

# **VII ENCONTRO VIRTUAL DO CONPEDI**

## **DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS II**

**EDSON RICARDO SALEME**

**YURI NATHAN DA COSTA LANNES**

**RONALDO FENELON SANTOS FILHO**

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

**Diretoria - CONPEDI**

**Presidente** - Profa. Dra. Samyra Haydêe Dal Farra Naspolini - FMU - São Paulo

**Diretor Executivo** - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

**Vice-presidente Norte** - Prof. Dr. Jean Carlos Dias - Cesupa - Pará

**Vice-presidente Centro-Oeste** - Prof. Dr. José Querino Tavares Neto - UFG - Goiás

**Vice-presidente Sul** - Prof. Dr. Leonel Severo Rocha - Unisinos - Rio Grande do Sul

**Vice-presidente Sudeste** - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio - Rio de Janeiro

**Vice-presidente Nordeste** - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

**Representante Discente:** Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

**Conselho Fiscal:**

Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão

Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais

Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte

Prof. Dr. Fernando Passos - UNIARA - São Paulo

Prof. Dr. Edinilson Donisete Machado - UNIVEM/UENP - São Paulo

**Secretarias**

**Relações Institucionais:**

Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná

Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia

Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal

**Comunicação:**

Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina

Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul

Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe

**Relações Internacionais para o Continente Americano:**

Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do sul

Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão

Prof. Dr. Felipe Chiarello de Souza Pinto - UPM - São Paulo

**Relações Internacionais para os demais Continentes:**

Profa. Dra. Gina Vidal Marcílio Pompeu - UNIFOR - Ceará

Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul

Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI - Santa Catarina

**Eventos:**

Prof. Dr. Yuri Nathan da Costa Lannes - FDF - São Paulo

Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina

Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

**Membro Nato** - Presidência anterior Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

D597

Direito, governança e novas tecnologias II [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Edson Ricardo Saleme; Ronaldo Fenelon Santos Filho; Yuri Nathan da Costa Lannes – Florianópolis: CONPEDI, 2024.

Inclui bibliografia

ISBN: 978-85-5505-891-2

Modo de acesso: [www.conpedi.org.br](http://www.conpedi.org.br) em publicações

Tema: A pesquisa jurídica na perspectiva da transdisciplinaridade

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Nacionais. 2. Direito. 3. Governança e novas tecnologias. VII Encontro Virtual do CONPEDI (1: 2024 : Florianópolis, Brasil).

CDU: 34



## **VII ENCONTRO VIRTUAL DO CONPEDI**

### **DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS II**

---

#### **Apresentação**

O VII ENCONTRO VIRTUAL DO CONPEDI, realizado em parceria com a Faculdade de Direito de Franca (FDF), com a Universidade UNIGRANRIO - Afya, com o Portugalense Institute For Legal Research - IJP e a Facultad de Derecho da Universidad de la República Uruguaye, entre os dias 24 e 28 de junho de 2024, apresentou como temática central “A Pesquisa Jurídica na Perspectiva da Transdisciplinaridade”. Esta questão suscitou intensos debates desde o início e, no decorrer do evento, com a apresentação dos trabalhos previamente selecionados, fóruns e painéis que no ambiente digital ocorreram.

Os trabalhos contidos nesta publicação foram apresentados como artigos no Grupo de Trabalho “DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS II”, realizado no dia 27 de junho de 2024, que passaram previamente por no mínimo dupla avaliação cega por pares. Encontram-se os resultados de pesquisas desenvolvidas em diversos Programas de Pós-Graduação em Direito, que retratam parcela relevante dos estudos que têm sido produzidos na temática central do Grupo de Trabalho.

As temáticas abordadas decorrem de intensas e numerosas discussões que acontecem pelo Brasil, com temas que reforçam a diversidade cultural brasileira e as preocupações que abrangem problemas relevantes e interessantes, a exemplo do direito digital, proteção da privacidade, crise da verdade, regulamentação de tecnologias, transformação digital e Inteligência artificial, bem como políticas públicas e tecnologia.

Espera-se, então, que o leitor possa vivenciar parcela destas discussões por meio da leitura dos textos. Agradecemos a todos os pesquisadores, colaboradores e pessoas envolvidas nos debates e organização do evento pela sua inestimável contribuição e desejamos uma proveitosa leitura!

Prof. Dr. Edson Ricardo Saleme – UNISANTOS

Prof. Dr. Yuri Nathan da Costa Lannes – FDF

Prof. Dr. Ronaldo Fenelon Santos Filho

# **A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL BASEADA NO RISCO: LIMITAÇÕES E PROPOSTAS COMPLEMENTARES PARA A PROTEÇÃO DOS DIREITOS HUMANOS**

## **RISK-BASED REGULATION OF ARTIFICIAL INTELLIGENCE: LIMITATIONS AND COMPLEMENTARY PROPOSALS FOR THE PROTECTION OF HUMAN RIGHTS**

**Marco Antonio Compassi Brun** <sup>1</sup>  
**Fernanda Viero da Silva** <sup>2</sup>  
**Tamara Cossetim Cichorski** <sup>3</sup>

### **Resumo**

O presente artigo objetiva estudar a regulação da inteligência artificial (IA) baseada no risco, conforme proposta do Regulamento da União Europeia para a Inteligência Artificial (EU AI Act). Parte da hipótese de que a regulação baseada no risco é um modelo que oferece suporte contributivo nesta complexa tarefa de regular os sistemas de IA, mas que, todavia, possui limitações, falhas e lacunas em relação aos danos individualizados e contextuais, o que pode aprofundar violações aos direitos humanos a partir do uso e aplicação desta nova tecnologia. A partir disso, exploram-se os avanços e os riscos produzidos pelos sistemas de IA ao longo das últimas décadas. Assim como, se analisa a regulação baseada no risco, sua contribuição como modelo regulatório da IA, a proposta oferecida pela União Europeia e as suas limitações que demandam propostas complementares. Metodologicamente, trata-se de pesquisa exploratória, com procedimento hipotético-dedutivo, abordagem qualitativa e transdisciplinar e técnica de pesquisa de revisão bibliográfica.

