noemicajournal.org

e-ISSN: 3086-7037

ARTIGO

Governança da inteligência artificial e sua integração com os princípios ESG em corporações latino-americanas

Artificial intelligence governance and its integration with ESG principles in Latin American corporations

Gobierno de la inteligencia artificial y su integración con los principios ESG en corporaciones latinoamericanas

Lincon Lopes

[1] Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP), São Paulo, SP, Brasil. ORCID: https://orcid.org/0000-0002-6281-1226

SUBMETIDO	ACEITO	PUBLICADO	DOI
17/03/2025	25/10/2025	05/12/2025	em processamento

RESUMO

A adoção corporativa da inteligência artificial apresenta dilemas éticos e de governança que afetam diretamente o desempenho ambiental, social e de governança (ESG) das organizações. Este trabalho examina como as empresas da América Latina incorporam, ou deixam de incorporar, marcos de governança algorítmica capazes de garantir transparência, mitigação de vieses e prestação de contas. Por meio de uma revisão teórica integrativa da literatura sobre governança corporativa, ética da inteligência artificial e sustentabilidade empresarial, constrói-se um modelo conceitual que articula três capacidades organizacionais: supervisão algorítmica, explicabilidade técnica e participação das partes interessadas. O argumento central sustenta que a inteligência artificial só gera valor sustentável quando subordinada a diretrizes éticas explícitas e a uma supervisão institucional independente da área que desenvolve ou adquire a tecnologia. Discutem-se ainda as particularidades do contexto latino-americano, marcado por marcos regulatórios heterogêneos, estruturas de propriedade concentradas e uma desigualdade estrutural que amplifica os riscos de exclusão algorítmica. O trabalho conclui com uma agenda de pesquisa empírica e recomendações práticas para conselhos de administração e comitês de sustentabilidade.

Palavras-chave: governança de IA, ESG, ética organizacional, criação de valor, sustentabilidade corporativa

ABSTRACT

The corporate adoption of artificial intelligence raises ethical and governance dilemmas that directly affect the environmental, social, and governance (ESG) performance of organizations. This paper examines how companies in Latin America incorporate, or fail to incorporate, algorithmic governance frameworks capable of ensuring transparency, bias mitigation, and accountability. Through an integrative theoretical review of the literature on corporate governance, artificial intelligence ethics, and corporate sustainability, a conceptual model is built around three organizational capabilities: algorithmic oversight, technical explainability, and stakeholder engagement. The central argument holds that artificial intelligence only generates sustainable value when it is subordinated to explicit ethical guidelines and to institutional oversight independent from the unit that develops or acquires the technology. The paper also discusses the particularities of the Latin American context, marked by heterogeneous regulatory frameworks, concentrated ownership structures, and structural inequality that amplifies the risks of algorithmic exclusion. The study closes with an agenda for empirical research and practical recommendations for boards of directors and sustainability committees.

Keywords: AI governance, ESG, organizational ethics, value creation, corporate sustainability

RESUMEN

La adopción corporativa de la inteligencia artificial plantea dilemas éticos y de gobernanza que afectan de forma directa el desempeño ambiental, social y de gobernanza (ESG) de las organizaciones. Este trabajo examina cómo las empresas de América Latina incorporan, o dejan de incorporar, marcos de gobernanza algorítmica capaces de garantizar transparencia, mitigación de sesgos y rendición de cuentas. Mediante una revisión teórica integrativa de la literatura sobre gobernanza corporativa, ética de la inteligencia artificial y sostenibilidad empresarial, se construye un modelo conceptual que articula tres capacidades organizacionales: supervisión algorítmica, explicabilidad técnica y participación de las partes interesadas. El argumento central sostiene que la inteligencia artificial solo genera valor sostenible cuando queda subordinada a directrices éticas explícitas y a una supervisión institucional independiente del área que desarrolla o adquiere la tecnología. Se discuten además las particularidades del contexto latinoamericano, marcado por marcos regulatorios heterogéneos, estructuras de propiedad concentradas y una desigualdad estructural que amplifica los riesgos de

exclusión algorítmica. El trabajo concluye con una agenda de investigación empírica y con recomendaciones prácticas para consejos de administración y comités de sostenibilidad.

