

INTELIGÊNCIA ARTIFICIAL: Governança e Transparência?

ARTIFICIAL INTELLIGENCE: Governance and Transparency?

*Manuela Betiele Aude Lovatto*¹

RESUMO

Este artigo aborda um tema importante para as empresas na era digital: governança e transparência dos dados utilizados pela Inteligência Artificial para analisar, automatizar e tomar decisões que podem variar de contratações a concessões de crédito. Novos desafios surgem diariamente em relação a quais algoritmos estão sendo usados, quais dados estão alimentando a IA para obter esses resultados e os potenciais vieses gerados. Tais vieses podem levar à discriminação, violações de privacidade e falhas regulatórias devido à falta de auditorias e à ausência de ética algorítmica implementada, impactando a reputação, a cultura organizacional e as operações das empresas, inclusive aquelas com altos padrões de governança e compliance. Amazon, Apple, Clearview AI e Uber enfrentaram esses desafios e servem como estudos de caso neste artigo. Serão apresentadas soluções práticas e recomendações éticas, como o uso de ferramentas de explicabilidade, auditorias internas e relatórios de conformidade para melhorar a transparência e a segurança. O estudo também destacará brevemente algumas iniciativas regulatórias, como o AI Act da União Europeia e os Projetos de Lei nº 210 de 2024 e 2.338 de 2024, que visam promover o uso ético da IA no país.

Palavras-chave: Inteligência Artificial. Governança. Transparência. Ética Algorítmica. Explicabilidade.

ABSTRACT

This article addresses a crucial topic for companies in the digital age: governance and transparency of data used by Artificial Intelligence to analyze, automate, and make decisions that can range from hiring to credit approvals. New challenges arise daily regarding which algorithms are being used, what data is feeding the AI to achieve these results, and the potential biases generated. Such biases can lead to discrimination, privacy violations, and regulatory failures due to the lack of audits and the absence of implemented algorithmic ethics, impacting companies' reputations, organizational culture, and operations, even those with high standards of governance and compliance. Amazon, Apple, Clearview AI, and Uber have faced these challenges and serve as case studies in this article. Practical solutions and ethical recommendations will be presented, such as the use of explainability tools, internal audits, and compliance reports to improve transparency and security. The study will also briefly highlight some regulatory initiatives, such as the European Union's AI Act and Brazil's Bill No. 210 of 2024, which aims to promote the ethical use of AI in Brazil.

Keywords: Artificial Intelligence. Governance. Transparency. Algorithmic Ethics. Explainability.

¹ E-mail: mb.audelovatto@gmail.com

1 INTRODUÇÃO

Nos dias de hoje, a Inteligência Artificial (IA) se tornou uma aliada estratégica indispensável para o ambiente empresarial. Sua capacidade de processar enormes volumes de dados, identificar padrões e automatizar decisões complexas vem transformando a forma como as empresas operam em áreas críticas como recrutamento, finanças e atendimento ao cliente. No recrutamento, agiliza a triagem de currículos e até realiza entrevistas preliminares, permitindo que os departamentos de recursos humanos economizem tempo e recursos. No setor financeiro possibilitam a previsão preditiva de tendências de mercado e uma gestão de riscos mais eficiente, resultando em decisões financeiras mais rápidas e assertivas. Além disso, no atendimento ao cliente, o uso de *chatbots* e assistentes virtuais equipados com IA garante suporte 24 horas por dia, o que aumenta a satisfação do cliente e otimiza os serviços.

Contudo, a crescente utilização da IA também impõe desafios que não podem ser ignorados. Um dos maiores problemas é a ausência de ética algorítmica e transparência nos sistemas de IA, que ao se basearem em dados históricos, sem filtros, geram vieses que podem resultar em injustiças. A opacidade desses sistemas, muitas vezes conhecidos como “caixas-pretas” ou “*black-box*”, dificulta a compreensão de como decisões são tomadas, o que levanta questões de responsabilidade, governança e transparência. Casos envolvendo gigantes como *Amazon*, *Apple Card*, *Clearview Inc.* e *UBER* ilustram as consequências desses riscos. Os episódios expuseram vulnerabilidades de governança de IA e reforçaram a importância de transparência de dados e demonstraram que muitas empresas estão desalinhadas entre suas diretrizes de conduta e base de criação de algoritmos.