**Palavras-chave:** Inteligência artificial, Direitos humanos, Regulação baseada no risco, Regulação da ia, Teoria do risco

### **Abstract/Resumen/Résumé**

The aim of this article is to study the risk-based regulation of artificial intelligence (AI), as proposed in the European Union Regulation on Artificial Intelligence (EU AI Act). It starts from the hypothesis that risk-based regulation is a model that offers contributory support in this complex task of regulating AI systems, but which nevertheless has limitations, flaws and

---

<sup>1</sup> Mestrando em Direitos Humanos (PPGD-UNIJUI) – Bolsa Gratuidade. Graduado em Direito pela UNIJUI. Especialista em Proteção de Dados: LGPD & GDPR pelo FMP e Universidade de Lisboa. E-mail: marcoantonio\_brun@outlook.com

<sup>2</sup> Doutoranda em Direito pelo PPGD da UNIJUI. Bolsista CAPES/PROSUC e integrante do Grupo de Pesquisa Direitos Humanos, Democracia e Tecnologias de Informação e Comunicação. Email: fefeviero@gmail.com.

<sup>3</sup> Mestranda em Direitos Humanos do PPGD - UNIJUI, com Bolsa Gratuidade. Pós graduada em Direito Imobiliário. Endereço eletrônico: tamaracossetim@gmail.com.

gaps in relation to individualized and contextual damage, which can deepen violations of human rights from the use and application of this new technology. It explores the advances and risks produced by AI systems over the last few decades. It also analyzes risk-based regulation, its contribution as a regulatory model for AI, the proposal offered by the European Union and its limitations, which call for complementary proposals. Methodologically, this is exploratory research, with a hypothetical-deductive procedure, a qualitative and transdisciplinary approach and a literature review research technique.

**Keywords/Palabras-claves/Mots-clés:** Artificial intelligence, Human rights, Risk-based regulation, Regulation of ai, Risk theory

## 1 INTRODUÇÃO

Os sistemas de inteligência artificial (IA) se tornaram presença constante na vida em sociedade. A sua capacidade de produzir, decidir, recomendar, elaborar previsões e resolver problemas em alta velocidade, com eficiência, com considerável acurácia e, em muitos casos, com redução de custos, impulsionou a utilização e aplicação de ferramentas baseadas em técnicas de IA no mercado.

No entanto, o rápido avanço da IA também levanta questões éticas, sociais e, inclusive, de direitos humanos. Especialmente diante das falhas observadas com o crescente uso e aplicação da tecnologia. As quais têm sido respondidas através — mas não apenas — da regulação baseada no risco, como a notória proposta da União Europeia.

Não obstante, apesar de a regulação baseada no risco possibilitar um modelo preventivo aos sistemas de IA, esta apresenta limitações quanto a proteção contra danos não quantificáveis, subjetivos e contextuais, os quais são geralmente representados pelos mais sensíveis e íntimos, como com discriminações e violações à privacidade e à dignidade da pessoa humana. Diante disso, surge o problema deste artigo a partir do questionamento: como dar andamento ao processo de regulação da IA baseada no risco e, ao mesmo tempo, proteger os direitos humanos?

Diante disso, em hipótese, entende-se que a regulação baseada no risco é um modelo que oferece suporte contributivo nesta complexa tarefa de regular os sistemas de IA, visto que fornece medidas de gerenciamento dos riscos e dos danos tanto para a proteção coletiva, quanto para o controle e responsabilização de *big techs* e corporações envolvidas direta e indiretamente nos processos de construção e fornecimento de máquinas inteligentes. Entretanto, o modelo regulatório comumente proposto possui limitações na sua precisão e, conseqüentemente, na prevenção contra violações individualizadas e contextuais dos direitos humanos, como as discriminações algorítmicas. Reconhecer essas deficiências é primordial, assim como aceitar que determinadas ferramentas de IA não devem ser toleradas, especialmente as lesivas ao direito à privacidade e ao princípio da dignidade da pessoa humana; ainda que o fomento à inovação e ao desenvolvimento tecnológico são objetivos a serem perseguidos.

O artigo fixou, como objetivo geral, a intenção de analisar a regulação da IA baseada no risco, bem como as suas limitações de precisão e de abrangência, com a construção de possíveis estruturas alternativas que impeçam a violação individualizada e contextual contra os direitos humanos. Já para o desenvolvimento do tema, há dois objetivos específicos, com um capítulo correspondente para cada. Desse modo, o primeiro capítulo aborda o estágio

tecnológico alcançado pelos sistemas de IA, seus avanços e seus riscos. Enquanto, o segundo capítulo discorre sobre a regulação baseada no risco do Regulamento da União Europeia para a Inteligência Artificial (*EU AI Act*), com propostas complementares para este modelo.

A metodologia aplicada é a de pesquisa exploratória, com método hipotético-dedutivo no procedimento, abordagem qualitativa e transdisciplinar, bem como revisão bibliográfica como técnica de absorção do conteúdo pesquisado.

## 2 O AVANÇO E OS RISCOS DOS SISTEMAS DE IA

Os sistemas de inteligência artificial (IA) são, desde muito tempo, uma das mais frequentes e relevantes pautas do avanço tecnológico, da ciência e dos mundos da fantasia e da ficção. Há quase um século, por exemplo, Fritz Lang, com *Metropolis* (1927), inaugurava a temática na sétima arte, a qual se tornou um tópico presente no cinema. O seu desenvolvimento criou, inclusive, figuras marcantes como o robô<sup>1</sup> HAL 9000, na adaptação da obra de Arthur C. Clarke *2001: A Space Odyssey* (1968), feita pelo diretor Stanley Kubrick.

Em mesmo sentido, a literatura também contribuiu com a imaginação popular e, principalmente, de entusiastas do futuro sobre a convivência entre máquinas inteligentes e humanos. Na obra “Eu, robô”, de Isaac Asimov (2014), esse cenário é ricamente detalhado e trabalhado, ainda que de modo puramente ficcional. O enredo aborda, dentre uma série de outros eventos, uma época em que a tecnologia de sistemas de IA atingiram níveis suficientemente capazes de tornar o robô um ente integralmente ativo na sociedade, com leis próprias e bem definidas, na intenção primordial de cooperar com o homem.

Esses elementos culturais e muitos outros inspiraram — e ainda inspiram — a vida real. Nesse contexto, ainda nos anos 50 do século XX, McCarthy *et al.* (1955) trouxeram as discussões para o campo da pesquisa científica, com estudos relativos à capacidade de reprodução das redes neurais humanas em máquinas. A abordagem ocorreu pela publicação do artigo “Uma Proposta para o Projeto de Pesquisa de Verão de Dartmouth sobre Inteligência Artificial”<sup>2</sup> e revelou um amplo campo do conhecimento a ser explorado e preenchido nas décadas seguintes.

---

<sup>1</sup> A palavra robô, na presente pesquisa, será utilizada no sentido de *softwares* e *hardwares* inteligentes, ou, para simplificar, máquinas dotadas — ou com utilização de métodos em pelo menos um nível — de inteligência artificial.

