução do Parlamento Europeu, de 19 de Maio de 2021, sobre a inteligência artificial na educação, na cultura e no sector audiovisual"(EUROPEU, 2021c), o Brasil ainda não fez nenhum movimento sobre o tema, embora a comunidade artística já tenha iniciado os debates com os membros do congresso brasileiro (BOMFIM, 2023).

4 CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho, foi possível observar as correntes de pensamento, debates e documentos da UE, Estados Unidos, China, UNESCO e do Brasil. Foi possível observar que existe uma correlação entre essas informações e os receios que as motivaram. As 1 e 2 demonstram essas correlações entre os blocos e países estudados por esse trabalho.

Tabela 1 – Princípios Éticos em Diferentes Contextos Internacionais

Princípios	UNESCO	UE	EUA	China	Brasil
Proporcionalidade e não causar dano	X				
Segurança, proteção	X	X	X	X	X
Justiça e não discriminação	X	X	X	X	X
Sustentabilidade	X	X			
Direito à privacidade e proteção de dados	X	X	X	X	X
Supervisão humana	X				
Transparência	X	X	X	X	X
Responsabilidade e prestação de contas	X	X	X	X	X
Sistemas auditáveis	X	X	X	X	X
Conscientização e educação acerca das IA	X				X
Governança	X	X	X	X	X
Gerencia e supervisão humana	X	X			
Bem-estar social		X	X	X	
Preocupação com direitos de mídia		X		X	
Confiabilidade nos sistemas e processos		X	X	X	
Aumentar o conhecimento ético sobre a inteligência artificial					X
Criar ambientes seguros e auditáveis para testes		X			X

Tabela 2 – Receios Relacionados à Inteligência Artificial em Diferentes Contextos Internacionais

Receios	UNESCO	UE	EUA	China	Brasil
Violação dos direitos humanos e das liberdades	X	X	X	X	X
Desigualdade entre povos devido ao uso de inteligência artificial	X				
Controle do conhecimento e uso da inteligência artificial por países e regiões	X				
Uso bélico da inteligência artificial	X				
Falta de transparência e governança	X	X	X	X	X
Técnicas de controle subliminares		X	X	X	
Controle governamental		X	X		
Preconceito algorítmico	X	X	X	X	X
Descontrole energético	X	X			
Sistemas não auditáveis e com segurança fraca		X	X	X	X
Não explicação do destino dos dados		X	X	X	X
Dependência excessiva e perda do controle humano	X		X		
Impacto nos empregos				X	
Fim da estabilidade social				X	
Cibercrimes					X

A partir dos documentos usados durante o trabalho e pelas informações nas tabelas 1 e 2, foi possível perceber uma correlação entre as preocupações dos países em comum, sendo elas a promoção de políticas de redução da desigualdade, respeito a privacidade, garantia de sistemas robustos contra preconceitos, promover o respeito aos direitos humanos e sistemas auditáveis e seguros. Entretanto, apenas a China apresenta preocupações no sentido de preservar a estabilidade social, não promovendo demissões massivas com o uso de IA descritas em seus documentos. Este tópico tende a ser mais debatido pela população (COX, 2023). Observamos que na China não existe nos documentos o receio de controle governamental e

existindo o de manipulação de opinião (subliminar ou não), justamente favorecendo a perpetuação do PCC. Outro ponto que podemos observar é que diferente da UNESCO, os países não se propuseram a diminuir diferenças de desenvolvimento e uso de IA entre países e regiões, visto que não é interessante no ponto de vista bélico ou em questões de segurança nacional. As tabelas também evidenciam que o Brasil procurou mais formalizar o desenvolvimento e a regulamentação do mesmo do que com os impactos que causariam a sociedade.

A segurança digital é um dos principais pontos levantado nos textos e observado nas tabelas 1 e 2, não somente pela dificuldade de anonimizar dados ou não inferir as pessoas pelos dados resultantes, como também pelos diversos usos maliciosos que sistemas munidos de IAs podem apresentar. Analistas de segurança da informação também deve considerar que a IA pode ser usada como arma, uma vez que pode ser utilizada não somente para perfilar pessoas, identificando suas vulnerabilidades e preferencias, como também pode ser usada para criação de dicionários personalizados de senhas, através de processamento de dados das redes sociais ou servindo de base para ataques de *phishing* personalizados. Programas como o Maltego (Maltego, 2008) usado para criar relações entre comportamentos digitais e físicos para perfilar pessoas e entidades, podem ser usados para fazer análises em redes sociais (HAI-JEW, 2017) que posteriormente podem servir de entrada para dados em IAs mal intencionadas. As próprias redes sociais também podem fazer uso de atividades não éticas com IA. O TikTok, rede social criada na China e banida de uso na mesma (YEUNG, 2023), utiliza IA para personificar o *feed* de vídeos para cada usuário (KANG; LOU, 2022), tornando assim altamente viciante. Esse vício fez com que o TikTok fosse acusado de ser uma arma por países ocidentais como os Estados Unidos (SHEPARDSON, 2023).