Por outro lado, há exemplos de boas práticas que mostram o caminho para uma governança sólida de IA. A IBM, por exemplo, implementou um sistema de auditorias regulares e ferramentas de *Explainable AI* (XAI), que permitem monitorar os sistemas de IA, identificar e mitigar vieses, e garantir a conformidade com as regulamentações. O *Google* também tem investido em iniciativas para melhorar a explicabilidade de seus algoritmos, desenvolvendo programas que visam eliminar discriminações algorítmicas antes que seus sistemas entrem em operação.

Diante da indiscutível iminência dos riscos gerados pelo uso da inteligência artificial tornou-se imprescindível a criação de regulamentação e diretrizes, a União Europeia criou o *AI Act* da União Europeia e o Brasil o Projeto de Lei nº 210, de 2024 que complementa as brechas deixadas pela LGDP. Essas regulamentações exigem que os sistemas de IA sejam transparentes, éticos e seguros, garantindo que seus impactos sociais sejam mitigados. Os

Projetos de Lei nº 210 de 2024 e 2.338 de 2023, por exemplo, impõe diretrizes rigorosas sobre a transparência e a responsabilidade no uso de dados para treinar algoritmos, buscando evitar a discriminação e a falta de clareza nas decisões automatizadas.

Este artigo, utiliza abordagem qualitativa, tem como objetivo explorar esses desafios e propor soluções práticas para aumentar a transparência e a segurança no uso de IA por empresas. A análise inclui a revisão de casos práticos e da literatura sobre governança e explicabilidade, buscando alternativas técnicas e éticas para sugerir práticas que possam ajudar as empresas a mitigar riscos, além de promover segurança para as empresas que usam a IA.

Ao adotar a explicabilidade as empresas não só atendem às exigências regulatórias nacionais e internacionais, como também constroem um ambiente de negócios mais ético e transparente, além de expor de maneira hialina sua cultura organizacional. Esse posicionamento é essencial para que as corporações mantenham sua competitividade e principalmente protejam sua reputação em um mercado cada vez mais dinâmico.

2 INTELIGÊNCIA ARTIFICIAL: GOVERNANÇA E TRANSPARÊNCIA

O Regulamento Geral de Proteção de Dados da Europa, no ano de 2018, avisou as empresas para proteger os dados dos consumidores e sua transferência digital ou enfrentar sanções pesadas e danos à reputação. Como resultado, a maioria das multinacionais implementou um sistema de governança de dados. Ocorre que o desafio tornou-se maior com o uso da Inteligência Artificial para ajudar na tomada de decisões, com isso foi gerada uma nova necessidade, a de *framework* legal sobre questões de qualidade de dados, transparência, supervisão humana e responsabilidade a chamada Governança de IA.

A governança de Inteligência Artificial no ambiente empresarial se refere ao conjunto de ferramentas, políticas, práticas e processos éticos nos quais os algoritmos e dados inseridos servem para a IA basear sua para tomada de decisões assegurando conformidade com as futuras regulamentações e antevendo possíveis irregularidades. Segundo *Duan et al. (2024)*, a governança robusta de IA é essencial para mitigar riscos associados a vieses algorítmicos e à falta de explicabilidade nas decisões automatizadas, garantindo que seus sistemas de operem de maneira alinhada a seus princípios de cultura organizacional.

A governança de IA em escala empresarial traz benefícios sutis mas que serão críticos no longo prazo. Um desses benefícios é o treinamento de seus modelos de IA, indiscutivelmente eles são caros mas retreiná-los por descumprir a lei ou políticas de uma empresa, são ainda mais significativos. Em pesquisa elaborada pela *The state of AI da McKinsey & Company*, em 2021

constatou que 38% das "empresas com melhor desempenho em IA" têm um framework claro para a governança de IA que abrange o processo de desenvolvimento de modelos.