<sup>2</sup> Traduziu-se, no original: “A Proposal for the Dartmouth Summer Project on Artificial Intelligence”.

Ainda assim, apesar de o avanço teórico, ficcional e até mesmo filosófico ter iniciado há considerável tempo, os sistemas de IA apenas tiveram sua aplicação prática nas últimas décadas. Isto é, a partir do momento em que a tecnologia disponível — especialmente de *hardware*<sup>3</sup> — se tornou suficiente para processar as informações, os dados e, em essência, viabilizar o funcionamento e a programação das técnicas que compõem o campo da IA. Desse modo, a evolução, o alcance e as possibilidades oferecidas através das ferramentas inteligentes estão intimamente ligadas com o aprimoramento do poder computacional, o qual acontece com cada vez mais velocidade e potência.

Com a intensiva inserção da IA no cotidiano humano dos últimos anos, também surgiram as demandas de conceituação da referida área do conceito. Nesse sentido, segundo Müller (2023, p. 2), é possível definir o termo IA por duas visões distintas: i) a IA clássica; e ii) a IA técnica. A primeira (i), desse modo, deriva diretamente da pesquisa de Dartmouth, em 1956, e se refere à IA como um “[...] programa de pesquisa para criar agentes computadorizados que tenham inteligência”<sup>4</sup> (Müller, 2023, p. 2). Enquanto na ciência da computação, tecnicamente (ii), a IA é citada como um “[...] conjunto de métodos da ciência da computação para percepção, modelagem, planejamento e ação [...]”<sup>5</sup> (Müller, 2023, p. 2).

No caso do Regulamento da União Europeia para a Inteligência Artificial (*EU AI Act*)<sup>6</sup>, discutido e proposto entre 2020 e 2021 e já aprovado pelo Parlamento Europeu (Lopes, 2023), seu Artigo 3º estabelece que o sistema de IA pode ser definido como um programa informático desenvolvido por uma ou várias técnicas e abordagens. Com capacidade, através de um conjunto de objetivos determinados por seres humanos, criar resultados, previsões, recomendações ou decisões.

As conceituações exploradas demonstram, ainda que por diferentes perspectivas, que a IA é um grande campo de estudo. Isto é, um sistema que absorve diversas técnicas computacionais, as quais, por múltiplos métodos, conseguem reproduzir — mesmo que parcialmente — atuações que, até então, apenas eram possíveis por humanos. Especificamente, pelo cérebro humano. Inclusive, conforme Aguirre-Celis e Miikkulainen (2023, p. 416), o funcionamento do cérebro é, desde o princípio da IA, a principal fonte de inspiração para o

---

<sup>3</sup> *Hardware*s se referem aos equipamentos. Ou seja, aos itens físicos e tangíveis de um dispositivo (Antunes, 2019, p. 143), os quais são compostos de peças e componentes para hospedar o funcionamento dos softwares.

<sup>4</sup> Traduziu-se, no original: “[...] AI is a research program to create computer-based agents that have intelligence”.

<sup>5</sup> Traduziu-se, no original: “[...] set of computer-science methods for perception, modelling, planning, and action [...]”.

<sup>6</sup> Em inglês, *European Union Artificial Intelligence Act*.

desenvolvimento dos seus mecanismos. E isso fez com que, atualmente, redes neurais artificiais consigam performar em mesmo nível, ou melhor que a mente do homem em tarefas específicas. Sobretudo, em atividades de repetição, com regras bem definidas e um propósito de resultado objetivamente claro para a máquina (Aguirre-Celis; Miikkulainen, 2023, p. 416).

No entanto, isso não significa que a IA e o cérebro humano funcionem de maneiras similares. Pelo contrário, Aguirre-Celis e Miikkulainen (2023, p. 416) esclareceram que o comportamento maquínico é consideravelmente simples em comparação ao funcionamento cerebral de uma pessoa. Assim como nas maneiras de aprender e de pensar, tornando o humano o único capaz de refletir criticamente, possuir compreensões gerais, atuar com complexidade em variadas funções, realizar ponderações (Nieva-Fenoll, 2018, p. 116), entre muitas outras características — especialmente as de natureza subjetiva e com graus emocionais.

Em contrapartida, os robôs superam os humanos em atribuições de caráter repetitório e objetivo, como a classificação e identificação de imagem (incluindo exames médicos), em jogos como Go e Xadrez — para citar algumas (Aguirre-Celis; Miikkulainen, 2023, p. 416). Nesses contextos, as máquinas não apenas são mais céleres, como a acurácia é, geralmente, maior que a de um especialista na área. Além de que os sistemas de IA são isentos de cansaço, estresse e vários outros fatores físicos e psicológicos que tornam o trabalho exercido por pessoas mais desgastante e, no comparativo, menos eficiente.

Essas possibilidades alcançadas pela IA são garantidas a partir de suas técnicas, dentre as quais estão o aprendizado de máquina (*machine learning*) e o aprendizado profundo de máquina (*deep learning*). No aprendizado de máquina, permite-se o processamento de grandes volumes de dados em pouco tempo (Bharadiya; Thomas; Ahmed, 2023, p. 87). Assim, esses algoritmos aprendem a partir das informações coletadas e, rapidamente, conseguem inferir, determinar e reconhecer padrões (Mourby; Cathaoir; Collin, 2021, p. 2; Nicholas, 2020, p. 714), facilitando a localização de soluções e saídas para diferentes problemáticas.

Já para a identificação de amostras e conceitos em camadas ainda mais profundas de dados e em mais de um nível hierárquico, utiliza-se a técnica do aprendizado profundo de máquina. Nesses casos, há o emprego das redes neurais artificiais para, com isso, empreender sucesso no acúmulo de experiência e conhecimento a partir do banco de informações que lhe foi alimentado. É, essencialmente, por meio desses recursos — e de outros não listados — que são contemporaneamente reconhecidos rostos de pessoas em segundos, com reconhecimento facial (Zalnierute, 2021, p. 290).

Válido mencionar, ainda, outra subárea da IA chamada de processamento de linguagem natural (*natural language processing – NLP*), a qual estuda a capacidade de compreensão, fala,

leitura e escrita da língua humana. Potencializado pelos modelos de linguagem *LLM* (*large language model*), que permitem a complexa geração de texto a partir de poucas informações ou intervenções pelo usuário humano (Coeckelbergh; Gunkel, 2023). Desse modo, foram esses os modelos que proporcionaram o desenvolvimento de ferramentas como o *ChatGPT* e de demais aplicações disponíveis para a elaboração e criação de imagens, desenhos artísticos, vídeos, músicas e diversas outras funcionalidades. Tecnologias de IA conhecidas como generativas (Epstein *et al.*, 2023, p. 3).

