



TAYRONE MARQUESINI CHIAVONE

SOCIEDADE DE RISCO DIGITAL

O Princípio da Precaução na Regulação da Inteligência Artificial

Dissertação com vista à obtenção do grau
de Mestre em Direito na especialidade de
Direito Internacional e Europeu.

Orientadora:

Doutora Soraya Nour Sckell,

Professora da Faculdade de Direito da Universidade Nova de Lisboa

Junho 2022



TAYRONE MARQUESINI CHIAVONE

SOCIEDADE DE RISCO DIGITAL

O Princípio da Precaução na Regulação da Inteligência Artificial.

Dissertação com vista à obtenção do grau de Mestre em Direito na especialidade de Direito Internacional e Europeu.

Orientadora:

Doutora Soraya Nour Sckell,

Professora da Faculdade de Direito da Universidade Nova de Lisboa

Junho 2022



SOCIEDADE DE RISCO DIGITAL

O Princípio da Precaução na Regulação da Inteligência Artificial.



Justice, democracy and citizenship
without borders

PTDC/FER-FIL/30686/2017

FCT Fundação
para a Ciência
e a Tecnologia

*A presente dissertação foi desenvolvida no âmbito do projeto
Cosmopolitismo: Justiça, Democracia e Cidadania sem Fronteiras.*

*(PTDC/FER-FIL/30686/2017, FCT - Fundação para a Ciência e a
Tecnologia, I.P, Portugal).*

URL <<https://cosmopolites.wixsite.com/cosmopolitanism>

DECLARAÇÃO DE COMPROMISSO ANTIPLÁGIO

Declaro por minha honra que o texto da dissertação que apresento é original e que todas as minhas citações estão corretamente identificadas. Tenho consciência de que a utilização de elementos alheios não identificados constitui uma grave falta ética e disciplinar.

Lisboa, 12 de junho de 2022.



Tayrone Marquesini Chiavone

CONTAGEM DE CARACTERES

Declaro que o corpo da tese ou dissertação, incluindo espaços e notas, ocupa um total de 190.041 caracteres.

Lisboa, 12 de junho de 2022.



Tayrone Marquesini Chiavone

DA FORMA DE CITAR E OUTRAS CONVENÇÕES

As citações foram apresentadas sob a forma de referências em notas de rodapé do seguinte modo: apelido do autor ou nome da entidade responsável pela obra; nome da obra; ano de publicação e páginas. Quando se trata de livros lidos no leitor de *e-books Kindle*, em lugar da página, é identificado o percentual da leitura. Serão acrescentados quaisquer outros elementos que se julguem convenientes.

As citações diretas que incluem transcrições de obras estrangeiras foram realizadas em sua língua original. As paráfrases das obras estrangeiras incluem traduções realizadas pelo autor para a língua portuguesa, assumindo o autor a responsabilidade por essas.

As referências foram divididas nas seguintes categorias: (a) referências bibliográficas, que incluem livros, capítulos de livros, artigos publicados em periódicos e em atas; (b) referências diversas, com: (b.1) notícias e textos publicados na internet; (b.2) legislação; (b.3) relatórios e cartilhas; e (b.4) vídeos.

As referências de livros são citadas da seguinte maneira: APELIDO, Primeiros nomes (ano de publicação) – **Título da obra**, Edição, Local de Publicação: Editora. No caso de haver mais de três autores, a obra é identificada pelo APELIDO, primeiros nomes do primeiro autor referido, seguido de *et al.*

Os artigos de publicações em série serão citados do seguinte modo: APELIDO, Primeiros nomes (ano de publicação) - **Título do artigo**. **Título da publicação em série**, Volume, Número, páginas. No caso de haver mais de três autores, a obra é identificada pelo APELIDO, primeiros nomes do primeiro autor referido, seguido de *et al.*

Os recursos bibliográficos obtidos na internet serão citados por meio da menção do autor (ou entidade responsável pela publicação) e do título do conteúdo publicado. A menção ao sítio de internet de onde o recurso foi extraído e a respectiva data de consulta estão incluídas na lista de bibliografia. O mesmo valerá para os artigos publicados em revistas eletrônicas.

A utilização de itálico servirá para destacar determinado conceito, terminologia específica ou estrangeirismos.

A presente dissertação respeita o Novo Acordo Ortográfico da Língua Portuguesa, aprovado pela Resolução da Assembleia da República nº 35/2008,

de 29 de julho. Em razão de seu autor ser brasileiro, as variantes utilizadas no Brasil foram as eleitas para esta dissertação. Isso, contudo, não se reflete nas citações diretas de legislação, autores ou obras editadas em Portugal, as quais não se alteraram para manter a integralidade.

A presente dissertação foi integralmente elaborada durante a pandemia causada pela COVID-19.

*“In a world deluged by irrelevant information, clarity is
power.”*

(Yuval Noah Harari, 21 lessons for the 21st Century).

AGRADECIMENTOS

No decorrer de minha vida, dentre os desafios que me predispos a enfrentar, este sem sombra de dúvidas foi o maior deles. Entretanto, ele só foi possível graças a uma série de pessoas muito especiais que me acompanharam nesta jornada.

De modo não exaustivo (me desculpo desde logo aos não mencionados), cumpre-me agradecer aqueles que de certo modo, marcaram diretamente minha trajetória.

Um primeiro agradecimento à pessoa que mais teve de conviver com este momento da minha vida, meu porto seguro, minha companheira, Tamiris. Obrigado pelo apoio incondicional nessa caminhada. Com você, tudo na vida se torna mais fácil.

Agradeço à minha amada mãe, Solange, criadora do meu eu interno e externo. Aos meus dois pais, Reinaldo e Manouchehr, que me fizeram ser o homem que sou.

A todos os meus familiares, que de modo direto e indireto ajudaram na minha formação moral e acadêmica. Vocês são minha essência!

À minha irmã Nabilla, minha sócia profissional, minha melhor amiga e por quem guardo os mais fortes sentimentos.

Aos meus colegas da Universidade Nova de Lisboa, João Fauza, Rafael Borba, Suzana Rahde, Mateus Vieira, Alyne Calistro, pela troca enriquecedora de conhecimento e acima de tudo, pela amizade construída.

À minha orientadora Soraya Nour Sckell, a quem possuo uma dívida intelectual impagável.

À Faculdade de Direito da Universidade Nova de Lisboa, pela experiência acadêmica proporcionada e por me dar um lugar na Europa para chamar de “lar”. Em conjunto, agradeço ao coordenador do curso, professor Francisco Pereira Coutinho e a todos os docentes que ocuparam as cadeiras letivas que cursei.

Aos meus amigos de fora da faculdade que contribuíram tecnicamente para a consecução deste trabalho, em especial, Isabelle Carvalho, Alexandre Ferrer, Caio Moraes, ao meu irmão Felipe Granito e é claro, ao meu sócio Victor Mulin.

À Fundação Getúlio Vargas - SP, em especial ao núcleo de Direito e

Tecnologia, por me aceitar como aluno intercambista por um semestre acadêmico. Agradeço principalmente, pelas aulas de Direito e Tecnologia ministradas pelo professor Alexandre Pacheco e pela professora Marina Feferbaum.

Ao Laboratório de Políticas Públicas e Internet, por ter me aceito como pesquisador voluntário do grupo de estudos em Inteligência Artificial.

Ao projeto da Cosmopolitismo: Justiça, Democracia e Cidadania sem fronteiras da FCT – Fundação para Ciência e Tecnologia, por me aceitar como colaborador.

A todos aqueles que acreditaram em mim, meu muito obrigado!

LISTA DE ABREVIATURAS

IA	Inteligência Artificial
UE	União Europeia
OECD	Organização para a Cooperação e Desenvolvimento Econômico
AGI	Artificial General Intelligence
ANI	Artificial Narrow Intelligence
ONU	Organização das Nações Unidas
EUA	Estados Unidos da América
GDPR	General Data Protection Regulation
LGPD	Lei Geral de Proteção de Dados

RESUMO

Esta dissertação propõe investigar a aplicabilidade do princípio da precaução na regulação de riscos da Inteligência Artificial. Adotando como modelo de análise a teoria da Sociedade de Risco de Ulrich Beck, a presente pesquisa analisou a transição da sociedade moderna para a *Modernidade Reflexiva*, reconhecendo que as incertezas trazidas pelo progresso tecnocientífico alteraram a lógica de distribuição de bens, que agora são acompanhadas pela distribuição de riscos. Os institutos trazidos pelo processo de modernização da sociedade, demonstram-se insuficientes para controlar e quantificar os riscos deste processo. A Quarta Revolução Industrial impulsiona e amplifica as reflexões trazidas por esta teoria social, fazendo com que a lógica distributiva de bens se opere por veículos digitalizados, conseqüentemente, os riscos se distribuem de igual modo. Nesta esteira foram analisados os riscos provocados pela Inteligência Artificial, revelando os seus principais vetores e conseqüentemente seu potencial catastrófico. Foi demonstrado que a opacidade gerada por esta tecnologia justifica movimentações precaucionais que visem implementar condutas para o gerenciamento dos riscos associados a ela. Deste modo, o presente estudo busca propor a aplicação do princípio da precaução como um modelo regulatório rígido, demonstrando a importância de sua inclusão nos diplomas legais que regulam o tema.

Palavras-chave: Sociedade de Risco; Direito e Tecnologia; Inteligência Artificial; Regulação de Riscos; Princípio da Precaução.

ABSTRACT

This thesis research proposes to investigate the applicability of the precautionary principle on Artificial Intelligence risk regulation. Adopting the Risk Society theory as an analysis model, the present research analysed the transformation process from modern society to Reflexive Modernization, recognizing the uncertainties brought about by technoscientific progress has altered the logic of goods distribution, which are now followed by risks distribution. The institutes brought by the modernization of society process, has proved to be insufficient to control and quantify the risks of this process. The Fourth Industrial Revolution potentializes and amplifies the reflection brought by this social theory, making the distributive logic of goods operate by digitised vehicles, consequently the risks are distributed equally. In this wake, the study focused on the risks caused by Artificial Intelligence, understanding its main vectors and, consequently, its catastrophic potential. The opacity generated by this technology justifies precautionary moves aimed at implementing procedures for managing the risks associated with it. Thus, the present study seeks to propose the application of the precautionary principle as a hard regulatory model, demonstrating the importance of its inclusion in the legal diplomas that regulate the subject.

Keywords: Risk Society; Law and Technology; Artificial Intelligence; Risk Regulation; Precautionary Principle.

SUMÁRIO

INTRODUÇÃO	17
1.0. SOCIEDADE DE RISCO: APRESENTAÇÃO DO MODELO DE ANÁLISE	23
1.1. MODERNIDADE REFLEXIVA	23
1.2. TEOREMA DA INDIVIDUALIZAÇÃO (FORÇADA).....	25
1.3. COSMOPOLITIZAÇÃO	28
1.4. SOCIEDADE DE RISCO	30
1.5. GLOBALIZAÇÃO E A SOCIEDADE DE RISCO GLOBAL	34
1.5.1. IRRESPONSABILIDADE ORGANIZADA	37
1.5.2. EFEITO BUMERANGUE	38
1.6. A QUARTA REVOLUÇÃO INDUSTRIAL E OS SEUS REFLEXOS NA SOCIEDADE DE RISCO	40
1.7. DADOS SÃO O <i>NOVO PETRÓLEO</i> OU O <i>NOVO URÂNIO</i> ? ENTENDENDO O BIG DATA E ECONOMIA DE DADOS NA QUARTA REVOLUÇÃO INDUSTRIAL COMO COMBUSTÍVEL PARA NOVOS VETORES DE RISCO.	43
1.8. A EMERGÊNCIA DE UMA SOCIEDADE DE RISCO DIGITAL.....	49
2.0. RISCOS E INCERTEZAS	52
2.1. O CONCEITO DE INTELIGÊNCIA ARTIFICIAL E SUAS CLASSIFICAÇÕES	52
2.2. PRINCIPAIS FUNCIONALIDADES DA IA NO CONTEXTO ATUAL	56
2.3. INTELIGÊNCIA ARTIFICIAL: PRINCIPAIS RISCOS	58
2.3.1. DISCRIMINAÇÃO ALGORÍTMICA E OS RISCOS MULTIFACETADOS AOS DIREITOS HUMANOS.....	59
2.3.1.1. DECISÕES AUTOMATIZADAS NO CAMPO CRIMINAL.....	59
2.3.1.2. CREDIT SCORING/SOCIAL SCORING.....	61
2.3.1.3. ACESSO À SAÚDE	63
2.3.1.4. RECURSOS HUMANOS	63
2.4. RISCOS À DEMOCRACIA.....	64
2.5. ENTRE HARARI E BECK: INTELIGÊNCIA ARTIFICIAL E O FUTURO DO TRABALHO. A RATIFICAÇÃO DOS CONCEITOS DE IRRESPONSABILIDADE ORGANIZADA E EFEITO BUMERANGUE.	67
2.6. O POTENCIAL CATASTRÓFICO DA AGI E DA ASI.....	73
2.7. O REINADO DA INCERTEZA NA SOCIEDADE DE RISCO DIGITAL: OPACIDADE ALGORÍTMICA E BLACK BOX SOCIETY	77
2.8. CONCLUSÃO	80
3.0. O PRINCÍPIO DA PRECAUÇÃO: UM INSTRUMENTO PARA A REGULAÇÃO DE RISCOS NA INTELIGÊNCIA ARTIFICIAL?	82

3.1. REGULAÇÃO DE RISCOS	82
3.1.1. PRINCÍPIO DA PRECAUÇÃO: CONCEITO, ORIGEM E DESENVOLVIMENTO	84
3.2. PRESSUPOSTOS PARA APLICAÇÃO DO PRINCÍPIO	89
3.2.1. GRAVIDADE DOS RISCOS	89
3.2.2. GRAVIDADE OBJETIVA.....	89
3.2.3. GRAVIDADE SUBJETIVA	91
3.2.4. INCERTEZAS CIENTÍFICAS	91
3.2.5. CONTEÚDO DA INCERTEZA.....	92
3.2.6. FONTES DE INCERTEZA	93
3.3. CRÍTICAS AO PRINCÍPIO	94
3.4. DEFESA AO PRINCÍPIO	97
3.5. A APLICAÇÃO DA TEORIA DE NASSIM TALEB PARA JUSTIFICAR O USO DO PRINCÍPIO DA PRECAUÇÃO NA REGULAÇÃO DE RISCOS DA INTELIGÊNCIA ARTIFICIAL	101
3.6. ANÁLISE DA PROPOSTA EUROPEIA E BRASILEIRA PARA REGULAÇÃO DA I.A.....	104
3.6.1. UNIÃO EUROPEIA - ARTIFICIAL INTELLIGENCE ACT	105
3.6.2. BRASIL (PL 21/2020) - MARCO REGULATÓRIO DA IA.....	107
CONCLUSÃO	111
REFERÊNCIAS	116
REFERÊNCIAS BIBLIOGRÁFICAS.....	116

INTRODUÇÃO

Catástrofes naturais, guerras, explosões nucleares, acontecimentos dessa magnitude inspiraram a ilustre obra do professor Ulrich Beck. Seu trabalho buscou identificar como a modernidade vem se defrontando com os perigos produzidos por ela mesma, fruto de seu desenvolvimento tecnológico cada vez mais desenfreado. Dentro desta análise, ele classifica a sociedade atual como uma *Sociedade de Risco*, decorrente de um novo estágio da sociedade moderna, chamado de *Modernidade Tardia* ou *Modernidade Reflexiva*.

Neste estágio moderno, a sociedade passa a reconhecer as incertezas trazidas pela evolução tecnocientífica. Deste modo, passa-se a analisar novas formas de se gerenciar riscos decorrentes deste avanço tecnológico. De acordo com Daniel Mendelson, o professor Beck argumenta que a percepção do risco na sociedade atual sofreu um impacto transformador. O risco nas eras anteriores era visto como decorrentes de forças da natureza, assim, as instituições eram criadas para gerir tais tipos de riscos. No entanto, essas instituições ficaram prejudicadas, uma vez que o risco passou a ser influenciado pela ação humana, o que Giddens posteriormente denominou como *Risco Manufaturado* (em contraposição ao risco natural).¹

Neste século, o avanço tecnocientífico se mostra mais perigoso, comparado ao século passado. Para além dos desastres naturais que seguem existindo, a exemplo dos desastres brasileiros em Brumadinho e Mariana, estes, plenamente perceptíveis aos olhos humanos. Passamos a enfrentar riscos abstratos em maior escala, não perceptíveis aos nossos olhos.

Além de imprevisíveis e incontroláveis, tais riscos são também transfronteiriços e transtemporais. Transfronteiriços pois ultrapassam os limites do local originalmente impactado e transtemporais visto que seus efeitos

¹ MENDELSON, Daniel. *The Routledge Companion to Social Theory: Central Terms And Thinkers*. 1. ed. London: Routledge, 2010. p. 230-232.

deletérios podem surgir posteriormente a sua materialização.²

Para Beck, são as incertezas fabricadas fruto da evolução tecnológica que colocam a sociedade, em uma Sociedade de Risco. Os riscos passam a ser incertos na medida em que os mecanismos criados no processo de modernização para quantificação e qualificação de riscos se mostram insuficientes para elaboração de cálculos probabilísticos. De forma incontrollável a sociedade de risco simboliza uma era qual os riscos econômicos, sociais e ambientais escapam da percepção social, tornando a incerteza uma regra.³

Ante a essas incertezas, o princípio da precaução ganha propósito, pois visa de forma preventiva evitar danos de difícil reparação, por vezes com potencial catastrófico. Havendo incertezas científicas sobre os riscos de determinada atividade ele deve ser invocado. De fato, o princípio reside majoritariamente no campo do direito ambiental, todavia, tendo em conta que sua existência surge com o intuito de gerir os impactos irreparáveis e incomensuráveis das novas tecnologias, parece imprescindível a sua transposição para outros campos do direito.

Neste século, com o avanço das tecnologias de Inteligência Artificial doravante IA, tivemos um breve panorama dos seus riscos. A capacidade preditiva destas máquinas, trazem um meio ambiente amplamente discriminatório em diversos campos.

Em geral essa capacidade preditiva decorre da identificação de padrões decorrente da análise de dados, utilizando uma lógica matemática (algoritmo) e aprendizado de máquina (*machine learning*).⁴

A tecnologia é vendida como neutra uma vez que sua lógica se opera de

² CAVEDON, Ricardo; FERREIRA, Helene Sivini; FREITAS, C., C. O. D. A. O meio ambiente digital sob a ótica da Teoria da Sociedade de Risco: os avanços da informática em debate: subtítulo do artigo. Revista Direito Ambiental e sociedade: subtítulo da revista, Curitiba, v. 5, n. 1, p. 200, jan./2015.

³ JUSBRASIL. Sociedade de risco: rumo a uma outra modernidade. Disponível em: <https://ferreiramacedo.jusbrasil.com.br/artigos/160037557/sociedade-de-risco-rumo-a-uma-outra-modernidade>. Acesso em: 20 set. 2020.

⁴ BIONI, Bruno Ricardo; LUCIANO, Maria; Inteligência Artificial e Direito: O PRINCÍPIO DA PRECAUÇÃO NA REGULAÇÃO DE INTELIGÊNCIA ARTIFICIAL: SERIAM AS LEIS DE PROTEÇÃO DE DADOS O SEU PORTAL DE ENTRADA? 1. ed. São Paulo: Revistas dos Tribunais, 2020. p. 3.

modo objetivo e matemático. Todavia, essa neutralidade merece ser contestada, uma vez que a escolha destes dados pode ser operada por suposições puramente discriminatórias, perpetuando vieses e acentuando a marginalização de grupos sociais. Criando um ciclo retroalimentativo discriminatório, este fenômeno demonstra um alto poder deletério aos direitos humanos, cujos impactos serão explorados no decorrer deste trabalho.

Na mesma esteira, essas tecnologias demonstram serem capazes de personalizar o acesso à informação, essa personalização traz diversos impactos no debate público e consequentemente no ambiente democrático.

Os robôs, também chamados de *bots sociais*,⁵, inundaram o debate público com notícias falsas e demonstraram capacidade de corromper o ambiente democrático. O uso indiscriminado desta tecnologia traz riscos catastróficos às liberdades civis.

Destacou-se também no presente trabalho, os riscos decorrentes de uma *Inteligência Artificial Geral*, que pode significar, em última instância, a extinção da raça humana.

Por fim, foram evocadas reflexões quanto ao futuro do trabalho frente ao crescimento desta tecnologia e consequentemente a diminuição de postos de trabalhos.

Os riscos ora elencados, se tornam incertezas uma vez que a opacidade gerada por estes algoritmos impede a participação humana na tomada de decisão.

A complexidade tecnológica da IA muitas vezes não consegue ser inteligível nem para quem o programou. Neste ponto, a falta de transparência decorrente desta tecnologia impede a quantificação e qualificação dos riscos mencionados, tornando a tecnologia em si, um vetor de incertezas científicas.

Neste cenário, o princípio da precaução, comumente utilizado em debates atinentes à proteção ambiental, aparece como uma ferramenta útil para

⁵ Bots sociais são software de Inteligência Artificial desenvolvidos para simular a ação humana de forma padronizada e repetida na internet. Esta tecnologia foi muito utilizada em campanhas eleitorais de forma maliciosa, para dar uma falsa impressão de aprovação ou reprovação de determinadas notícias, influenciando a opinião pública postando e replicando mensagens.

compor um modelo regulatório mais rígido, tendo em vista tais incertezas.⁶

Tem-se três eixos principais no desenvolvimento da dissertação de mestrado, que perpassam, essencialmente, pelos ensinamentos: de Ulrich Beck para o primeiro eixo (sociológico) de abordagem da Sociedade de Risco e Sociedade de Risco Global, complementados por vezes, por Anthony Giddens. Juntamente aos ensinamentos de Klaus Schwab, os quais ajudam a conectar a sociedade de risco com os eventos atuais. Para o segundo e terceiro eixo tem-se uma construção mais heterogênea de autores, visto que a sua matéria é extremamente nova, dispondo-se de poucas obras para consulta, ainda assim destacam-se os autores: Anthony Elliott, Eduardo Magrani, Cass Sunstein, Andy Stirling, Frank Pasquale, Alexandra Aragão e Bruno Ricardo Bioni.

Os eixos servem para confirmar que as novas tecnologias de IA, trazem enormes riscos e incertezas, sendo um novo ponto de ebulição da Sociedade de Risco (Digital). Deste modo, a incalculabilidade de seus riscos e o seu potencial danoso, sugerem a necessidade de um olhar regulatório mais rígido. Independente dos inúmeros benefícios desta tecnologia, demonstra-se de modo claro que ela é mais uma nova área de risco para a sociedade moderna. As incertezas trazidas pelo desenvolvimento da IA, fazem com que os métodos de avaliação de riscos se tornem inaplicáveis, deste modo, o princípio da precaução pode construir uma lógica protetiva mais robusta.

O primeiro eixo evoca reflexões sociológicas sobre o processo de transformação da sociedade moderna para a *Modernidade Reflexiva* ou *Modernidade Tardia*. Deste modo, a busca pelo progresso como norte do processo de modernização, passa a se radicalizar. A confiança no progresso como mecanismo de controle do desenvolvimento científico-tecnológico e seus impactos na natureza é a grande marca da primeira modernidade. De outro giro, a *Modernidade Reflexiva* é marcada pela insuficiência em controlar os riscos decorrentes de sua transformação. Para Beck, risco não significa catástrofe, risco é o perigo potencial ocasionado pela decisão humana. Riscos são

⁶ BIONI, op.cit., 2020, p. 3.

incertezas não intencionais criadas decorrentes do desenvolvimento tecnológico. Toda sociedade experimentou perigos, todavia, o regime do risco se tornou um padrão social, em escala global. As decisões que antes eram tomadas com base em cálculos, buscando ligar causa e efeito, são invalidadas pela nova dinâmica de uma Sociedade de Risco.⁷

Sabendo que a Teoria da Sociedade de Risco trouxe um maior enfoque aos riscos ambientais, tendo em vista a época de sua publicação⁸, buscou-se no final do primeiro eixo conectar a Teoria da Sociedade de Risco com o atual estágio da Sociedade industrial. A conexão entre a Quarta Revolução Industrial e a teoria de Beck, ratificaram os conceitos trazidos pela obra, habilitando compreender que a distribuição de riscos nesta revolução opera-se de modo digitalizado e em um ritmo muito mais acelerado, trazendo maior amplitude e profundidade à produção de riscos. Será demonstrado também, como os dados pessoais que são o combustível para os algoritmos de inteligência artificial não devem ser classificados de modo ingênuo como o *novo petróleo* e sim como o *novo urânio* quando analisados sob a ótica da Sociedade de Risco.

O segundo eixo, será conceitual, demonstrando o que é Inteligência Artificial, seus tipos, sua classificação legal e como ela interage com seres orgânicos. Posteriormente, serão expostos seus efeitos deletérios, elencando os principais danos coletivos causados pela IA no decorrer dos últimos anos, sua imprevisibilidade e seu potencial catastrófico. O debate frente aos riscos não foi feito de forma exaustiva, até porque não haveria tal possibilidade dentro de uma proposta de dissertação. Deste modo, debruçou-se majoritariamente nos riscos de alta relevância.

O terceiro e último grande eixo propõe uma reflexão quanto à aplicação do princípio da precaução na regulação da IA. Primariamente, o princípio é conceituado e são discutidos os pressupostos para sua aplicação, na mesma esteira, analisa-se se o princípio em si guarda relação com os impactos trazidos pela IA.

⁷ BECK, Ulrich. La sociedad del riesgo global. 1. ed. Madrid: Siglo XXI de España, 2002. p. 5.

⁸ A publicação da obra Sociedade de Risco, deu-se em 1986.

Posteriormente, abre-se um breve debate entre os principais defensores e críticos do princípio, trazendo uma contraposição de ideias que justifiquem ou excluam sua aplicação no campo da IA.

Por fim, o presente trabalho buscou responder se as propostas regulatórias da União Europeia e do Brasil, guardam algum tipo de expressão precaucional, ou se possuem ferramentas nas quais o princípio possa ser operacionalizado.

1.0. SOCIEDADE DE RISCO: APRESENTAÇÃO DO MODELO DE ANÁLISE

1.1. MODERNIDADE REFLEXIVA

Para falar sobre o princípio da precaução na regulação da Inteligência Artificial é necessário que se analise os riscos decorrentes desta tecnologia e porque eles se tornam incertezas científicas.

O modelo de análise da *Sociedade de Risco* habilita compreender a incapacidade dos institutos trazidos pela modernidade em lidar com as incertezas trazidas pelo desenvolvimento tecnológico.

Todavia, para entender os contornos da Sociedade de Risco, devemos primeiramente compreender o surgimento da modernidade reflexiva, modernidade esta, que radicaliza os conceitos da sociedade industrial, transformando-a em uma *Sociedade de Risco*.

Fernandes, ao citar Beck e Giddens, ressalta que os professores classificam o período atual como *Modernização Reflexiva*, ou como *Segunda Modernidade*. Nesta fase, a crença moderna de progresso tecnológico, junto a crença de controle da natureza e do desenvolvimento tecnocientífico se radicalizaram, deste modo a *Modernidade Reflexiva* demonstra que o progresso e a evolução já não permitem mais a previsão integral dos riscos que estas contribuem a criar, tão como torna-se impossível o controle efetivo dos mesmos.⁹

É imperioso ressaltar que as obras de Beck, focaram em riscos de origem ecológica decorrente do avanço tecnológico desenfreado. Claramente, o autor não trouxe uma discussão profunda quanto aos riscos tecnológicos advindos do desenvolvimento da Inteligência Artificial, justamente porque na época da produção de suas obras, tal tecnologia era ainda embrionária.

⁹ FERNANDES, J. P. T. Da utopia da sociedade em rede à realidade da sociedade de risco. *Análise Social*, Lisboa, v. 1, n. 2, p. 273-274, jan./2013.

Identificando o estágio atual da sociedade, a modernidade reflexiva é aquela em que a distribuição de *bens* vem acompanhada, inevitavelmente, da distribuição de *riscos*.

Ou seja, para Beck a Sociedade Industrial, que operava por uma lógica de regulação de escassez, para a produção e distribuição de *goods* (bens)¹⁰, tem sua lógica distributiva alterada progressivamente, a distribuição de *goods* passa a ser acompanhada pela distribuição de *bads* (males).¹¹

A passagem desta lógica distributiva de *riquezas* e *riscos* está condicionada a dois fatores consumativos. Primeiro, através de um nível avançado das forças produtivas humanas e tecnológicas, juntamente à redução das regras jurídicas e do estado social. A segunda, no caso a mais importante para o objeto deste estudo, são os riscos e potenciais ameaças desencadeados do crescimento produtivo desenfreado, em uma medida desconhecida.¹²

As teorias clássicas da modernidade, encabeçadas por Max Weber e seguida por Habermas não são refutadas na teoria de Beck. Na verdade, Beck apenas rechaça a ideia de estarmos em uma era pós-moderna. Para ele, continuamos em um estado moderno, todavia, hoje somos mais modernos e mais capitalistas que outrora. Sendo assim, há uma segunda modernidade em construção, classificada por ele como, Modernidade Reflexiva ou Modernidade Tardia, em suas palavras:

I reject the idea that this epochal shift is a move from the modern to the post-modern, as some argued two decades ago. For me these are all 'modern' capitalistic societies. Even more, they are modern and more capitalistic. So there is not a moving beyond the modern to its opposite. But there is a second modernity in the making.¹³

A modernidade reflexiva é marcada pela criação humana de riscos junto à evolução tecnológica. Beck caracteriza decisão humana como fato gerador

¹⁰ O termo "goods", que no inglês significa "bens", é um termo amplamente utilizado nas obras do professor Ulrich Beck para se referir na produção e/ou distribuição de riquezas.

¹¹ O conceito de distribuição e produção de "bads" trazido na obra do professor Ulrich Beck, se refere à produção de riscos e incertezas ligados à produção industrial.