A transparência na IA é particularmente relevante para garantir a confiança dos usuários, conforme ressaltado por Sanches Vilalta (2024), que afirma que “a transparência de algoritmos contribui diretamente para a confiança do consumidor e para a credibilidade das empresas que utilizam IA”. Quando as decisões automatizadas são explicáveis, os usuários e consumidores finais passam a entender como e por que as escolhas foram feitas, deixando expressa redução de riscos legais, vieses e alucinações algorítmicas, demonstrando inexistência de decisões discriminatória e evitando processos judiciais além de sanções regulatórias e crises reputacionais.

As empresas que adotam governança de IA e práticas transparentes conseguem promover a inovação ética e segura, se tornando referência, Correia e Água (2023) destacam que “a transparência e a governança proativa em IA ajudam as empresas a inovar com segurança, criando um ambiente competitivo e alinhado às melhores práticas de mercado” ou seja resulta na famosa inovação sustentável, garantindo que o seu uso seja benéfico tanto para as organizações quanto para a sociedade em geral e abrindo portas para serem referência no que o mercado atualmente procura.

Uma boa governança de IA deverá sempre se basear em quatro pilares: a) rastreável: as organizações devem ser capazes de rastrear as origens dos modelos de IA, bem como os dados que os treinaram; b) explicável: capacidade de justificar como e por que um modelo de AI produziu determinados resultados para quando um auditor de conformidade, um cliente ou até mesmo um stakeholder interno quiser saber de decisão específica tomada; c) confiável: isso significa que além de sustentar a segurança algorítmica, engloba também monitorar o desempenho da IA ao longo do tempo para detectar "desvios" ou alterações não intencionais; d) responsável: a cultura de IA responsável é capaz de promover uma compreensão em toda a organização e demonstrar o alinhamento com os princípios da empresa para o uso da IA.

Operacionalizar a governança da IA é extremamente difícil por englobar diferentes profissionais, estruturas, protocolos, tecnologias entre outros fatores, por isso geralmente é utilizada uma abordagem em três etapas: 1ª etapa: é criar a base para a supervisão, nela é verificada a situação real, pontos fortes, fracos e capacidade de implementação e gerenciamento de riscos além de analisar se as prioridades e metas estão alinhadas a necessidade do negócio; 2ª etapa: documentar sua ética, os valores éticos da empresa deverão ser refletivos dentro do modelo de IA a ser utilizada, ou seja, será treinada em sua base com nas diretrizes da organização a qual pertence; 3ª etapa: Adaptar as estruturas de governança existentes à IA,

alguns elementos da governança de IA se sobrepõem aos programas de governança existentes, e as empresas devem avaliar oportunidades para incorporar a governança de IA em suas práticas atuais, como gestão de riscos de terceiros, aquisições, arquitetura empresarial, jurídico, privacidade e segurança da informação, criando eficiência e gerenciamento de riscos.

3 DESAFIOS TÉCNICOS E ÉTICOS GERADOS PELA BLACK BOX

O maior desafio técnico está no fato de que maioria dos modelos de IA são formados por redes neurais profundas (*deep learning*), e geram as chamadas "caixas-pretas" (*black box*), ou seja, embora possam produzir resultados precisos, é praticamente impossível entender e explicar como os resultados foram alcançados se tornando um sistema impenetrável e por consequência que não pode ser auditado, cabendo discussão sobre a confiabilidade, justiça e ética dessas decisões.

Um estudo de *Harvard* aponta que modelos de *machine learning* com *black box* são aqueles algoritmos em que “os humanos, mesmo aqueles que os projetam, não podem entender como as variáveis estão sendo combinadas para fazer previsões.”

Um mapeamento feito por Mittelstadt *et al.* (2016) que revela existência de pelo no mínimo sete tipos de problemas éticos que devem ser considerados fundamentais quando se analisa a criação, o funcionamento e o uso dos algoritmos: falibilidade, opacidade, viés, discriminação, autonomia, privacidade e responsabilidade.

Estes desafios se tornam graves em setores sensíveis onde decisões errôneas ou mal justificadas podem gerar implicações diretas e prejudiciais para empresas, indivíduos e grupos. A falta de explicabilidade impede auditorias adequadas e compromete a supervisão regulatória, dificultando o cumprimento de normas éticas e legais.