Pode-se notar também que a auditoria e intervenção humana, bem como o entendimento de como os dados são decididos por sistemas com IA, são um dos pontos mais levantados entre os documentos. Existe a preocupação de que os sistemas nasçam ou se tornem enviesados ao longo do tempo, prejudicando assim pessoas que precisam desses sistemas. Afinal, seria extremamente problemático, por exemplo, juntar questões financeiras para decidir uma possível pessoa a receber um transplante de órgão, ou decidir prender ou não uma pessoa baseada no seu tom de pele. Entretanto, cabe-se ressaltar que a intervenção humana não precisa ser

usada em casos simples, pois poderia gerar uma sobrecarga no tempo de decisão. Com base nas informações adquiridas e nas tabelas 1 e 2, selecionamos um conjunto de princípios que foram escolhidos através das informações adquiridas, que são comuns a todos os casos estudados:

- Equidade, respeito e não discriminação a classes, povos e indivíduos: Para evitar o preconceito algorítmico e abusos em favorecimento de um determinado grupo;
- Regras e registros claros: Para evitar resultados nebulosos e permitir a auditoria das decisões;
- Privacidade dos dados: Para proteger a privacidade e a identidade dos dados coletados e seus donos;
- Segurança durante todo o processo de vida dos dados usados: Para evitar ataques que têm como objetivo alterar resultados ou coletar dados;
- Responsabilidade sobre o uso: Para evitar o uso indevido e irresponsável, bem como decidir se um determinado sistema deve ou não ser desenvolvido;

Utilizando estes princípios como base, acreditamos que uma IA pode não apresentar problemas ou minimamente fornecer dados sobre eventuais pontos de falha para auxiliar novos desenvolvimentos a não cometê-los. O futuro das IAs deverá ser construído na base de erros e acertos, adquirindo assim experiência para os futuros desenvolvedores e usuários na construção e uso de IAs mais justas.

Uma possível extensão deste trabalho seria investigar mais a fundo o impacto de IA na sociedade. O trabalho apresenta as leis e documentos usados como base nos governos mas não sabemos se as IAs apresentam mais problemas após a publicação destes documentos. Outro possível trabalho seria a identificação de IAs que não seguem os princípios definidos pelos próprios países os blocos que os criaram. Também é possível estudar os impactos dos blocos e nações vizinhos as determinações, como no caso do Brasil que pertence ao MERCOSUL e pode impactar outros países do bloco. Também não está no escopo do trabalho quais punições a quem desenvolve uma IA com problemas e preconceitos e se o empregador da

mesma também deve ser responsabilizado, sendo também uma sugestão de trabalho futuro.

A inteligência artificial é uma ferramenta que pode auxiliar a humanidade em questões como aumento de produtividade, melhoras na qualidade do trabalho, avanço científico e auxílio a populações, entretanto é imprescindível um olhar atento de Governos e empresas para que essa importante ferramenta não se torne um problema, criando mais disparidade entre povos e permitindo abusos por parte de Governos ditatoriais, empresas mal intencionadas e cibercriminosos.

REFERÊNCIAS

- BISMARCK, E. Projeto de lei n° 21, de 2020. 2020.
- BOARD, D. I. Ai principles: recommendations on the ethical use of artificial intelligence by the department of defense: supporting document. **United States Department of Defense**, 2019.
- BOMFIM, C. Lira quer retomar votação do pl das fake news e destravar debate sobre direitos autorais. **G1**, 2023. Disponível em: <https://g1.globo.com/politica/blog/camila-bomfim/post/2023/08/03/lira-quer-retomar-votacao-do-pl-das-fake-news-e-destravar-direitos-autorais.ghtml>.
- BRASIL, R. d. U. no. **Declaração universal dos direitos humanos**. Brasília, 1998. Disponível em: <https://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos>.
- CHACÓN, J. C. de la F. La inteligencia artificial y su aplicación en el mundo militar. In: INSTITUTO ESPAÑOL DE ESTUDIOS ESTRATÉGICOS. **La inteligencia artificial, aplicada a la defensa**. [S.l.], 2019. p. 69–98.
- CONGRESS, U. S. **U.S. Code Title 10 - Copyrights: Limitations on exclusive rights: Fair use**. 1956. Disponível em: <https://uscode.house.gov/browse/prelim@title10&edition=prelim>.
- CONSTANT, S. A. The computer fraud and abuse act: A prosecutor’s dream and a hacker’s worst nightmare-the case against aaron swartz and the need to reform the cfaa. **Tul. J. Tech. & Intell. Prop.**, HeinOnline, v. 16, p. 231, 2013.
- COTRIM, G.; FERNANDES, M. Fundamentos da filosofia. **São Paulo: Saraiva**, 2013.
- COX, J. Lira quer retomar votação do pl das fake news e destravar debate sobre direitos autorais. **BBC**, 2023. Disponível em: <https://www.bbc.com/portuguese/articles/cn4rvelpmejo>.
- DEFENSE, G. C. of the Department of. Law of war manual. 2023. Disponível em: <https://media.defense.gov/2023/Jul/31/2003271432/-1/-1/0/DOD-LAW-OF-WAR-MANUAL-JUNE-2015-UPDATED-JULY%202023.PDF>.