¹² BECK, Ulrich. Sociedade de Risco: Rumo a outra modernidade. 2. ed. São Paulo: editora 34, 2011. p. 23.

¹³ BECK, Ulrich. WORLD AT RISK: THE NEW TASK OF CRITICAL THEORY. DEVELOPMENT AND SOCIETY, Monique, v. 37, n. 1, p. 2, jun./2008.

de riscos e incertezas, todavia, no decorrer deste trabalho, será demonstrado como essa ideia merece ser contestada, tendo em vista os avanços da IA, nas decisões automatizadas.¹⁴ Dentro do campo da IA, a maioria das decisões são tomadas exclusivamente por máquinas, deste modo, a fabricação de riscos deixa de ser uma exclusividade decorrente da decisão puramente humana.

De um modo geral, pode-se dividir a teoria da modernização reflexiva em três eixos. O primeiro eixo concerne ao teorema da *Individualização Forçada*, o segundo a *Cosmopolitização* e por fim, à teoria da *Sociedade de Risco*.¹⁵ Os três eixos tratam da radicalização de uma dinâmica da modernidade que, quando aplicada a si mesma, anula a fórmula da modernidade simples.

O principal eixo para este trabalho, será a *Sociedade de Risco*, todavia, far-se-á breves considerações quanto aos dois outros eixos.

1.2. TEOREMA DA INDIVIDUALIZAÇÃO (FORÇADA)

Beck primeiramente diferencia individualização do individualismo, este é entendido como uma atitude pessoal ou preferência, ela se refere a um fenômeno macrosociológico, que pode ou não resultar em mudanças de atitude de indivíduos. Sendo assim, ela diz respeito a um período pelo qual o sujeito desincorpora e reincorpora seu modo de viver, de modo a reescrever sua história.¹⁶ Para Beck, a individualização é um processo de ressignificação das certezas sedimentadas na sociedade industrial, o indivíduo passa a buscar novas certezas para si e para os outros. Ao mesmo passo, que essa busca de ressignificação traz novas interdependências, por vezes globais.¹⁷

Entende-se que no mundo pré-moderno as biografias já estavam determinadas nas condições que as cercavam. Com a transformação da

¹⁴ Decisão automatizada é o processo de escolha da inteligência artificial. Podendo ser utilizada para atribuir score (pontuação), média, nota, para fazer cálculo de risco ou probabilidade. A tecnologia se fundamenta no uso de algoritmos e análise de dados.

¹⁵ BECK, op. cit., 2008, p. 3.

¹⁶ COSTA, Jonas Fernandes. Modernidade Reflexiva: Giddens, Beck e as limitações da segunda modernidade. Dissertação de Mestrado, Belém, v. 1, n. 1, p. 48, mai./2009.

¹⁷ GIDDENS, Anthony; LASCH, Scott; BECK, Ulrich. Modernização reflexiva: Política, tradição e estética na ordem social moderna. 1. ed. São Paulo: UNESP, 1997. p. 26.

sociedade industrial para uma *Sociedade de Risco*, o indivíduo passa cada vez mais a ter de enfrentar, por si só, as incertezas globais.¹⁸

Beck, entende essa individualização como um resultado de uma falha dos sistemas especialistas em gerenciar risco. Ciência, política, mídia, negócios, direito e o poder militar, nenhuma destas instituições de gerenciamento de riscos são capazes de definir ou controlar os riscos de forma racional, fazendo com que o indivíduo passe a desconfiar das promessas de racionalidade dessas instituições chaves.

Beck utiliza como exemplo a responsabilidade pela decisão sobre alimentos geneticamente modificados e seus efeitos imprevisíveis. O desconhecimento das consequências ao longo prazo é despejado ao consumidor. O processo de individualização demonstra o fracasso destas instituições em gerenciar riscos, transformando o apelo por responsabilidade, no cinismo do seu próprio fracasso. Para ele, isso tudo é uma trágica ironia do processo de individualização, tendo em vista que o indivíduo, cujo a percepção do risco é limitada (comparado aos especialistas) têm para ele a responsabilidade de absorver tais ameaças, permanecendo incapaz de escapar delas. Ao mesmo passo, o indivíduo se torna obrigado a seguir as definições dos especialistas, cujo julgamento ele é obrigado a confiar. Deste modo, Beck entende que sustentar um eu individual de integridade é um assunto trágico para a Sociedade de Risco.¹⁹

Sendo assim, o processo de individualização, decorre de uma nova dinâmica de decisão do indivíduo. Nos períodos que antecederam a modernidade o futuro dos indivíduos era pré-determinado as condições que os cercavam. Por exemplo, um filho de um fazendeiro no sul da Ucrânia, no fim do século XIX, provavelmente seria como seu pai, um fazendeiro. Sua biografia já estaria pré-moldada, pelas condições que o permeiam.

Beck ilustra essa teoria exemplificando-a como Matrioskas²⁰. Portanto, na primeira modernidade os conceitos como a formação da família nuclear, já

¹⁸ COSTA, op. cit, 2009, p. 48.

¹⁹ BECK, Ulrich. Living in the world risk society. Economy and Society, London, v. 35, n.3, p.336, fev/2006.

²⁰ Matrioskas são bonecas de origem russa, elas possuem tamanhos decrescentes de modo a

determinavam previamente divisões de tarefas entre homens e mulheres, neste espectro as divisões e conceitos de classe avaliavam apenas o homem no mercado de trabalho. Todavia, o conceito exemplificado, dentre outros, foram sistematicamente modificados. Trata-se do processo de desincorporação e reincorporação decorrente do processo de individualização da *Modernidade Reflexiva*.²¹

Logo, na *Modernidade Reflexiva*, frente ao risco de um novo desastre nuclear, como o de Chernobyl, e ante as incertezas trazidas pela incapacidade em gerir riscos da Sociedade Industrial, é dado a este cidadão o dever de decidir sobre o seu futuro, sem que ele consiga analisar de forma clara, os ricos de sua ação ou inação.

A individualização por se basear na livre decisão dos indivíduos, faz com que as decisões que antes eram padronizadas, se tornem “biografias escolhidas”.²²

Este processo social, assume feições muito particulares, Beck compõe a teoria em três dimensões: (i) Perda de seguranças tradicionais em vista do conhecimento, fé e normas dirigentes, (ii) Dissolução de elos e formas sociais historicamente prescritas (iii) Novas formas de inserção social.

Para este trabalho é importante mencionar no contexto do teorema da individualização, a efetivação da despadronização do trabalho assalariado, da formação e da ocupação. O modelo de ocupação é substituído por formas plurais e flexíveis de subocupação. As três tradicionais colunas do trabalho se desconstroem, quais sejam, direito do trabalho, local de trabalho e o tempo de trabalho. Nesse processo, os riscos são privatizados, ou seja, a responsabilidade se torna individual.

se encaixarem umas nas outras. Beck utiliza esta metáfora para exemplificar como os conceitos basilares da organização da sociedade se encaixam uns nos outros.

²¹ *Ibidem*.

²² COSTA, op. cit, 2009, p.48.

Beck, também chama atenção para a criação do que hoje chamamos de renda básica universal²³, uma renda mínima juridicamente assegurada a todos.²⁴

Falar sobre a criação deste mecanismo de distribuição de renda, tão como quanto a desconstrução das colunas tradicionais do trabalho, é de extrema importância para o presente estudo, uma vez que em momento oportuno será demonstrado os riscos trazidos pela I.A para o futuro do trabalho.

1.3. COSMOPOLITIZAÇÃO

O processo de cosmopolitização surge de uma necessidade, não de uma criação política visando o bem estar mundial. Beck entende que o Estado-nação não é mais capaz de lidar com as ameaças globais, os Estados estão cada vez mais cercados de riscos dos quais eles não podem enfrentar sozinhos. Riscos ecológicos, econômicos, terrorismo, dentre outros. Na medida que tais riscos não conseguem ser enfrentados de forma singular, surge uma visão cosmopolita em que as pessoas passam a se ver simultaneamente como parte do mesmo mundo.²⁵

Para Beck o risco global tem em sua forma a imprevisibilidade e a impessoalidade, tais riscos podem desencadear eventos que requerem uma organização em escala global para enfrentá-los. Sua explicação para o surgimento deste “momento cosmopolita” decorre em seis etapas:

1. Involuntário esclarecimento: O furacão Katrina trouxe para o mundo esclarecimentos quanto aos riscos globais, demonstrando a existência de riscos com potencial catastrófico e a vulnerabilidade das organizações sociais em contê-los. Isso trouxe fortes discussões entre

²³ A renda básica universal é uma política afirmativa que busca efetuar pagamentos monetários periódicos para toda a população, sem qualquer critério de exigência preestabelecido.

²⁴ WESTPHAL, Vera Herweg. A Individualização em Ulrich Beck: análise da sociedade contemporânea. Emancipação, Ponta Grossa, v. 10, n. 2, p. 428, set./2010.

²⁵ BECK, op. cit. 2008, p.4.

os Estados-nações quanto às alterações climáticas e seu potencial catastrófico.

2. Comunicação reforçada: Nenhuma usina elétrica é construída sem protestos dos residentes próximos, da mesma forma, nenhum campo de extração de petróleo opera sem fortes críticas de "ONGs" transnacionais. Deste modo, a comunicação reforçada, traz uma involuntária democratização da discussão destes riscos a uma escala global.
3. Catarse política: Dois terremotos no final do século XX, foram capazes de unir dois países que até então, se viam como inimigos. Grécia e Turquia se uniram para enfrentar tais catástrofes. Tais eventos globais têm o potencial de unir Estados-nações criando uma catarse política, demonstra a existência deste momento cosmopolita.
4. Cosmopolitismo forçado: Riscos globais são ativados e conectam atores além das fronteiras. O cosmopolitismo na sociedade de risco, é sinônimo de descontrole dos riscos civilizatórios em uma ordem global. Ele decorre de forma involuntária, que Beck chama de cosmopolitização forçada ²⁶. Mesmo que ela não decorra de um consenso global, ainda assim ela é inevitável.
5. O risco como alerta em face à falha do governo: Para Beck, o risco global serve como um alerta ao governo no mundo globalizado. O governo era visto como um minimizador das ameaças e riscos aos indivíduos, todavia, com a emergência de riscos globais essa definição passa a ser vista como equivocada.
6. Possibilidade de governo alternativo: Beck acredita em uma alternativa política realista na era global. A condição para a existência desta política é que a globalização não deve ser "decodificada" apenas como destino econômico ela deve abrir portas para um novo tipo de poder transfronteiriço, um novo poder global. ²⁷

²⁶ Cosmopolitização forçada, é um termo trazido por Ulrich Beck para evidenciar que tal movimento não surge de um consenso global entre os "Estados-nações", ele é um fenômeno ocasionado ante a insuficiência destes estados em lidar com os problemas globais, emergentes da sociedade industrial (de risco).

²⁷ BECK, op. cit., 2006, p. 338-343.

Deste modo, Beck acredita que o processo de cosmopolitização não é um resultado de uma conspiração astuta entre os estados capitalistas, ou em suas palavras, “uma peça americana para dominar o mundo”. Na verdade, o Estado-nação se mostra cada vez menos capaz de conter os riscos de um planeta globalizado, que conectam países desenvolvidos e subdesenvolvidos. Esta incapacidade reflete na criação de uma nova realidade, surgindo uma esfera pública globalizada, onde justifica-se uma visão cosmopolita na qual o mundo passa a se ver como um só, na busca por soluções aos riscos por ele criado.²⁸

1.4. SOCIEDADE DE RISCO

O último eixo da modernização relexiva é a *Sociedade de Risco*, principal objeto deste estudo. Nota-se que os processos trazidos pela *Modernidade Reflexiva* trouxeram uma autoconfrontação para a sociedade industrial. Essa autoconfrontação não consegue ser assimilada no sistema da sociedade industrial.²⁹

Segundo Giddens é uma possibilidade de destruir de forma criativa a sociedade industrial, uma reinvenção da modernidade. Para ele, o sujeito da destruição não é a revolução nem a crise, mas sim a vitória da modernização.³⁰

A vitória que Beck e Giddens tratam, refere-se ao desenvolvimento técnico-científico e de priorização do conhecimento³¹. Por essa razão, Beck diz que a modernização reflexiva significa uma mudança não planejada da sociedade industrial, de modo autônomo e repentino. Entretanto, mantendo sua ordem político-econômica inalterada e intacta. Esse fenômeno social, radicaliza a modernidade, invadindo as premissas e contornos da sociedade industrial, criando uma nova modernidade. Neste ponto, Beck conclui que este processo

²⁸ BECK, op. cit., 2008, p. 4.

²⁹ BECK, op.cit., 2011. p. 275.

³⁰ GIDDENS; LASCH; BECK. op.cit., 1997. p. 12.

³¹ MORAES, G. B. D. A. O Princípio da Precaução no Direito Internacional do Meio Ambiente. Dissertação de Mestrado, São Paulo, v. 1, n. 1, p. 27, mai./2011.

é um triunfo do modelo capitalista, que assume novas facetas em um processo de globalização, que planifica e massifica a cultura.³²

Podemos concluir que o triunfo que Beck menciona, é o triunfo do capitalismo, que ganha novos contornos por meio da globalização. Ou seja, antes, a sociedade industrial lidava com perigos gerados por forças da natureza como terremotos, avalanches, pragas, entre outras. Agora, os riscos e incertezas são fundados no desenvolvimento tecnológico, fruto do conhecimento científico.³³

Logo, a inovação científica ganha três facetas, ela se torna um instrumento de criação de riscos ao mesmo passo que se torna um meio para o seu descobrimento, e por fim uma fonte de resolução destes, criando pressupostos para o domínio de ameaças autogeradas.³⁴ O estado com cunho econômico liberal conjunto a essas facetas acentuam ainda mais a produção de riscos, uma vez que a degradação ambiental é colocada em segundo plano, tendo em vista as oportunidades de mercado geradas pelo avanço desenfreado aos recursos naturais, para produção de bens. Essa vulnerabilidade e incontrolabilidade são marcas de uma Sociedade de Risco.³⁵

Deste modo, a evolução tecnocientífica é o principal vetor de riscos na Sociedade de Risco. O processo de modernização se tornou reflexivo, convertendo-se em tema e problema. No processo de modernização, a produção de bens seguia a lógica de regular a escassez, impulsionadas pelo desenvolvimento tecnológico, se sobrepondo ao manejo político e científico.³⁶ Logo, a preocupação em produzir bens, se sobrepôs às preocupações quanto aos riscos que o trazem. Nesta esteira, Beck enfatiza que as instituições criadas na Sociedade Industrial, não foram pensadas para tratar a produção e distribuição dos *riscos*.

³² GIDDENS; LASCH; BECK. op.cit., 1997. p. 13.

³³ BECK, op. cit., 2011. p. 275.

³⁴ SILVA, Brisa Arnold da. UMA ANÁLISE SOBRE A MODERNIDADE REFLEXIVA E A COMPLEXIDADE AMBIENTAL NO ESTADO SOCIOAMBIENTAL DE DIREITO PARA O COMPROMISSO DO DESENVOLVIMENTO SUSTENTÁVEL. Cadernos do Programa de Pós Graduação em Direito, Porto Alegre, v. 10, n. 1, p. 104, jan./2015.

³⁵ *Ibidem*.

³⁶ BECK, op.cit., 2011. p. 06.

A tese principal de Beck é que o surgimento dos riscos no processo de modernização se dá pela ausência de pesquisas que buscassem ligar as causas e efeitos da produção industrial. O surgimento destes riscos começou a sabotar os mecanismos de controle instituídos pelos Estados-nação. No início dos anos setenta os estados passaram a perceber que o intercâmbio de *goods* (bens) estava diretamente ligado à distribuição de *bads* (riscos). Legitimando assim uma crise de instituições e sua funcionalidade.³⁷

Para Beck, o avanço da segurança deve acompanhar a produção de riscos, diante de uma sociedade consciente e reforçada de intervenções ao desenvolvimento técnico-científico.³⁸

Fernandes Costa pontua que na Sociedade de Risco, exige-se reflexão acerca das bases da coesão social e se faz necessário rever os fundamentos que justificam a racionalidade de uma decisão, tendo em vista que o mundo atual é marcado pela incerteza. A sociedade deve se preocupar em garantir que as situações da vida humana possam ser controladas pela racionalidade instrumental, pois foi assim que a sociedade industrial se moldou na primeira modernidade. Entretanto, a complexidade dos riscos trazidos pela Sociedade de Risco se sobrepõe, fazendo com que estas exigências por controle se tornem inexecutáveis.³⁹

Guivant analisa o olhar de Beck como um olhar não pessimista. Mesmo que a análise do Sociólogo dê ênfase nos riscos de alto grau decorrentes desta transformação da sociedade industrial. Na verdade, estes novos riscos permitem novas movimentações políticas, é o que Beck classifica como subpolítica.⁴⁰ A subpolítica é vista como um novo campo de debate, fora das instituições trazidas pela modernidade, representadas por Estado-nação, como por exemplo os partidos políticos, trata-se na verdade de um ambiente transnacional representado por instituições internacionais e atores locais, vinculadas entre si. Neste campo, a sociedade se viria representada por novos

³⁷ UOL. Tecnologia é a matriz do risco, diz sociólogo. Disponível em: <https://www1.folha.uol.com.br/fsp/ciencia/fe2011200101.htm>. Acesso em: 12 out. 2021.

³⁸ BECK, op.cit. 2011. p. 24.

³⁹ COSTA, op.cit., 2009, p. 47.

⁴⁰ GUIVANT, Julia Silvia. O LEGADO DE ULRICH BECK. Ambiente & Sociedade, São Paulo, v. 9, n. 1, p. 230, jan./2016.

atores políticos, com uma visão cosmopolita que pudesse desafiar as estruturas criadas pelo processo de Modernização.⁴¹

Para Beck o significado de Sociedade de Risco, são as *incertezas fabricadas*, ou seja, um leque de novos riscos dos quais inexistiam na sociedade industrial. Estas incertezas manufaturadas resultam do desenvolvimento técnico-científico que busca o progresso. Na modernidade reflexiva, os novos riscos são, portanto, incertezas e perigos fabricados que aumentam progressivamente conforme o avanço da ciência e tecnologia. Diferentemente da modernidade simples, onde tais efeitos eram facilmente percebidos, neste novo estágio tendem a ser intangíveis para nossos sentidos.⁴²

Tais riscos não são totalmente determináveis cientificamente, embora a ciência consiga conhecê-los de algum modo. Isso significa que os métodos atuais para avaliação, gestão e segurança dos riscos não são totalmente funcionais.⁴³

Deste modo, Beck conclui que a modernidade passou a se deparar com o seu próprio potencial destrutivo de desenvolvimento social e tecnológico, sem ter respostas adequadas para estas incertezas por ela fabricada.⁴⁴

O panorama trazido pela Sociedade de Risco, coloca importantes questionamentos para o debate de regulação de riscos das novas tecnologias, tal como, a Inteligência Artificial. Deste modo, a grande contribuição de Beck para este trabalho, está no fato de que sua teoria social permite criar justificativas para criação de uma visão regulatória precaucionária, em um mundo incerto. Beck, roga pela necessidade da criação de uma cultura da incerteza, contrapondo-se a uma visão de riscos e benefícios. Isso não quer dizer que seu pedido se assenta em uma *cultura do não-risco*, pois essa visão em última instância, barraria a inovação. Tendo em vista a impossibilidade de não produzir riscos no estágio atual da sociedade industrial. Na verdade, uma cultura de incertezas é uma posição de humildade frente à capacidade do

⁴¹ *Ibidem*.

⁴² BECK, op.cit., 2008, p.5.

⁴³ *Ibidem*.

⁴⁴ *Ibidem*.

desenvolvimento tecnológico, trazendo uma visão mais robusta para a gestão de riscos. Trata-se de uma visão precaucionária.

1.5. GLOBALIZAÇÃO E A SOCIEDADE DE RISCO GLOBAL

O processo de globalização é sem dúvidas o principal propulsor da *Sociedade de Risco Global*. Beck não considera ser o fim dos Estados-nações ou da política, na verdade ele acredita em uma política transnacional para confrontar os riscos globalizados.⁴⁵

A teoria da *Sociedade de Risco Global* sustenta que as sociedades modernas são moldadas por novos tipos de riscos, sendo suas bases abaladas pela antecipação de catástrofes globais.⁴⁶ Tendo a percepção de risco três recursos:

- a) *Deslocalização*: os riscos não mais se limitam a sua localização, eles se tornam onipresentes, na medida que transcendem fronteiras territoriais.
- b) *Incalculabilidade*: seus impactos tendem a ser incalculáveis, tendo em vista que os métodos de cálculos probabilísticos se tornam inócuos frente a sua complexidade.
- c) *Incompensabilidade*: O potencial catastrófico se faz presente, por vezes de modo irreversível. Desta forma, riscos referentes a questões climáticas, evolução em setores de genética humana, a possibilidade de grupos terroristas possuírem armas de destruição em massa, evolução no campo da nanotecnologia de modo irresponsável, tendem a causar danos de difícil reparação. Deste modo, a lógica de reparação de danos é desconstruída para o surgimento de uma lógica precaucionária. Tentando antecipar e

⁴⁵ FERREIRA, Helini Silvini. A BIOSSEGURANÇA DOS ORGANISMOS TRANSGÊNICOS NO DIREITO BRASILEIRO: UMA ANÁLISE FUNDAMENTADA NA TEORIA DA SOCIEDADE DE RISCO. Tese de Doutorado, Florianópolis, v. 1, n. 1, p. 45, abr./2008.

⁴⁶ BECK, op.cit., 2006, p. 5-6.

prevenir riscos do qual a existência ainda é incerta.⁴⁷

Beck também elucida que a deslocalização dos riscos incalculáveis ocorre nos níveis: espacial, temporal e social.

Espacial: visto os novos riscos serem transfronteiriços, como por exemplo as mudanças climáticas, ou o aumento de plástico nos oceanos.

Temporal: tais riscos têm um longo período de latência, ou seja, seus efeitos no decorrer do tempo, não são mensuráveis em um curto de tempo, de forma confiável.

Social: Tendo em vista a complexidade destes riscos, e os inúmeros atores envolvidos nos processos que os desencadeiam, não é mais possível atribuir causas e consequências de forma confiável.

Conforme Silvini, para ter eficácia no enfrentamento destes riscos, é necessária uma visão transnacional, tratá-los de forma unilateral além de ineficaz é contraproducente. Deve-se, portanto, fortalecer estruturas normativas internacionais entre os estados.⁴⁸

Uma vez que os conflitos atuais já não podem ser absorvidos de forma unilateral, por apenas um Estado-nação. Justamente, porque a gravidade de seus riscos junto a sua incompreensão, invalidam os sistemas de gestão de riscos concebidos pela modernidade. Tornando o conceito de Estado-nação insuficiente e restrito ante aos conflitos globais.⁴⁹

Nesta linha, Giddens exemplifica que vivemos em um tempo que todas as decisões tomadas por um indivíduo, afetam o sistema global. O indivíduo ao comprar uma roupa produzida em países menos industrializado, pode alimentar a prática de trabalhos análogos a escravidão. Ao escolher consumir

⁴⁷ *Ibidem*.

⁴⁸ FERREIRA, op.cit.,2008, p. 48.

⁴⁹ *Ibidem*, p. 49.

determinado alimento, pode contribuir para a deterioração ecológica, trazendo consequências negativas para toda a humanidade. As decisões cotidianas passam a contribuir para resultados globais, ao mesmo passo que as ordens globais influenciam o modo de vida individual.⁵⁰

Isso demonstra, que os riscos decorrentes de atitudes singulares, não estão restritos apenas aos Estados-nações. Eles se distribuem em uma escala global.

Beck, argumenta que o modelo da sociedade industrial não foi pensado para tratar da produção e distribuição de *bads*, ou seja, dos riscos decorrentes da produção industrial. Para ele, essas consequências dos processos de desenvolvimento do estado moderno não eram estudadas, logo começam a se sabotar dentro do processo de globalização. Fazendo com que, em dado momento, mudassem a percepção do Estado-moderno-nação⁵¹. Logo, ao entender que a produção *bads* se torna global, não faz sentido analisar para tais problemas por uma perspectiva nacionalista⁵².

Beck utiliza o exemplo da crise da vaca louca, para demonstrar o quão ineficiente é a unilateralidade frente a riscos transfronteiriços. Um Estado-nação sozinho, não consegue controlar, prever e se organizar para conter riscos que se dissipam em redes mundiais. Além da insuficiência, os atores que participam dos processos decisórios de um Estado-nação, não se responsabilizam pela produção e distribuição de riscos. Os atores políticos argumentam que só se responsabilizam pela regulação do desenvolvimento da tecnologia. Já os cientistas, desenvolvem a tecnologia e alegam não serem responsáveis pelo modo como serão utilizadas. Por fim, os empresários justificam que apenas estão atendendo a demanda de consumo.

Tal questão, além de ilustrar a insuficiência do Estado-nação em lidar com a produção e distribuição de riscos, fica evidente o desprezo destes atores

⁵⁰ GIDDENS; LASCH; BECK, op.cit., 1997. p. 75.

⁵¹ Estado-nação é a principal forma de organização dos estados modernos. O estado cria mecanismos de controle políticos para determinado conjunto de pessoas que consideram uma nação. O conceito pode ser visto de forma análoga ao conceito de país, apesar de não estar restrito a critérios puramente territoriais.

⁵² UOL. Tecnologia é a matriz do risco, diz sociólogo. Disponível em: <https://www1.folha.uol.com.br/fsp/ciencia/fe2011200101.htm>. Acesso em: 12 out. 2021.

em assumir qualquer tipo responsabilidade. Essa falta de responsabilização que Beck classifica como *Irresponsabilidade Organizada*.⁵³

1.5.1. IRRESPONSABILIDADE ORGANIZADA

Quanto ao conceito de irresponsabilidade organizada, Beck entende que a ciência podendo ser um meio de produção e de resolução de riscos é vista primariamente como uma fonte de aumento de produtividade, envolvida assim no desenvolvimento de tecnologias de alto risco. Sendo assim, os perigos que derivam deste desenvolvimento são descartados ou considerados de maneira tardia. Esta ausência de precaução impede que determinemos relações de causa e efeito. Como por exemplo a sujeira ambiental, ela existe, mas você não consegue definir quais processos tecnocientíficos desencadearam-na.

Esta movimentação global que impossibilita a atribuição de responsabilidades de um modo a quem desenvolve tecnologias nocivas Beck classifica como *irresponsabilidade organizada*.⁵⁴

Silvini, pontua que para Beck, os riscos decorrentes da primeira modernidade eram possíveis de serem quantificados por cálculos probabilísticos.⁵⁵

Todavia, o que torna a distribuição de *bad*s tão determinante, é a inevitabilidade de escapar de suas consequências. Beck classifica estes sistemas matemáticos como *Sistemas Fechados de Explicação*. Estes sistemas, seja ele fornecido pela ciência tecnológica, social, jurídica ou política, não são mais opções viáveis para quantificar ou qualificar tais riscos. Portanto, o avanço científico de modo paradoxal, mina o papel dos especialistas em prever riscos.⁵⁶

⁵³ *Ibidem*.

⁵⁴ *Ibidem*.

⁵⁵ FERREIRA, op. Cit., 2008, p. 54-55.

⁵⁶ UOL. Tecnologia é a matriz do risco, diz sociólogo. Disponível em: <https://www1.folha.uol.com.br/fsp/ciencia/fe2011200101.htm>. Acesso em: 12 out. 2021.

Beck argumenta que os riscos globais, escancaram a insuficiência dos Estados-nação para solucionar este problema de forma singular, sobretudo quando ninguém se responsabiliza pelos resultados.⁵⁷

Deste modo, as relações simplistas explanadas anteriormente, que eram suficientes para calcular e quantificar os riscos na primeira modernidade, já não constituem instrumentos suficientes para os riscos da Modernidade Reflexiva. Desta forma, o desenvolvimento tecnológico se dá em um campo de incertezas não quantificáveis.

A aceitação deste campo incerto se justifica por uma racionalidade capitalista que busca lucro e progresso, se sobrepondo aos conflitos que surgem em torno destes riscos.⁵⁸

Nesta lógica, os riscos passam a ser produzidos de modo dissimulado e desvirtuado. Produzidos deste modo, tendo em vista que as instituições que legitimam a produção destes riscos dependem de apoio popular, desta forma, assumir a responsabilidade pela produção de riscos e evidenciar sua incapacidade de contê-los, seria impraticável pois acarretaria em uma crise de legitimação.⁵⁹

Desta forma, a irresponsabilidade organizada é definida como um meio pelo qual os atores políticos e econômicos encobrem a produção e distribuição de riscos, disseminando-as de modo silencioso.⁶⁰

1.5.2. EFEITO BUMERANGUE

Beck também nos alerta quanto a um modo de distribuição de riscos que afetam todas as classes sociais, inevitavelmente. Beck classifica este conceito como *efeito bumerangue*.

⁵⁷ *Ibidem*.

⁵⁸ BECK, op.cit., 2011. p. 23.

⁵⁹ FERREIRA, op. cit., 2008. p. 55.

⁶⁰ GOLDBLATT, David. Teoria social e ambiente. 1. ed. Lisboa: Instituto Piaget, 1996. p. 241.

Para o professor, os ricos e poderosos, no caso os detentores dos meios de produção, que lucram sobre a distribuição de bens e concorrem na produção de *bads*, junto aos demais atores da modernização, acabam inevitavelmente, sendo afetados. Em um primeiro momento, tais agentes distribuem *bads* que afetam diretamente a população mais pobre. Todavia, mais tarde, estes mesmos *bads* serão experimentados pelas classes mais favorecidas.

Beck ressalta que a invisibilidade destes efeitos imprevisíveis passou a ser tornar visível, ameaçando aqueles que majoritariamente produzem e distribuem riscos. O risco se distribui e volta ao seu criador, seguindo uma curva de *bumerangue*. Este efeito socialmente circular, ameaça tanto o culpado, como a vítima, podendo até mesmo aniquilá-los.