Entre os riscos éticos mais evidentes da caixa-preta algorítmica está a discriminação algorítmica e qualidade de dados. Ocorre quando sistemas de IA são treinados com dados históricos que contêm vieses, a reprodução e amplificação resulta em decisões discriminatórias. Isso pode acontecer em diferentes contextos, como recrutamento, crédito e até mesmo em processos judiciais, afetando desproporcionalmente grupos historicamente marginalizados. Segundo Barocas e Selbst (2016), algoritmos podem "codificar a discriminação presente em padrões históricos", tornando essencial o desenvolvimento de medidas de governança que mitiguem esses riscos.

Além disso, a falta de transparência e explicabilidade gera um ambiente em que é difícil atribuir responsabilidade legal pelas decisões algorítmicas. Se um algoritmo de IA comete um erro — como negar um empréstimo injustamente ou rejeitar um candidato

qualificado a uma vaga de emprego — a quem cabe a responsabilidade? A ausência de clareza na rastreabilidade das decisões dificulta a localização de erros e prejudica a capacidade de impor responsabilidade, afinal em determinadas vezes nem mesmo as empresas entendem o fato que resultou em uma decisão tomada pelo algoritmo.

4 CASOS PRÁTICOS: DA APPLE AO UBER

A corrida pelo uso Inteligência Artificial se tornou algo tão infrene nos últimos anos que até mesmo grandes empresas, consolidadas mercado mundial como *Apple*, *Amazon*, *Clearview AI* e *UBER* não estavam devidamente preparadas para quando os vieses algoritmos aconteceram e enfrentaram graves repercussões devido a falhas em seus sistemas de governança de dados e IA.

4.1 APPLE CARD: DISCRIMINAÇÃO DE GÊNERO NOS LIMITES DE CRÉDITO

Em 2019, surgiram acusações contra o *Apple Card*, as decisões geradas pela IA supostamente ofereciam limites de crédito menores para mulheres em comparação com homens, mesmo quando seus perfis financeiros eram semelhantes. Essas alegações geraram grande repercussão e levantaram acusações de preconceito de gênero pelo algoritmo.

Embora a empresa tenha negado qualquer intenção de discriminação, ela não forneceu detalhes sobre como o algoritmo tomava essas decisões, expondo um problema de falta de transparência no sistema e investigações por parte dos reguladores, levando o *Goldman Sachs*, parceiro da *Apple*, a revisar o sistema de concessão de crédito. Em março de 2021, após uma investigação detalhada, o Departamento de Serviços Financeiros de Nova York concluiu que não havia evidências de discriminação.

O caso do *Apple Card* mostrou a importância de implementar governança e explicabilidade em sistemas de IA, especialmente em áreas sensíveis como finanças. A ausência de supervisão adequada pode levar à perpetuação de vieses inconscientes, resultando em discriminação e danos à reputação.

4.2 AMAZON: O CASO DO RECRUTAMENTO

Em 2014, a *Amazon* desenvolveu um algoritmo para automatizar seu processo de recrutamento. A ideia era criar uma ferramenta que pudesse agilizar a seleção de currículos, identificando os melhores candidatos para posições técnicas, como engenheiros de software. No entanto, o algoritmo começou a apresentar viés de gênero, desfavorecendo as candidatas. Currículos que mencionaram termos como "mulheres" ou que incluíam experiências em faculdades femininas eram penalizados, enquanto currículos de homens eram privilegiados e o

motivo disso foi a forma como o sistema havia sido treinado, com dados históricos dominados por perfis masculinos.

Ao perceber o problema, a *Amazon* tentou ajustar o sistema, mas o viés persistiu. A empresa decidiu então descontinuar a ferramenta em 2018, revertendo ao processo tradicional de recrutamento. Essa experiência evidenciou para o setor de tecnologia em geral que algoritmos não são inerentemente neutros. Ao contrário, eles podem amplificar preconceitos se forem treinados com dados que refletem desigualdades passadas.

Os algoritmos de recrutamento precisam ser projetados com transparência e critérios de imparcialidade desde o início, além de passarem por monitoramento e auditorias contínuas para garantir que promovam a diversidade e inclusão.