DEFENSE, U. S. D. of. Dod adopts ethical principles for artificial intelligence. 2020. Disponível em: <https://www.defense.gov/News/Releases/release/article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/>.

EUROPEIA, C. Proposta de quadro regulamentar em matéria de inteligência artificial. 2021. Disponível em: <https://digital-strategy.ec.europa.eu/pt/policies/regulatory-framework-ai>.

EUROPEU, P. Regulamento (ue) 2016/679 do parlamento europeu e do conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a diretiva 95/46/ce (regulamento geral sobre a proteção de dados). 2016. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A32016R0679>.

EUROPEU, P. Regime relativo aos aspetos éticos da inteligência artificial, da robótica e das tecnologias conexas. 2020. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_PT.html.

EUROPEU, P. Resolução do parlamento europeu, de 20 de outubro de 2020, que contém recomendações à comissão sobre o regime de responsabilidade civil aplicável à inteligência artificial. 2020. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_PT.html.

EUROPEU, P. Resolução do parlamento europeu, de 20 de outubro de 2020, sobre os direitos de propriedade intelectual para o desenvolvimento de tecnologias ligadas à inteligência artificial. 2020. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_PT.html.

EUROPEU, P. European parliament resolution of 20 january 2021 on artificial intelligence: questions of interpretation and application of international law in so far as the eu is affected in the areas of civil and military uses and of state authority outside the scope of criminal justice. 2021. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0009_EN.html.

EUROPEU, P. Meps call for an ethical framework to ensure artificial intelligence respects eu values. 2021. Disponível em: <https://www.europarl.europa.eu/news/pt/press-room/20210517IPR04135/meps-call-for-an-ethical-framework-to-ensure-ai-respects-eu-values>.

EUROPEU, P. Resolução do parlamento europeu, de 19 de maio de 2021, sobre a inteligência artificial na educação, na cultura e no sector audiovisual (2020/2017(ini)). 2021. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0238_PT.html.

EUROPEU, P. Resolução do parlamento europeu, de 6 de outubro de 2021, sobre a inteligência artificial no direito penal e a sua utilização pelas autoridades policiais e judiciárias em casos penais. 2021. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_PT.html.

FIGUEIREDO, A. M. Ética: origens e distinção da moral. **Saúde, Ética & Justiça**, Universidade de Sao Paulo, Agencia USP de Gestao da Informacao Academica (AGUIA), v. 13, n. 1, p. 1, jun 2008.

FILHO, M. S. M.; JUNQUILHO, T. A. Projeto victor: perspectivas de aplicação da inteligência artificial ao direito. **Revista de Direitos e Garantias Fundamentais**, Sociedade de Ensino Superior de Vitoria, v. 19, n. 3, p. 218–237, dec 2018.

FORNASIER, M. D. O.; BECK, C. CAMBRIDGE ANALYTICA: ESCÂNDALO, LEGADO e POSSÍVEIS FUTUROS PARA a DEMOCRACIA. **Revista Direito em Debate**, Editora Unijui, v. 29, n. 53, p. 182–195, may 2020.

GARCIA, A. C. Ética e inteligencia artificial. **Computação Brasil**, n. 43, p. 14–22, 2020.

HAI-JEW, S. Real-time sentiment analysis of microblog messages with the maltego “tweet analyzer” machine. In: **Social Media Listening and Monitoring for Business Applications**. [S.l.]: IGI Global, 2017. p. 316–337.

HINE, E.; FLORIDI, L. Artificial intelligence with american values and chinese characteristics: a comparative analysis of american and chinese governmental AI policies. **AI & SOCIETY**, Springer Science and Business Media LLC, jun 2022.

HOUSE, T. W. Blueprint for an ai bill of rights. 2022. Disponível em: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>.

IP, P. K. Is confucianism good for business ethics in china? **Journal of Business Ethics**, Springer Science and Business Media LLC, v. 88, n. 3, p. 463–476, aug 2009.