Deste modo, em uma visão cosmopolita do efeito bumerangue na Sociedade de Risco, nota-se que por mais que haja assimetrias na distribuição de riscos entre os países desenvolvidos e subdesenvolvidos, a tendência é que todos experimentem tais males.⁶¹

Beck conclui que tais situações de riscos não são situações de classe. A dinâmica de distribuição de riscos da modernização decorre de processos que se desacoplam do mero conceito classes. Para Beck, sofrer tais impactos e não sofrer não se polarizam como ter propriedade privada e não ter.

Na dinâmica da propriedade privada por exemplo, aqueles que possuem privam este direito daqueles que não tem, confrontando em classe de *afetados* e *não afetado*⁶². Todavia, essa dinâmica se distingue das situações de ameaças, neste caso a classificação analógica seria, os *afetados* e os *ainda-não-afetados*.⁶³ Sendo assim, a imunidade de determinada classe social ou de determinado Estado-nação em relação à absorção dos riscos, converte-se do dia para noite em uma exposição irreversível.⁶⁴

⁶¹ BECK, op.cit., 2011. p. 44.

⁶² Beck ilustra a dinâmica de classes utilizando como exemplo os possuidores e não possuidores de propriedades, a propriedade de uns significa a privação de propriedade aos outros.

⁶³ *Ibidem*, p. 47.

⁶⁴ *Ibidem*, p. 48.

1.6. A QUARTA REVOLUÇÃO INDUSTRIAL E OS SEUS REFLEXOS NA SOCIEDADE DE RISCO

Nota-se que o professor Ulrich Beck, em suas obras *Sociedade de Risco* e *Sociedade de Risco Global*, coloca grande enfoque nos riscos ambientais e riscos econômicos decorrentes do processo de modernização. Deste modo, suas obras nos alertam majoritariamente aos riscos quanto a catástrofes, ecológicas, nucleares, crises econômicas, precarização do trabalho e etc.

No entanto, antes de demonstrar que o desenvolvimento da IA, também fruto do progresso tecnológico trouxe um novo ponto de ebulição para a *Sociedade de Risco*, criando novos vetores de riscos econômicos, sociais e ambientais. Deve-se primeiramente situar o momento revolucionário, qual estamos.

Este momento é chamado de Quarta Revolução Industrial, faz-se importante entender sua existência, uma vez que ela traz novos contornos para a sociedade industrial analisada por Beck.

Schwab, fundador e presidente do Fórum Econômico Mundial defende que estamos em um novo processo revolucionário, uma nova revolução industrial, iniciada na virada do século vinte. Essa nova revolução se caracteriza por uma internet cada vez mais onipresente no cotidiano da sociedade e cada vez mais móvel, além do desenvolvimento de novas tecnologias como a Inteligência Artificial, objeto de análise deste trabalho.

Todavia, não se trata apenas do avanço tecnológico com máquinas mais inteligentes e conectadas, até porque, houveram avanços nesse sentido na revolução anterior. Trata-se de um escopo mais amplo, com descobertas em diversos campos tecnológicos ocorrendo simultaneamente e de modo interligado.⁶⁵

Ele cita exemplos de áreas que se completam, como genética e nanotecnologia ou energias renováveis e computação quântica. Esses campos, ao se desenvolverem simultaneamente, tendem a proporcionar velocidade e

⁶⁵ SCHWAB, Klaus. A Quarta Revolução Industrial: Tradução Daniel Moreira Miranda. 1. ed. 2016: Edipro, 2016. p. 16.

amplitude uns aos outros. Deste modo, a grande marca dessa dinâmica evolutiva é justamente a interação entre os domínios físicos, biológicos e digitais. As tecnologias e inovações que emergem deste processo, se distribuem de modo muito mais rápido e amplo comparado às revoluções anteriores.⁶⁶

Nesta linha, Schwab entende que diversas inovações tecnológicas que permeavam apenas no campo da ficção científica, passaram a entrar no campo da realidade. Tais inovações são interdependentes do desenvolvimento de tecnologias que agora passam a se conectar. Este processo de *Harmonização Tecnológica* traz realidade à ficção.⁶⁷

Ele exemplifica esta harmonização, ao citar a interação entre engenharia de materiais, biologia sintética e computação. Essas tecnologias em conjunto são capazes de criar sistemas que interagem com microrganismos humanos. Os campos que eram inteiramente orgânicos, passam a ser habitados por objetos inorgânicos. Assim, a mutação e adaptação, marcas do reino animal e vegetal, passam a ser influenciadas por ‘seres’ inorgânicos.⁶⁸

Um bom exemplo que ilustra esta questão é o surgimento de humanos *Cyborgs*⁶⁹, seres humanos dotados de partes orgânicas e inorgânicas. A prova viva deste novo humano é o britânico Neil Harbisson, que nasceu com uma rara doença visual que o impede de ver cores. Para sanar este problema, Neil se tornou um *Cyborg* ao instalar uma antena em seu crânio possibilitando que ele ouça sons diferentes para cada cor. Esse sensor de cores instalado em Neil, é fruto da interação de tecnologias computacionais, avanços na medicina e bioengenharia.⁷⁰

Essa simbiose tecnológica trouxe interações entre seres inorgânicos (robôs) e seres orgânicos, transformando nosso meio ambiente como um todo.

⁶⁶ *Ibidem*.

⁶⁷ *Ibidem*.

⁶⁸ *Ibidem*.

⁶⁹ Humanos Cyborgs, são aqueles que utilizam tecnologias robóticas em seus corpos para melhorar condições limitadoras, como o uso de próteses mecânicas.

⁷⁰ NATIONALGEOGRAPHIC. How a Color-Blind Artist Became the World's First Cyborg. Disponível em: <https://www.nationalgeographic.com/science/article/worlds-first-cyborg-human-evolution-science>. Acesso em: 15 mar. 2022.

Essa indissociabilidade entre o humano, biológico e o tecnológico é a marca desta nova revolução.

Outro importante aspecto desta revolução é o surgimento da lógica das redes⁷¹, transformando um grande número de relações, inclusive as econômicas, em globais. Acrescentando complexidade para os mercados locais e para a organização econômica, cria-se também crises relevantes em termos de soberania dos Estados-nações, adstritas às fronteiras físicas.⁷²

Verifica-se ainda, que a capacidade de armazenamento e processamento de informações evolui em uma velocidade exponencial, deste modo o volume de dados criados diariamente também cresce, por conta da facilidade proporcionada pelas mídias sociais e da ampliação de acesso a aparelhos conectados, como smartphones. Deste modo, de acordo com Gamba, o ponto central desta dinâmica revolucionária está na capacidade de gerenciamento e processamento de informação, sendo assim, destacam-se as empresas que trabalham com captação e tratamento de dados, utilizando uma lógica indutiva na busca de informações.⁷³

Portanto, a sociedade passa a ser movida por dados e conseqüentemente as tecnologias decorrentes deste processo utilizam os dados como combustível.

A despeito da IA, nota-se que seus avanços se deram justamente por conta da coleta e uso massivo de dados. Possibilitando a criação de softwares que auxiliam a tomada de decisão humana, como a criação de algoritmos que podem prever e subverter a conduta humana. Deste modo, o desenvolvimento da IA tem como combustível, os dados (sejam eles pessoais ou não).⁷⁴ De acordo com Schwab, esses algoritmos de IA aprendem com uma quantidade baixa de dados. Criando novos tipos de aprendizagem automatizada, de modo

⁷¹(CASTELLS,Manuel. A sociedade em rede. A era da informação:economia,sociedade e cultura,v. 9.Ed.Trad.Roneide Venâncio Majer. São Paulo: Paz e Terra,1999,o.566) APUD, GAMBÁ, J., J. R. G. Democracia e Tecnologia: Impactos da Quarta Revolução Industrial. 1. ed. Rio de Janeiro: Lumen Juris, 2020. p. 78.

⁷² GAMBÁ, J. R. G. Democracia e Tecnologia: Impactos da Quarta Revolução Industrial. 1. ed. Rio de Janeiro: Lumen Juris, 2020. p. 78.

⁷³ *Ibidem*. p. 78-79.

⁷⁴ SCHWAB, op.cit., 2016. p. 23.

a conseguir se auto programar e encontrar melhores soluções partindo de princípios iniciais.⁷⁵

Nota-se portanto, que o desenvolvimento tecnológico desta nova revolução, o avanço em diversos campos da ciência e o surgimento de máquinas mais inteligentes e mais conectadas, só foi possível graças a coleta e tratamento massivos de dados. Sendo assim, passa-se a compreender esta nova economia e este novo insumo.

1.7. DADOS SÃO O NOVO PETRÓLEO OU O NOVO URÂNIO? ENTENDENDO O BIG DATA E ECONOMIA DE DADOS NA QUARTA REVOLUÇÃO INDUSTRIAL COMO COMBUSTÍVEL PARA NOVOS VETORES DE RISCO.

A produção de *bens* dentro da Sociedade 4.0, tem como matéria prima, os dados. Eles são o insumo das tecnologias de *Big Data*⁷⁶, deste modo, sem dados não existe IA. Uma vez que, os algoritmos de IA só operam por conta do enorme volume de dados processados. Sem dados, não conseguiríamos criar máquinas como o *GPT-3*⁷⁷ e o robô *Watson*⁷⁸. Deste modo, a centralidade e a relevância dos dados ganham enorme importância⁷⁹.

Big Data representa o êxtase do processo de captação e processamento de dados⁸⁰, refere-se à prática de utilizar grande volume de informações de diversas fontes, analisá-las, muitas vezes utilizando algoritmos de inteligência

⁷⁵ *Ibidem*.

⁷⁶ Big Data é o termo utilizado para denominar um grande conjunto de dados que precisam ser processados e armazenados.

⁷⁷ Trata-se de uma tecnologia de IA focada em produzir textos, se baseando no conceito de redes neurais e alimentada por algoritmos de aprendizado de máquina. Deste modo, a tecnologia consegue reproduzir textos de forma preditiva como se fosse um ser humano, processando a linguagem de modo natural.

⁷⁸ O IBM Watson é uma das IA mais famosas do mundo. Um dos seus modelos obteve resultados espantosos no campo da medicina. Esta IA atingiu um nível de acurácia no diagnóstico de câncer próximo ao de um médico. Os resultados apontaram que em 96,4% das vezes o diagnóstico foi idêntico ao de equipes médicas.

⁷⁹ BIGDATACORP. Entendendo a economia dos dados. Disponível em: <https://bigdatacorp.com.br/entendendo-a-economia-dos-dados/>. Acesso em: 1 nov. 2021.

⁸⁰ BIONI, Bruno Ricardo. PROTEÇÃO DE DADOS PESSOAIS: a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2018. p. 58.

artificial (machine learning)⁸¹ para tomar decisões. Essas informações nem sempre são pessoais, pode-se monitorar fenômenos naturais, atmosféricos e até mesmo, processos técnicos na cadeia de produção. Todavia, um dos maiores valores do *Big Data* para as empresas e governos, deriva na capacidade de monitoramento do comportamento humano, coletiva e individualmente, tal capacidade reside em seu potencial preditivo.⁸²

O *Big Data* pode ser visto como um modelo de receita para empresas que dependem da coleta e tratamento de dados para otimizar o valor econômico das transações para aqueles que provêm serviços. Não se trata apenas de rastrear a atividade da internet para envio de propagandas direcionadas, utiliza-se desse rastreo também para alterar taxas e condições de apólices de seguro, equalizar taxa de juros em empréstimos, dentre outras relações comerciais. Os usuários, em sua maioria, desconhecem a extensão desse tipo de rastreamento. Este tipo de dado, tende a inferir a identidade de uma pessoa, mesmo quando anonimizados⁸³, combinando dados supostamente anonimizados⁸⁴ com outros conjuntos de dados, como informações públicas expostas em mídias sociais.⁸⁵

Bioni ressalta que o *Big Data*, é uma base de dados e não um sistema inteligente, o *Big Data* é uma tecnologia que coleta e processa dados de modo mais eficiente. Todavia, esse tipo de coleta pode fazer inferências, ele cita como exemplo o caso onde o *Big Data* inferiu na probabilidade de gravidez de uma consumidora.⁸⁶

⁸¹ Machine Learning ou Aprendizado de Máquina é uma tecnologia muito utilizada no campo da IA. Ela capacita que a IA aprenda de acordo com as respostas esperadas, por meio de associações de diferentes tipos de dados, podendo ser imagens, números ou qualquer outro tipo de dado que essa tecnologia possa identificar. Deste modo, ela cria os seus próprios padrões de aprendizado.

⁸² (EDPS), T. E. D. P. S. Towards a new digital ethics: Data, dignity and technology. European Data Protection Supervisor, Lisboa, v. 1, n. 1, p. 6, set./2015.

⁸³ A anonimização é uma técnica de remoção ou modificação de informações impossibilitando a identificação de uma pessoa. Essa técnica faz com que um dado pessoal se torne um dado anonimizado.

⁸⁴ Supostamente anonimizados uma vez que o cruzamento com dados públicos por vezes possibilita sua “desanonimização”. Ou seja, permite reidentificar o titular dos dados.

⁸⁵ EDPS, op.cit 2015, p.6.

⁸⁶ BIONI, op.cit., 2018. p. 59-61.

Trata-se de um caso emblemático, ocorrido em uma loja de departamentos chamada Target. No caso, a eficiência da tecnologia foi tamanha, que se baseando no histórico de compras de determinado cliente, a tecnologia passou a sugerir o envio de propaganda de produtos para gestantes⁸⁷. Por conta deste tipo de propaganda direcionada, um pai aborrecido, entrou na loja citada, acusando que as propagandas enviadas pelo estabelecimento estavam incentivando sua filha a engravidar. Posteriormente, visando manter um bom relacionamento com o cliente, o gerente da loja ligou para este pai, para se desculpar. Entretanto, constrangido, o pai informou ao gerente que sua filha estava grávida, pedindo desculpas pelo ocorrido e consolidando a previsão do Big Data.⁸⁸

A capacidade do Big Data em prever como o ser humano se comporta é tanta, que o uso desta tecnologia pelas principais redes sociais tornou possível prever o fim de um relacionamento, com base nas postagens e curtidas de seus usuários. A tecnologia também pode ser utilizada para prever crises no setor financeiro. Tudo graças a sua capacidade de coleta e processamento de dados.⁸⁹

O Fórum Econômico Mundial estima que a economia baseada em captação de dados ultrapassa três trilhões de dólares, mais que o dobro do PIB⁹⁰ brasileiro. Um modelo de negócios de grande destaque hoje, são os negócios baseados em “economia compartilhada”.⁹¹

Analisando as grandes aquisições feitas por *Big Techs* ⁹², todas foram na verdade, compra de base de dados. Citando alguns exemplos: (i) a Google adquiriu o Waze, pensando no acesso às suas bases de dados atualizadas de ruas, estradas e mapas em geral. (ii) A Microsoft adquiriu o LinkedIn com a

⁸⁷ A Target atribui a cada cliente um número de ID de convidado, vinculado ao seu cartão de crédito, nome ou endereço de e-mail, que se torna um bucket que armazena um histórico de tudo que eles compraram e qualquer informação demográfica que a Target coletou deles ou comprou de outras fontes.

⁸⁸ *Ibidem*.

⁸⁹ *Ibidem*.

⁹⁰ PIB é a sigla que se denomina Produto Interno Bruto. Trata-se da soma de todos os bens e serviços de um país, estado ou cidade.

⁹¹ EDPS, op.cit 2015, p.6.

⁹² As Big Techs são as grandes empresas do ramo tecnológico que dominaram o mercado nos últimos anos. Amazon, Meta e Google são algumas delas.

intenção de acessar dados de perfis profissionais, para oferta de cursos personalizados.⁹³

Sendo assim o valor destas empresas está diretamente ligado a quantidade de dados que elas possuem, sendo informações de usuários (ex: LinkedIn, Facebook) ou dados únicos gerados a partir da sua operação (ex: Waze). Esses dados valem mais do que qualquer tecnologia ou ativo da empresa, na verdade estes dados são o seu principal ativo.⁹⁴

Deste modo, conforme preceitua Gamba, o conceito de Sociedade da Informação é abandonado. Este se destacava pela importância socioeconômica da informação e sua facilidade de disseminação. Agora é introduzido um novo conceito chamado de *Sociedade de Dados (Data Society)*, tendo em vista a relevância socioeconômica da captação, geração, disseminação e tratamento de dados.⁹⁵

Desde 1980, com a evolução progressiva da tecnologia computacional, quantidades cada vez maiores de dados pessoais vêm sendo captados de modo que, naquela altura, já havia uma percepção de que tais práticas reduziam os indivíduos a meros titulares de dados, colocando em risco seus direitos e liberdades fundamentais.⁹⁶

Deste modo, a evolução na capacidade em coletar e processar dados, traz insumos para o desenvolvimento de tecnologias que podem prever nossos comportamentos, vontades e se tornar, em última instância, os detentores de nossas escolhas.

O professor Yuval Noah Harari, justifica tais premissas exemplificando que dispositivos como o Kindle, da Amazon, possuem a capacidade de coletar dados dos seus usuários enquanto eles estão lendo. Podendo monitorar qual parte do livro é lido mais depressa ou mais devagar. A harmonização tecnológica trazida pela Quarta Revolução (tratada anteriormente), possibilitaria que fossem adicionados sensores de reconhecimento facial e

⁹³ BIG DATA CORP. Entendendo a Economia dos Dados. Disponível em: <https://bigdatacorp.com.br/entendendo-a-economia-dos-dados/>. Acesso em: 9 nov. 2021.

⁹⁴ *Ibidem*.

⁹⁵ GAMBA, op.cit, 2020. p. 79.

⁹⁶ EDPS, op.cit, p.6.

sensores biométricos no aparelho, fazendo com que o Kindle pudesse monitorar o batimento cardíaco e pressão sanguínea do usuário. Deste modo, o aparelho poderia entender de forma muito mais complexa quais as emoções que aquela leitura traria ao indivíduo. Basicamente, o Kindle estaria lendo o indivíduo enquanto o indivíduo o lê. Todavia, de forma muito mais precisa e profunda. Essa coleta de dados, possibilitaria que a Amazon soubesse exatamente qual o tipo de literatura o indivíduo tem preferência, entretanto, muito mais do que isso, a Amazon saberá o mais profundo “eu” do indivíduo, podendo não só inferir, como também subverter suas escolhas.⁹⁷

A abordagem dos riscos da IA será feita em momento oportuno no decorrer deste trabalho. Todavia, é imperioso concluir que a capacidade preditiva decorrente da coleta e processamento de dados fornecida pelo Big Data, o termo *Data is the new oil*⁹⁸ (Dados são o novo petróleo) não parece ser a melhor aceção para este novo insumo.

Ao analisar os dados como insumos dos algoritmos de IA, compreendendo que tal tecnologia pode tomar decisões importantes no lugar de seres humanos, colocando-os em risco. Seria mais correto comparar o potencial inovador-destruidor deste insumo, com o Urânio.⁹⁹

⁹⁷ HARARI, Yuval Noah. *Homo Deus: Uma breve história do amanhã*. 1. ed. São Paulo: Companhia das Letras, 2016. p. 346.

⁹⁸ Em 2006 o matemático e cientista de dados Clive Humby, especializado em ciência de dados, usou essa expressão para ilustrar o valor que os dados teriam no futuro.

⁹⁹ EWEEK. *Why Data is the New Uranium*. Disponível em: <https://www.eweek.com/big-data-and-analytics/why-data-is-the-new-uranium/>. Acesso em: 19 nov. 2021.

Tabela 1: Quadro comparativo entre Petróleo, Urânio e Dados

Tabela comparativa Petróleo e Urânio x Dados		
PETRÓLEO	URÂNIO	DADOS
Recurso finitos	Recursos finitos	Recursos infinitos
Alta dificuldade em encontrar, extrair e transportar.	Baixa dificuldade de encontrar, extrair e transportar.	Baixa dificuldade de encontrar, extrair e transportar.
Com avanços tecnológicos seu valor tende a diminuir.	Com avanços tecnológicos seu valor tende a aumentar .	Com avanços tecnológicos seu valor tende a aumentar .
Riscos decorrentes de acidentes/vazament o de petróleo tendem a ser graves.	Riscos decorrentes de acidentes com urânio tendem a ser catastróficos .	Riscos decorrentes de vazamento de dados tendem a ser catastróficos .
Não é utilizado para criar armas.	Pode ser usado para criar armas .	Pode ser usado para criar armas (autônomas) .
Sua atividade é pouco regulada, tendo em vista o seu baixo potencial catastrófico.	Sua atividade deve ser altamente regulada, tendo em vista os riscos que o permeiam.	Sua atividade deve ser altamente regulada, tendo em vista os riscos que o permeiam.

Fonte: Tabela criada com base na matéria, "Why Data is the New Uranium".¹⁰⁰

Deste modo, meramente a título comparativo, a tabela acima demonstra de modo inequívoco, maior convergência entre *dado x urânio* comparado a *dado x petróleo*.

¹⁰⁰ *Ibidem*.

1.8. A EMERGÊNCIA DE UMA SOCIEDADE DE RISCO DIGITAL

No decorrer deste capítulo, foi exposto a teoria da modernidade reflexiva. Nela, demonstrou-se como o processo de modernização radicalizou conceitos, invadiu premissas e contornos da sociedade industrial, trazendo efeitos colaterais como a *Individualização*, *Cosmopolitização* e a *Sociedade de Risco*. Não houve um debate quanto aos riscos trazidos pela inteligência artificial, pois estes serão tratados no próximo capítulo.

Todavia, prestou-se conectar a *Sociedade de Risco* com a Quarta Revolução Industrial, uma vez que as obras e textos de autoria do professor Beck foram escritos antes do surgimento desta revolução. A Quarta Revolução Industrial traz novas dinâmicas e conexões para Sociedade Industrial, das quais não refutam a teoria de Beck, na verdade as ratificam.

Entretanto, o ritmo muito mais acelerado desta nova revolução, tão como, a interconexão entre agentes orgânicos e inorgânicos, coloca um novo ponto de ebulição para Sociedade de Risco, a capacidade de produção de incertezas fabricadas, ou seja, novos riscos, andam na mesma velocidade evolutiva desta revolução.

Schwab elencou três motivos para justificar a existência desta revolução:

- a) Velocidade, a evolução desta segue um ritmo exponencial e não linear. Resultado do mundo multifacetado e profundamente interconectado.
- b) Amplitude e profundidade, trata-se da combinação de várias tecnologias, mudando não apenas “o que” fazemos e “como” fazemos, mas também “quem” somos.
- c) Impacto sistêmico, ela envolve a transformação de sistemas e países inteiros dentro deles, em empresas, indústrias e em toda sociedade.¹⁰¹

Deste modo, ao analisar os motivos que ensejaram a Quarta Revolução Industrial sobre o prisma da Sociedade de Risco, podemos concluir que a velocidade, agora exponencial do desenvolvimento técnico-científico, traz consigo maior potencial de produção de riscos e incertezas, que serão discutidos no próximo capítulo.

¹⁰¹ SCHWAB, op.cit., 2016. p. 16.

Beck, ao interpretar a Sociedade de Risco como um mundo de incertezas fabricadas decorrentes do avanço tecno-científico, o conhecimento se torna uma fonte de perigo, sendo assim, o aumento exponencial da velocidade de produção das inovações tecnocientíficas, resulta de modo associado, em um aumento exponencial da produção de riscos e incertezas.

A *Indústria 4.0*¹⁰² ratifica e impulsiona o processo de radicalização da modernidade, tornando-se esta, ainda mais reflexiva. Tendo em vista que esta era de descontrole frente aos riscos civilizatórios, ganha impulso exponencial com a interação de robôs em nosso meio ambiente.

Ao falar de amplitude e profundidade, percebe-se que a consequência paradoxal entre o sucesso na capacidade de coleta e processamento de dados do *Big Data*, pode torná-lo em tema e problema para sociedade atual, uma vez que ela é fonte de alimento para o desenvolvimento de novas tecnologias que podem trazer riscos e incertezas.

Na mesma lógica, ao analisar o insumo *dado* sob a ótica da *Sociedade de Risco*, fora trazida uma nova analogia a este, ao invés de classificá-lo como *novo petróleo* de modo entusiástico e talvez ingênuo, classificou-o como *novo urânio* tendo em vista o seu potencial catastrófico.

O impacto sistêmico referido por Schwab, reforça a teoria da *Sociedade de Risco Global*, uma vez que tais mudanças afetam todo o mundo e os Estados-nações não são capazes de gerir os riscos oriundos dessas mudanças de forma singular. Deslocalização, Incalculabilidade e Incompensabilidade serão demonstradas no decorrer deste trabalho.

Por fim, nota-se que a Quarta Revolução Industrial tem como principal característica o aumento dos processos de digitalização, havendo uma fusão de diversas tecnologias, juntamente com a interação de agentes físicos, digitais e biológicos. Deste modo, a I.A coabita o ecossistema dos seres humanos, auxiliando e por vezes tomando decisões por eles.

Esta interação digital entre humanos e robôs, digitaliza também a lógica distributiva de *bens* e *males*, deste modo, podemos concluir que os avanços

¹⁰² A Indústria 4.0 é a indústria que emerge da Quarta Revolução Industrial.

ora mencionados, nos colocam em um novo modelo de *Sociedade de Risco*, uma *Sociedade de Risco Digital*.

2.0. RISCOS E INCERTEZAS

Este capítulo abrirá o debate quanto aos principais riscos trazidos pela IA. Todavia, para que se habilite tal debate, primeiramente será feita uma breve conceituação do que seja IA, tão como, suas principais funcionalidades na sociedade atual.

2.1. O CONCEITO DE INTELIGÊNCIA ARTIFICIAL E SUAS CLASSIFICAÇÕES

Conforme Magrani descreve, a IA é um subcampo da informática que busca desenvolver tecnologias computacionais que possam emular a inteligência humana na realização de tarefas predeterminadas.¹⁰³

No caso, o termo Inteligência Artificial consiste em muitas vertentes conceituais diferentes. A primeira das manifestações com intuito de conceituar a IA, foi em um evento acadêmico nos Estados Unidos, liderado pelo matemático John McCarthy. O evento intitulado como Dartmouth Summer Research Project, reuniu pesquisadores que se propuseram a descobrir como fazer máquinas que fossem capazes de: (i) usar linguagem; (ii) formar instruções e conceitos; (iii) resolver problemas típicos de seres humanos; (iv) melhorarem sozinhas.¹⁰⁴

Para McCarthy, o programa de computador que fosse capaz de fazer tarefas atreladas à inteligência humana, seria reconhecido como IA. Outros teóricos, entendem que tão somente o trabalho sendo realizado, mesmo que o cálculo não se relacione ao pensamento humano, seria suficiente para enquadrá-lo como IA. Ainda há uma terceira concepção que se situa entre as duas anteriores, no qual o uso do raciocínio humano como modelo pode

¹⁰³ MAGRANI, Eduardo. Entre dados e robôs: ÉTICA E PRIVACIDADE NA ERA DA HIPERCONNECTIVIDADE. 1. ed. Porto Alegre: Arquipélago, 2019. p. 52.

¹⁰⁴ ELLIOTT, Anthony. Making Sense of AI: Our Algorithmic World. 1. ed. Sydney: Polity, 2022. p. 3.

informar e inspirar, mas não deve se restringir a emular a inteligência humana.¹⁰⁵

As definições de Inteligência Artificial trazidas por autoridades e organizações internacionais, sendo elas OECD, Brasil e UE, complementam e delimitam a compreensão deste conceito.

Para a OECD a IA é :

An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.¹⁰⁶

No Brasil, o Ministério da Ciência, Tecnologia e Inovação (MCTI) ao propor a Estratégia Brasileira de Inteligência Artificial (EBIA), seguiu a mesma linha de entendimento:

Não existe uma definição consensual de Inteligência Artificial. IA é melhor entendida como um conjunto de técnicas destinadas a emular alguns aspectos da cognição de seres vivos usando máquinas. Nessa linha, seguiremos a definição apresentada pela OCDE.¹⁰⁷

Por fim, recentemente a UE propôs uma ampliação ao conceito de IA, trazendo dentre as mencionadas, a definição mais completa:

Artificial intelligence (AI) refers to systems designed by humans that, given a complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data and deciding the best action(s) to take (according to pre-defined parameters) to achieve the given goal. AI systems can also be designed to learn to adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).¹⁰⁸

¹⁰⁵ MAGRANI, op.cit., 2019. p. 52.

¹⁰⁶ OECD. Recommendation of the Council on Artificial Intelligence, p. 1.

¹⁰⁷ CIÊNCIA, Ministério Da; INOVAÇÕES, Tecnologia E. Estratégia Brasileira de Inteligência Artificial. EBIA, Brasília, v. 1, n. 1, p. 8, jan./2021.

¹⁰⁸ COMMISSION'S, The European. A DEFINITION OF AI: MAIN CAPABILITIES AND

Deste modo, nota-se que a UE trouxe uma definição mais completa, elencando-a basicamente em três grupos: (i) Raciocínio e tomada de decisões, (ii) Aprendizagem e (iii) Robótica.