4.3 CLEARVIEW AI: VIOLAÇÕES DE PRIVACIDADE

A *Clearview Inc.*, é uma empresa fornecedora de serviços de reconhecimento facial, que foi investigada pela Autoridade de Proteção de Dados Holandesa (DPA) por coletar ilegalmente mais de 30 bilhões de imagens de redes sociais, inclusive de crianças, para criar uma base de dados global. O serviço oferecido permite que autoridades investigativas busquem imagens, sem distinção de localização geográfica ou nacionalidade, atingindo cidadãos da UE e de outros países.

Em decisão publicada de setembro de 2024, a *Clearview* foi multada em €30,5 Milhões por violação do Regulamento geral de proteção de dados (GDPR) e obrigada a parar o tratamento de dados biométricos na Holanda, a empresa também foi instruída a cumprir as medidas em até três meses, sob pena de multa de €250.000 por mês para cada uma das medidas corretivas não atendidas.

Esse caso apresenta inúmeras curiosidades, mas o mais interessante diz respeito à aplicabilidade territorial. A *Clearview Inc.* é uma empresa com sede nos Estados Unidos mas o DPA holandês determinou que o Regulamento Geral de Proteção de Dados deveria ser utilizado por dois motivos: 1) a empresa monitora o comportamento de titulares de dados na UE e 2) sua política de privacidade permitia que cidadãos do Espaço Econômico Europeu (EEA) apresentassem reclamações às autoridades competentes, demonstrando o monitoramento de indivíduos na UE.

A *Clearview Inc.* foi condenada como controladora nas atividades de tratamento, já que é ela quem define os meios e finalidades do tratamento, enquanto seus clientes não influenciam o banco de dados ou o algoritmo. Outra questão diz respeito ao tratamento de dados biométricos e a aplicabilidade territorial do GDPR, concluindo que, ao utilizar algoritmos para

converter fotos em vetores com o objetivo de identificar pessoas de maneira única, se enquadram como dados biométricos.

4.4 UBER: EXPLORAÇÃO LOGARÍTMICA

A UBER redefiniu a mobilidade urbana mas que também se tornou sinônimo de uma cultura empresarial questionável, no quesito governança de IA também apresentou falhas e seus algoritmos de gestão começaram a ganhar destaque em 2017. Nesse ano, investigações jornalísticas e processos judiciais trouxeram à tona práticas internas da empresa, na qual algoritmos eram utilizados para estabelecer tarifas e determinar as condições de trabalho de maneira que favorecesse a redução de custos, incluindo o controle sobre os ganhos dos motoristas e o incentivo a jornadas de trabalho mais longas.

Além disso, surgiram críticas ao uso de práticas como *surge pricing* (preços dinâmicos) e *gamificação* (recompensas virtuais e incentivos de metas) para manipular o comportamento dos motoristas, os incentivando a trabalhar mais sem um aumento proporcional de ganhos. Essas práticas levantaram questionamentos sobre ética e transparência, além de gerarem discussões sobre o impacto da gestão algorítmica nas condições de trabalho e remuneração dos trabalhadores de plataformas, que acabam por ficar em uma posição de dependência econômica sem proteção adequada.

A *Uber* é frequentemente citada em discussões sobre a importância da transparência e explicabilidade de algoritmos e a necessidade de normas e regulamentações mais claras sobre o uso de IA e algoritmos a fim de evitar manipulação econômica e garantir um equilíbrio justo entre eficiência empresarial e proteção de direitos dos trabalhadores.

Esses casos ilustram a importância da explicabilidade e transparência nos sistemas de IA. Tanto *Amazon*, *Apple*, *Clearview AI* e principalmente a *UBER* enfrentaram repercussões sérias por não anteciparem como os vieses nos dados ou na programação poderiam levar a decisões prejudiciais e discriminação. Mas qual qual seria a solução? Como eliminar/abrir essa “caixa-preta” algorítmica?

5 EXPLICABILIDADE DA INTELIGÊNCIA ARTIFICIAL (XAI) E MECANISMOS DE MITIGAÇÃO DE RISCOS

A solução consiste na chamada: explicabilidade em algoritmos de Inteligência Artificial. A explicabilidade se refere à capacidade de um sistema de IA de fornecer explicações claras e compreensíveis sobre como ele chega a decisões, o que é essencial em um contexto em

que as decisões algorítmicas afetam diretamente negócios e a vida dos indivíduos, como pode ser observado nos casos acima.