Por conseguinte, a partir das técnicas apresentadas — e outras —, bem como o potencial na produção de resultados pela aplicação de sistemas de IA, a tornaram ubíqua. Desse modo, a tecnologia tem se inserido nas mais diversas áreas do conhecimento e do trabalho nos últimos anos, com a automatização de atividades repetitivas e com o auxílio, apoio e recomendação em tarefas mais densas e delicadas. Campos como o da medicina, do direito, do design, da arquitetura, das engenharias e das ciências naturais e da computação, para citar alguns, já possuem considerável índice de utilização de ferramentas baseadas em IA. Não necessariamente com a substituição do profissional pela máquina, mas, principalmente, como mecanismo de assistência — direta ou indireta — nas tomadas de decisões e demais funções que até pouco tempo atrás eram exercidas apenas por humanos.

Ademais, para além das áreas do conhecimento, os sistemas de IA também têm revolucionado o cenário dos negócios e, como apontam Bharadiya, Thomas e Ahmed (2023, p. 86), modificado as interações e os relacionamentos sociais e econômicos. Em razão disso, Kaminski (2023, p. 104) observa que a utilização de máquinas inteligentes tem crescido em companhias e governos em uma variada gama de aplicações, *e.g* no sistema de saúde, de moradias, oportunidades educacionais e muito mais.

Entretanto, a natural movimentação de adaptação a evolução tecnológica no máximo de setores possíveis, para, essencialmente, aproveitar a oferta de eficiência, baixos custos e recomendações precisas nas tomadas de decisões por robôs, tem levantado preocupações por conta dos riscos que podem ser cometidos pela IA.

Kaminski destaca (2023, p. 110), a partir disso, que sistemas de IA não só podem falhar, como, comprovadamente, falham. Um *software*<sup>7</sup> pode apresentar erros, resultados podem ser imprevisíveis e irracionais, bem como, um robô pode performar com excelência em uma situação, mas fracassar gravemente em outra (Kaminski, 2023, p. 104), com danos e consequências de diferentes magnitudes.

---

<sup>7</sup> *Softwares* são programas de computador que integram a estrutura física (*hardware*) (Antunes, 2019, p. 143).

A fim de exemplificar a abrangência, a gravidade e a proporção dos danos gerados através das mais diversas formas de uso de sistemas de IA, Kaminski (2023, p. 110) designa três casos de naturezas distintas: i) IA e segurança: veículos autônomos; ii) IA e recrutamento de funcionários; e iii) IA e saúde pública: acesso a cuidados e prescrições.

O primeiro relato (i) narra o primeiro acidente de trânsito fatal cometido por um veículo autônomo. Em março de 2018, uma SUV da Volvo estava em movimento para testes na cidade de Tempe, Arizona (EUA), com uma operadora de segurança que trabalhava para a *Uber*. A motorista humana acompanhava de dentro do veículo para agir em situações de correção de percurso e de movimento ou em emergências. Apesar disso, o programa instalado no automóvel falhou em detectar uma pedestre de 49 anos a tempo de evitar o impacto que ocasionou a morte daquela. A pessoa responsável por acompanhar o veículo também não percebeu por distração momentânea e excesso de confiança na máquina (Kaminski, 2023, p. 111).

O segundo caso de estudo (ii) aborda a ferramenta automatizada de recrutamento de funcionários que tem se popularizado entre grandes empresas como o *McDonald's*, entre outras. No entanto, esses sistemas de IA, da mesma forma, apresentaram problemas. Kaminski (2023, p. 114) detalha que, em 2014, a *Amazon* desenvolveu um programa com essas características para “caçar talentos”. Todavia, no ano seguinte, a equipe notou que o robô era tendencioso contra mulheres e desclassificava os currículos de candidatas femininas. Isso ocorria, pois, a máquina replicava o viés do conjunto de dados com o qual foi treinada, visto que o histórico de contratações da empresa para cargos bem-sucedidos era majoritariamente masculino.

Esses exemplos práticos refletem o que se conhece por viés algorítmico. Ou seja, quando há ausência de neutralidade nessas ferramentas, ao contrário do que se espera de decisões automatizadas. Isso ocorre, pois, tanto os dados que servem de alimentação para o seu efetivo funcionamento, quanto os modelos matemáticos desenvolvidos por programadores, foram gerados e/ou elaborados por humanos (Schwede, 2023, p. 36-37). Logo, carregam e reproduzem suas características e valores, inclusive as de cunho discriminatório, racista, misógino, xenofóbico — para citar alguns. Assim, sistemas inteligentes tendem a reproduzir preconceitos historicamente perpetuados (Cebral-Loureda; Rincón-Flores; Sanchez-Ante, 2023, p. 2).

O terceiro apontamento de Kaminski (2023, p. 118) também retrata o enviesamento de decisões, porém no sistema de saúde (iii). Na situação trabalhada, pesquisadores identificaram, em 2019, que o algoritmo criado para identificar pacientes de alto risco para o direcionamento de recursos adicionais no gerenciamento de cuidados demonstrou erros no seu funcionamento. Em suas análises, o programa ignorava erroneamente pessoas negras, pois a máquina havia compreendido em seu treinamento com dados que, uma vez que aqueles geravam custos

menores em comparação com pacientes brancos — resultados da desigualdade social —, então automaticamente teriam menores riscos de saúde. Logo, podiam ser — equivocadamente — reduzidos pela filtragem computacional.

A menção aos cenários de danos já identificados nos últimos anos de crescente aplicação e uso dos sistemas de IA, além de muitos outros de natureza física e psicológica, com impactos que vão desde a segurança pública, até a privacidade e a dignidade da pessoa humana, alertam para as consequências da automatização, do uso excessivo de dados e da inserção de máquinas inteligentes sem prevenções ou delimitações.

Esses efeitos — junto dos danos — observados na transformação tecnológica promovida pela IA, integram, como um de seus elementos, o risco. Teoria que encontra um pilar de fundamentação no desenvolvimento feito por Beck (2011, p. 39). Assim, para o autor, os riscos vão além dos resultados deletérios percebidos entre o passado e o presente, eis que também são um “ainda-não” evento, ou seja, um componente futuro. Sendo essa uma das principais razões pelas quais os riscos não se limitam apenas aos danos e efeitos já ocorridos.