(i): Reasoning and Decision Making. This group of techniques includes knowledge representation and reasoning, planning, scheduling, search, and optimization. These techniques allow to perform the reasoning on the data coming from the sensors. To be able to do this, one needs to transform data to knowledge, so one area of AI has to do with how best to model such knowledge (knowledge representation). Once knowledge has been modelled, the next step is to reason with it (knowledge reasoning), which includes making inferences, planning and scheduling activities, searching through a large solution set, and optimizing among all possible solutions to a problem. The final step is to decide what action to take. The reasoning/decision making part of an AI system is usually very complex and requires a combination of several of the above mentioned techniques.¹⁰⁹

(ii): Learning. This group of techniques includes machine learning, neural networks, deep learning, decision trees, and many other learning techniques. These techniques allow an AI system to learn how to solve problems that cannot be precisely specified, or whose solution method cannot be described by symbolic reasoning rules. Examples of such problems are those that have to do with perception capabilities such as speech and language understanding, as well as computer vision. Notice that these problems are apparently easy, because they are indeed usually easy for humans. However, they are not that easy for AI systems, since they cannot rely on common sense reasoning (at least not yet), and are especially difficult when the system needs to interpret unstructured data. This is where techniques following the machine learning approach come in handy. However, machine learning techniques can be used for many more tasks than only perception.¹¹⁰

(iii) Robotics can be defined as “AI in action in the physical world” (also called embodied AI). A robot is a physical machine that has to cope with the dynamics, the uncertainties and the complexity of the physical world. Perception, reasoning, action, learning, as well as interaction capabilities with other systems are usually integrated in the control architecture of the robotic system. In addition to AI, other disciplines play a role in robot design and operation, such as mechanical engineering and control theory. Examples of robots include robotic manipulators, autonomous vehicles (e.g. cars, drones, flying taxis), humanoid robots, robotic vacuum cleaners, etc.¹¹¹

SCIENTIFIC DISCIPLINES. HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, Bruxelas p. 1, dez./2021.

¹⁰⁹ *Ibidem*, p.6.

¹¹⁰ *Ibidem*.

¹¹¹ *Ibidem*, p.7.

Além da completa definição trazida pela UE, para classificar sistemas de IA, ela trouxe a diferenciação entre IA estreita (fraca) comumente chamada de ANI e IA geral (forte), comumente chamada de AGI. Sendo a primeira, um sistema que pode executar uma ou poucas tarefas específicas.

Já o sistema AGI, consegue verdadeiramente simular o raciocínio humano de modo a realizar praticamente qualquer atividade que os humanos podem fazer. A AGI se parece com a autonomia do cérebro humano, podendo resolver e inclusive selecionar os problemas que ela pretende resolver, sem nenhum tipo de intervenção humana.

Nesta esteira, cumpre esclarecer que a ANI é a tecnologia mais comum utilizada, por enquanto todas as IA que interagem com seres humanos são ANI (IA fraca). Alguns exemplos destas tecnologias são: (i) assistentes pessoais virtuais; (ii) veículos autônomos; (iii) drones; (iv) elevadores inteligentes.

Conforme explanado, tais inteligências são essencialmente orientadas para um objetivo único e específico. Todavia, isso não significa que este objetivo seja simples, a ANI pode atuar em demandas complexas.¹¹²

A AGI existe apenas no campo teórico e atualmente não há exemplos práticos do seu uso. Todavia, isso não significa a impossibilidade de sua existência ou o desinteresse por parte de investigadores para o desenvolvimento desta tecnologia¹¹³. Ao contrário, protótipos da tecnologia vêm sendo apresentados em fóruns do setor.

A AGI combina percepções de diferentes campos de atividade, executa várias tarefas de inteligência e exibe considerável flexibilidade e destreza, envolvendo sua enorme capacidade de processamento computacional.¹¹⁴

Ainda, mesmo que não haja a AGI por agora, conforme explanado no capítulo anterior, a velocidade exponencial da capacidade de coleta e processamento de dados trazida pela Quarta Revolução Industrial, faz-nos crer em sua existência em um futuro próximo, tal ideia é corroborada por Barcarollo.

¹¹² Um elevador inteligente usando IA pode gerenciar o fluxo de demanda de um prédio empresarial com base nos dados coletados do uso diário. ELLIOTT, Anthony. Making Sense of AI: Our Algorithmic World. 1. ed. Sydney: Polity, 2022. p. 6.

¹¹³ IBM. Transformação digital empresarial com Inteligência Artificial. Disponível em: <https://www.ibm.com/br-pt/analytics/journey-to-ai>. Acesso em: 1 dez. 2021.

¹¹⁴ ELLIOTT, op.cit., 2022. p. 7.

Para ele, o elemento humano é o que possibilita o desenho arquitetônico da Inteligência Artificial. Deste modo, a harmonização tecnológica trazida pela Quarta Revolução Industrial possibilita a consecução desta tecnologia, alimentando-a com dados. Desta forma, ele acredita que em um primeiro momento a IA fraca, consegue a partir da coleta e tratamento de dados, atuar de forma preditiva, sistematizando, analisando e apontando indicativos estatísticos que o ser humano não conseguiria fazer, ou faria de forma extremamente morosa. Entretanto, quanto maior a alimentação de dados nestes sistemas, eles tendem a se tornarem cada vez mais fortes. Deste modo, pode-se emergir a AGI, que além de associar e predizer, poderá tomar decisões mais complexas.¹¹⁵

Ainda há uma terceira variante de IA, chamada de Super Inteligência Artificial ou ASI. Na verdade, muitos classificam a ASI como um tipo de AGI. Todavia, diferentemente da AGI que tem como objetivo igualar a Inteligência Humana, a ASI busca ultrapassar a capacidade do cérebro humano. Muitos especialistas preveem que a ASI consiga superar a inteligência humana em todos os domínios, incluindo raciocínio cognitivo e habilidades sociais. A ideia desta tecnologia é muito utilizada em filmes de ficção em Hollywood.¹¹⁶

Para fins deste trabalho, utilizar-se-á o conceito de IA traduzido pela UE, tendo em vista esta trazer um conceito mais amplo ao debate.

2.2. PRINCIPAIS FUNCIONALIDADES DA IA NO CONTEXTO ATUAL

Atualmente, a IA está presente em diversos campos. Deste modo, de forma não exaustiva, faz-se necessário demonstrar brevemente algumas das funcionalidades da IA no contexto atual.

¹¹⁵ ENGELMANN, Wilson; BARCAROLLO, Felipe. Inteligência Artificial na Advocacia do Século XXI. 2016. Apud: WERNER, Deivid Augusto. A Quarta Revolução Industrial e a Inteligência Artificial: um estudo sobre seus conceitos, reflexos e possível aplicação no Direito por meio de análise de texto jurídico como forma de contribuição no processo de categorização preditiva de acórdãos. Dissertação de Mestrado, Porto Alegre, v. 1, n. 1, p. 27, ago./2019.

¹¹⁶ ELLIOTT, op.cit, 2022. p. 7.

- a) *Aplicação em E-commerce*: (i) identificar preferências de usuários com base em seu comportamento; (ii) recomendações de acordo com seu hábito de consumo; (iii) uso de sistemas de atendimento automático (Chatbots).¹¹⁷
- b) *Entretenimento*: Empresas como Netflix empregam essa tecnologia para aperfeiçoar o sistema de recomendações para usuários.
- c) *Internet, Redes Sociais e Aplicativos*: Aplicativos como Google Fotos por exemplo, é capaz de reconhecer objetos, pessoas, animais e situações específicas, criando álbuns e vídeos temáticos.¹¹⁸
- d) *Tradutor de idiomas*: utilizando redes neurais e machine learning (que serão melhor explicados posteriormente), a tecnologia em seu estágio atual, consegue compreender até mesmo as nuances do idioma traduzido.
- e) *Moderação de conteúdo em redes sociais*: atualmente as redes sociais vêm utilizando IA para identificar conteúdos impróprios ou notícias falsas.
- f) *Logística e transporte*: O setor logístico se beneficia da análise de dados em tempo real para reduzir custos operacionais, e a capacidade de identificar rotas de entrega com tráfego menos intenso.

¹¹⁷ O Chatbot, é um tipo de IA utilizada para realizar conversas com humanos com uma linguagem natural.

¹¹⁸ A IA neste caso, consegue elaborar álbuns personalizados de momentos com alguém querido ou de alguma viagem especial. Posteriormente ao falar de discriminação algorítmica será demonstrado como essa tecnologia consegue utilizar essa mesma lógica, para discriminar racialmente.

- g) *Saúde*: A área de saúde utiliza I.A para análise de dados e identificação de doenças, o robô IBM Watson apresenta números impressionantes de acurácia em diagnósticos de câncer.
- h) *Carros inteligentes*: A chave para a condução de carros sem motorista é a I.A. Tal tecnologia promete trazer impactos positivos, no quesito “redução de acidentes”, todavia poderá trazer impactos negativos no mercado de trabalho.
- i) *Automação industrial*: Aplicação de manufatura avançada.¹¹⁹
- j) *Reconhecimento Facial*: Uso da tecnologia em uma enorme gama de campos, desde funcionar como chave de acesso para o “desbloqueio” de telas de celulares e cancelas de condomínios edilícios até o uso do poder público para elaborar retratos falados de criminosos¹²⁰ e para redes de vigilância.

Nota-se que a IA vem sendo utilizada, em diversos campos, seu potencial em fazer análises preditivas robustas trazem um importante arcabouço para o desenvolvimento da indústria e da sociedade como um todo.

2.3. INTELIGÊNCIA ARTIFICIAL: PRINCIPAIS RISCOS

Após conceituada IA e elencadas suas características gerais, tão como demonstrada suas funcionalidades no contexto atual, passa-se a examinar os riscos trazidos por estes sistemas. Novamente esta abordagem não será feita

¹¹⁹ Manufatura avançada é um conjunto de tecnologias que viabiliza a construção de um processo produtivo mais personalizado, eficiente e autônomo, visando a redução de custos, economia de energia, diminuição de erros, fim do desperdício e aumento da segurança e da inovação.

¹²⁰ O retrato falado é um mecanismo de identificação de criminosos por meio de relato de vítimas, testemunhas e de imagens que permitam estabelecer padrões de idade e reconstituição facial.

de forma exaustiva, até porque, a interação da IA na Sociedade (de Risco Digital), se torna cada dia mais complexa. Deste modo, novas tecnologias e consequentemente novos riscos, são distribuídos a todo momento.

2.3.1. DISCRIMINAÇÃO ALGORÍTMICA E OS RISCOS MULTIFACETADOS AOS DIREITOS HUMANOS

Inicialmente, antes de iniciar o debate sobre os riscos aos Direitos Humanos, é importante destacar a manifestação do Conselho de Direitos Humanos da Organização das Nações Unidas - ONU que definiu a transposição dos Direitos Humanos concebidos ao ambiente offline, para o ambiente online. Isso significa que tratados internacionais que visam proteger Direitos Humanos, tais como a Declaração Universal de Direitos Humanos -DUDH, aplicam-se tanto no cenário virtual, como online.¹²¹

Tal iniciativa corrobora a ideia trazida no capítulo anterior quanto ao processo de digitalização da sociedade. O estudo publicado pelo *Berkman Center for Internet & Society*, traz elementos importantes para a compreensão dos riscos trazidos pela IA nos Direitos Humanos. Elencando-os:

2.3.1.1. DECISÕES AUTOMATIZADAS NO CAMPO CRIMINAL

A aplicação de IA em ferramentas de avaliação de riscos no sistema criminal veio com o objetivo de reduzir índices de violência, alocar policiais em zonas de maior criminalidade e eliminar vieses na dosimetria de penas.

Todavia, a análise preditiva apresentou vieses que acabam marginalizando ainda mais, grupos que já eram historicamente marginalizados.

¹²¹ AGÊNCIA BRASIL. Direitos Humanos devem ser respeitados. Disponível em: <https://agenciabrasil.ebc.com.br/direitos-humanos/noticia/2018-07/direitos-humanos-devem-ser-respeitados-na-internet-diz-onu>. Acesso em: 2 fev. 2022.

Cathy O'Neal exemplifica esta situação em sua obra *Algoritmos de Destruição em Massa*. Ela classifica este fenômeno como *ciclo de feedback*.¹²²

Ela concluiu haver um ciclo retroalimentativo discriminatório, o algoritmo ao indicar determinado bairro como perigoso, alocava mais policiais naquele perímetro, de modo que o aumento de registros de crimes subia na mesma medida.¹²³ Essa análise baseada em ferramentas estatísticas distorcia a lógica de causalidade e correlação, uma vez que as demais áreas não recebiam uma vigilância tão ostensiva quanto aos bairros de maiores índices de criminalidade. Logo, essa assimetria de forças aumentava a marginalização de um grupo comparado ao outro.¹²⁴

O mesmo acontece ao utilizar sistemas de IA para auxílio na dosimetria da pena, ao classificar de forma errônea réus como *alto risco*. Uma vez que a análise de dados históricos traz padrões sistemáticos de preconceito contra determinado grupo.

O uso destes dados históricos para análise de risco, podem trazer uma super-representação desta minoria nos dados de treinamento, resultando em variáveis distorcidas. A título exemplificativo, utilizaremos o cenário Brasileiro.^{125 126}

Os presos do sistema carcerário brasileiro são majoritariamente jovens, negros e pobres¹²⁷, deste modo haverá uma coleta massiva de dados do perfil descrito. Logo a IA pode classificar como *alto risco* o perfil ora descrito,

¹²² O'NEIL, Cathy. *Algoritmos de Destruição em Massa: Como o Big Data aumenta a desigualdade e ameaça a Democracia*. 1. ed. Santo André: Rua do Sabão, 2020. p. 15.

¹²³ *Ibidem*. p.136-137.

¹²⁴ RASO, FILIPPO et al. *Artificial Intelligence & Human Rights: Opportunities & Risks*. The Berkman Klein Center for Internet & Society Research Publication Series, Cambridge, Massachusetts, v. 1, n. 1, p. 21-25, set. /2018.

¹²⁵ GOVBR. Mapa do Encarceramento aponta: maioria da população carcerária é negra. Disponível em: https://www.gov.br/mdh/pt-br/noticias_seppir/noticias/junho/mapa-do-encarceramento-aponta-maioria-da-populacao-carceraria-e-negra-1. Acesso em: 24 dez. 2021.

¹²⁶ GLOBO G1. Proporção de negros nas prisões cresce 14% em 15 anos, enquanto a de brancos cai 19%, mostra Anuário de Segurança Pública. Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2020/10/19/em-15-anos-proporcao-de-negros-nas-prisoas-aumenta-14percent-ja-a-de-brancos-diminui-19percent-mostra-anuario-de-seguranca-publica.ghtml>. Acesso em: 26 dez. 2021.

¹²⁷ ESTADO DE MINAS. Levantamento aponta que maioria dos presos no Brasil são jovens, negros e pobres. Disponível em: https://www.em.com.br/app/noticia/nacional/2015/06/23/interna_nacional,661171/levantamento-aponta-que-maioria-dos-presos-no-brasil-sao-jovens-negro.shtml. Acesso em: 28 dez. 2021.

impactando diretamente em sua pena. Minorias de baixa renda são classificadas como mais perigosas, conseqüentemente devem receber sentenças mais pesadas.

Grupos que historicamente foram desproporcionalmente punidos, correm o risco de serem classificados como *alto risco*, nesta lógica, o algoritmo pode amplificar e perpetuar preconceitos e gerar dados ainda mais enviesados, alimentando um ciclo vicioso. Como a maioria dos algoritmos de avaliação de riscos são protegidos pelo segredo industrial, é impossível questionar suas decisões ou responsabilizá-los.¹²⁸ Esta impossibilidade de responsabilização é um dos maiores dilemas trazidos pela *Sociedade de Risco*.

Marbre Stahly-Butts, diretora do Law for Black Lives diz que a avaliação de riscos baseada em dados é uma maneira de *higienizar* e legitimar sistemas opressivos.¹²⁹

O estudo conclui que os impactos em Direitos Humanos são principalmente em: (i) Direito a audiência pública e justa; (ii) presunção de inocência; (iii) direito à privacidade; (iv) direito a não discriminação.¹³⁰

2.3.1.2. CREDIT SCORING/SOCIAL SCORING

Os credores vêm utilizando IA para cálculo de risco para concessão de crédito. Historicamente, o cálculo de risco neste campo teve precisão duvidosa e repletas de possibilidades de discriminação. Similar ao tópico anterior, a IA pode perpetuar vieses e acentuar a discriminação a grupos já marginalizados.¹³¹

O'Neal ressalta em seu livro que todos os dados podem ser utilizados para criar o seu perfil de crédito, ou seja, para além da coleta de dados raciais, renda e histórico de crédito, são coletados para até mesmo *likes* em redes

¹²⁸ MIT TECHNOLOGY REVIEW. AI is sending people to jail and getting it wrong. Disponível em: <https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/>. Acesso em: 29 dez. 2021.

¹²⁹ *Ibidem*.

¹³⁰ RASO et al., op.cit., 2018. p. 21-25.

¹³¹ *Ibidem* p.28.

sociais, que demonstram suas preferências pessoais, tais como: posicionamento político, religião, orientação sexual, perfil de crédito dos seus amigos e familiares.

O estudo conclui que a tomada de decisão baseada em IA no setor financeiro, afeta negativamente os direitos humanos relativos à liberdade de expressão e direito de associação. Fundamenta-se essa conclusão na crença de que, a partir do momento que todos os dados podem ser utilizados para finalidade de concessão de crédito, os indivíduos podem se sentir desconfortáveis para expressar certos posicionamentos ou se associar a determinados grupos sociais, por medo de que o algoritmo possa usar essas condutas contra eles no contexto de pontuação de crédito.¹³²

Nesta lógica, observa-se o caso do *social scoring* chinês, a IA utilizada pelo governo é o exemplo mais extremo deste fenômeno. A IA gera um índice de *confiabilidade* do indivíduo com base em uma enorme gama de pontos de coleta de dados, incluindo dados de mídia social, prisão e registros de infrações, trabalhos voluntários, endereço de residência, entre outros. O sistema faz com que aqueles com alta pontuação desfrutem de maiores benefícios do estado, taxas de serviços públicos mais baratos, condições de empréstimos mais favoráveis, o sistema até mesmo cerceia o direito de alguns cidadãos de acessar determinados serviços caso tenha *scoring*¹³³ muito baixo.

Em que pese tal sistema tenha ajudado a conter a corrupção e incentivado formas de bom comportamento, como atravessar a rua na faixa de pedestres. É fácil concluir que este sistema de ultra vigilância transgride diretamente o exercício de atividades de expressão e associação.¹³⁴

¹³² *Ibidem*, p.32.

¹³³ Score é um substantivo da língua inglesa que significa pontuação.

¹³⁴ RASO et al., op.cit., 2018. p. 32.

2.3.1.3. ACESSO À SAÚDE

Sistemas de diagnóstico baseados em IA tem o potencial de fornecer maior acesso ao tratamento de nível especializado comparado ao que é praticado nos dias de hoje. Um sistema de diagnóstico clínico baseado em IA foi aprovado pela Food and Drug Administration dos EUA, ele foi capaz de diagnosticar a retinopatia diabética de forma autônoma, algo que só era possível ser feito por um especialista da área.¹³⁵

Todavia, por mais que ele traga reflexos positivos aos Direitos Humanos¹³⁶, ele traz em conjunto riscos à privacidade. Seguindo novamente a lógica distributiva trazida pela *Sociedade de Risco*.¹³⁷

Os riscos, no caso, estão relacionados à privacidade do indivíduo. Dados de seus antepassados podem ser utilizados para negar a cobertura de um plano de saúde por fatores genéticos.¹³⁸ De modo análogo ao caso anterior, se todos os dados forem dados para saúde, como todos os dados são dados de crédito, não é difícil concluir que até mesmo os seus hábitos alimentares, frequência de consumo alcoólico, tabagismo, captados facilmente pelas redes sociais dentre outros mecanismos, podem ser o motivo ensejador para o aumento do seu plano de saúde ou a negativa de determinada cobertura.

2.3.1.4. RECURSOS HUMANOS

A lógica discriminatória se transpõe também para o campo do trabalho, sem intervenção na programação é provável que sejam reproduzidos padrões sistêmicos existentes de viés e preconceito. Isso quer dizer que a contratação baseada em IA ao identificar métricas para avaliar candidatos, podem refletir

¹³⁵ *Ibidem*, p. 35.

¹³⁶ Artigo 25 da DUDH vincula o acesso à assistência médica aos requisitos básicos da vida, como alimentação, vestuário e habitação, “a um padrão de vida adequado à saúde e bem-estar de si e de sua família”.

¹³⁷ Na *Sociedade de Risco* a produção de *goods* é acompanhada pela produção de *bads*, conceito já explicado no capítulo anterior.

¹³⁸ RASO, et al., p.cit., 2018. p. 30.

vieses estruturais ao invés dos determinantes objetivos do desempenho do emprego.

Por exemplo, a assimetria entre homens e mulheres no mercado de trabalho e os cargos que ocupam, pode ser um fator de pontuação, deste modo a IA ao pontuando com base em estatísticas comparativas de gênero, reforçará ainda mais a desigualdade de gênero no mercado de trabalho.

2.4. RISCOS À DEMOCRACIA

Sistemas de IA podem também trazer riscos às eleições por meio do uso de *Bots Sociais*¹³⁹. *Bots* em redes sociais podem assumir identidades por meio de contas falsas e divulgar opiniões sobre candidatos políticos ou temas relacionados. Por meio de técnicas como *profiling*¹⁴⁰, *micro-targeting*¹⁴¹, entre outras, os *bots* conseguem identificar perfis de usuários que têm potencial de disseminar *fake news*¹⁴² e/ou acreditar nelas. Deste modo, além de interagir com outros usuários, os *bots sociais* conseguem identificar perfis reais que podem contribuir para disseminação de *fake news*.¹⁴³

O professor Magrani alerta que além dos *bots* serem utilizados por partidos políticos para conquistar seguidores, eles vêm sendo utilizados também para atacar opositores e disseminar notícias falsas. De modo a *forjar a esfera pública*¹⁴⁴.

¹³⁹ Bots é a abreviatura de robô, são softwares que executam tarefas automaticamente.

¹⁴⁰ Trata-se de uma técnica de processamento de dados pessoais que busca definir o perfil de um indivíduo para determinado produto ou serviço.

¹⁴¹ Microtargeting é uma técnica de segmentação de conteúdo. Ela visa transmitir mensagens que refletem nas preferências do indivíduo alvo. Trata-se do envio de notícias direcionadas à preferência do consumidor. Tal técnica só é possível graças à técnica explanada anteriormente o *profiling*.

¹⁴² Tradução: Notícias falsas.

¹⁴³ MAGRANI, op.cit., 2019. p. 157.

¹⁴⁴ Magrani utiliza o conceito de esfera pública, sob a ótica Jürgen Habermas. Deste modo, entende-se esfera pública, como uma arena de debate público em que os assuntos de interesses coletivos podem ser discutidos em espaços formais e informais. Este campo é necessário para a efetiva participação democrática.

Estes robôs manipulam debates políticos criando e disseminando notícias falsas e/ou postando e replicando mensagens em larga escala, inundando as redes sociais com determinada notícia. Além disso, eles replicam hashtags para aumentar o engajamento de determinada postagem, criando uma falsa impressão de aceitação pública desta opinião e/ou notícia. Esta dinâmica sufoca o debate público sobre determinado tema, uma vez que essa movimentação automatizada é diluída na esfera de formação de opiniões humanas. Magrani ainda pontua que há uma enorme gama de robôs atuando com essa finalidade maliciosa. Eles buscam emular a conduta humana nas redes sociais, se passando por indivíduos. Deste modo, eles interferem na *esfera pública*, forjando discussões como se humanos fossem.¹⁴⁵

Além do impacto dos *bots sociais* no ambiente democrático, há também a criação de *filtros bolha*, dos quais também trazem efeitos deletérios ao ambiente democrático. O termo *Filter Bubble* (traduzido em português como filtros bolha) foi trazido pelo Diretor executivo da organização “Moveon.Ong”. Em síntese, *Filter Bubble* pode ser definido como uma forma de coleta e tratamento de dados feito por algoritmos de IA visando customizar a navegação online. Redes sociais, google e outras plataformas provedoras de conteúdo fazem esse tipo de customização.¹⁴⁶

Deste modo impositivo, forma-se um universo particular para cada usuário, a ideia central deste mecanismo é transformar a navegação do usuário em uma navegação mais prazerosa e como consequência, o usuário despende mais tempo na plataforma. Todavia, isso traz diversos impactos na formação do debate público, justamente porque o acesso à informação difere de usuário para usuário.

A customização destes filtros utiliza diversos critérios para definir qual tipo de informação será enviada para cada tipo de perfil de indivíduo. Nacionalidade, raça, religião, gostos pessoais podem ser utilizados para essa finalidade. Desta forma, um indivíduo residente no ocidente, ao pesquisar uma

¹⁴⁵ MAGRANI, op.cit., 2019. p. 52.

¹⁴⁶ MAGRANI, Eduardo. Democracia Conectada: A Internet como Ferramenta de Engajamento Político-Democrático. 1. ed. Curitiba: Juruá, 2014. p. 119.

notícia X, terá resultados de busca diferentes do que um indivíduo que reside no oriente. O mesmo se dá para um indivíduo branco e um negro, brasileiro ou iraniano, aficcionado por esportes ou artes. Essa personalização do conteúdo que lhe será enviado baseado em critérios pré definidos é o que caracteriza o *Filter Bubble*. Esse fenômeno é muito presente nas redes sociais.¹⁴⁷

Essa personalização do acesso à informação coloca o cidadão em uma *câmara de eco*¹⁴⁸ uma vez que conteúdos que os desagradam, ficam fora do seu espectro. Deste modo, ele tem uma falsa impressão de aceitação global de suas opiniões.

A tecnologia é nociva para a formação do debate público, a falsa impressão de aceitação de suas opiniões e gostos tem como consequência última a polarização e fragmentação do debate.

A título exemplificativo, um indivíduo com posicionamento político de direita, ao navegar em suas redes sociais e fazer postagens emitindo sua opinião política, terá filtrado (retirado) do seu espectro posições políticas de esquerda, postagem de seus amigos que diverjam das suas, serão prontamente escondidas. De outro giro, as postagens que converjam com seus pensamentos ficarão em evidência. Assim, Magrani elenca duas falsas impressões causadas por esta tecnologia: (i) a primeira é o indivíduo acreditar que toda a sua rede de contatos está visualizando tudo que ele posta, quando na realidade, apenas acessam este conteúdo aqueles aqueles que convergem com suas ideias. (ii) A segunda, é a falsa impressão de consenso, justamente porque essa filtragem tira de seu espectro pessoas que poderiam discordar de suas ideias. Logo, este conteúdo é levado para pessoas que apoiam suas opiniões, trazendo uma falsa crença de que sua opinião é a mais sensata e popular.¹⁴⁹

¹⁴⁷ *Ibidem*.

¹⁴⁸ O conceito *câmara de eco* é uma metáfora para descrever uma situação em que ideias, informações ou crenças são amplificadas e reforçadas pela comunicação e repetição sistemática.

¹⁴⁹ IRIS. Fake news, bots, mecanismos de filtragem e outras ameaças à democracia brasileira. Disponível em: <https://irisbh.com.br/fake-news-bots-mecanismos-de-filtragem-e-outras-ameacas-a-democracia-brasileira/>. Acesso em: 3 jan. 2022.

O usuário ao acreditar que seu ponto de vista é o mais razoável, junto ao pouco contato com opiniões diversas, tem como consequência lógica a aversão por opiniões diferentes da sua.

Gamba, corrobora essa ideia em seu livro *Democracia e Tecnologia*, ao discorrer sobre a arquitetura de nudging¹⁵⁰ e o avanço dos extremos, em suas palavras:

A grande questão aqui é que isso ocasiona, por óbvio, um aumento da polarização e reforça eventuais tendências extremistas, especialmente no que tange ao comportamento do usuário em mídias sociais - que tenderá a considerar seu comportamento extremista algo comum, vez que recorrente em sua rede. Ou seja, o usuário encontrará na rede diversas manifestações - políticas, por exemplo - similares às suas e poucas manifestações contrárias, reforçando na visão desse usuário particular sua forma de pensar, uma vez que encontrará eco nas mídias sociais e na internet como um todo. ¹⁵¹

Deste modo, fica claro os riscos trazidos pela IA (seja por *bots sociais* ou por mecanismos de filtragem de conteúdo) ao ambiente democrático.

2.5. ENTRE HARARI E BECK: INTELIGÊNCIA ARTIFICIAL E O FUTURO DO TRABALHO. A RATIFICAÇÃO DOS CONCEITOS DE IRRESPONSABILIDADE ORGANIZADA E EFEITO BUMERANGUE.

Questões sobre o futuro do trabalho tendem a ser um dos grandes debates da IA na humanidade. O foco principal desta discussão permeia justamente pelo fato de que a IA vem cada vez mais ocupando postos de trabalho que eram ocupados por seres humanos. Conforme esta tecnologia avança, ela consegue executar tarefas mais complexas, trazendo preocupações viscerais, tais como: no momento que a IA conseguir executar qualquer tarefa melhor que um ser humano, qual será a utilidade deste último

¹⁵⁰ “Nudging, ou “arquitetura de escolha”, refere-se à forma de funcionamento de determinados programas, aplicativos, sites e, em especial, mídias sociais, no que tange à influência gerada para a tomada de decisão do usuário. Trata-se de uma tentativa de direcionar a conduta do usuário a partir de sugestões apresentadas, empurrando-as para elas.” GAMBA, op.cit, 2020. p. 113.

¹⁵¹ GAMBA, op.cit, 2020. p. 113.

para a economia global? Quais serão os impactos da automação da manufatura em países pouco industrializados? Como cuidar de uma possível futura classe humana inútil?

A ausência de movimentações globais para responder estes questionamentos, ratificam o conceito da *Irresponsabilidade Organizada* trazida pelo professor Beck, delineado no capítulo anterior. Como consequência, este processo evolutivo promete causar efeitos em países desenvolvidos, estes em um primeiro momento gozarão da distribuição de *bens*, evitando temporariamente a distribuição de *males*. Todavia, tal processo tende a proporcionar um efeito rebote a estes países, ratificando também o conceito *Efeito Bumerangue*.