A explicabilidade permite que desenvolvedores, auditores e usuários finais compreendam cada uma das lógicas por trás das decisões algorítmicas é contrapartida à *black box* e se relaciona à capacidade de explicação dos resultados gerados por uma IA por humanos, permitindo transparência à operacionalização da tecnologia até a origem, é nesse momento que o real alicerce ético e ideológico da empresa pode ser visto.

A adoção de práticas robustas de governança de IA e ferramentas como a *XAI*, inicialmente, envolve a criação de políticas e diretrizes para o uso ético da tecnologia. Um código de conduta claro define os limites aceitáveis e os compromissos da empresa em relação à ética, à transparência e à mitigação de vieses. As políticas de governança devem tratar de temas essenciais, como o uso responsável de dados, a prevenção de discriminação e a proteção de direitos individuais, assegurando que todos os envolvidos na operação dos sistemas de IA estejam alinhados com esses valores.

O monitoramento constante dos sistemas e a realização de auditorias regulares são práticas fundamentais para verificar a conformidade dos algoritmos com as diretrizes éticas e regulatórias. Auditorias, tanto internas quanto externas, permitem que as empresas identifiquem problemas ou desvios, ajustando os algoritmos antes que possíveis falhas resultem em impactos negativos. Empresas como a *IBM*, *Facebook* e o *Google*, já implementam auditorias frequentes para assegurar que seus sistemas estejam em conformidade com os padrões regulatórios e éticos, o que não apenas mitiga riscos legais, mas também fortalece a confiabilidade dos sistemas.

Outras medidas que podem ser implementadas pelas empresas para garantir a confiabilidade e segurança desses sistemas, como: obrigatoriedade de avaliações de conformidade pré-lançamento, monitoramento contínuo após a implementação e a criação de sistemas de gestão de qualidade robustos e vai além as empresas também devem manter registros detalhados e realizar auditorias regulares para comprovar a conformidade, além de fornecer informações acessíveis sobre o funcionamento e as limitações dos sistemas de IA para os usuários.

6 REGULAMENTAÇÃO DE INTELIGÊNCIA ARTIFICIAL: AI ACT E BRASIL

Ao analisar as regulamentações sobre o tema pode ser percebido que a mais pautada por diretrizes e mais completa é a *AI Act* da União Europeia, aprovada em 2023 e com entrada

em vigor prevista para todas as fases até 2027, é uma das legislações mais abrangentes no mundo, ela não somente regula Inteligência Artificial mas também promove o uso responsável.

A *AI Act* adota uma abordagem baseada em risco, classificando os sistemas de IA em quatro categorias: risco mínimo, risco limitado, alto risco e risco inaceitável, impondo obrigações mais rigorosas para aqueles considerados de alto risco, enquanto sistemas de risco inaceitável serão proibidos.

Além disso, introduziu requisitos específicos para modelos de IA de propósito geral e IA generativa, exigindo transparência na divulgação de conteúdo gerado e a adoção de medidas para evitar a criação de conteúdo ilegal. Outra inovação foi a criação do primeiro escritório europeu de IA para coordenar a conformidade, implementação e fiscalização.

As multas do *AI Act* podem chegar até 35 milhões de euros ou 7% do volume de negócios global das empresas no caso de IA proibidas; a 15 milhões de euros ou 3% para violações da lei e 7,5 milhões de euros ou 1,5% pelo fornecimento de informações incorretas. As empresas pequenas, médias e startups ficarão sujeitas a multas administrativas.

O Brasil ainda não possui uma regulamentação específica sobre o tema e acaba por analogia utilizando a LGDP, que dispõem exigências na transparência no tratamento de dados pessoais e requer que as empresas demonstrem uma base legal clara para o uso dos dados, particularmente é uma legislação padrão e tecnicamente omissa quando o assunto é governança de IA, diretrizes, mitigação, explicabilidade e ética algorítmica.