Nesse sentido, para Beck (2011, p. 39), riscos são uma extensão futura das consequências atualmente previsíveis. Por isso, é fundamental interpretar o seu conceito não apenas sob o aspecto de real, mas, também, de irreal — simultaneamente. Primeiramente, o real está associado aos danos identificados, aos impactos que fizeram ou fazem parte do contexto social. No caso dos sistemas de IA, são as falhas e violações já reportadas, como algumas das demonstradas por Kaminski, além de diversas outras relacionadas ao viés algorítmico; aos erros pontuais em tomadas de decisões delicadas que geram prejuízos aos usuários ou a um terceiro; às lesões físicas decorrentes de acidentes por problemas de operação ou programação; assim como na destinação abusiva de *softwares* de vigilância por governos ou empresas privadas; na substituição desarrazoada no mercado de trabalho; na transgressão da propriedade intelectual; na propagação de *fake news*, criações de *deep fakes* e métodos similares que atingem as democracias — dentre inúmeros outros.

Por outro lado, os riscos irreais são visualizados como ameaças projetadas no futuro, isto é, são os prognósticos e, por vezes, as próprias especulações e conjecturas, que cercam o cenário social. São nesses que, segundo Beck (2011, p. 40), reside a força do argumento do risco. Isso, pois, além de se esperar que os danos sejam escalonados e exponenciais e, portanto, quanto mais distantes, mais perigosos e destrutivos, também há a condição intrínseca de futuro. Em essência, este é virtual, uma construção “prognosticada”, porém fictícia. E, ainda que venha a se concretizar na sua totalidade, até que isso aconteça, os seus eventos são inexistentes, ou seja, fazem parte de um tempo ainda não alcançado, logo, sem reflexos materiais imediatos.

No entanto, a força social descrita por Beck (2011, p. 40), está presente nos riscos irreais justamente por estes representarem o catalisador que move as atuações no presente. Já que o teor da ameaça e, principalmente, a sua incalculabilidade e imprevisibilidade, moldam o núcleo da consciência do risco, refletindo nas investidas preventivas e de mitigação de danos. Tais como as propostas de legislações e regulações; os pactos, encontros e convenções internacionais; as declarações e alerta de especialistas — para citar alguns.

Inclusive, é a partir dessas premissas que a União Europeia tem se mobilizado nos últimos anos para construir uma estrutura regulatória que consiga gerenciar os riscos oferecidos pela IA (Schuett, 2023, p. 1), de modo a eliminar ou — ao menos — atenuar consequências deletérias.

### **3 A REGULAÇÃO BASEADA NO RISCO: PROPOSTAS COMPLEMENTARES**

O regulamento denominado de *EU AI Act* deve entrar em vigência em 2024 e, de acordo com Schuett (2023, p. 1), servirá de referência para outros países como Estados Unidos da América e Reino Unido, semelhante aos efeitos do Regulamento Geral de Proteção de Dados – RGPD (*General Data Protection Regulation – GDPR*) (2016/679). O qual, inclusive, foi uma das principais fontes de inspiração para a Lei Geral de Proteção de Dados Pessoais – LGPD (Lei n.º 13.709/2018), motivo pelo qual é provável que o cenário se repetirá acerca da IA. A qual possui incipiente cenário regulatório no Brasil, com destaques para o Projeto de Lei n.º 759/2023 e para o Projeto de Lei n.º 2338/2023, ambos em tramitação.

Especificamente sobre o *EU AI Act*, o modelo de regulação adotado foi o baseado no risco. Isso decorre da adequação dos sistemas de IA para intervenções preventivas, visando antever minimamente as consequências de resultados muitas vezes imprevisíveis ou extremamente complexos, características que geralmente são atendidas por essa estrutura regulatória (Kaminski, 2023, p. 127). Nesse sentido, o regulamento europeu categoriza os sistemas de IA em diferentes níveis de riscos, a partir de suas aplicações e utilizações práticas. Com isso, proíbe aqueles que demonstrem riscos inaceitáveis; impõe requisitos específicos aos que apresentem alto risco; e, para os de riscos baixos ou mínimos, permite maior liberdade, mediante orientações e diretrizes (Schuett, 2023, p. 4).

Dessa forma, as práticas de IA descritas no Título II do ainda não definitivo *EU AI Act* (2021) são expressamente proibidas, *e.g.* programas de reconhecimento facial em tempo real ou pontuação social (*social scoring*) — conforme Artigo 5º (Novelli *et al.*, 2023, p. 1). Já as de risco elevado, como máquinas destinadas ao uso educacional, empregatício, de migração, no

sistema de justiça, entre outros, devem cumprir uma série de requisitos estabelecidos no Artigo 8º e seguintes. Tais como a gestão de riscos, práticas de governança de dados, transparência, prestação de contas e supervisão humana — para citar apenas alguns.

O regulamento europeu para a IA possui diversas disposições para além das observadas que visam, justamente, abarcar os possíveis danos causados por robôs para impedi-los antes mesmo de sua consumação. Seja com a proibição, seja com a mitigação de sistemas propensos a causar prejuízos à saúde e à segurança física e psicológica dos usuários e de terceiros.

Apesar disso, Kaminski (2023, p. 152) detalha que modelos de regulação baseados no risco são destinados à proteção de danos quantificáveis. Isto é, consequências deletérias medidas por parâmetros objetivos. E esses fatores, embora obtenham êxito em muitos cenários, o tornam, por outro lado, limitado ao enfrentar a complexidade de alguns sistemas de IA. A autora aponta que isso pode levar a uma miopia regulatória, uma vez que os sistemas complexos têm maior probabilidade de sofrer com riscos imprevisíveis e catastróficos (Kaminski, 2023, p. 155).

O caráter imprevisível e incalculável do risco não é, no entanto, uma exclusividade da IA ou das novas tecnologias — apesar de ser amplificado por elas. Para Beck (2011, p. 32-33), essa condição está associada sistematicamente aos processos da modernização, os quais se estendem para a contemporaneidade. No entanto, pouco — ou não suficientemente — se vincula esses dois fatores. Não sem estabelecer uma explicação causal para que estes riscos sejam vistos “[...] como produto do modo de produção industrial [...]” (Beck, 2011, p. 33). Dessa forma, o autor (2011, p. 33) relata como os riscos socialmente reconhecidos estão conectados com o modelo de causa e efeito e que, a partir disso, tenta-se intrincar todo fenômeno nocivo com essa base, sem se considerar que algumas ameaças são completamente alheias em dimensão social, conteúdo, espaço ou tempo. Muito menos se sabe até onde podem surgir efeitos colaterais.