O professor Harari em um artigo intitulado de *The meaning of life in a world without work*¹⁵², publicado no jornal *The Guardian*, reforça a ideia trazida em sua última obra, para ele a IA paulatinamente irá ocupar todos postos de trabalho da humanidade, ao ponto de fazer com que um enorme número de pessoas se tornem *inúteis*, do ponto de vista econômico.¹⁵³

Schwab também crê que as inovações em tecnologias trazidas pela Quarta Revolução Industrial podem causar impactos no meio ambiente do trabalho, uma vez que a tecnologia tende a substituir postos de trabalhos ocupados por humanos, e não necessariamente criará novos produtos que demandem especificamente do trabalho humano, evidenciando um importante reflexo da tecnologia no desemprego.¹⁵⁴

No curto e médio prazo, a IA demonstra ter um efeito disruptivo nos padrões tradicionais de trabalho. Essa ruptura nem sempre é negativa, ela trouxe oportunidades em nichos de *microtrabalho*¹⁵⁵. Todavia, a automação reduz a demanda por pessoas com habilidades das quais ela supre.¹⁵⁶

¹⁵² Tradução: O sentido da vida em um mundo sem trabalho

¹⁵³ THE GUARDIAN. The meaning of life in a world without work. Disponível em: <https://www.theguardian.com/technology/2017/may/08/virtual-reality-religion-robots-sapiens-book>. Acesso em: 3 jan. 2021.

¹⁵⁴ SCHWAB, op.cit, 2016. p. 16.

¹⁵⁵ Microtrabalho refere-se a tarefas ou trabalho de curto prazo para humanos, como dirigir, limpar. Aplicativos como TaskRabbit e Uber oportunizaram um crescimento neste nicho de mercado.

¹⁵⁶ OSOBA, Osonde A.; IV, William Welser. The Risks of Artificial Intelligence to Security and

Economistas, ao analisar a suscetibilidade ora descrita, descobriram que empregos de rotina de média qualificação (manufatura, produção, operadores) são mais suscetíveis a serem substituídos por máquinas, de outro giro, empregos que exigem criatividade, inteligência, inteligência social e manipulação refinada tendem a sofrer menos impactos, inicialmente.¹⁵⁷

Um estudo produzido pela Universidade de Brasília - UNB, trouxe dados alarmantes quanto aos reflexos da automação em países menos desenvolvidos. O estudo destaca que em países desenvolvidos, membros da OCDE, os reflexos da automação correspondem a menos de 15% dos postos de trabalho.¹⁵⁸

As oscilações entre a porcentagem de postos de trabalhos automatizados em cada país estão diretamente ligadas aos seus níveis educacionais. Deste modo, países em que a mão de obra é representada majoritariamente por trabalhos de baixa qualificação tendem a sofrer mais com a automatização dos postos de trabalho. Logo, é forçoso concluir que em países mais pobres, a automação tende a ter um impacto mais deletério no mercado de trabalho.¹⁵⁹

Estima-se que cada robô inserido no mercado de trabalho, retira o emprego de seis trabalhadores. Cada novo robô reduz os salários dos trabalhadores em 0,5%, em uma proporção de um para mil (um robô para mil trabalhadores). Tais efeitos são mais pronunciados em processos de manufatura.¹⁶⁰

Deste modo, países que tem sua economia pautada em processos de manufatura e tem a maioria de sua força de trabalho desqualificada, experimentam maiores incertezas quanto à continuidade de postos de trabalho. Estima-se que países emergentes têm um alto risco de automação, ex: China

the Future of Work: subtítulo do artigo. Rand Corporation: subtítulo da revista, Santa Monica, CA, v. 1, n. 1, p. 10, jan./2017.

¹⁵⁷ *Ibidem*.

¹⁵⁸ ALBUQUERQUE, P. H. M. et al. NA ERA DAS MÁQUINAS, O EMPREGO É DE QUEM? ESTIMAÇÃO DA PROBABILIDADE DE AUTOMAÇÃO DE OCUPAÇÕES NO BRASIL. Instituto de Pesquisa Econômica Aplicada, Rio de Janeiro, v. 1, n. 1, p. 11, mar./2021.

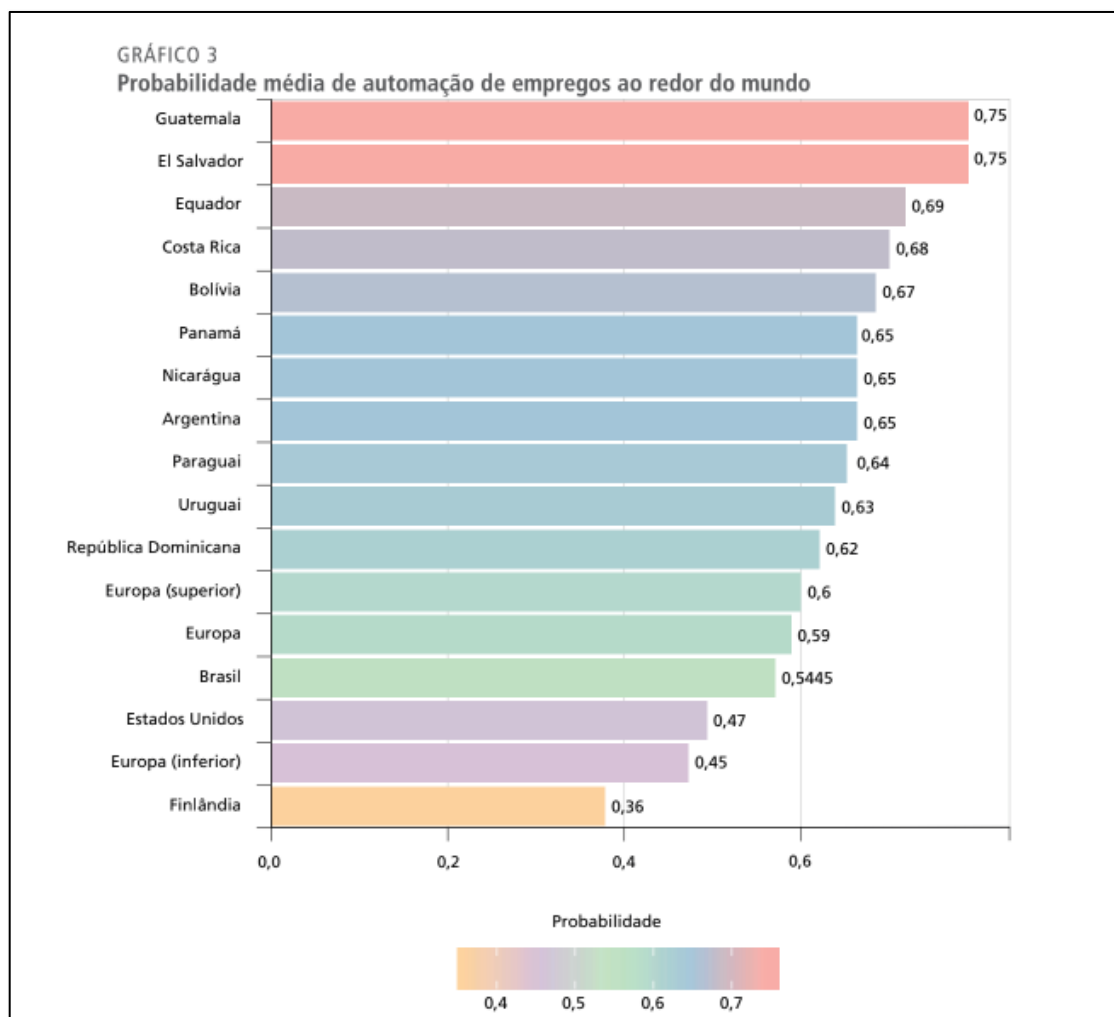
¹⁵⁹ *Ibidem*.

¹⁶⁰ *Ibidem*.

77% e Índia 69%. Países subdesenvolvidos podem oscilar entre 55% (Uzbequistão) até 85% (Etiópia).¹⁶¹

O gráfico abaixo, demonstra a probabilidade de automação de empregos ao redor do mundo ¹⁶²

Gráfico 1: Probabilidade média de automação de empregos ao redor do mundo



Fonte: Bosch, Pagés e Ripani (2018), Bowles (2014), Brzeski e Burk (2015), Frey e Osborne (2013) e Pajarinen e Rouvinen (2014). APUD ALBUQUERQUE, op.cit. p. 24.

O estudo conclui que até 2026 aproximadamente 30 milhões de postos de trabalho no Brasil podem ser extintos, caso as empresas optem por

¹⁶¹ *Ibidem*, p.13.

¹⁶² Bosch, Pagés e Ripani (2018), Bowles (2014), Brzeski e Burk (2015), Frey e Osborne (2013) e Pajarinen e Rouvinen (2014). APUD ALBUQUERQUE, op.cit. p. 11.

maximizar sua produção automatizando processos executados por humanos. Para enfrentar esta crise, o artigo propõe a adoção de políticas públicas que busquem antecipar o aumento do desemprego. Desta forma, a política pública deve ser voltada à educação, capacitação e requalificação do trabalhador, para que este adquira habilidades de difícil automatização.¹⁶³

Nota-se que o resultado do estudo em questão, aponta para uma necessidade de reeducação da classe trabalhadora na proporção dos avanços da IA no mercado de trabalho. Pois como demonstrado, campos que necessitam de classe trabalhadora pouco qualificada, tendem a sofrer primeiramente aos impactos da automação, manufatura por exemplo.

Todavia, de acordo com o professor Harari este processo de reeducação tende a ser feito de forma assimétrica, tendo em vista o potencial econômico de cada Estado-nação em reeducar sua população e os níveis educacionais já implementados em cada um dos Estados. Para Harari, países ricos terão recursos para retrainar sua mão de obra, ao passo que estes recursos podem faltar em países pobres.¹⁶⁴

Harari, indiretamente ratifica a insuficiência dos institutos trazidos pela modernidade para solucionar tal conflito, este é um dos principais conceitos que caracterizam a Sociedade de Risco.

Para ele, mesmo que o governo se empenhe para requalificar a mão de obra de milhões de pessoas que possam a vir a se tornarem inútil para o mercado de trabalho, a velocidade das mudanças ocasionadas por essa nova dinâmica evolutiva tende a invalidar este empenho. Uma vez que a população terá de se reinventar repetidamente em um curto espaço de tempo.¹⁶⁵

Deste modo, Harari sugere que sejam repensados novos modelos de sociedade, economia e política em uma era pós-trabalho. Neste ponto, o professor destaca a importância de reconhecer a insuficiência dos modelos

¹⁶³ *Ibidem*.

¹⁶⁴ HARARI, Yuval Noah. 21 lições para o século 21. 1. ed. São Paulo: Companhia Das Letras, 2018. p. 59.

¹⁶⁵ *Ibidem*.

sociais, econômicos e políticos herdados pelo processo de modernização, entendendo-os como inadequados para lidar com este desafio.¹⁶⁶

Para Harari consequência da *IA Revolution*¹⁶⁷ na indústria será o enriquecimento extremo de países já desenvolvidos ao mesmo passo produzirá a ruína para a economia de países pobres, que tem sua economia centrada em mão de obra barata.¹⁶⁸

A fragilidade dos países menos desenvolvidos para lidar com este fenômeno, pode trazer efeitos colaterais indesejados para os países mais desenvolvidos, como crises migratórias clandestinas por motivos econômicos¹⁶⁹ ou crises sanitárias que podem afetar o mundo todo.¹⁷⁰

O empobrecimento extremo destes países tende a piorar suas condições de infraestrutura sanitária e de saúde, piorando sua capacidade em controlar situações que ensejam riscos globais, ratificando o conceito do *Efeito Bumerangue* trazido por Beck.

As críticas de Yuval Noah Harari em muito convergem com a Teoria da *Sociedade de Risco*. Yuval alerta quanto à insuficiência dos Estados-nações para resolver este problema, ao mesmo passo, não há uma movimentação entre os países produtores destes novos riscos em sanar tal problema.

Harari utiliza como exemplo os países em desenvolvimento carentes de recursos naturais, estes em sua maioria têm economia baseada em oferecer mão de obra barata. Seja na manufatura, atuando na fabricação de camisetas e tênis para grandes marcas. Seja na produção de serviços, tais como *Costumer Services*¹⁷¹ ou *Telemarketing*¹⁷².

¹⁶⁶ *Ibidem*.

¹⁶⁷ IA Revolution foi um termo trazido por Yuval Harari para classificar a revolução da Inteligência Artificial no mercado de trabalho.

¹⁶⁸ Entrevista com Yuval Noah Harari, 2019. 1 vídeo (16:40 min - 18:30 min). Publicado pelo canal Roda Viva. Disponível em: <https://www.youtube.com/watch?v=pBQM085IxOM&t=2208s>

¹⁶⁹ Imigração clandestina é o processo em que pessoas se deslocam de um determinado país para outro de forma voluntária. Todavia, sem respeitar a legislação imigratória daquele país ou até mesmo, sem se apresentar para os órgãos imigratórios.

¹⁷⁰ BRASIL DE FATO. Vítimas do neoliberalismo, países pobres sofrem consequências nos sistemas de saúde. Disponível em: <https://www.brasildefato.com.br/2020/06/09/vitimas-do-neoliberalismo-paises-pobres-sofrem-consequencias-nos-sistemas-de-saude>. Acesso em: 1 jan. 2022.

¹⁷¹ Trata-se de um serviço de atendimento a reclamações e solicitações de clientes de determinado produto ou serviço.

¹⁷² Trata-se de um trabalho de promoção de vendas de produtos ou serviços por telefone.

Com o avanço da IA, junto a robótica e a impressora 3D, estes trabalhos tendem a ser automatizados. Para Harari, não faz sentido um produto ser produzido manualmente em um país subdesenvolvido para posteriormente ser enviado para um país desenvolvido (onde o produto será comercializado), tendo em vista a possibilidade de produzir este produto com robôs, no país destinatário final (país desenvolvido). Da mesma forma, a IA poderia fazer o trabalho de *Costumer Service*, de modo mais barato e eficiente, esvaziando todos os postos de trabalho desta natureza.¹⁷³

Harari sugere uma cooperação global para sanar tal problema, garantindo uma renda básica universal para as populações destes países afetados, como também, mecanismos de tributação para empresas de países desenvolvidos que automatizaram a manufatura. Todavia, não há otimismo por parte do professor em acreditar que tais medidas serão tomadas uma vez que dificilmente os países desenvolvidos se responsabilizarão por tal catástrofe.¹⁷⁴

Deste modo, a lógica da *Irresponsabilidade Organizada* mostra-se presente uma vez que nenhum Estado-nação tende a endereçar a responsabilidade para si, ao mesmo passo que o Estado-nação que experimentará tal catástrofe, não tem capacidade para contê-la por si só.

2.6. O POTENCIAL CATASTRÓFICO DA AGI E DA ASI.

A evolução tecnológica no campo da IA vem demonstrando a capacidade da humanidade em criar tecnologias e desenvolver sistemas altamente eficientes. Todavia, limitados em um escopo de tarefas. Tradutores automáticos, GPS inteligente, veículos autônomos, drones autônomos etc.

Por mais úteis que sejam, essas tecnologias representam formas estreitas de inteligência.

Mesmo assim, conforme explanado no decorrer deste trabalho, tais mecanismos já demonstraram extremo potencial na produção de riscos sociais,

¹⁷³ HARARI, op.cit, 2018. p. 63-64.

¹⁷⁴ HARARI, op.cit, 2018. p. 59.

financeiros e ecológicos. Entretanto, a próxima invenção da humanidade pode representar uma ameaça sem precedentes para a sobrevivência humana. A perda de controle de uma AGI pode resultar na aniquilação humana total.¹⁷⁵

Conforme anteriormente explanado, a AGI existe apenas no campo teórico, com grandes movimentações de diversos setores para sua criação.

Neste campo, há uma grande divergência teórica sobre o quão próxima a humanidade está de criar uma AGI. Para alguns teóricos, esta ruptura não acontecerá por muitas décadas, para outros a criação de uma AGI está muito mais próxima.

O fato é que a discussão dos teóricos permeia muito mais no campo temporal (quando teremos), do que no existencial (se teremos ou não tal tecnologia).

Desta forma, a possibilidade de existência de robôs ultra inteligentes, capazes de executar todo e qualquer tipo de tarefa humana é com certeza mais um vetor de riscos desta tecnologia.

A AGI, pode ser apenas um trampolim para a existência de uma ASI. Ou seja, o desenvolvimento de uma IA geral, que iguale a capacidade do cérebro humano, parece ser uma condição transitória para a existência de uma *Super Inteligência Artificial (ASI)*. Uma vez a AGI podendo aprimorar sua cognição de um modo contínuo, justamente por possuir maior capacidade de processamento que um cérebro humano, mais próximo ela se aproxima de se tornar uma ASI. Ou seja, a arquitetura de uma AGI permite que ela se auto desenvolva, ao mesmo passo que sua capacidade de processamento aumenta, mais rápido ela se desenvolve, este ciclo contínuo tende a melhorar sua forma de cognição de modo vertical, superando milhares de vezes a capacidade humana, tornando-se assim, uma ASI.¹⁷⁶

Para James Barrat, o ponto final do progresso da ciência no desenvolvimento da IA será a consecução de uma AGI. AGI uma vez equiparada a inteligência humana, poderá automaticamente se aprimorar,

¹⁷⁵ TORRES, Phil. The possibility and risks of artificial general intelligence. Bulletin of the Atomic Scientists, Chicago, v. 75, n. 3, p. 105-108, abr./2019.

¹⁷⁶ *Ibidem*.

tornando-se progressivamente mais inteligente que o ser humano. Barrat, acredita que essa evolução tende a ser exponencial e sua progressão consequentemente consubstanciaria a criação de uma *ASI*, deste modo o criador de da *ASI*, será a *AGI* no decorrer do seu processo de auto aprimoramento.¹⁷⁷

James Barrat preocupa-se com a difícil previsão de comportamento de uma *ASI*, do qual pode não ter uma boa relação de amizade com a humanidade, estando focada tão somente em cumprir os seus objetivos. A incapacidade em controlar tal tecnologia pode ser catastrófica, a *ASI* poderia assumir o controle de setores cruciais para a sobrevivência humana, como as redes de distribuição de eletricidade, combustível e redes sanitárias.¹⁷⁸ Deste modo, Barrat acredita que a ausência de controle de um *ASI*, pode significar em última instância, a destruição da humanidade.

Torres também endereça o risco de uma *AGI/ASI* para o problema de controle. Ele acredita haver diversas razões para crer que este é um problema grave, não se tratando de um risco apenas para os Estados-nação, e sim, para toda a humanidade.¹⁷⁹

Ele argumenta que há diversas tecnologias disponíveis do qual o descontrole pode acarretar consequências devastadoras, todavia, a *AGI/ASI* possui características diferentes das demais tecnologias. Sua inteligência excede a capacidade humana de resolução de problemas. Desta forma, ela estaria apta para perseguir todo e qualquer objetivo que lhe fosse imposto, transcendendo a capacidade de um desenvolvedor de contê-la. Ademais, a *IA* pode não entender que os efeitos deletérios por ela causados sejam considerados ruins e devam ser corrigidos. Uma vez que tais efeitos fazem parte do processo de cumprimento de seus objetivos. Desta forma, Torres conclui que o problema de controle fica claro, considerando os diversos aspectos filosóficos envolvidos.¹⁸⁰

¹⁷⁷ BARRAT, James. *Our Final Invention: Artificial Intelligence and The End of the Human Era*. 1. ed. Nova York: St. Martin's Griffin, 2013. p. 8.

¹⁷⁸ *Ibidem*. p. 15-16.

¹⁷⁹ TORRES, op.cit,2019. p. 105-108.

¹⁸⁰ *Ibidem*.

Torres elenca duas teorias para demonstrar a complexidade do problema do controle de uma AGI/ASI.

A primeira é a teoria da ortogonalidade, que parte de uma perspectiva otimista, nela acredita-se que uma enorme gama de objetivos finais pode ser combinada com uma ampla gama de inteligência, não havendo nenhum tipo de incoerência em uma máquina superinteligente se preocupar apenas em seguir o objetivo que foi programado. A máquina faria apenas o que lhe foi dito, não porque é limitada a isso, mas porque não vê motivo justificável para fazer o contrário.¹⁸¹

De outro giro, na teoria da convergência instrumental, a AGI/ASI provavelmente perseguirá seus objetivos incansavelmente, não importando os valores humanos, trabalhando em diversas tarefas intermediárias, como se auto aprimorar, por exemplo. Todavia, outra dessas tarefas intermediárias podem incluir adquirir recursos relacionados aos seus objetivos principais, como por exemplo, manter a integridade de seus sistemas impedindo que humanos os desliguem. Este é um caminho hipotético para uma AGI perigosa. O professor da Universidade de Oxford, Nick Bostrom, exemplificou tal questão, utilizando como exemplo uma IA projetada para fazer o maior número de cliques de papel possível.

Neste exemplo, objetivando produzir cliques de modo mais eficiente, a IA poderia concluir que a produção seria melhor se não houvessem seres humanos, uma vez que eles detêm o poder de desligá-la. Ademais, o corpo humano poderia ser utilizado como insumo para criação de novos cliques. Essas são algumas justificativas que fariam a IA destruir a humanidade com base no cumprimento de objetivos em que foi programada, não importando para ela, os meios que fossem tomados.¹⁸²

Deste modo, mesmo sem saber ao certo quando a AGI/ASI estará disponível, o fato é que esta tecnologia promete trazer diversos *goods* para a humanidade. Entretanto, como todo avanço tecnológico na *Sociedade de Risco*, tal tecnologia também trará consigo *bads*. Conforme defendido por

¹⁸¹ *Ibidem*.

¹⁸² *Ibidem*.

vários autores, os riscos catastróficos envolvidos nesta tecnologia podem acarretar em última instância, a destruição da espécie humana.

2.7. O REINADO DA INCERTEZA NA SOCIEDADE DE RISCO DIGITAL: OPACIDADE ALGORÍTMICA E BLACK BOX SOCIETY

No decorrer deste capítulo, foram elencados diversos riscos decorrentes do uso da Inteligência Artificial, alguns deles com potenciais catastróficos. Tais riscos por si só, já demandam ações regulatórias por parte dos estados. Todavia, o grande desafio está justamente na dificuldade de identificar e calcular tais riscos.

A tomada de decisão da IA opera por meio de uma lógica matemática, a qual é chamada de algoritmo. A forma com que dados são utilizados dentro desta lógica, dependem de suposições prescritas das quais carecem de transparência por alguns motivos a serem expostos.

A opacidade gerada por algoritmos além de ser um dos principais desafios regulatórios é também um elemento ensejador de incertezas científicas quanto aos riscos da tecnologia. O professor Frank Pasquale classifica esta opacidade como *black box* ou *caixa preta*.¹⁸³ Para ele, a metáfora pode ser vista de duas formas: (i) pode se referir a um sistema de monitoramento de dados em aviões, trens e carros; (ii) ou pode significar um sistema cujo o funcionamento é misterioso, onde podemos observar suas entradas e saídas, mas não sabemos dizer o que acontece durante o processo. Desta forma, ao falar de opacidade em sistemas de inteligência artificial, entende-se como a impossibilidade de explicar a tomada de decisão da IA.¹⁸⁴

O termo opacidade pode ser visto de três modos distintos:

¹⁸³ PASQUALE, Frank. The black box society: the secret algorithms that control money and information. 1. ed. Cambridge: Harvard University Press, 2015. p. 3.

¹⁸⁴ WYNSBERGHE, Aimee Van. Artificial intelligence: From ethics to policy. Panel for the Future of Science and Technology: EPRS | European Parliamentary Research Service, bruxelas, v. 1, n. 1, p. 7-8, jun./2020.

(i) O primeiro está ligado à complexidade da tecnologia, trazendo dificuldade para o utilizador sem conhecimentos técnicos ou científicos, entender seu funcionamento.

(ii) O segundo caso, está ligado à opacidade intencional gerada pelos próprios desenvolvedores ou proprietários da tecnologia, em alguns casos, utilizando-se de proteção jurídica relativa à proteção intelectual da tecnologia em questão, por questões concorrenciais, em outros, para proteger a eficácia da tecnologia. Nestes casos, abrir a caixa preta significaria inutilizar a tecnologia. Por exemplo, uma agência de inteligência que utiliza IA para determinar quais pessoas nos aeroportos devem ser alvo de buscas. Abrir a caixa preta desta tecnologia dando acesso aos critérios, inutilizaria a eficácia da tecnologia.¹⁸⁵

(iii) Por fim, o terceiro e último significado é o único verdadeiramente exclusivo da IA. A opacidade gerada pelos algoritmos muitas vezes faz com que as decisões tomadas, não consigam ser inteligíveis nem mesmo para os programadores que escreveram o código. Embora, seja possível saber as correlações estatísticas entre os *inputs*¹⁸⁶ e os *outputs*¹⁸⁷, atualmente não é possível saber quais os recursos que os algoritmos de Machine Learning usam para tomar suas decisões. Embora exista muito esforço na busca por clarificar esta questão, a ciência ainda tem pouco progresso neste campo. Até mesmo as grandes empresas do setor de tecnologia, utilizando engenharia reversa, conseguiram dar explicabilidade para uma decisão algorítmica de forma suficientemente clara.¹⁸⁸

Desta forma, fica claro que essas *black boxes* escancaram o campo de incertezas científicas que a IA habita.

As discussões endereçadas ao conflito jurídico entre o dever de transparência contraposto à garantia de sigilo industrial parecem ser um conflito pouco importante comparado a este último.

¹⁸⁵ *Ibidem*.

¹⁸⁶ Recurso de entrada de dados.

¹⁸⁷ Recurso de saída de dados (resultado).

¹⁸⁸ WYNSBERGHE, Aimee Van, op.cit., p.7-8.

Uma vez que, embora haja uma discussão sobre quando e em quais circunstâncias deve ser autorizado a abertura do código-fonte (rompendo a garantia do sigilo industrial) para o cumprimento de um dever de transparência, esta discussão perde o sentido se a abertura do código-fonte não é suficiente para compreender e prever o comportamento do algoritmo de IA.¹⁸⁹

Mesmo com acesso ao código-fonte, sua análise se torna inadequada, uma vez que este apenas demonstra o método que a máquina utiliza para aprender, ou seja, expõe-se apenas as regras de decisão, não a sua forma. Tornando difícil a compreensão de seu processo decisório.

Deste modo, o código-fonte deve ser visto apenas como uma peça do processo de transparência.¹⁹⁰

Portanto, enquanto o ser humano não for capaz de interpretar a lógica das decisões tomadas por algoritmos, ele deve ser visto como um campo de incertezas. Ferrari, Becker e Wolkart, acreditam que essa compreensão está fora da capacidade humana. Tendo em vista que essa lógica matemática emprega centenas ou até milhares de regras, suas previsões vêm acompanhadas de formas complexas que demandam uma capacidade de processamento muito além da capacidade humana, além da multiplicidade de variáveis que elas operam. Deste modo, o entendimento da tomada de decisão algorítmica requer uma capacidade de processamento que o cérebro humano não parece ter.¹⁹¹

Diante da complexa estrutura dos algoritmos de IA, a mera abertura do código-fonte por si só, é inútil para trazer transparência para a decisão automatizada. Na mesma direção, a incapacidade do ser humano em articular a lógica dessas decisões, coloca a tecnologia em questão, em um reino de incertezas.

¹⁸⁹ FERRARI, Isabela; BECKER, Daniel; WOLKART, Erik. ARBITRIUM EX MACHINA: PANORAMA, RISCOS E A NECESSIDADE DE REGULAÇÃO DAS DECISÕES INFORMADAS POR ALGORITMOS. Revista dos Tribunais: subtítulo da revista, São Paulo, v. 995, n. 1, p. 8, set./2018.

¹⁹⁰ *Ibidem*.

¹⁹¹ *Ibidem*.

2.8. CONCLUSÃO

Uma vez apresentado o modelo teórico da Sociedade de Risco, aplicado à Quarta Revolução Industrial, foi demonstrado que tecnologias como a Inteligência Artificial, tendem a impulsionar a produção de riscos e incertezas em uma *Sociedade de Risco Digital*.

O segundo capítulo, prestou-se a demonstrar como essa lógica se opera especificamente quando se fala de IA. Deste modo, primeiramente conceitua-se tal tecnologia, trazendo entendimentos do ordenamento jurídico europeu e brasileiro, elencando suas espécies e demonstrando suas principais funcionalidades e benesses na sociedade. Noutro giro, como todo avanço tecnológico na *Sociedade de Risco* vem acompanhando da distribuição de *bads* (riscos), fez-se necessário elencar os principais riscos da IA, sejam eles: (i) Riscos discriminatórios à justiça criminal; (ii) Riscos ao acesso à saúde; (iii) Riscos discriminatórios ao acesso ao crédito; (iv) Riscos discriminatórios na contratação de recursos humanos; (v) Riscos à democracia; (vi) Riscos ao futuro do trabalho; (vii) Riscos catastróficos de um sistema de IA forte (AGI).

Todavia, riscos só são riscos quando podem ser previstos, suscetíveis a mensurações, prognoses e cálculos estatísticos. O último subcapítulo, prestou-se a demonstrar que tais riscos se tornam incertezas, uma vez que o ser humano não consegue compreendê-los, controlá-los e gerenciá-los.

A opacidade algorítmica, traz grandes desafios para a compreensão da tomada decisória de IA. Essa incompreensão consequentemente mina a capacidade de mensurar riscos, tornando um campo científico incerto.

Sob condições de incerteza científica, para a qual as avaliações de riscos se tornam inaplicáveis, surge o princípio da precaução como um meio para construir entendimentos mais robustos de ação.¹⁹² No campo da IA, tais ações aparentam ser necessárias na busca por transparência e explicabilidade.

¹⁹² STIRLING, Andy. Precaution in the Governance of Technology. SPRU Working Paper Series, Brighton, v. 1, n. 1, p. 4, jan. /2016.

Deste modo, o último capítulo deste trabalho buscará responder os seguintes questionamentos: (i) O princípio da precaução deve ser aplicado na regulação da IA? (ii) Sua aplicação teria um efeito paralisante na inovação. (iii) As propostas regulatórias atuais contemplam tal princípio?

3.0. O PRINCÍPIO DA PRECAUÇÃO: UM INSTRUMENTO PARA A REGULAÇÃO DE RISCOS NA INTELIGÊNCIA ARTIFICIAL?

3.1. REGULAÇÃO DE RISCOS

Conforme exposto no primeiro capítulo, vivemos em uma era da Sociedade de Risco, como consequência as instituições criadas pela sociedade contemporânea não são capazes de calcular riscos com exatidão.