No Senado está em tramitação do PL da Inteligência Artificial (Projeto de Lei 2338/2023) e já sofreu 145 emendas, ele é baseado na *AI Act* e visa regulamentar o tema e explorar, em alguns de seus artigos, a necessidade de transparência destas soluções, principalmente em casos de alto risco.

Os principais pontos de regulamentação do PL da Inteligência Artificial são: a definição de princípios para o desenvolvimento e uso da IA, criação de uma agência nacional de inteligência artificial, estabelecimento de requisitos para o uso de IA em diferentes setores, regulamentação da coleta e do uso de dados e a promoção da pesquisa e desenvolvimento em IA.

Existe também o Projeto de Lei nº 210/2024, em trâmite no Senado Federal, que busca dispor sobre os princípios para uso da tecnologia de inteligência artificial no Brasil, com foco em transparência, responsabilidade e mitigação de riscos no uso de IA. Ele estabelece diretrizes para empresas e órgãos que desenvolvem e implementam sistemas de IA, propondo classificações de risco e exigências específicas para IA de alto impacto, como no setor de saúde, educação e segurança pública.

7 CONSIDERAÇÕES FINAIS

A governança da Inteligência Artificial nas empresas se estabelece como um componente estratégico essencial para garantir o uso ético, seguro e responsável da tecnologia, especialmente considerando os desafios éticos e técnicos que ela impõe. Ao adotar práticas claras e sólidas de governança, transparência e explicabilidade as organizações ganham não apenas em conformidade regulatória, eficiência operacional e confiança com *stakeholders* e consumidores mas também a capacidade de rastrear decisões algorítmicas, mitigação eficaz de riscos e principalmente garantir que seus modelos de IA operem de maneira alinhada a seus valores corporativos o que contribui para uma vantagem competitiva sustentável no mercado.

A implementação de explicabilidade IA nas empresas se torna ferramenta central para proteção dos dados, algoritmos e a mitigação de riscos. Utilizá-las em conjunto com auditorias regulares permite a identificação, correção de vieses éticos e técnicos e redução da opacidade dos sistemas “caixa-preta” criando um diferencial estratégico inovador ético e sustentável.

Os casos práticos discutidos exemplificam os riscos inerentes a ausência de ética algorítmica. As acusações de discriminação de gênero contra o *Apple Card* e do sistema de recrutamento da *Amazon* ilustram como o viés de discriminatório mal supervisionados podem replicar e amplificar desigualdades. A *Clearview*, com sua coleta massiva de dados biométricos sem consentimento, sendo um dos casos com a maior sanção aplicada sobre o assunto, e a *Uber*, com mais uma de suas práticas questionáveis, nesse caso algorítmicas, no controle das condições de trabalho, mostram que a ausência de explicabilidade e governança pode trazer sérias repercussões regulatórias e de imagem para as empresas.

A criação do *AI Act* na União Europeia tornou-se o marco legal do tema no mundo, trazendo uma visão inovadora porém rígida nas regras e sanções. No Brasil ainda não existe uma regulamentação clara, mas ambos os Projetos de Lei em tramitação trarão diretrizes que não podem ser ignoradas.

As empresas que já começaram a implementar governanças de IA alinhadas ao seu DNA, com estrutura transparentes, explicáveis e éticas, se tornaram referência, estabeleceram vantagem competitiva, e atenderam uma exigência de mercado, fortalecendo a confiança do consumidor e protegendo sua reputação.

REFERÊNCIAS

ÁGUA, P. B., & CORREIA, A. **A research agenda on de-biasing the board.** In G. M. Mantovani, A. Kostyuk, & D. Govorun (Eds.), *Corporate governance: Theory and practice* (pp. 16–21). Virtus Interpress. Disponível em: <<https://doi.org/10.22495/cgtapp2>>. Acesso em: 15 set. 2024.

ALAIN, GUILLAUME, AND YOSHUA BENGIO. **Understanding intermediate layers using linear classifier probes.** 2016.

BOSTROM, N.; YUDKOWSKY E. **A ética da inteligência artificial. Fundamento – Rev. de Pesquisa em Filosofia**, v. 1, n. 3 – maio – ago. 2011.