A imprevisibilidade dos muitos dos efeitos deletérios — agravada pelo aspecto temporal do risco — leva à análise contextual para um estágio de mera suposição causal, teórica, na qual é preciso muito mais se presumir o risco como verdadeiro, do que o perceber deste modo. O que reforça a condição invisível ou irreal do risco (Beck, 2011, p. 33). E, acima de tudo, o isola ao âmbito técnico-científico, enquanto se desprende da consciência social. Consequentemente, fortalece-se o que Beck (2011, p. 69) descreve como “[...] monopólio de racionalidade da definição científica do risco”. Cenário agravado pela equivocada pressuposição, baseada no argumento de autoridade da área do conhecimento — ou de líder político, ou de CEO de uma

*big tech* —, de que riscos podem ser objetivamente especificados, enquanto se ignora as suas minúcias e os seus reflexos na sociedade.

Esse isolamento técnico-científico com a perda de consciência social tem se revelado no campo da IA no decorrer de sua implantação e com o crescimento de sua aplicação. O repercutido caso COMPAS<sup>8</sup>, denunciado pela *ProPublica* (2016), permite analisar comparativamente o ponto teórico de Beck e a utilização prática da IA e dos algoritmos. Isso, pois, o COMPAS, sistema utilizado na justiça criminal em alguns locais dos Estados Unidos da América para detectar os riscos e probabilidades de um criminoso voltar a cometer um crime, a partir de questionamentos pessoais e com resultados capazes de influenciar diretamente a decisão do juiz, se mostrou falho, diante do viés algorítmico.

A constatação ocorreu mediante pesquisa investigativa realizada pelo jornalismo independente da *ProPublica*, na qual Angwin *et al.* (2016) demonstraram que réus negros recebiam pontuação elevada de risco de cometimento de novos crimes em relação aos réus brancos, ainda que em situações similares ou de menor reprobabilidade de conduta dos primeiros.

No entanto, ao integrar esses fatos aos apontamentos de Beck (2011, p. 69), realça-se não apenas o notório viés algorítmico que deve ser incisivamente enfrentado, mas, também, o fracasso da autoridade técnico-científica — e, aqui, representada, principalmente, por CEOs de *big techs*<sup>9</sup>, os quais detêm o monopólio do desenvolvimento de sistemas de IA — em perceber riscos. Parte disso ocorre pela insuficiência do modelo de divisão de trabalho ultraspecializado, abstinência de *práxis* e desligamento da realidade social; e, a outra parte, pela justificativa do avanço tecnológico (Beck, 2011, p. 55), a qual normaliza danos em camadas mais frágeis da sociedade em prol do crescimento e da produção.

Nesse sentido, segundo Beck (2011, p. 69), a ciência possibilita a estipulação de riscos, ainda que sem conseguir, singularmente, quantificá-los. Entretanto, é a população quem percebe riscos — e quem primeiro sofre com seus efeitos, a começar por aqueles historicamente discriminados. O caso COMPAS reflete esse cenário ao destacar que o preconceito transmitido pela máquina apenas foi considerado após a investigação social pela *ProPublica*. Isto é, o Poder Judiciário, enquanto utilizava a ferramenta, assim como a *Northpointe* — empresa que criou o *software* e treinou os algoritmos — assentiram, ainda que tacitamente, ao seu uso

---

<sup>8</sup> Sigla para *Correctional Offender Management Profiling for Alternative Sanctions*. Em português, Perfil de Gerenciamento de Infratores Correcional para Sanções Alternativas

<sup>9</sup> *Big techs* são as grandes empresas de tecnologia de informação e comunicação que dominam o mercado do desenvolvimento e da inovação, tais como: *Google, Apple, Meta, Microsoft e Amazon*.

preconceituoso. Com argumentos de soberania tecnológica e matemática da máquina, apesar da recusa da companhia em divulgar detalhes do seu sistema (Corberth-Davies *et al.*, 2016).

Essa espécie de indiferença em relação aos riscos por aqueles que estão no topo da cadeia hierárquica da sociedade, reflete a dinâmica econômica e de mercado. Assim, além de as consequências, efeitos e danos de uma nova ferramenta como a IA atingirem, inicialmente, as camadas mais baixas, o aparente desinteresse em mitigar riscos ocorre, também, por ser estratégico. Em essência, pois, os riscos também são oportunidades de mercado (Beck, 2011, p. 56). Assim, um sistema de IA — utilizado para vigilância, por exemplo —, por mais deletério que possa ser, ainda terá proveito com a coleta e venda de dados de usuários pelas *big techs*. Desse modo, segundo Beck (2011, p. 56), cria-se uma dinâmica de oposição que cresce proporcionalmente ao avanço da sociedade de risco, no qual alguns são afetados pelos riscos, enquanto outros lucram com eles.

Portanto, percebe-se uma relação estreita entre a produção de riqueza, típica da sociedade industrial, e a produção social de riscos (Beck, 2011, p. 23). Dessa forma, se na sociedade industrial ou de classes está o conceito de como distribuir riqueza de maneira desigual, sem perder a sua legitimidade. Na sociedade de riscos, o paradigma gira em torno do como distribuir e isolar efeitos colaterais decorrentes do processo de modernização e da produção acelerada sem que se cruze o limite do “aceitável” e “tolerável” — cientificamente falando.

A distribuição, desse modo, seja de riqueza, seja de riscos, obedece, ainda assim, o esquema de classes. A diferença sistemática, conforme aponta Beck (2011, p. 41) reside na inversão da acumulação, ou seja, as riquezas acumulam-se em cima, enquanto os riscos, embaixo. Por consequência, realça-se a atração entre pobreza extrema e riscos extremos. Logo, para Beck (2011, p. 41), os riscos não revogam, mas reforçam a sociedade de classes.

Nesse contexto, contemporaneamente, aqueles que detêm o poder de produção de sistemas de IA, conseqüentemente, são os últimos a serem atingidos por seus danos. O que garante a continuidade de um estágio de aprimoramento tecnológico acelerado e agressivo. Todavia, os riscos tendem a, cedo ou tarde, se sobrepôr à produção de riquezas, atingindo cada vez mais a população, em níveis proporcionalmente mais altos. Inclusive, aqueles que produziram e lucraram com os riscos, o que é nomeado por Beck (2011, p. 44) de “efeito bumerangue”.

Por conta disso, a sociedade de risco, apesar das relações e da necessária compreensão conjunta, não se limita ao conceito de sociedade de classe (Beck, 2011, p. 44). Justamente pela dinâmica evolutiva da primeira, a qual ultrapassa as fronteiras socioeconômicas e impõe à

humanidade a necessidade de se congregarem — ainda que minimamente —, para mitigar e atenuar os impactos advindos do desenvolvimento.