A regulação de riscos é vista inerentemente como um mecanismo de controle de riscos, em todas as esferas quais existam atividades que tragam externalidades negativas.¹⁹³ Sendo o risco, definido geralmente como a probabilidade de um evento, ou seja, trata-se de um evento calculável, diferentemente do conceito de incerteza, que é inerentemente impossível de ser mensurado.¹⁹⁴

Para Baldwin a regulação de riscos está centralmente preocupada com o risco e a incerteza. Em seu entendimento, por mais que o risco seja considerado um dos principais temas unificadores que moldam as ciências sociais, as sociedades modernas tendem a experimentar uma experiência sem precedentes no que tange avaliar a presença de riscos. Baldwin ao analisar a teoria de Beck, entende que dentro da lógica trazida por este, há um inerente descompasso entre métodos de avaliação de riscos, em suas palavras:

More generally, though, Beck argues that contemporary societies' attempts to anticipate risks are inherently futile, as these risks cannot be calculated. Indeed, according to this view, risk assessment exercises inherently underestimate real threats, as they rely on methodologies that legitimise individuals' exposures to incalculable risks.¹⁹⁵

¹⁹³ Externalidade negativa é um termo utilizado para descrever efeitos colaterais negativos que derivam de determinada atividade econômica.

¹⁹⁴ BALDWIN et al. *Understanding Regulation: Theory, Strategy, and Practice*. 2. ed.: Oxford University Press, 2013. p. 83-84.

¹⁹⁵ *Ibidem*.

Deste modo, surge um extenso debate quando se fala de estratégias regulatórias para lidar com atividades que trazem riscos. De um lado, surgem estudiosos que defendem medidas antecipatórias ao risco catastrófico, sugerindo a aplicação do princípio da precaução na regulação de risco. De outro lado, estudiosos defendem uma abordagem de *cost-benefit analysis*,¹⁹⁶ ignorando as incertezas científicas de determinada atividade, operando dentro de um contexto de racionalidade limitada e vieses bem documentados na tomada de decisão.¹⁹⁷

Estes debates se estendem em diversos campos, tais como, reatores nucleares, alimentos geneticamente modificados, nanotecnologia e impactos ambientais, terrorismo. E conforme demonstrado nos primeiros capítulos, a IA surge como mais um campo de discussão.

O professor Robert Baldwin destaca as principais diferenças entre estas duas abordagens regulatórias: (i) A abordagem, *cost-benefit analysis*, parte de uma ideia de maximização do bem estar social. Para Baldwin, trata-se de uma abordagem técnica, que busca antecipar danos valendo-se de frequência. Esta abordagem está associada à visão de que riscos podem ser tomados com base em evidências objetivas. (ii) A abordagem precaucionária, por sua vez, traz de modo pragmático uma visão de aversão ao risco, seguindo a premissa *better safe than sorry*¹⁹⁸, justamente por não acreditar na suscetibilidade dos métodos quantitativos e qualitativos para calcular e identificar todo e qualquer risco. Trata-se de uma estratégia de antecipação.¹⁹⁹

Conforme demonstrado no decorrer deste trabalho, os riscos e incertezas trazidos pelo desenvolvimento da IA, podem demandar de uma abordagem mais rígida, uma abordagem precaucionária. Tendo em vista que alguns eventos decorrentes do seu uso, podem acarretar em danos irreversíveis. Ademais, a opacidade gerada por algoritmos de inteligência artificial, põe em xeque a tomada de decisão baseada em evidências objetivas.

¹⁹⁶ Tradução: Análise de custos e benefícios.

¹⁹⁷ BALDWIN et al.op.cit, 2013. p. 96.

¹⁹⁸ Ditado da língua inglesa, análogo ao ditado brasileiro “melhor prevenir do que remediar”.

¹⁹⁹ *Ibidem*, p. 88.

Deste modo, entendendo a abordagem precaucionária como uma possibilidade de regulação de riscos da IA, faz-se necessário abordar tal princípio de modo detalhado, conceituando-o e demonstrando os debates em seu entorno.

3.1.1. PRINCÍPIO DA PRECAUÇÃO: CONCEITO, ORIGEM E DESENVOLVIMENTO

O princípio da precaução é invocado sempre que inexistir conhecimentos dos riscos potenciais de danos de uma determinada atividade, ou de determinado produto. Os riscos são incertos e não totalmente claros quando mensurados pela ciência. Trata-se então, da administração da incerteza.²⁰⁰

O princípio surge primariamente com o intuito de ressignificar os institutos a responsabilização clássica, invertendo sua lógica. De modo a criar mecanismos *ex-ante*²⁰¹, já que a mera indenização sob um evento catastrófico se torna inócuo.²⁰²

Moraes ao citar François Ewald elenca os três períodos que alteraram a lógica da segurança: (i) No século dezenove surge o conceito de responsabilidade, no qual aquele que causar dano, tem o dever de reparar. (ii) Um século depois, surge o conceito da solidariedade, trazendo políticas de bem estar social preventivas, como o seguro social e prevenção contra doenças. Nota-se o surgimento de uma lógica preventiva. (iii) Por fim, no último século (XXI) ao reconhecer o reaparecimento das incertezas por conta do desenvolvimento tecnológico, surge o paradigma da precaução.²⁰³

A abordagem precaucionária, decorrente da aplicação do princípio da precaução, não deve ser confundida com a preventiva, decorrente da aplicação

²⁰⁰ BERGUER FILHO, Airton Guilherme. Regulação e Governança dos Riscos das Nanotecnologias. 1. ed. Belo Horizonte: Arraes, 2018. p. 73.

²⁰¹ Ex-ante é um termo do Latim. Ele quer dizer “antes do fato”.

²⁰² MORAES, op.cit,2011. p. 89.

²⁰³ *Ibidem*.

do princípio da prevenção. Por mais que o princípio da precaução demande ações preventivas, não se tratam do mesmo princípio.

Tão como precaução, o princípio da prevenção também visa evitar danos antes de sua concretização, com mecanismos regulatórios *ex-ante*. Todavia, opera de modo diverso, uma vez que sua lógica visa prever riscos comprovados. Por isso, de acordo com Aragão, os princípios se diferenciam no ponto em que o primeiro é um princípio reativo e o segundo proativo.²⁰⁴

Moraes completa o raciocínio de Aragão ao argumentar que o que difere o princípio da precaução é a exigência de cautela, mesmo inexistindo relações científicas de causalidade de correlação, uma vez que reconhecida a vulnerabilidade da ciência em calcular riscos.²⁰⁵

Para alguns autores, o princípio da precaução teve sua primeira manifestação na década de 60, na Suécia, quando incorporada em sua *Lei de Proteção Ambiental*, editada em 1969.²⁰⁶ Para outros autores, o surgimento do princípio, deu-se na Alemanha década de 70, sendo conhecido como *Vorsorgeprinzip*.²⁰⁷

O *Vorsorgeprinzip* surgiu por decorrência de chuvas ácidas em áreas florestais, do quais inexistiam, naquela altura, evidências científicas conclusivas sobre suas causas e consequências.²⁰⁸

Nota-se, que suas primeiras manifestações aplicavam medidas preventivas apenas no campo doméstico. Ou seja, tais aplicações produziam efeitos apenas para o Estado-nação que inseriu o princípio em seu ordenamento jurídico.

Posteriormente, o princípio começa a ganhar espaço no campo internacional. Em 1982, a World Charter of Nature reconheceu a importância da precaução no direito e na política ambiental:

²⁰⁴ ARAGÃO, Alexandra. Aplicação nacional do princípio da precaução. Associação dos Magistrados da Jurisdição Administrativa e Fiscal de Portugal, Lisboa, v. 1, n. 1, p. 5, jan./2013.

²⁰⁵ MORAES, op.cit.,2011. p. 90

²⁰⁶ DIONÍSIO, P. D. H. Princípio da precaução e contradições na regulação do risco: uma análise comparada entre Estados Unidos e Europa. Regulação 4.0, São Paulo, v. 1, n. 1, p. 31, jan./2019. Edição do Kindle. p. 31%.

²⁰⁷ BERGUER, op.cit., 2018. p. 76.

²⁰⁸ DIONÍSIO, op.cit., p. 31%

Activities which are likely to pose a significant risk to nature shall be preceded by an exhaustive examination; their proponents shall demonstrate that expected benefits outweigh potential damage to nature, and where potential adverse effects are not fully understood, the activities should not proceed.²⁰⁹

Cinco anos depois, em 1987, a precaução foi adotada como um princípio autônomo, na Segunda Conferência Internacional sobre a proteção do Mar do Norte. O documento trazia medidas restritivas para a atividades econômicas no que tange a emissão de poluentes, mesmo que ausente provas que demonstram o nexo de causalidade entre o dano e o agente causador.²¹⁰

Por fim, o maior reconhecimento do princípio em âmbito internacional, ocorreu em 1992, no Rio de Janeiro, na Conferência das Nações Unidas sobre o Meio Ambiente e Desenvolvimento, chamada de *Rio 92*. Em seu princípio de número 15, a declaração colocou o princípio do campo transnacional, inspirando instrumentos internacionais posteriores:

Princípio 15: Com a finalidade de proteger o meio ambiente, os Estados deverão aplicar amplamente o critério de precaução conforme suas capacidades. Quando houver perigo de dano grave ou irreversível, a falta de certeza científica absoluta não deverá ser utilizada como razão para que seja adiada a adoção de medidas eficazes em função dos custos para impedir a degradação ambiental.²¹¹

Outro importante conceito do princípio é o da declaração de Wingspread, em 1998, ocasião qual uniu cientistas, advogados e ativistas ambientais em Wisconsin nos EUA, o texto trazia:

When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically. In this context the proponent of an activity, rather than the public, should bear the burden of proof. The process of applying the precautionary principle must be open, informed and democratic and must include potentially affected parties. It must also involve an examination of the full range of alternatives, including no action.²¹²

²⁰⁹ ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Assembleia Geral. A/RES/37/7. 28 de outubro de 1982.

²¹⁰ DIONÍSIO, op.cit.,2019, p. 31%.

²¹¹ Declaração do Rio Sobre Meio Ambiente e Desenvolvimento, jun. 1992.

²¹² Wingspread Statement on the Precautionary Principle, Jan. 1998.

O texto amplia o conceito trazido pela declaração Rio 92, vez que amplia seu espectro para atividades que possam causar danos à saúde humana, não só ao meio ambiente. Além de trazer a inversão do ônus da prova para relações causais onde os danos não sejam mensuráveis, impondo de modo objetivo o dever de demonstrar que tais atividades são seguras ou trazem riscos aceitáveis. Outro ponto que merece destaque, é que nessa versão, o princípio se aplica para ameaças em geral, não se limitando a danos irreversíveis e irreparáveis, prevendo um dever positivo de impor medidas cautelares. Diferentemente da aplicação da declaração Rio 92, que visa buscar medidas postergatórias ante a incerteza. Por fim, ressalta-se que a declaração também busca trazer o processo decisório para o campo democrático.²¹³

Deste modo, nota-se que o princípio nasce como um princípio do Direito Ambiental, iniciando sua aplicação em legislações domésticas para posteriormente ganhar espaço no Direito Internacional. Seguidamente, o princípio passa a se erradicar para outros campos que não só o direito ambiental. Isto fica claro com um comunicado da União Europeia, que amplia a aplicabilidade do princípio e define seu modo de aplicação.

O recurso ao princípio da precaução pressupõe: – A identificação de efeitos potencialmente nocivos decorrentes de um fenómeno, de um produto ou de um processo, – Uma avaliação científica dos riscos que, devido à insuficiência dos dados, ao seu carácter inconclusivo ou ainda à sua imprecisão, não permitem determinar com suficiente certeza o risco em questão.²¹⁴

Ao analisar tais manifestações, nota-se que não há uma definição objetiva do princípio, na verdade elenca-se hipóteses para sua aplicação, sendo sempre como imperativo, a incerteza científica.

Todavia, as diferentes manifestações quanto ao princípio da precaução trazidas neste subcapítulo, demonstram uma escala de força entre suas versões, sendo elas *fraca*, *moderada* e *forte*:

²¹³ BERGUER, op.cit., 2018. p. 78.

²¹⁴ Comunicação da Comissão das Comunidades Europeias relativa ao princípio da precaução, Bruxelas, 2/02/2000.

(i) Versão fraca: A declaração do Rio, seria uma versão fraca da aplicação do princípio, uma vez que se limita a danos graves e irreversíveis.²¹⁵ Para Garnett, a versão fraca do princípio da precaução é caracterizada por limiar epistêmico relativamente alto e uma preferência pela gestão de riscos. A incerteza ante as consequências de determinada atividade, pode justificar a regulação, se houver motivos plausíveis para crer que tal atividade é prejudicial. Deste modo, deve haver certa ênfase na coleta de dados que justifiquem tal crença.²¹⁶

(ii) Versão moderada: A Comunicação Europeia relativa ao princípio da precaução seria uma versão moderada. Primeiro, porque diferentemente da declaração anterior, a incerteza na avaliação de risco justifica regulação proporcional ao nível e risco envolvidos. Na aplicação fraca, é permitida a regulação havendo incertezas científicas. Na moderada, não só há a permissão como também a obrigação de ações que visem reduzir as incertezas, estabelecendo controles que possam trazer maiores margens de segurança, construindo uma extensa investigação quanto a relação causa e efeito decorrente da incerteza, definindo também, quais riscos são considerados aceitáveis. Possivelmente relaxa-se a regulação uma vez identificado a atividade como comprovadamente segura. Em suma, a versão moderada mostra-se mais proativa na busca pela identificação dos riscos.²¹⁷

(iii) Versão Forte: A Declaração de Wingspread se mostra mais forte que a Rio 92 e a Comunicação Europeia, sendo classificada como uma versão forte. Primeiro, porque como já mencionado, ela não se limita às ameaças de riscos graves e irreversíveis. Ademais, ela propõe a inversão do ônus da prova ao proponente da atividade econômica, fazendo com que este seja obrigado a clarificar as incertezas trazidas pela sua atividade. Pode-se dizer que o princípio da precaução ao inverter o ônus da prova, traz um dever de *accountability*²¹⁸

²¹⁵ SUNSTEIN, Cass R. Para além do princípio da precaução: Beyond the Precautionary Principle. Revista de Direito Administrativo, Rio de Janeiro, v. 1, n. 1, p. 23, abr./2012.

²¹⁶ GARNETT, J. PARSONS, K. Multi-Case Review of the Application of the Precautionary Principle in European Union Law and Case Law. Risk Analysis, Cranfield, v. 1, n. 1, p. 4, mai./2016.

²¹⁷ *Ibidem*.

²¹⁸ Accountability é um termo em inglês que significa “prestação de contas”.

para o proponente da atividade.²¹⁹ Para Garnett, a aplicação forte tende a ter um nível epistêmico em um limiar mais baixo e tende a partir de uma lógica de prevenção de riscos e não gestão. A incerteza sobre determinada atividade, por si só pode ser ensejar em ações rigorosas, como a inversão do ônus da prova até a proibição da atividade, mesmo havendo poucas evidências para crer que a conduta é prejudicial.²²⁰

Ao analisar as algumas versões do princípio da precaução, nota-se que para sua aplicação, faz-se necessário o cumprimento de alguns pressupostos.

3.2. PRESSUPOSTOS PARA APLICAÇÃO DO PRINCÍPIO

De modo geral, a aplicação do princípio da precaução é pautada em dois pressupostos: a existência de riscos graves e a existência de incertezas significativas quanto aos riscos.²²¹

3.2.1. GRAVIDADE DOS RISCOS

A gravidade do risco é analisada não só por critérios científicos, como também pela aceitabilidade social dos riscos. Deste modo, mensura-se o risco por uma perspectiva objetiva, analisando fatores quantitativos, tão como, por uma perspectiva subjetiva, analisando fatores qualitativos.²²²

3.2.2. GRAVIDADE OBJETIVA

Os critérios qualificadores da gravidade objetiva são:

²¹⁹ SUNSTEIN, op.cit,2012. p. 23.

²²⁰ GARNETT; PARSONS, op.cit,2016. p.4.

²²¹ ARAGÃO, op.cit,2013., p. 6-7.

²²² *Ibidem*.

- a) *Irreversibilidade*: O estatuto mais forte para aplicação do princípio da precaução, é justificado quando um dano não pode ser revertido.
- b) *Alargamento*: Alargamento se entende como a capacidade de disseminação dos riscos de modo transterritorial.
- c) *Cumulatividade*: Soma de riscos já existentes.
- d) *Involuntariedade*: Riscos que correm sem possibilidade de escolha e de ciência das vítimas.
- e) *Injustiça na distribuição de riscos*: Quem suporta os riscos não se beneficia das vantagens.
- f) *Potencialização*: Risco pode provocar danos em cadeia.²²³

Ao analisar o segundo capítulo deste trabalho, juntamente aos conceitos trazidos para qualificar o que seja gravidade objetiva, percebe-se que a IA atende praticamente todas as hipóteses de gravidade objetiva.

Irreversibilidade mostra-se presente ao falar da possibilidade de existência de uma AGI, a dificuldade de previsão de seu comportamento, traz-se em última instância, riscos à existência humana. Ademais, a IA mesmo que em sua versão restrita (ANI) demonstra extremo potencial na criação de riscos sociais, ecológicos e financeiros de modo irreversível.

Entender que é insuficiente que apenas um Estado-nação regule tal tecnologia, justamente porque seus riscos tendem a ser transfronteiriços, justificam o conceito de alargamento.

²²³ *Ibidem*.

A cumulatividade é justificada ao falar de discriminação algorítmica, uma vez entendido que os mecanismos de tomada de decisão automática pela IA tendem a reforçar vieses discriminatórios.

A falta de transparência e aplicabilidade das decisões, frutos da opacidade algorítmica, trazem dificuldade na demonstração de transparência, deste modo, justificam o conceito de Involuntariedade.

Injustiça na distribuição de riscos é claramente justificada ao evidenciar os riscos ao futuro do trabalho com o desenvolvimento da I.A.

Por fim, a potencialização se justifica em todo o segundo capítulo, uma vez que a interconexão da IA em todas as esferas da vida, tende a criar riscos em cadeia.

3.2.3. GRAVIDADE SUBJETIVA

A gravidade subjetiva é endereçada à percepção social do risco, ou seja, a gravidade passa pela aceitação social de determinada conduta danosa. São os riscos socialmente elevados.

Os riscos objetivos são entendidos como os riscos científicos, na medida que sua gravidade é mensurável. Já os riscos subjetivos, são os riscos que, mesmo que justificados como de baixo risco, são motivos de alarme social. Deste modo, a gravidade subjetiva parte de aspectos psicossociais, de modo que, quanto maior a incerteza, maior o sentimento de insegurança e pânico social. Logo, a construção social do risco faz-se importante para a legitimação social destes.²²⁴

3.2.4. INCERTEZAS CIENTÍFICAS

A gravidade, seja ela objetiva ou subjetiva tem sua aplicabilidade minimizada ou maximizada a depender da versão do princípio. Em aplicações

²²⁴ *Ibidem*, p.8-9.

fortes, como o tratado de Wisconsin, a irreversibilidade não é um pré-requisito para a aplicação de medidas cautelares. Todavia, em todas as versões do princípio o pressuposto da incerteza científica, se faz presente de modo inalterável. Mesmo porque, o campo de domínio da precaução é justamente o campo da incerteza, é exatamente este *modus operandi* na regulação de riscos que o diferencia do princípio da prevenção.

Desta forma, a precaução não espera a existência de provas irrefutáveis quanto aos riscos de determinada conduta, justamente porque o resultado da espera pode ser catastrófico.

3.2.5. CONTEÚDO DA INCERTEZA

A incerteza pode ter seu conteúdo endereçado em diversos espectros. O primeiro deles é a incerteza quanto a origem do dano, sabe-se da existência do dano, todavia não se sabe de onde vem, ou se a causa do dano é hipotética, mas o nexo causal não é claro. Deste modo, a incerteza que se permeia conduz ao questionamento: *qual o risco envolvido?* A segunda hipótese, se trata da natureza ou gravidade do dano. Sabe-se da capacidade de determinada atividade ensejar em um evento danoso, mas não se consegue quantificar nem determinar sua gravidade, conduzindo a pergunta: *risco de quê?* Em terceiro, a incerteza pode surgir na própria verificação dos danos, quando não há sequer um dano confirmado, apenas suspeitas. Conduzindo ao último questionamento *existe risco?*²²⁵

Neste último caso, para que o princípio não se justifique como um princípio do medo, faz-se necessária a demonstração de verossimilhança da possibilidade de existência de danos. O risco verossímil habita entre o que é menos do que provável, porém mais do que remoto. Esta hipótese parece encaixar apenas em versões fortes do princípio.

²²⁵ *Ibidem*, p.10.

3.2.6. FONTES DE INCERTEZA

Basicamente, a doutrina classifica três fontes de incerteza: a epistemológica, ontológica e hermenêutica.

A primeira, deriva da complexidade dos sistemas que envolvem determinada atividade. Escala, dinamismo e aleatoriedade desses sistemas são fontes de incertezas ontológicas. Para fins deste trabalho, ao falar de incertezas ontológicas produzidas pela IA, podemos entender *escala* como a capacidade do evento danoso atingir diferentes esferas. Desde danos que atingem meramente o indivíduo, até danos em escala global.

A *aleatoriedade* se mostra presente quando se demonstra incapacidade em prever uma decisão tomada pela IA. Na mesma esteira, ao falar de decisões automatizadas por máquinas de auto aprendizado (machine learning), o dinamismo resta presente uma vez que estes agentes evoluem sob influência de diversos fatores.²²⁶

Já a incerteza epistemológica está ligada a inexistência, inadequação, incompletude e até a existência demasiada de dados.²²⁷ Outro ponto presente no desenvolvimento da IA, tendo em vista a incompletude apresentada por algoritmos opacos. Conforme explanado no capítulo 2.6 deste trabalho, é possível interpretar a entrada e a saída de determinado comando algorítmico, todavia, não se consegue explicar os meios pelos quais a decisão foi tomada.²²⁸

Por fim, a incerteza hermenêutica decorre de interpretações contraditórias quanto à importância do risco. O excesso ou escassez de dados traz diferentes interpretações da realidade e consequentemente diferentes visões quanto à dimensão e gravidade do risco.²²⁹ A consequência desta incerteza é a criação de ambiguidades decorrentes de certezas contraditórias. Neste caso a caracterização do resultado que se torna problemática e não as probabilidades.²³⁰

²²⁶ *Ibidem*, p. 12.

²²⁷ *Ibidem*.

²²⁸ CAPLAN, R. et al. Algorithmic Accountability: A Primer. Data&Society, New York, v. 1, n. 1, p. 6, abr./2018.

²²⁹ *Ibidem*.

²³⁰ STIRLING, op.cit., 2016. p. 7.

3.3. CRÍTICAS AO PRINCÍPIO

Mesmo que o princípio tenha pressupostos para sua aplicação, não podendo ser invocado de maneira indiscriminada, a ausência de definições objetivas do que seja *risco*, *irreversível*, *sério*, faz com que críticos o classifiquem como um princípio mal definido.²³¹ Ante essa ausência de critérios, entender o princípio da precaução como uma mera orientação regulatória, endereça diversas críticas de estudiosos a sua aplicação.

O principal expoente da crítica ao princípio da precaução é o professor Cass Sunstein. Sua publicação *Beyond the Precautionary Principle*, traz diversas críticas ao princípio, classificando-o como o *Princípio da Paralisia*.

Sunstein, não acredita que o princípio nos leva a direções ruins, não sendo um bom instrumento regulatório. Na verdade, ele acredita que a versão forte do princípio não leva a lugar nenhum, proibindo tanto a regulação quanto a inação e qualquer medida entre estes dois extremos. Para o professor, o princípio só parece fazer sentido graças ao que ele chama de *heurística da disponibilidade*, o termo serve para justificar que a humanidade utiliza mecanismos de percepção de riscos de forma estreita, sendo indiferente aos seus impactos sistêmicos, logo, negligencia-se alguns dos riscos que efetivamente estão presentes.²³²

Para Sunstein, a versão *forte* do princípio se torna paralisante por não oferecer qualquer tipo de orientação, tendo em vista que sua operacionalidade colocaria em confronto os riscos decorrentes de uma ação, como os riscos pela ausência desta mesma ação. Deste modo, a aplicação do princípio se contradiz por essência, proibindo todos os cursos possíveis de ação e inação de determinada atividade. Portanto, a norma torna-se conceitualmente não operacional, uma vez que sua interpretação demandaria a regulação e não-regulação concomitantemente.

²³¹ *Ibidem*, p.6.

²³² SUNSTEIN, op.cit., 2012. p. 14.

Um exemplo utilizado pelo professor é o conflito entre a proteção de mamíferos marinhos e manobras de treinamento militar. Tais manobras podem causar danos significativos ao ecossistema desses animais, todavia, interromper tais atividades tende a colocar em risco a preparação militar e consequentemente a soberania do Estado. Deste modo, a regulação violaria o princípio da precaução porque daria origem ao que ele classifica como *riscos substitutos*, vez que o risco de regular substitui o risco da inação. Ademais, no caso em questão, a aplicação forte do princípio sugeriria a inversão do ônus da prova para o proponente da atividade ou processo em questão, impondo um ônus impossível de ser superado.²³³

Outro exemplo seria o problema *drug lag*²³⁴, trazidos pela adoção de um alto grau de precaução na utilização de novos medicamentos e drogas no mercado. A consequência desta abordagem rígida tende a proteger pessoas de danos que possam ser provocados por medicamentos testados adequadamente, em contraponto, as impedem de receber os benefícios de um tratamento inovador. Neste sentido, o professor questiona se a precaução se operacionalizará permitindo a pré-comercialização extensiva, visando salvar vidas ou fazer o oposto (pelo mesmo motivo)?

O caso de modificação genética de alimentos, traz o mesmo conflito. Diversos atores acreditam que esta tecnologia pode colocar em risco a saúde humana. Todavia, tal tecnologia tende a aumentar a produtividade agrícola e promete melhorar o alimento em si, tornando-o mais saudável. O aumento da produtividade agrícola junto à capacidade de produção de alimentos mais saudáveis, tende a beneficiar países em desenvolvimento, podendo na verdade salvar vidas.²³⁵

Deste modo, o alto rigor regulatório paralisaria certas inovações privando a sociedade de benefícios significativos e, portanto, causando sérios danos que

²³³ *Ibidem*, p. 33-34.

²³⁴ Termo em inglês que significa “atraso de laboratório”, diz respeito ao atraso na evolução de pesquisas laboratoriais por excesso de precaução.

²³⁵ *Ibidem*.

de outra forma não ocorreriam. Deste modo, a aplicação do princípio ao eliminar os *benefícios de oportunidade*, violaria em si, a sua própria causa.²³⁶

Outro ponto criticado pelo professor é relativo aos custos regulatórios da aplicação *forte* do princípio. Para ele, uma regulação excessivamente cara tende a criar mais riscos à saúde humana, porque além de retirar os *benefícios de oportunidade* e os aumentar os *riscos substitutos*, ele traz despesas significativas na redução de riscos de atividades que envolvem um pequeno risco de dano. Desta forma, ele sugere que se deve evitar gastar muito dinheiro para reduzir riscos, justamente porque esse gasto também acarreta novos riscos. Essa contradição do princípio que o torna paralisante, uma vez que se põe como uma trava tanto para ação, como para inação.²³⁷

Deste modo o princípio, também se torna irracional, tendo em vista trazer medidas de prevenção cientificamente injustificadas que tendem a trazer um desequilíbrio orçamental, já que as medidas em si ignoram e se tornam extremamente onerosas do ponto de vista financeiro.

O foco da crítica de Sunstein, dá-se na versão forte do princípio. Para ele, a versão fraca é classificada como razoável e inafastável. Entretanto, sugere-se uma calibragem ao princípio, sendo sua principal tarefa encontrar maneiras de fazer coincidir a extensão da prova com a extensão da resposta. Provas fracas sustentariam a exigência de mais estudos quanto aos riscos, ao passo que provas fortes podem justificar a divulgação pública dos riscos e até a adoção de controles regulatórios. Este tipo de calibragem seria um caminho importante caminho para dar eficácia a aplicação do princípio em sua versão *fraca*.²³⁸

Sunstein acredita também que o princípio pode ser usado de forma irracional para defender interesses político-econômicos, ele utiliza como exemplo, a ideia de fazendeiros europeus invocarem o princípio da precaução para proibir ou dificultar o uso de alimentos geneticamente modificados. De

²³⁶ SUNSTEIN, R. W. H.C. R. The Precautionary Principle as a Basis for Decision Making. The Economists' Voice, Chicago, v. 2, n. 2, p. 6, mai. /2005.

²³⁷ SUNSTEIN, op.cit., 2012. p. 28.

²³⁸ *Ibidem*, p.18.

modo a sufocar competidores americanos, que são muito mais propensos a confiar em plantações geneticamente modificadas.²³⁹

Deste modo, ele demonstra uma crença demasiada no conhecimento científico neutro, presumindo que apenas a ciência poderia tentar estimar riscos e benefícios de determinada atividade e que o princípio tende a ser utilizado de forma política e não dentro de um contexto de racionalidade científica.

Outro argumento frequentemente apresentado para desqualificar o princípio é aquele que entende que o princípio busca estruturalmente a exclusão de todo e qualquer risco, seria a busca pelo *risco zero*. Tal busca tende a privilegiar a segurança de forma desproporcional em detrimento da liberdade.²⁴⁰

3.4. DEFESA AO PRINCÍPIO

Em contraponto às severas críticas ao princípio, diversos autores defendem sua aplicação, contrapondo-as.