KANG, H.; LOU, C. AI agency vs. human agency: understanding human–AI interactions on TikTok and their implications for user engagement. **Journal of Computer-Mediated Communication**, v. 27, n. 5, p. zmac014, 08 2022. ISSN 1083-6101. Disponível em: <https://doi.org/10.1093/jcmc/zmac014>.

LOPES, A. Parlamento europeu aprova "eu ai act", primeiro marco regulatório de ia do mundo. **Exame**, 2023. Disponível em: <https://exame.com/inteligencia-artificial/parlamento-europeu-aprova-eu-ai-act-primeiro-marco-regulatorio-de-ia-do-mundo/>.

MADIEGA, T. A. Artificial intelligence act. **European Parliament: European Parliamentary Research Service**, 2021.

Maltego: Open Source Intelligence and Forensics Software. 2008. <https://www.maltego.com/>. Acessado em 29 de abril de 2024.

MASSENO, M. D.; MARTINS, G. M.; JÚNIOR, J. L. de M. F. A segurança na proteção de dados: Entre o RGPD europeu e a LGPD brasileira. **Revista do CEJUR/TJSC: Prestação Jurisdicional**, Editora Alumniin, v. 8, n. 1, p. e346, jan 2020.

MCCLUSKEY. Polônia anuncia que deixará de fornecer armas para a ucrânia. set. 2023. Disponível em: <https://www.cnnbrasil.com.br/internacional/polonia-anuncia-que-deixara-de-fornecer-armas-para-a-ucrania/>.

MORAIS, F. S. de. O uso da inteligência artificial na repercussão geral: Desafios teóricos e éticos. **Direito Público**, Instituto Brasiliense de Direito Publico, v. 18, n. 100, jan 2022.

PACHECO, R. Projeto de lei nº 2338, de 2023. **SF/23833.90768-16**, 2023.

PEIXOTO, F. H. **Direito e Inteligência artificial - Referenciais Básicos: com comentários à Resolução CNJ 332/2020**. [S.l.]: Fabiano Hartmann Peixoto, 2020.

PEIXOTO, F. H. Projeto victor: relato do desenvolvimento da inteligência artificial na repercussão geral do supremo tribunal federal. **Revista Brasileira de Inteligência Artificial e Direito-RBIAD**, v. 1, n. 1, p. 1–22, 2020.

PEREIRA, M. S. d. C.; PERDONA, G. d. S. C. Do algoritmo ao preconceito: um balanço literário. **Anais**, 2022.

POVO, G. do. Colégio católico no rio investiga estudantes que usaram ia para forjar nudes de colegas. 2023. Disponível em: <https://www.gazetadopovo.com.br/vida-e-cidadania/colégio-rio-investigacao-montagem-nudes/>.

RÊGO, V. V. do. Projeto de lei nº 872, de 2021. **SF/21733.11077-00**, 2021.

REPÚBLICA, P. da. Lei geral de proteção de dados - lei nº 13.709, de 14 de agosto de 2018. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm.

RICARDO, I. Exposição de crianças em redes sociais é porta de entrada para criminosos. 2022.

ROSSETTI, R.; ANGELUCI, A. Ética algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação. **Galáxia (São Paulo)**, FapUNIFESP (SciELO), n. 46, 2021.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. Upper Saddle River, NJ, 2021.

SCIENCE, M. of; CHINA, T. of the People's Republic of. New generation artificial intelligence code of ethics. set. 2021. Disponível em: https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html.

SHEPARDSON, R. A. D. Tiktok congressional hearing: Ceo shou zi chew grilled by us lawmakers. **Reuters**, 2023. Disponível em: <https://www.reuters.com/technology/tiktok-ceo-face-tough-questions-support-us-ban-grows-2023-03-23/>.

SOVRANO, F. et al. Metrics, explainability and the european AI act proposal. **J, MDPI AG**, v. 5, n. 1, p. 126–138, feb 2022.

TOSCANO, B. M. Inteligencia artificial en el espacio ultraterrestre:¿ un nuevo desafío para la otan? Universidad de Sevilla, 2023.

UNESCO. Recomendação sobre a Ética da inteligência artificial. p. 45, 2021. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por.

VALENTIM, S. Projeto de lei n° 5051, de 2019. **SF/19687.15646-23**, 2019.

YEUNG, S. W. J. Tiktok is owned by a chinese company. so why doesn't it exist there? **Reuters**, 2023. Disponível em: <https://edition.cnn.com/2023/03/24/tech/tiktok-douyin-bytedance-china-intl-hnk/index.html>.

ZANCHETTA, C. The united states in and out of unesco: the turbulent relationship between the country and the organisation from its creation until today. 2023.