BRASIL. Projeto de Lei nº 210, de 2024. Dispõe sobre os princípios para uso da tecnologia de inteligência artificial no Brasil. **Senado Federal.**

BRASIL. Projeto de Lei nº 2.338 de 2023. BRASIL. Projeto de Lei nº 210, de 2024. Dispõe sobre os princípios para uso da tecnologia de inteligência artificial no Brasil. **Senado Federal.**

COLBACK, LUCY. **The Impact of AI on Business and Society.** *FT Future Forum*, 16 Oct. 2020. Disponível em: <<https://www.ft.com/content/e082b01d-fbd6-4ea5-a0d2-05bc5ad7176c>>. Acesso em 20 set. 2024.

COLE, S. **Amazon Pulled the Plug on an AI Recruitment Tool That Was Biased Against Women.** 2018.

COMISSÃO EUROPEIA. **Proposta de Regulamento do Parlamento Europeu e do Conselho relativo a regras harmonizadas sobre inteligência artificial (Regulamento EU AI Act).** Disponível em: <<https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A52021PC0206>>. Acesso em 17 out. 2024.

DUAN, Y., EDWARDS, J. S., & DWIVEDI, Y. K. **Governança de IA e mitigação de riscos: Estratégias organizacionais para reduzir vieses algorítmicos.** *International Journal of Information Management.* 2024.

EDELMAN, B. G., & GERADIN, D. **Efficiencies and regulatory shortcuts: How should we regulate companies like Airbnb and Uber?** *Stanford Technology Law Review*, 19, 293. 2016.

FIRTH-BUTTERFIELD, KEY. **Impact of Artificial Intelligence and Business Concerns.** *World Economic Forum*, 20 Sept. 2021, <https://www.wipo.int/tech_trends/en/artificial_intelligence/ask_the_experts/techtrends_ai_firth.html>. Acesso em: 16 set. 2024.

GALASKI, JASCHA. **AI Regulation: Present Situation And Future Possibilities.** *LIBERTIES*, 8 Sept. 2021. Disponível em: <<https://www.liberties.eu/en/stories/ai-regulation/43740>>. Acesso em: 16 set. 2024.

LAKSHMANAN, RAIVE. **Clearview AI Faces €30.5M Fine for Building Illegal Facial Recognition Database,** 4 set. 2024. Disponível em: <<https://thehackernews.com/2024/09/clearview-ai-faces-305m-fine-for.html>>. Acesso em: 15 out. 2024.

KHMARA, ANNA. **The Impact of Artificial Intelligence (AI) on Business.** *ITChronicles*, 15 Mar. 2022. Disponível em: <<https://itchronicles.com/artificial-intelligence/the-impact-of-artificial-intelligence-ai-on-business/>>. Acesso em 20 set.2024.

McKINSEY & COMPANY. **The state of AI in 2021.** 2021. Disponível em: <<https://www.mckinsey.com/%7E/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/Global%20survey%20The%20state%20of%20AI%20in%202021/Global-survey-The-state-of-AI-in-2021.pdf>>. Acesso em: 26 set. 2024..

PIERRO, B. **O mundo mediado por algoritmos.** Revista Pesquisa Fapesp, abril de 2018. Disponível em: <<https://revistapesquisa.fapesp.br/o-mundo-mediado-por-algoritmos/>>. Acesso em 10 out. 2024.

MARR, BERNARD. **What Are the Negative Impacts of Artificial Intelligence (AI)?** Bernard Marr & Co, 13 July 2021. Disponível em: <<https://bernardmarr.com/what-are-the-negative-impacts-of-artificial-intelligence-ai/>>. Acesso: 20 set. 2024.

O'SHAUGHNESSY, MATT. **One of the Biggest Problems in Regulating AI Is Agreeing on a Definition.** *Carnegie Endowment for International Peace*, 6 Oct. 2022, Disponível em:<<https://carnegieendowment.org/2022/10/06/one-of-biggest-problems-in-regulating-ai-is-agreeing-on-definition-pub-88100>>. Acesso em 17 set.2024.

UNIÃO EUROPEIA **Lei de Inteligência Artificial da.** Disponível em: <https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.pdf> Acesso em 12 out. 204.

WACHTER, S., & MITTELSTADT, B. **A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI.** *Columbia Business Law Review*, 2019.