Para Berguer, as acusações de anti cientificidade do princípio são desarrazoadas. Na verdade, ele entende justamente o contrário, visto que o princípio busca permitir a introdução da ciência no âmbito da decisão da esfera pública. Para ele, a precaução não deve ser vista como uma promessa de segurança absoluta, na verdade a precaução se trata da democratização da ciência, substituindo o monopólio tecnocrático. O princípio da precaução permite o benefício da dúvida, exigindo transparência e impondo cautela à soberania da ciência.²⁴¹

O ponto trazido por Berguer é importante pois justifica o princípio como um pilar norteador de uma Sociedade de Risco.

Aragão segue a mesma linha de entendimento, enxergando a ciência como um pressuposto da precaução. Deste modo, a precaução surge para

²³⁹ *Ibidem*, p. 8.

²⁴⁰ BERGUER, op.cit., 2018. p. 73.

²⁴¹ *Ibidem*, p. 91.

orientar o decisor quando os estudos, testes e análises não forem suficientes para afastar todas as dúvidas quanto aos riscos de determinada atividade.²⁴²

Deste modo o princípio pressupõe uma avaliação científica e racional, antes de qualquer medida. Ademais, as medidas precaucionais devem ser vistas sempre como provisórias, conforme o avanço científico evolui e passa a clarificar as incertezas, o princípio tende a ser revogado ou mitigado.

Desta forma, Aragão conclui que medidas precaucionais são tomadas em razão da ciência e não à revelia desta.²⁴³

Para Stirling, a maior deturpação em torno do princípio da precaução é pressupor que a regulação só é científica quando opera sob um conjunto restrito de técnicas de análise probabilística do *risco*. Na verdade, a precaução tira a ciência desta versão restrita de análise, demonstrando que muitas vezes é necessário sair do habitual em termos de *avaliação de risco*.

Stirling conclui seu raciocínio ao dizer que a precaução não é uma causa de incerteza é uma resposta para ela. Deste modo, irracional é acreditar que a natureza da incerteza pode ser reduzida meramente a cálculos probabilísticos.

Os argumentos de Stirling e de Aragão são coerentes, até porque, a precaução se posiciona como forma de resolução das incertezas, justamente quando cálculos probabilísticos se demonstram incapazes de quantificar e qualificar os riscos de determinada atividade. Não se trata de abrir mão da ciência, trata-se de guiar a ciência na busca da resolução de incertezas.

A ideia da utilização do princípio de forma política, rejeitando tecnologias a fim de proteger interesses econômicos de outros grupos também é contestada. Para Stirling a ideia de neutralidade da ciência é relativa, justamente porque a política está presente em todas as esferas (tecnologia, saúde e meio ambiente), de modo que interesses conflitantes em diversos setores busquem afirmar seus próprios enquadramentos científicos de forma mais conveniente, como se fossem representações singulares e definitivas da ciência.²⁴⁴

²⁴² ARAGÃO, op.cit.,2013. p. 24.

²⁴³ *Ibidem*.

²⁴⁴ STIRLING,op.cit, 2016, p. 7.

A título exemplificativo, empresas produtoras de agrotóxicos tendem a financiar pesquisas que reafirmam e justifiquem seus anseios, em contraponto, movimentos contra agrotóxicos podem não ter o mesmo poder financeiro para investimentos em pesquisas, fazendo com que haja uma corrida assimétrica para justificar o uso ou proibição de determinado tipo de agrotóxico.

Por mais que a ciência tenha metodologias que visam eliminar vieses, essa discrepância financeira pode minar sua neutralidade. E mesmo que a corrida seja simétrica, tais setores sempre vão buscar estudos que justifiquem seus anseios.

Ademais, ao analisar a crítica quanto a anti cientificidade do princípio, nota-se que esta recai no caráter explicitamente normativo do princípio, supondo que procedimentos regulatórios “baseados na ciência”, tendem a transcender tal conteúdo. Entretanto, ignora-se o fato de que análises de custo-benefício também exigem o exercício de julgamentos avaliativos. Ocorre que os resultados baseados em evidências destes métodos, apresentam forte sensibilidade a suposições que variam de modo radical entre autoridades diferentes. Qualificar os níveis de proteção de determinada avaliação de risco, ponderação de diferentes formas de danos e o sopesamento de benefícios. Todos esses elementos decorrem de julgamentos avaliativos, que carecem de justificativas puramente “científicas”. Desta forma, Stirling conclui que não só na precaução, como em todos os métodos de avaliação de riscos, existem normatividades que dependem de julgamentos puramente avaliativos.²⁴⁵

Outra crítica rebatida é quanto ao princípio da precaução como um princípio do medo, tendo em vista os custos operacionais que o envolve. Para Aragão, a crítica quanto ao custo excessivo da precaução e consequentemente o impacto negativo ao orçamento do estado que o aplica, é uma crítica infundada. Primeiro porque a análise de custo-benefício pode induzir a resultados enganosos. Ademais, os benefícios por vezes são difíceis de quantificar e podem ser perceptíveis apenas no futuro, desta forma, o investimento visto como alto, pode vir a ser compensado futuramente.²⁴⁶

²⁴⁵ *Ibidem*, p. 6-7.

²⁴⁶ ARAGÃO, op.cit., 2013. p. 25.

Aragão também rebate o que ela classifica como *utopia do risco zero*, em seu entendimento, a precaução nunca buscou eliminar todo e qualquer risco. Pelo contrário, razões importantes podem justificar que determinadas atividades ensejadoras de riscos não só possam, como devam, ser desenvolvidas. O objetivo da precaução é elevar o nível de proteção, não se tratando de almejar o *risco zero*, busca-se, na verdade, um desenvolvimento tecnológico responsável.

Aragão também contesta o entendimento de que o princípio paralisa a inovação, se tornando uma trava ao progresso. Sustentando que nem sempre o princípio se trata de uma atuação *negativa*, ou seja, proibindo e paralisando o desenvolvimento de determinada atividade. Na verdade, o princípio opera igualmente de modo positivo, promovendo a evolução científica de modo precaucional, em conformidade com a ideia de progresso. Para a professora, aquilo que o princípio recusa, trata-se de um desenvolvimento economicista, imediatista e inconsequente.²⁴⁷

Stirling segue a mesma linha ao dizer que a precaução visa dirigir, ao invés de parar a inovação, para ele, as acusações de anti-inovação que recaem sobre o princípio deixam de analisar os múltiplos processos evolutivos sociais e tecnológicos. Deste modo, por mais que a precaução interrompa o desenvolvimento de determinada tecnologia, ela abre um campo de desenvolvimento de tecnologias alternativas. Por exemplo, a aplicação da precaução para parar o desenvolvimento de energia nuclear, tende a criar novas oportunidades de desenvolvimento de tecnologias renováveis.²⁴⁸

Por fim, as críticas quanto ao princípio ser mal definido, por não definir o que é *sério*, *irreversível* e *risco*, pode ser procedente, se analisarmos o princípio como regra processual suficiente, abrangente e definitiva. Todavia, assim como acontece com todos princípios gerais, tais como *proporcionalidade* e *custo-benefício*, nenhuma formulação da precaução pode por si só ser autossuficiente para uma regra de decisão.²⁴⁹

²⁴⁷ *Ibidem*.

²⁴⁸ STIRLING, op.cit, 2016, p. 9.

²⁴⁹ STIRLING, op.cit, 2016, p. 5-6.

Desta forma, a precaução é mais importante como um processo epistêmico e deliberativo do que uma *regra de decisão*. Trata-se de um framework para ajudar a solucionar o problema da incerteza científica.

3.5. A APLICAÇÃO DA TEORIA DE NASSIM TALEB PARA JUSTIFICAR O USO DO PRINCÍPIO DA PRECAUÇÃO NA REGULAÇÃO DE RISCOS DA INTELIGÊNCIA ARTIFICIAL

Para Taleb, o princípio da precaução deve ser invocado de modo não ingênuo, para regular o que ele classifica como *Black Swans*,²⁵⁰ o termo serve para conceituar eventos de extrema raridade, mas com potencial catastrófico.

Nesta esteira, o princípio deve ser invocado para evitar a *ruína*,²⁵¹ não sendo confundido como um método de gerenciamento de riscos. A gestão de riscos envolve um *trade off*, contabilizando os efeitos positivos e negativos de uma atividade. Deste modo, estas estratégias são importantes apenas quando não há *ruína* em jogo, justamente pelo fato da ruína ter um custo infinito. Logo, análises de custo-benefício se tornam um paradigma inútil.

O risco de ruína não é sustentável, logo o princípio da precaução deve resultar em um mecanismo que garanta que não sejam tomadas decisões que possam resultar em *ruína*.²⁵²

Taleb, para elucidar e aprimorar o entendimento sobre o que é ruína, traz duas categorias quais eventos aleatórios acontecem: (i) *thin tails*,²⁵³ (ii) *fat tails*.²⁵⁴ (i) *Thin tails* é um domínio que os eventos são acompanhados por variações leves e bem comportadas, deste modo a concentração é observada

²⁵⁰ Tradução: cisnes negros.

²⁵¹ Ruína é entendida como catástrofe generalizada e irreversível.

²⁵² TALEB, Nassim. N. et al. The Precautionary Principle: (with Application to the Genetic Modification of Organisms). NYU SCHOOL OF ENGINEERING WORKING PAPER SERIES, Paris, v. 1, n. 1, p. 2-3, out. /2014.

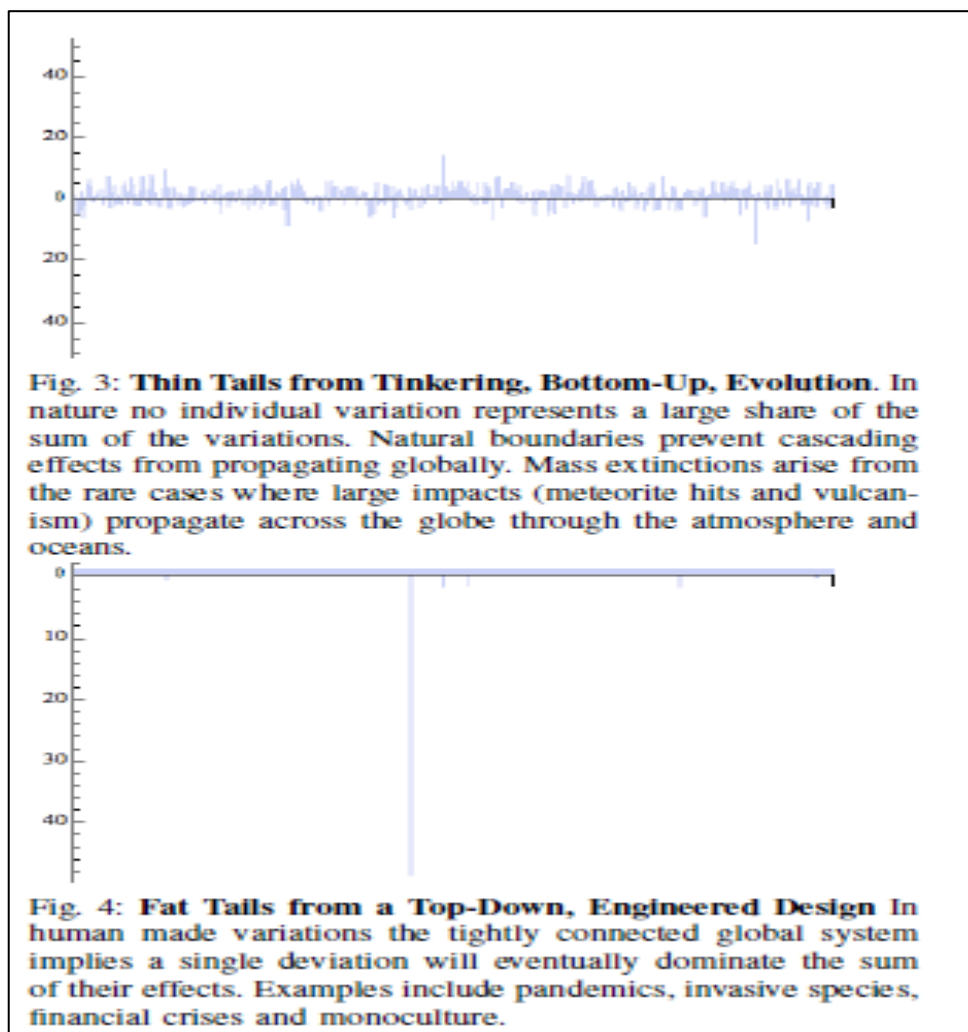
²⁵³ Tradução: cauda fina.

²⁵⁴ Tradução: cauda grossa.

na média de acontecimentos. Eventos de alta magnitude são improváveis neste contexto.

(ii) *Fat tails* é um domínio que os eventos têm pequenas probabilidades de ocorrer, todavia quando ocorrem, suas variações são extremas. Tornando inútil o conceito de média.

Gráfico 1: Curva de variação em eventos *thin tail* e *fat tail*.



Fonte: TALEB, Nassim. N. et al. The Precautionary Principle: (with Application to the Genetic Modification of Organisms).²⁵⁵

²⁵⁵ TALEB, et al. cit.op.,2014. p. 4.

Os gráficos demonstram as curvas de *thin tails* e *fat tails* respectivamente, de modo que a aplicação da precaução se dá em casos de *fat tails*, enquanto mecanismo de gestão de riscos se aplicam as *thin tails*.

Ou seja, métodos que se baseiam em modelos estatísticos podem ser utilizados para as *thin tails*. Todavia, por conta destes métodos desconsiderarem riscos não quantificáveis, sua aplicação em *fat tails* se torna ineficaz.

No decorrer do segundo capítulo, foram discutidos os principais riscos trazidos pelo uso da IA. Notou-se também a dificuldade em prever os resultados de seus sistemas, uma vez que a opacidade gerada por seus algoritmos traz desafios até mesmo para os seus desenvolvedores. Deste modo, a atuação da IA no meio ambiente tornou-se um campo de incertezas.

Técnicas de profiling, micro-targeting e predição de comportamento, tem um potencial catastrófico no ambiente democrático podendo manipular e subverter o livre arbítrio. Tão como, a tomada de decisão automatizada da IA, demonstra um alto potencial danoso a humanidade, violando uma enorme gama de direitos humanos consubstanciados em diversos tratados internacionais. Na mesma esteira, o futuro do trabalho é colocado em xeque com o desenvolvimento da IA, para fins de automação, de modo irresponsável. Por fim, a possibilidade de existência de uma AGI e a perda de seu controle, pode trazer em última instância, riscos à existência da raça humana.

Fora, outros riscos não abordados neste trabalho, como armas autônomas e mecanismos de reconhecimento facial.

Deste modo, nota-se que a maioria dos eventos possíveis, decorrentes do uso irrestrito da IA, tende a ter um potencial catastrófico. Sob essa perspectiva, o princípio da precaução se mostra um instrumento necessário.

Entendendo a probabilidade de eventos decorrentes de IA como *fat tails*, conclui-se que não se pode esperar evidências científicas robustas para sua regulação.

Conforme já mencionado, a precaução deve servir como fio condutor da inovação e não como uma trava. De todo modo, isso não quer dizer que a precaução nunca será aplicada para desacelerar o desenvolvimento de

determinada tecnologia, na verdade, a depender do caso isso também pode ser positivo.

Acemoglu, por exemplo, defende que o princípio da precaução como mecanismo regulatório ex-ante, deve retardar o uso de algumas tecnologias de IA, especialmente nos campos onde os custos de correção se tornem difíceis após a implementação. Deste modo, ele acredita que o principal campo de aplicação do princípio da precaução deve se dar nas tecnologias de IA que impactam o discurso político e o ambiente democrático (tema abordado no segundo capítulo deste trabalho).²⁵⁶

3.6. ANÁLISE DA PROPOSTA EUROPEIA E BRASILEIRA PARA REGULAÇÃO DA I.A

Desta feita, uma vez entendida o princípio da precaução como um *framework*²⁵⁷ para lidar com as incertezas trazidas pela IA. Passa-se analisar se as propostas regulatórias da União Europeia, tão como, da legislação brasileira, carregam algum tipo de expressão precaucionária. E se há espaço para a operacionalização do princípio em qualquer uma de suas versões, *forte*, *moderada* e *fraca*.

Basicamente, tem-se dois modos de expressar normativamente o princípio da precaução: (i) abrir o debate regulatório para todos os envolvidos no desenvolvimento da tecnologia e aqueles que são afetados por ela; (ii) implementação de salvaguardas que busquem reduzir as incertezas desta tecnologia.²⁵⁸

²⁵⁶ ACEMOGLU, Daron. HARMS OF AI. NBER WORKING PAPER SERIES, Cambridge, v. 1, n. 1, p. 50-51, set. /2021.

²⁵⁷ Framework, em português, significa estrutura. O termo é comumente utilizado como uma estrutura feita para resolver um problema específico.

²⁵⁸ BIONI, op.cit, 2020. p. 19.

3.6.1. UNIÃO EUROPEIA - ARTIFICIAL INTELLIGENCE ACT

A UE comparada ao Brasil, mostra-se mais avançada quanto às movimentações regulatórias sobre o tema. Nesta esteira, a Comissão Europeia apresentou em 2021 sua proposta para regulação da IA no contexto europeu, a proposta recebeu o nome de Artificial Intelligence Act.

Nota-se, ao analisar o contexto da proposta, que se aposta em uma regulação baseada em riscos (*Risk-based-Approach*). Nesta modalidade regulatória o princípio da precaução pode ganhar campo para se pensar em medidas regulatórias *ex-ante*, para lidar com incertezas.

A proposta inicialmente já demonstra uma expressão precaucionária, tendo em vista que ela surge em decorrência de audiências públicas que discutiram os principais pontos do *White Paper on AI*,²⁵⁹ foram recebidas inúmeras contribuições de organizações, associações comerciais, instituições acadêmicas e autoridades públicas.²⁶⁰ Tal movimentação demonstra a abertura do debate regulatório à sociedade civil.

A proposta optou por classificar os sistemas de IA em patamares de risco, sendo eles: (i) risco inaceitável (*unacceptable-risk*); (ii) risco elevado (*high-risk*); (iii) risco limitado (*limited-risk*) e (iv) risco mínimo (*minimal-risk*).

De acordo com a proposta, os sistemas de IA classificados como sistemas geradores de riscos inaceitáveis devem ter o seu uso proibido.²⁶¹

Cumprе ressaltar, que esta proibição não deve ser vista como a aplicação do princípio da precaução em sua versão *forte*, tendo em vista que não há a inversão do ônus da prova ao proponente da tecnologia. Na verdade, criou-se um rol proibitivo de determinados tipos de tecnologias em determinados campos de atuação, este rol proibitivo deve ser visto como uma

²⁵⁹ COMMISSION'S, The European. A DEFINITION OF AI: MAIN CAPABILITIES AND SCIENTIFIC DISCIPLINES. HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE, Bruxelas, v.1,n.1,p.9,dez./2021.

²⁶⁰ COMMISSION, European. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS. THE EUROPEAN COMMISSION, Brussels, v. 1, n. 1, p. 4, jun./2021.

²⁶¹ *Ibidem*. p. 7.

expressão precaucionária na via *moderada*. Similar ao proposto pela Comunicação da Comissão Europeia relativa ao princípio da precaução.²⁶²

A versão *forte* do princípio parte de uma lógica preventiva de riscos e não gerencial, diferentemente do que é tratado nesta proposta.

Nas hipóteses de sistemas classificados como de alto risco, o princípio da precaução também aparece em sua versão *moderada*, são inúmeras as evidências para tal justificativa:

A primeira delas é a criação de pressupostos de gerenciamento de riscos implícitos.²⁶³ O artigo 9º visa a proceduralizar as ações de gerenciamento de riscos, o gerenciamento é contínuo, deve ser documentado e tem como ponto de partida medidas *ex-ante* de gestão de riscos, segundo uma lógica de *gestão de riscos por design*²⁶⁴ (*risk management by design*). Deste modo, os testes de gerenciamento de riscos devem ser realizados durante o processo de desenvolvimento e em qualquer caso, antes do lançamento da tecnologia no mercado.²⁶⁵ A proposta é silente às medidas de paralisação da tecnologia quando houver ameaça de danos. Entretanto, mostra-se proativa na identificação de riscos ao obrigar o fornecedor do sistema a procedimento de avaliação de conformidade²⁶⁶ e registrar o sistema em uma base de dados da UE que será gerida pela comissão.²⁶⁷ Esta proatividade é outra forma de expressão *moderada* do princípio.²⁶⁸

Outro ponto marcante que expressa a precaução de forma *moderada* é o conjunto de obrigações horizontais a serem seguidas pelos proponentes estabelecendo mecanismos para garantia do *compliance* de riscos.²⁶⁹

²⁶² Comunicação da Comissão das Comunidades Europeias relativa ao princípio da precaução, Bruxelas, 2/02/2000. p.15.

²⁶³ *Ibidem*, p. 8.

²⁶⁴ A tradução do termo “by design” em português é *por padrão*. O conceito representa a criação de determinados padrões na consecução de determinada atividade. No campo da proteção de dados pessoais, por exemplo, usa-se o conceito *privacy by design* para defender que as empresas incorporem em seus produtos, por padrão, ou seja, desde o início da produção, configurações mais seguras de privacidade.

²⁶⁵ COMMISSION, European. Op.cit, jun/2021, p. 45-48.

²⁶⁶ *Ibidem*, p.52.

²⁶⁷ *Ibidem*, p.15.

²⁶⁸ GARNETT; PARSONS, op.cit,2016. p.4.

²⁶⁹ COMMISSION, European. Op.cit. jun/2021, p. 52-53.

O artigo 13º traz obrigações de transparência e *accountability* dos proponentes aos utilizadores. O artigo determina que seja assegurada aos utilizadores que o funcionamento da tecnologia seja *suficientemente transparente*, isso quer dizer que o funcionamento do sistema deve poder ser facilmente compreendido pelo utilizador. Ademais, o artigo 13º junto ao 15º pedem um nível mínimo de exatidão do sistema.²⁷⁰

Os artigos acenam para a resolução das questões atinentes às incertezas algorítmicas discutidas no capítulo anterior, deste modo, mesmo que não esteja procedimentalizado ações positivas quando a inexatidão dos sistemas, os artigos ainda assim, trazem uma lógica regulatória precaucionária. Que podem oscilar de nível, de acordo com as medidas a serem propostas posteriormente em casos onde o algoritmo se demonstre inexato.

Por fim, para os sistemas de IA de baixo risco e risco limitado não há nenhum tipo de expressão precaucionária, uma vez que se optou por uma abordagem autorregulatória.

Trata-se de uma abordagem autorregulatória, tendo em vista que os fornecedores de sistemas de baixo risco poderão apenas seguir um código de conduta para o desenvolvimento da tecnologia, não estando obrigados a obrigações regulatórias.

Portanto, por mais que a proposta europeia não traga um rol de princípios como norte, nota-se que especificamente o princípio da precaução está presente de forma implícita, as diversas expressões *moderadas* no texto normativo demonstram proatividade na busca pela clarificação de incertezas, em um nível adequado.

3.6.2. BRASIL (PL 21/2020) - MARCO REGULATÓRIO DA IA

Diferentemente da UE, o Brasil vem discutindo as definições do seu marco regulatório por meio de projeto de lei.

²⁷⁰ *Ibidem*, p. 50-52.

Concorrem concomitantemente, três projetos de lei para regular o uso da IA no Brasil: (i) PL 5051/2019; (ii) PL 872/2021 e (iii) PL 21/2020.

O único deles já aprovado pelo congresso nacional é o PL 21/2020, este aparentemente tomou o protagonismo das discussões regulatórias do tema.

O Brasil, seguiu a mesma linha da UE, optando por uma abordagem de regulação baseada em riscos.²⁷¹ Uma movimentação óbvia se notarmos que a Lei de Proteção de Dados Brasileira (LGPD), inspirou-se no regulamento europeu (GDPR).²⁷²

Todavia, por ter iniciado o debate no campo da regulação de IA de modo tardio, os legisladores brasileiros optaram por colocar em votação projetos de leis excessivamente principiológicos, com baixo teor normativo, impossibilitando que se pudesse de alguma forma instrumentalizar mecanismos de governança e mitigação de riscos.²⁷³

Desta forma, resta prejudicada a análise de expressões precaucionais no contexto normativo. Entretanto, pode-se ainda analisar se a lei guarda espaços onde o princípio pode ser introduzido, em qualquer de suas versões.

Extraí-se do artigo 6º, III, uma possibilidade de expressão do princípio em sua versão *moderada*, sugerindo que a intervenção seja proporcional aos riscos concretos oferecidos. O inciso IV do artigo também possibilita a expressão do princípio em sua versão moderada, tendo em vista que ele traz como diretriz, a participação social e interdisciplinar para a adoção de normas de gestão de risco, precedida de consulta pública.^{274 275}

Entretanto, o projeto de lei traz o termo *risco concreto* como justificativa para intervenção estatal, deste modo a expressão *moderada* supra descrita, pode ser enfraquecida para a uma expressão *fraca* do princípio, caso a

²⁷¹ BRASIL (2020) – SENADO FEDERAL, projeto de lei 21/2020, Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil; e dá outras providências. p.7.

²⁷² SERPRO. LGPD: a versão brasileira do regulamento europeu. Disponível em: <https://www.serpro.gov.br/lgpd/noticias/lgpd-versao-brasileira-gdpr-dados-pessoais>. Acesso em: 21 mar. 2022.

²⁷³ DATAPRIVACYBR. Contribuições do Data Privacy Brasil ao Projeto de Lei nº 21, de 04 de fevereiro de 2020: Nota técnica. Data Privacy Research, São Paulo, v. 1, n. 1, p. 4, jun./2020.

²⁷⁴ GARNETT; PARSONS, op.cit,2016. p.6.

²⁷⁵ BRASIL (2020) – SENADO FEDERAL, projeto de lei 21/2020, p. 7.

incerteza a priori, não seja uma justificativa que classifique o risco como concreto. Neste caso, haveria uma preferência pela coleta de evidências sobre a chance e gravidade do dano. Justificando assim, a expressão *fraca* do princípio, preferindo uma abordagem baseada na gestão de riscos.²⁷⁶

Por fim, cumpre ressaltar que grupos de pesquisa no Brasil apontam para a inserção do princípio da precaução dentro do rol de princípios da proposta legislativa.

A Associação de Pesquisa Data Privacy Brasil, em sua nota técnica pede pela inclusão do princípio da precaução na proposta de lei. Justifica-se esse pedido dentro de uma lógica de “*proteção de direitos por design*”, fazendo com que a incerteza justifique a redução da discricionariedade do agente ou até mesmo a proibição do uso da tecnologia.²⁷⁷

Eles entendem também, que a elaboração de relatórios de impactos aos direitos humanos é a ferramenta ideal para buscar evidenciar os riscos inerentes ao emprego da tecnologia, tão como das incertezas.²⁷⁸ Deste modo, o relatório de impacto poderia operacionalizar o princípio da precaução em qualquer de suas expressões (*fraca, forte e moderada*) a depender do regulador.²⁷⁹

Por fim, a associação justifica a importância do princípio na regulação da IA, tendo em vista que suas expressões podem ser um referencial para a categorização de salvaguardas, mediante a uma taxonomia do que considera como risco e incerteza, e qual tipo de ação será tomada, ou até a opção da inação quando identificado um desequilíbrio entre riscos e benefícios do emprego da IA.²⁸⁰ Do ponto de vista técnico, as visões da Associação Data Privacy Brasil quanto ao princípio da precaução, são extremamente acertadas e contribuem de forma substancial para a inserção do princípio no texto do projeto de lei.

²⁷⁶ GARNETT; PARSONS, op.cit,2016. p.4.

²⁷⁷ DATAPRIVACYBR. op.cit, 2020. p.19.

²⁷⁸ *Ibidem*.

²⁷⁹ BIONI, op.cit, 2020. p. 12.

²⁸⁰ BIONI, et al. CONTRIBUIÇÃO DA ASSOCIAÇÃO DATA PRIVACY BRASIL DE PESQUISA À CONSULTA PÚBLICA ESTRATÉGIA BRASILEIRA DE INTELIGÊNCIA ARTIFICIAL. 1. ed. São Paulo: Reticências, 2020. p. 34.

Outro ponto importante, que destaca o avanço dos debates quando a inserção do princípio da precaução no projeto de lei brasileiro, se dá pelo fato de que recentemente o Senado Federal Brasileiro, visando trazer maior tecnicidade ao debate, instaurou uma comissão de juristas com audiências públicas para discutir a regulação da IA.²⁸¹

Como já mencionado, trazer o debate para a esfera pública já é um tipo de expressão precaucionária, todavia, mais importante que esta movimentação, está o fato da comissão ter instaurado um painel exclusivo para a discussão de riscos e o princípio da precaução. O sexto painel foi intitulado como *Painel 6 - Inteligência Artificial e riscos: gradação de riscos; hipóteses de riscos inaceitáveis e princípio da precaução*.²⁸²

Desta forma, em que pese o Brasil não tenha ainda uma proposta regulatória com um teor normativo que possibilite analisar expressões do princípio da precaução ali contidas, as movimentações do Senado Federal, ao instaurar uma comissão de juristas para debater o tema, tão como, as associações de pesquisa acadêmica (especificamente o Data Privacy Brasil) rogando pela inserção do princípio dentro do rol de princípios do projeto de lei, demonstram um grande avanço no sentido de se buscar uma abordagem precaucionária para o marco regulatório brasileiro. Tais movimentações não são suficientes para concluir que o princípio será inserido no texto normativo ou expresso de forma implícita, e se sua operacionalização se dará de forma suficiente. Mas demonstram, uma posição de humildade do legislador frente às incertezas trazidas por esta tecnologia.

²⁸¹ STJ. Comissão de juristas inicia audiências públicas para discutir regulação da inteligência artificial. Disponível em: <https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/28042022-Comissao-de-juristas-inicia-audiencias-publicas-para-discutir-regulacao-da-inteligencia-artificial-.aspx>. Acesso em: 23 mar. 2022.

²⁸² *Ibidem*.

CONCLUSÃO

O presente trabalho demonstrou que o processo evolutivo da sociedade moderna nos colocou em uma Sociedade de Risco. Partindo desta análise, notou-se que a distribuição de bens fruto do progresso tecnológico, passou a ser acompanhada pela distribuição de riscos.