É nesse cenário de ampliação de riscos de magnitudes praticamente incalculáveis, que são construídas propostas como as do *EU AI Act*. Visto que as ameaças, inicialmente justificadas em razão do progresso tecnológico, perdem em tolerância com a sua crescente abrangência, não apenas contra diferentes classes, mas, também, em face de instituições e da democracia.

Apesar disso, embora a regulação oferecida pela União Europeia objetive controlar diferentes tipos de IA a partir de diferentes medidas, proporcionais ao seu grau ou potencial lesivo (Kaminski, 2023, p. 157), os parâmetros e os padrões técnicos do *EU AI Act* demonstram generalização excessiva e isso, de acordo com Novelli *et al.* (2023, p. 2), pode levar a uma avaliação com regras muito brandas — ou até muito rígidas. Sobretudo, para regular a IA generativa e técnicas de *LLM* (grandes modelos de linguagem), marcadas pela versatilidade e por respostas inesperadas.

A fim de atenuar os problemas aventados, Novelli *et al.* (2023, p. 2) propõem diferentes abordagens na teoria do risco inserida no regulamento europeu. No caso, trata-se da importação da teoria do risco ecológico, essencialmente dos modelos IPCC<sup>10</sup> (Painel Intergovernamental sobre Mudanças Climáticas) para riscos de mudanças climáticas. A justificativa advém das similaridades entre ambas as áreas, posto que sua complexidade é crescente, seus impactos são — em várias ocasiões — imprevisíveis e há alta dependência do contexto específico e das partes afetadas para a verificação dos fatores e magnitudes do risco.

Nesse sentido, o modelo IPCC classifica os riscos de mudanças climáticas como consequências de três componentes: (A) ameaças (*hazards*); (E) exposição (*exposure*); e (V) vulnerabilidade (*vulnerability*). As ameaças — também citadas como “perigos” — representam as fontes potenciais de danos. A exposição está vinculada ao que pode ser afetado pela fonte de perigo. Enquanto a vulnerabilidade associa-se aos atributos e circunstâncias que tornam os elementos expostos suscetíveis aos danos (Novelli, *et al.*, 2023, p. 3). Sugere-se, ainda, uma quarta determinante classificada como (R) resposta (*response*), a qual se refere às medidas que podem neutralizar ou atenuar os riscos. A partir desse esquema, entende-se que o risco geral é o resultado da interação entre as determinantes, os condutores e os riscos extrínsecos (Novelli, *et al.*, 2023, p. 3).

---

<sup>10</sup> A sigla “IPCC” significa *Intergovernmental Panel on Climate Change*, e foi traduzido para Painel Intergovernamental sobre Mudanças Climáticas, em português. Contudo, sigla original foi mantida.

Segundo Novelli *et al.* (2023, p. 3), o modelo supera a principal deficiência do *EU AI Act*, o qual considera tão somente o estágio estático do risco e suas determinantes, sem, todavia, as interações dos condutores (riscos intersetoriais). Com a aplicação do IPCC, consegue-se verificar não apenas os fatores, como os resultados dos seus cruzamentos, os quais podem ser agregados, compostos ou em cascata.

Para exemplificar, os autores aplicam o modelo em sistemas de IA generativa, ou seja, as de *LLM*, popularizadas por ferramentas como o *ChatGPT* (Novelli, *et al.*, 2023, p. 3). Nesse contexto, as ameaças (A) seriam representadas por *e.g.* opacidade, o tamanho do conjunto de dados para treinamento e a má qualidade destes. Por si só, apesar dos riscos técnicos demonstrados, os exemplos não conduzem a uma violação explícita de um direito fundamental. Contudo, quando esses fatores são combinados e se entrelaçam pelos seus condutores, a máquina pode gerar resultados discriminatórios — viés algorítmico (Novelli, *et al.*, 2023, p. 5). Consequência que se torna mais grave se se considerar, também, a determinante da vulnerabilidade (V) do indivíduo, com base em sua raça, etnia, gênero, classe, idade, etc. (Novelli *et al.*, 2023, p. 5-6).

Ainda na situação hipotética, os condutores de exposição (E) representam os valores que possivelmente seriam lesados com a aplicação de *LLMs*. Como a violação de direitos autorais dos dados de treinamento ou a privacidade dos titulares dos dados pessoais utilizados. Por conseguinte, para avaliar riscos dessa natureza, seria necessário ponderar e sopesar os direitos em perigo apontados com os benefícios oferecidos pelo uso da tecnologia (Novelli *et al.*, 2023, p. 6). Com isso, as medidas de resposta (R), como a governança, a qualidade na coleta de dados, a transparência nos procedimentos, o exame prévio de viés, supervisão humana, a privacidade diferencial, entre outras, se tornariam mais precisas no combate ao(s) dano(s), com maior probabilidade de êxito, por abranger as determinantes, os condutores e as suas interações setoriais e intersetoriais (Novelli *et al.*, 2023, p. 6).

#### **4 CONCLUSÃO**

A presente pesquisa abordou sobre os avanços e os riscos apresentados com o crescente uso e aplicação dos sistemas de IA. Nesse direcionamento, surgem, como resposta, regulações que objetivam o controle desta nova tecnologia. Dentre as propostas já concretizadas, o Regulamento da União Europeia para a Inteligência Artificial (*EU AI Act*) tem se mostrado a principal, em termos de completude e abrangência.

Apesar disso, o *EU AI Act* tem em seu modelo, a regulação baseada no risco, a qual, embora demonstre características que fazem sentido em uma tentativa de regular a IA, também demonstram falhas e lacunas em relação aos possíveis danos individualizados e contextuais, capazes de, inclusive, ferir direitos humanos.

Diante disso, realizadas as análises e o questionamento central, entende-se que a regulação baseada no risco pode ser continuada, se compreender suas falhas e lacunas, corrigindo-as a partir de alternativas viáveis e já demonstradas. Desse modo, conclui-se pela confirmação da hipótese com o reconhecimento da necessidade de proteção de direitos humanos mesmo com a presente proposta regulatória.

Conforme as pesquisas realizadas, pode-se observar a existência de alternativas e caminhos auxiliares para as limitações existentes na regulação baseada no risco. A demanda regulatória é urgente e a complicada tarefa de aplicação no âmbito tecnológico apenas expõe a complexidade dos sistemas de IA. Os quais, por um lado, solucionam problemas e aprimoram meios de solução, enquanto, por outro, como consequência, tornam ainda mais sensível e frágil a proteção de direitos e de garantias em face das grandes corporações de tecnologia (*big techs*), as quais detêm o monopólio da criação, produção e desenvolvimento de tais máquinas.

Portanto, é essencial que o ordenamento jurídico, na construção de uma estrutura regulatória de sistemas de IA, especialmente quando propenso a incorporar outras normas, como deve ser o caso do Brasil com o *EU AI Act*, esteja preparado não apenas para criar proteções de mercado e limitações industriais para as grandes empresas. Mas, principalmente, para construir arquiteturas alinhadas aos princípios e às garantias constitucionais, assim como aos direitos humanos. Para isso, é necessário que a regulação baseada no risco obtenha complementos de precisão, com a capacitação na identificação e avaliação de riscos individualizados e contextuais.

## 5 REFERÊNCIAS

2001: *A Space Odyssey*. Direção: Stanley Kubrick. Produção: Stanley Kubrick. [S.l.]: Metro-Goldwyn-Mayer. 1968.

AGUIRRE-CELIS, N.; MIIKKULAINEN, R. What AI can do for Neuroscience: understanding how the brain represents word meanings. In: CEBRAL-LOUREDA, M.; RINCÓN-FLORES, E. G.; SANCHEZ-ANTE, G. **What AI can do**: strengths and limitations of artificial intelligence. Boca Raton: CRC Press, Cap. 20, 2023.

ANGWIN, J. *et al.* Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks. **ProPublica**, 2016. Disponível em:

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 30 jan. 2024.

ANTUNES, H. S. Inteligência artificial e responsabilidade civil: enquadramento. **Revista de Direito da Responsabilidade**, Coimbra, Ano 1, p. 139-154, 2019.. Disponível em: <https://revistadireitoresponsabilidade.pt/2019/inteligencia-artificial-e-responsabilidade-civil-enquadramento/>. Acesso em: 03 nov. 2022.

ASIMOV, I. **Eu, robô**. Tradução de Aline Storto Pereira. 1ª. ed. São Paulo: Aleph, 2014.

BECK, U. **Sociedade de risco**: rumo a uma outra modernidade. Tradução de Sebastião Nascimento; inclui uma entrevista inédita com o autor. 2ª ed. São Paulo: Editora 34, 2011.

BHARADIYA, J. P.; THOMAS, R. K.; AHMED, F. Rise of artificial intelligence in business and industry. **Journal of Engineering Research and Reports**, Hoogly; London, 25, n. 3, 2023. 85-103. Disponível em: [https://www.researchgate.net/publication/371307024\\_Rise\\_of\\_Artificial\\_Intelligence\\_in\\_Business\\_and\\_Industry](https://www.researchgate.net/publication/371307024_Rise_of_Artificial_Intelligence_in_Business_and_Industry). Acesso em: 25 nov. 2023.

CEBRAL-LOUREDA, M.; RINCÓN-FLORES, E. G.; SANCHEZ-ANTE, G. **What AI can do**: strengths and limitations of artificial intelligence. Boca Raton: CRC Press, 2023.

COECKELBERGH, M.; GUNKEL, D. J. ChatGPT: deconstructing the debate and moving it forward. **AI & Society**, London, 2023. Disponível em: <https://link.springer.com/article/10.1007/s00146-023-01710-4#citeas>. Acesso em: 08 jul. 2023.

COMISSÃO EUROPEIA. Regulamento do Parlamento Europeu e do Conselho que Estabelece Regras Harmonizadas em Matéria de Inteligência Artificial (Regulamento Inteligência Artificial) e Altera Determinados Atos Legislativos da União. **COMISSÃO EUROPEIA**, 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206>. Acesso em: 24 nov. 2023.

CORBETT-DAVIES, S. *et al.* A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear. **The Washington Post**, 2016. Disponível em: <https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/>. Acesso em: 01 fev. 2024.

EPSTEIN, Z. et al. Art and the science of generative AI: a deeper dive. **Science Perspectives**, Washington DC, p. 1-23, 2023. Disponível em: <https://arxiv.org/abs/2306.04141>. Acesso em: 23 jul. 2023.

KAMINSKI, M. E. Regulating the risks of AI. **Boston University Law Review**, Boston, 103, 2023. 101-165. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4195066](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4195066). Acesso em: 24 nov. 2023.

LOPES, A. Parlamento Europeu aprova "EU AI Act", primeiro marco regulatório de IA do mundo. **Exame**, 2023. Disponível em: <https://exame.com/inteligencia-artificial/parlamento-europeu-aprova-eu-ai-act-primeiro-marco-regulatorio-de-ia-do-mundo/>. Acesso em: 08 jul. 2023.

MCCARTHY, J. et al. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. **Dartmouth Summer Research Project on Artificial Intelligence**, n. 1, p. 1-13, 1955. Disponível em: <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>. Acesso em: 23 set 2020.

METROPOLIS. Direção: Fritz Lang. Produção: Universum Film. [S.l.]: [s.n.]. 1927.

MOURBY, M.; CATHAOIR, K. Ó.; COLLIN, C. B. Transparency of machine-learning in healthcare: the GDPR & European health law. **Computer Law & Security Review**, Amsterdã, v. 43, p. 1-14, 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0267364921000844>. Acesso em: 06 mar. de 2022.

MÜLLER, V. C. Philosophy of AI: a structured overview. In: SMÜHA, N. **Cambridge handbook on the law, ethics and policy of Artificial Intelligence**. Cambridge: Cambridge University Press, p. 1-24, 2023.

NICHOLAS, G. Explaining Algorithmic Decisions. **Georgetown Law Technology Review**, Washington D.C., v. 4, n. 2, p. 711-730, 2020. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3523456](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3523456). Acesso em: 30 mar. 2022.

NIEVA-FENOLL, J. **Inteligencia artificial y proceso judicial**. Madrid: Marcial Pons, 2018.

NOVELLI, C. et al. Taking AI risks seriously: a proposal for the AI Act. **SSRN**, Rochester, p. 1-8, 2023. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4447964](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4447964). Acesso em: 08 jul. 2023.

SCHUETT, J. Risk management in the artificial intelligence act. **European Journal of Risk Regulation**, Cambridge, p. 1-19, 2023. Disponível em: <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/risk-management-in-the-artificial-intelligence-act/2E4D5707E65EFB3251A76E288BA74068>. Acesso em: 25 nov. 2023.

SCHWEDE, M. A. **Inteligência artificial tendenciosa: discriminação racial pelo viés algorítmico**. Londrina: Thoth, 2023.

ZALNIERIUTE, M. Burning bridges: the automated facial recognition technology and public space surveillance in the modern state. **The Columbia Science & Technology Law Review**, New York, v. 22, n. 2, p. 284-307, 2021. Disponível em: <https://journals.library.columbia.edu/index.php/stlr/article/view/8666>. Acesso em: 05 mar. de 2022.