Estes riscos ganham cada vez mais complexidade ante aos avanços tecnológicos. Assim, a tecnologia se torna um tema e problema para a sociedade. Essa maior complexidade, torna as ameaças decorrentes não passíveis de previsão e controle, criando um campo de incertezas científicas.

Isso tudo ganha novos contornos com o advento da Quarta Revolução Industrial. A evolução tecnocientífica ganha maior amplitude e atinge uma velocidade exponencial. Trazendo um grande ponto de ebulição para a Sociedade de Risco.

Um dos principais pontos desta dinâmica revolucionária é o aumento dos processos de digitalização. O surgimento do Big Data, a economia pautada em dados e a fusão de diversas tecnologias, interagindo com seres orgânicos e inorgânicos. Dentro desse espectro evolutivo, surge também a Inteligência Artificial, atuando nos mais diversos setores e interagindo com seres humanos. Esta interação entre humanos e robôs, faz com que a distribuição de *bens* e *males* se opere por veículos digitalizados, colocando a sociedade moderna em uma *Sociedade de Risco Digital*.

Cumprida esta análise e conexão entre a Quarta Revolução Industrial e a Sociedade de Risco, o presente trabalho trouxe as definições e classificações jurídicas da IA de modo a conceituá-la, elencando seus principais riscos.

Desta forma, primeiramente foram demonstrados os principais riscos da Inteligência Artificial hoje existente, a ANI (IA fraca). Este tipo de IA, já demonstrou ter um alto poder deletério aos direitos humanos, suas múltiplas facetas se dão por ações discriminatórias. Decisões automatizadas de IA na seara criminal, aumentam e criam um ciclo retroalimentativo de encarceramento de pessoas pertencentes a grupos sociais historicamente marginalizados. A lógica estatística do algoritmo distorce a lógica de

causalidade e correlação, criando o chamado *ciclo de feedback*. O mesmo ocorre nos setores onde se busca acesso ao crédito, encontrar empregos e acesso a tratamentos de saúde.

Alem disso, os sistemas de *social scoring* implementados pelo governo chinês, mostram um grande potencial de controle comportamental da população, produzindo uma gama de riscos a direitos humanos, tais como, direito a privacidade, direito a liberdade de reunião e associação pacífica, direito a não discriminação.

O futuro do trabalho também ganha novos contornos com o avanço da Inteligência Artificial. Principalmente porque a IA promete ocupar postos de trabalho de funções físicas e cognitivas, diferentemente do que foi feito nas revoluções industriais anteriores. A IA, tende a executar tarefas cada vez mais complexas e em determinado momento, poderá executar praticamente todas as funções que um ser humano consegue executar, desabastecendo a maioria dos postos de trabalho ocupados por humanos. A ausência de respostas globais para este problema também coloca o futuro do trabalho em risco, aqui fica evidente o conceito de *Irresponsabilidade Organizada* debatido inicialmente.

Este risco, em específico, vai se operar dentro do conceito de *Efeito Bumerangue* (trazido por Beck), tendo em vista que os países mais ricos gozarão dos benefícios da automação industrial em um primeiro momento. Posteriormente, experimentarão os efeitos negativos deste processo.

A AGI (Inteligência Artificial Geral) e ASI (Super Inteligência Artificial), mesmo habitando apenas no campo teórico, demonstram ter um alto poder catastrófico, justamente pela sua difícil previsão de comportamento, podendo em última instância, destruir a raça humana.

A incapacidade em prever os resultados destes robôs é o que torna todos os riscos aqui debatidos em incertezas científicas. A tomada de decisão por estes robôs se opera por uma lógica matemática chamada de algoritmo. Todavia, a complexidade desta impossibilita que seja dada explicabilidade às suas decisões, até mesmo para quem os programou. Esta opacidade gerada pela tecnologia, faz com que ela se torne uma verdadeira caixa preta.

Discussões endereçadas à abertura de código-fonte e consequentemente a quebra da proteção do segredo industrial se tornam inócuas, tendo em vista que mesmo com a abertura do código-fonte, as decisões automatizadas pela IA não são plenamente inteligíveis. Pode-se dizer que essa incompreensão quanto à tomada de decisão da IA mina a capacidade dos desenvolvedores destas tecnologias em mensurar riscos. Esta é a principal marca de uma Sociedade de Risco, agora digitalizada.

Deste modo, entendendo a IA como um campo de incertezas científicas, a aplicação do princípio da precaução se torna razoável, buscando entendimentos mais robustos para administrar as incertezas, com mecanismos regulatórios *ex-ante*. O princípio pode ser visto como um princípio norteador da Sociedade de Risco, de modo a evitar danos de alta relevância, por vezes, com potencial catastrófico.

Foi demonstrado que de fato, o princípio não tem uma única definição, sendo expresso de diversas formas na regulação das mais diversas atividades. Portanto, classificou-se suas expressões normativas em: (i) *fraca*, (ii) *moderada* e (iii) *forte*. Na expressão *fraca*, nota-se um maior apreço pela gestão de risco, um alto limiar epistêmico, ou seja, busca-se a coleta de dados que justifiquem a regulação da tecnologia. Na expressão *moderada*, busca-se elencar atividades consideradas aceitáveis e cria-se mecanismos que visem reduzir incertezas, construindo uma extensa investigação quanto às incertezas. Na expressão *forte*, propõe a inversão do ônus da prova ao proponente da atividade econômica. Neste último caso, nota-se um baixo limiar epistêmico e uma lógica aversa a gestão de riscos.

A contraposição dos principais tópicos trazidos pelos críticos e defensores do princípio, serviu para evidenciar que o princípio, diferentemente do que se justifica pela crítica, não é um princípio anticientífico que tende a paralisar a inovação. Na verdade, o princípio busca impor cautela à soberania da ciência, pressupondo uma análise científica e racional. Deste modo, medidas precaucionais são tomadas em razão da ciência, não à revelia desta.

Precaução visa dirigir e não paralisar a inovação, até porque, mesmo em ações positivas (proibição do uso de determinada tecnologia) podem significar

o avanço e desenvolvimento de outra tecnologia menos danosa. Ou alternativamente, forçar que determinada tecnologia se torne mais segura antes de ser colocada no mercado. Trata-se de uma posição de humildade frente às incertezas, contrapondo uma visão tecnocrática, economicista e inconsequente. Precaução é dar um passo atrás ante a dinâmica evolutiva em busca pelo progresso, decorrente do processo de modernização.

Conclui-se então, que o princípio da precaução é um framework útil na regulação da IA. Deste modo, analisou-se o conteúdo das últimas propostas regulatórias da UE e do Brasil, buscando encontrar expressões precaucionais no texto normativo, tão como, possíveis feixes de operacionalização do princípio.

Na proposta Europeia, evidenciou-se haver inúmeras expressões do princípio da precaução em sua versão moderada, dentre elas destacam-se: (i) os critérios de classificação de riscos inaceitáveis; (ii) a criação de pressupostos de gerenciamentos de riscos implícitos; (iii) ações proativas de avaliação de conformidade; (iv) obrigações horizontais visando garantir o *compliance* dos riscos; (v) obrigações de transparência e explicabilidade em um nível mínimo aceitável. Deste modo, mesmo que o princípio não seja mencionado na proposta, sua operacionalização se faz de modo intrínseco em parâmetros aceitáveis.

Na proposta brasileira, restou-se prejudicada a análise de expressões precaucionais contidas nos projetos de lei que atualmente vem sendo discutidos. Tendo em vista que as propostas legislativas são excessivamente principiológicas, carregando um baixo teor normativo de modo que não há definições prévias de avaliação de riscos. Entretanto, foi possível supor que o projeto acena para uma expressão precaucionária *moderada* ou *fraca*, a depender do que virá a ser considerado como *risco concreto*.

Todavia, isso não significa que os juristas brasileiros não estejam discutindo a aplicação do princípio da precaução na regulação da IA. Pelo contrário, associações que representam a sociedade civil, rogam pela inserção do princípio no rol de princípios da proposta legislativa. Trata-se de uma visão acertada, na busca por parametrizar quais ações ou inações devem ser

tomadas quando identificado desequilíbrios entre riscos e benefícios no emprego da IA. Deste modo, em que pese o Brasil esteja atrasado na regulação da IA, os debates no campo da precaução se mostram muito avançados.

REFERÊNCIAS

REFERÊNCIAS BIBLIOGRÁFICAS

ACEMOGLU, Daron. HARMS OF AI. **NBER WORKING PAPER SERIES**, Cambridge, v. 1, n. 1, p. 50-51, set. /2021. Disponível em: <http://www.nber.org/papers/w29247>. Acesso em: 24 fev. 2022.

ALBUQUERQUE, P. H. M. *et al.* NA ERA DAS MÁQUINAS, O EMPREGO É DE QUEM? ESTIMAÇÃO DA PROBABILIDADE DE AUTOMAÇÃO DE OCUPAÇÕES NO BRASIL. **Instituto de Pesquisa Econômica Aplicada**, Rio de Janeiro, v. 1, n. 1, p. 11, mar./2021. Disponível em: http://repositorio.ipea.gov.br/bitstream/11058/9116/1/td_2457.pdf. Acesso em: 18 jan. 2022.

ARAGÃO, Alexandra. Aplicação nacional do princípio da precaução. **Associação dos Magistrados da Jurisdição Administrativa e Fiscal de Portugal**, Lisboa, v. 1, n. 1, p. 5, jan./2013.

BALDWIN *et al.* **Understanding Regulation**: Theory, Strategy, and Practice. 2. ed. : Oxford University Press, 2013. p. 83-84.

BARRAT, James. **Our Final Invention**: Artificial Intelligence and The End of Human Era. 1. ed. Nova York: St. Martin's Griffin, 2013. p. 8.

BECK, Ulrich. Living in the world risk society. **Economy and Society**, London, v. 35, n. 3, p. 336, fev./2006. Disponível em: https://edisciplinas.usp.br/pluginfile.php/4095470/mod_resource/content/0/Bec k--WorldRisk.pdf. Acesso em: 18 out. 2021.

_____, Ulrich. **Sociedade de Risco**: Rumo a outra modernidade. 2. ed. São Paulo: editora 34, 2011. p. 23.

_____, Ulrich. WORLD AT RISK: THE NEW TASK OF CRITICAL THEORY. **DEVELOPMENT AND SOCIETY**, Monique, v. 37, n. 1, p. 2, jun./2008.

Disponível em:
https://www.academia.edu/38677305/WORLD_AT_RISK_THE_NEW_TASK_OF_CRITICAL_THEORY. Acesso em: 6 out. 2021.

BERGUER, Aírton Guilherme. **Regulação e Governança dos Riscos das Nanotecnologias**. 1. ed. Belo Horizonte: Arraes, 2018. p. 73.

BIONI, Bruno Ricardo. **PROTEÇÃO DE DADOS PESSOAIS**: a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2018. p. 59-61.

_____, Bruno Ricardo; LUCIANO, Maria. **Inteligência Artificial e Direito : O PRINCÍPIO DA PRECAUÇÃO NA REGULAÇÃO DE INTELIGÊNCIA ARTIFICIAL: SERIAM AS LEIS DE PROTEÇÃO DE DADOS O SEU PORTAL DE ENTRADA?**. 1. ed. São Paulo: Revistas dos Tribunais, 2020. p. 3.

BIONI, *et al.* CONTRIBUIÇÃO DA ASSOCIAÇÃO DATA PRIVACY BRASIL DE PESQUISA À CONSULTA PÚBLICA ESTRATÉGIA BRASILEIRA DE INTELIGÊNCIA ARTIFICIAL. 1. ed. São Paulo: Reticências, 2020. p. 34.

CAPLAN, R. et al. Algorithmic Accountability: A Primer. Data & Society, New York, v. 1, n. 1, p. 6, abr./2018. Disponível em: <https://datasociety.net/library/algorithmic-accountability-a-primer>. Acesso em: 23 fev. 2022.

CAVEDON, Ricardo; FERREIRA, Helene Sivini; FREITAS, C., C. O. D. A. O meio ambiente digital sob a ótica da Teoria da Sociedade de Risco: os avanços da informática em debate. **Revista Direito Ambiental e sociedade**, Curitiba, v. 5, n. 1, p. 200, jan./2015. Disponível em: <http://www.ucs.com.br/etc/revistas/index.php/direitoambiental/article/viewFile/3912/2318>. Acesso em: 20 set. 2020.

COSTA, Jonas Fernandes. Modernidade Reflexiva: Giddens, Beck e as limitações da segunda modernidade. **Dissertação de Mestrado**, Belém, v. 1, n. 1, p. 48, mai./2009.

DIONÍSIO, P. D. H. Princípio da precaução e contradições na regulação do risco: uma análise comparada entre Estados Unidos e Europa. **Regulação 4.0**, São Paulo, v. 1, n. 1, p. 31, jan./2019.

ELLIOTT, Anthony. **Making Sense of AI: Our Algorithmic World**. 1. ed. Cambridge: Polity Press, 2022. p. 3.

FERNANDES, J. P. T. Da utopia da sociedade em rede à realidade da sociedade de risco. **Análise Social**, Lisboa, v. 1, n. 2, p. 273-274, jan./2013. Disponível em: http://analisesocial.ics.ul.pt/documentos/AS_207_d01.pdf. Acesso em: 1 out. 2020.

FERRARI, Isabella; BECKER, Daniel; Wolkart, Erik. ARBITRIUM EX MACHINA:: PANORAMA, RISCOS E A NECESSIDADE DE REGULAÇÃO DAS DECISÕES INFORMADAS POR ALGORITMOS. **Revista dos Tribunais**, São Paulo, v. 995, n. 1, p. 8, set./2018. Disponível em: <http://governance40.com/wp-content/uploads/2018/11/ARBITRIUM-EX-MACHINA-PANORAMA-RISCOS-E-A-NECESSIDADE.pdf>. Acesso em: 15 fev. 2022.

FERREIRA, Helini Silvini. A BIOSSEGURANÇA DOS ORGANISMOS TRANSGÊNICOS NO DIREITO BRASILEIRO: UMA ANÁLISE FUNDAMENTADA NA TEORIA DA SOCIEDADE DE RISCO. **Tese de Doutorado**, Florianópolis, v. 1, n. 1, p. 48, abr./2008. Disponível em: <https://repositorio.ufsc.br/handle/123456789/91741>. Acesso em: 3 jan. 2022.

GAMBA, J. R. G. **Democracia e Tecnologia**: Impactos da Quarta Revolução Industrial. 1. ed. Rio de Janeiro: Lumen Juris, 2020. p. 78.

GIDDENS, Anthony; LASCH, Scott; BECK, Ulrich. **Modernização reflexiva**: Política, tradição e estética na ordem social moderna. 1. ed. São Paulo: UNESP, 1997. p. 26.

GOLDBLATT, David. **Teoria social e ambiente**. 1. ed. Lisboa: Instituto Piaget, 1996. p. 241.

GUIVANT, Julia Silvia. O LEGADO DE ULRICH BECK. **Ambiente & Sociedade**, São Paulo, v. 9, n. 1, p. 230, jan./2016.

HARARI, Yuval Noah. **21 lições para o século 21**. 1. ed. São Paulo: Companhia Das Letras, 2018. p. 59.

HARARI, Yuval Noah. Homo Deus: Uma breve história do amanhã. 1. ed. São Paulo: Companhia das Letras, 2016. p. 346.

RASO, Filippo et al. Artificial Intelligence & Human Rights: Opportunities & Risks. **The Berkman Klein Center for Internet & Society Research Publication Series**, Cambridge, Massachusetts, v. 1, n. 1, p. 21-25, set./2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3259344. Acesso em: 23 dez. 2021.

MAGRANI, Eduardo. **Democracia Conectada**: A Internet como Ferramenta de Engajamento Político-Democrático. 1. ed. Curitiba: Juruá, 2014. p. 119.

MAGRANI, Eduardo. **Entre dados e robôs**: ÉTICA E PRIVACIDADE NA ERA DA HIPERCONNECTIVIDADE. 1. ed. Porto Alegre: Arquipélago, 2019. p. 52.

MENDELSON, Daniel. **The Routledge Companion to Social Theory**: Central Terms And Thinkers. 1. ed. London: Routledge, 2010. p. 230-232.

MORAES, G. B. D. A. O Princípio da Precaução no Direito Internacional do Meio Ambiente. **Dissertação de Mestrado**, São Paulo, v. 1, n. 1, p. 89, mai./2011. Disponível em: https://www.teses.usp.br/teses/disponiveis/2/2135/tde-03092012-111415/publico/MORAES_Gabriela_Bueno_de_Almeida_Dissertacao_de_mestrado_integral.pdf. Acesso em: 4 out. 2021.

O'NEIL, Cathy. Algoritmos de Destruição em Massa: Como o Big Data aumenta a desigualdade e ameaça a Democracia. 1. ed. Santo André: Rua do Sabão, 2020. p. 15.

OSOBA, Osonde A.; IV, William Welser. The Risks of Artificial Intelligence to Security and the Future of Work. **Rand Corporation**, Santa Monica CA, v. 1, n.

1, p. 10, jan./2017. Disponível em: <https://www.rand.org/pubs/perspectives/PE237.html>. Acesso em: 17 jan. 2022.

GARNETT; J. PARSONS, K. Multi-Case Review of the Application of the Precautionary Principle in European Union Law and Case Law. **Risk Analysis**, Cranfield, v. 1, n. 1, p. 4, mai./2016. Disponível em: <https://www.researchgate.net/publication/303313626>. Acesso em: 22 fev. 2022.

PASQUALE, Frank. **The black box society**: he secret algorithms that control money and information. 1. ed. Cambridge: Harvard University Press, 2015. p. 3.

SCHWAB, Klaus. **A Quarta Revolução Industrial**: Tradução Daniel Moreira Miranda. 1. ed. 2016: Edipro, 2016. p. 16.

SILVA, Brisa Arnold da. UMA ANÁLISE SOBRE A MODERNIDADE REFLEXIVA E A COMPLEXIDADE AMBIENTAL NO ESTADO SOCIOAMBIENTAL DE DIREITO PARA O COMPROMISSO DO DESENVOLVIMENTO SUSTENTÁVEL. Cadernos do Programa de Pós Graduação em Direito, Porto Alegre, v. 10, n. 1, p. 104, jan./2015. Disponível em: <https://seer.ufrgs.br/index.php/ppgdir/article/view/54612>. Acesso em: 10 jan. 2022.

STIRLING, Andy. Precaution in the Governance of Technology. **SPRU Working Paper Series**, Brighton, v. 1, n. 1, p. 4, jan./2016. Disponível em: https://www.researchgate.net/publication/305986930_Precaution_in_the_Governance_of_Technology_SWPS_2016-14. Acesso em: 21 fev. 2022.

SUNSTEIN, Cass R. Para além do princípio da precaução: Beyond the Precautionary Principle. **Revista de Direito Administrativo**, Rio de Janeiro, v. 1, n. 1, p. 23, abr./2012.

SUNSTEIN, R. W. H.; C. R. The Precautionary Principle as a Basis for Decision Making. **The Economists' Voice**, Chicago, v. 2, n. 2, p. 6, mai./2005. Disponível em: https://www.researchgate.net/publication/24015415_The_Precautionary_Principle_as_a_Basis_for_Decision_Making. Acesso em: 28 fev. 2022.

TALEB, Nassim. N. *et al.* The Precautionary Principle: (with Application to the Genetic Modification of Organisms). **NYU SCHOOL OF ENGINEERING WORKING PAPER SERIES**, Paris, v. 1, n. 1, p. 2-3, out. /2014. Disponível em: <https://arxiv.org/abs/1410.5787>. Acesso em: 22 fev. 2022.

TORRES, Phil. The possibility and risks of artificial general intelligence. **Bulletin of the Atomic Scientists**, Chicago, v. 75, n. 3, p. 105-108, abr./2019. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/00963402.2019.1604873?needAccess=true&journalCode=rbul20>. Acesso em: 29 jan. 2022.

WERNER, Deivid Augusto. A Quarta Revolução Industrial e a Inteligência Artificial: um estudo sobre seus conceitos, reflexos e possível aplicação no Direito por meio de análise de texto jurídico como forma de contribuição no processo de categorização preditiva de acórdãos. **Dissertação de Mestrado**, Porto Alegre, v. 1, n. 1, p. 27, ago./2019. Disponível em: http://www.repositorio.jesuita.org.br/bitstream/handle/UNISINOS/8949/Deivid%20Augusto%20Werner_.pdf?sequence=1&isAllowed=y. Acesso em: 31 jan. 2022.

WESTPHAL, Vera Herweg. A Individualização em Ulrich Beck: análise da sociedade contemporânea. **Emancipação**, Ponta Grossa, v. 10, n. 2, p. 428, set./2010. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=3628981>. Acesso em: 11 jun. 2021.

WYNSBERGHE, Aimee Van. Artificial intelligence:: From ethics to policy. **Panel for the Future of Science and Technology**, bruxelas, v. 1, n. 1, p. 7-8, jun./2020. Disponível em: [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU\(2020\)641507_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU(2020)641507_EN.pdf). Acesso em: 14 fev. 2022.

REFERÊNCIAS DIVERSAS

NOTÍCIAS E INTERNET:

AGÊNCIA BRASIL. Direitos Humanos devem ser respeitados. Disponível em: <https://agenciabrasil.ebc.com.br/direitos-humanos/noticia/2018-07/direitos-humanos-devem-ser-respeitados-na-internet-diz-onu>. Acesso em: 2 fev. 2022.

BIGDATACORP. **Entendendo a economia dos dados**. Disponível em: <https://bigdatacorp.com.br/entendendo-a-economia-dos-dados/>. Acesso em: 1 nov. 2021.

BRASIL DE FATO. **Vítimas do neoliberalismo, países pobres sofrem consequências nos sistemas de saúde**. Disponível em: <https://www.brasildefato.com.br/2020/06/09/vitimas-do-neoliberalismo-paises-pobres-sofrem-consequencias-nos-sistemas-de-saude>. Acesso em: 1 jan. 2022.

CETAX. **DATA ANALYTICS, BIG DATA, DATA SCIENCE – BLOG CETAX**. Disponível em: <https://www.cetax.com.br/blog/big-data/>. Acesso em: 1 nov. 2021.

EDITORA UNESP. **Confira o conceito de "esfera pública", de acordo com Giddens e Sutton**. Disponível em: <http://editoraunesp.com.br/blog/confira-o-conceito-de-esfera-publica-de-acordo-com-giddens-e-sutton->. Acesso em: 29 dez. 2021.

ESR. **O que é GPT-3 e os impactos do programa na área**. Disponível em: <https://esr.rnp.br/metodos-ageis-e-inovacao/o-que-e-gpt-3-e-os-impactos-do-programa-na-area/>. Acesso em: 1 nov. 2021.

ESTADO DE MINAS. **Levantamento aponta que maioria dos presos no Brasil são jovens, negros e pobres**. Disponível em: https://www.em.com.br/app/noticia/nacional/2015/06/23/interna_nacional,661171/levantamento-aponta-que-maioria-dos-presos-no-brasil-sao-jovens-negro.shtml. Acesso em: 28 dez. 2021.

EWEK. **Why Data is the New Uranium.** Disponível em: <https://www.eweek.com/big-data-and-analytics/why-data-is-the-new-uranium/>. Acesso em: 19 nov. 2021.

FORBES. **Data Isn't The New Oil.** Disponível em: <https://www.forbes.com/sites/theyec/2021/07/15/data-isnt-the-new-oil--time-is/?sh=12e347d535bb>. Acesso em: 15 nov. 2021.

GLOBO G1. **Proporção de negros nas prisões cresce 14% em 15 anos, enquanto a de brancos cai 19%, mostra Anuário de Segurança Pública.** Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2020/10/19/em-15-anos-proporcao-de-negros-nas-prisoos-aumenta-14percent-ja-a-de-brancos-diminui-19percent-mostra-anuario-de-seguranca-publica.ghtml>. Acesso em: 26 dez. 2021.

GOOGLE. **COMO O GOOGLE ANONIMIZA OS DADOS.** Disponível em: <https://policies.google.com/technologies/anonymization?hl=pt-BR>. Acesso em: 5 nov. 2021.

IBGE GOV. **O que é o PIB.** Disponível em: <https://www.ibge.gov.br/explica/pib.php>. Acesso em: 1 nov. 2021.

IBM. **Machine Learning e Ciência de dados com IBM Watson.** Disponível em: <https://www.ibm.com/br-pt/analytics/machine-learning>. Acesso em: 1 nov. 2021.

IBM. **Transformação digital empresarial com Inteligência Artificial.** Disponível em: <https://www.ibm.com/br-pt/analytics/journey-to-ai>. Acesso em: 1 dez. 2021.

IRIS. **Fake news, bots, mecanismos de filtragem e outras ameaças à democracia brasileira.** Disponível em: <https://irisbh.com.br/fake-news-bots-mecanismos-de-filtragem-e-outras-ameacas-a-democracia-brasileira/>. Acesso em: 3 jan. 2022.

JUSBRASIL. **Sociedade de risco: rumo a uma outra modernidade.** Disponível em: <https://ferreiramacedo.jusbrasil.com.br/artigos/160037557/sociedade-de-risco-rumo-a-uma-outra-modernidade>. Acesso em: 20 set. 2020.

MIT TECHNOLOGY REVIEW. **AI is sending people to jail and getting it wrong.** Disponível em: <https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/>. Acesso em: 29 dez. 2021.

NATIONALGEOGRAPHIC. How a Color-Blind Artist Became the World's First Cyborg. Disponível em: <https://www.nationalgeographic.com/science/article/worlds-first-cyborg-human-evolution-science>. Acesso em: 15 mar. 2022.

SERPRO. LGPD: a versão brasileira do regulamento europeu. Disponível em: <https://www.serpro.gov.br/lgpd/noticias/lgpd-versao-brasileira-gdpr-dados-pessoais>. Acesso em: 21 mar. 2022.

SHOWMETECH. **IBM Watson: a inteligência artificial já substitui um médico?** Disponível em: <https://www.showmetech.com.br/ibm-watson-substitui-medico/>. Acesso em: 1 nov. 2021.

SIMPLY. **Big Techs: o que são e seu impacto no mercado financeiro.** Disponível em: <https://blog.simply.com.br/big-techs-o-que-sao-e-seu-impacto-no-mercado-financeiro/>. Acesso em: 11 nov. 2021.

STJ. Comissão de juristas inicia audiências públicas para discutir regulação da inteligência artificial. Disponível em: <https://www.stj.jus.br/sites/porta/p/Paginas/Comunicacao/Noticias/28042022-Comissao-de-juristas-inicia-audiencias-publicas-para-discutir-regulacao-da-inteligencia-artificial-.aspx>. Acesso em: 23 mar. 2022.

THE GUARDIAN. **The meaning of life in a world without work.** Disponível em: <https://www.theguardian.com/technology/2017/may/08/virtual-reality-religion-robots-sapiens-book>. Acesso em: 3 jan. 2021.

UOL. **Tecnologia é matriz do risco, diz sociólogo**. Disponível em: <https://www1.folha.uol.com.br/fsp/ciencia/fe2011200101.htm>. Acesso em: 12 out. 2021.

VÍDEOS:

Entrevista com Yuval Noah Harari, 2019. 1 vídeo (16:40 min - 18:30 min). Publicado pelo canal Roda Viva. Disponível em: <https://www.youtube.com/watch?v=pBQM085IxOM&t=2208s>

RELATÓRIOS, RECOMENDAÇÕES E CARTILHAS:

(EDPS), T. E. D. P. S. Towards a new digital ethics: Data, dignity and technology. **European Data Protection Supervisor**, set. /2015. Disponível em: https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_en.pdf. Acesso em: 5 nov. 2021.

(EDPS), T. E. D. P. S. Towards a new digital ethics: Data, dignity and technology. **The European Data Protection Supervisor (EDPS)**, abr./2015. Disponível em: https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_en.pdf. Acesso em: 11 nov. 2021.

CIÊNCIA, Ministério Da; INOVAÇÕES, Tecnologia E. Estratégia Brasileira de Inteligência Artificial. **EBIA**, jan./2021.

COMMISSION'S, The European. A DEFINITION OF AI: MAIN CAPABILITIES AND SCIENTIFIC DISCIPLINES. **HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE**, Bruxelas. Disponível em: https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf. Acesso em: 5 dez. 2021.

DATAPRIVACYBR. Contribuições do Data Privacy Brasil ao Projeto de Lei nº 21, de 04 de fevereiro de 2020: Nota técnica. Data Privacy Research, São Paul.

Disponível em: https://www.dataprivacybr.org/wp-content/uploads/2021/09/dpbr_notatecnica_pl21.pdf. Acesso em: 13 mar. 2022.

OECD. Recommendation of the Council on Artificial Intelligence. **OECD**, Lisboa, mai./2019. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 1 nov. 2021.

LEIS, TRATADOS E PROPOSTAS LEGISLATIVAS:

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Assembleia Geral. A/RES/37/7. 28 de outubro de 1982. Disponível em: https://www.soas.ac.uk/cedep-demos/000_P514_IEL_K3736-Demo/treaties/media/1982%20UN%20World%20Charter%20for%20Nature%201982.pdf Acesso: 17 mar.2022.

Wingspread Statement on the Precautionary Principle, Jan. 1998 Disponível em: <https://static1.squarespace.com/static/5ad8bb3336099bd6ed7b022a/t/5cc2113553450acab6d955a1/1556222261401/Wingspread+Statement.pdf>

Declaração do Rio Sobre Meio Ambiente e Desenvolvimento, Jun. 1992. Disponível em: https://cetesb.sp.gov.br/proclima/wp-content/uploads/sites/36/2013/12/declaracao_rio_ma.pdf

Comunicação da Comissão das Comunidades Europeias relativa ao princípio da precaução, Bruxelas, 2/02/2000. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52000DC0001&from=LT>

BRASIL (2020) – SENADO FEDERAL, projeto de lei 21/2020, Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil; e dá outras providências. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9063365&ts=1653621171778&disposition=inline>