Palabras clave: gobernanza de IA, ESG, ética organizacional, creación de valor, sostenibilidad corporativa

INTRODUÇÃO

A incorporação acelerada de tecnologias baseadas em inteligência artificial no núcleo das operações corporativas mudou os termos da estratégia competitiva. Em um cenário econômico marcado pela hiperconectividade digital e por mercados cada vez mais voláteis, a transformação digital costuma se apresentar quase como uma condição de sobrevivência para qualquer empresa que aspire a permanecer relevante. No entanto, esse avanço tecnológico ocorre paralelamente a uma exigência institucional, social e regulatória crescente em torno da sustentabilidade corporativa, sintetizada nos critérios ambientais, sociais e de governança, conhecidos pela sigla em inglês ESG (Environmental, Social and Governance). Enquanto a adoção da inteligência artificial costuma ser justificada por sua capacidade de melhorar a eficiência operacional e a rentabilidade financeira, seu efeito sobre as dimensões ESG das organizações é ambivalente e, em não poucos casos, fonte de riscos sistêmicos que os órgãos de direção executiva nem sempre avaliam com o rigor necessário.

O problema central abordado neste estudo é a desconexão, tanto operativa quanto institucional, que costuma ser observada entre as estratégias de transformação digital baseadas em inteligência artificial e os marcos de governança e responsabilidade social corporativa. Na América Latina, as corporações enfrentam o desafio de adotar tecnologias avançadas em ambientes institucionais caracterizados por marcos regulatórios fragmentados e por uma alta sensibilidade social diante da desigualdade estrutural que já afeta a região há décadas (La Porta, López-de-Silanes, Shleifer e Vishny, 2000). A ausência de um modelo integrado de governança da inteligência artificial e de princípios ESG expõe as empresas a falhas éticas severas, à perda de confiança de investidores institucionais e a danos reputacionais que podem comprometer sua continuidade operacional em médio prazo.

Convém precisar, desde o início, o que se entende aqui por governança da inteligência artificial. Seguindo Wirtz, Weyerer e Geyer (2019) e Cihon, Maas e Kemp (2020), entende-se como o conjunto de estruturas, processos, normas e mecanismos de controle que uma organização estabelece para

dirigir, supervisionar e auditar o design, a aquisição e o uso de sistemas algorítmicos, de modo que suas decisões sejam explicáveis, auditáveis e compatíveis com os valores declarados da empresa. Essa definição se distingue da simples gestão de riscos tecnológicos porque incorpora uma dimensão normativa: não basta que um sistema funcione de maneira eficiente, ele também precisa poder se justificar diante das partes afetadas por suas decisões.

A literatura sobre administração estratégica e economia institucional aponta, há algum tempo, que a transformação digital não é um processo neutro nem desprovido de valores. Pelo contrário, ela reconfigura as relações de poder no interior da organização, altera as estruturas de trabalho e modifica a distribuição de externalidades negativas sobre comunidades locais e ecossistemas (Zuboff, 2019). As corporações que implementam sistemas algorítmicos sem salvaguardas éticas suficientes tendem a priorizar ganhos de eficiência de curto prazo em detrimento da equidade social e da integridade institucional, o que corrói, com o tempo, os alicerces da sustentabilidade corporativa.

Este trabalho se apoia nas contribuições de Agrawal, Gans e Goldfarb (2018) sobre a economia da previsão, nos marcos de governança sociotécnica propostos por Gasser e Almeida (2017), e na literatura sobre ética aplicada a sistemas algorítmicos desenvolvida por Mittelstadt, Allo, Taddeo, Wachter e Floridi (2016) e por Jobin, Ienca e Vayena (2019). Sustenta-se que a governança algorítmica não pode se limitar ao cumprimento normativo superficial; ela precisa se constituir como uma capacidade dinâmica capaz de reconfigurar as rotinas organizacionais diante das demandas mutáveis das partes interessadas (Teece, 2007). Essa perspectiva coincide com a advertência de Whittlestone, Nyrup, Alexandrova e Cave (2019), que assinalam que os princípios éticos declarativos, sem mecanismos concretos de implementação, tendem a permanecer no plano retórico e a não modificar as práticas organizacionais reais.

A pergunta que orienta esta pesquisa é a seguinte: como as corporações latino-americanas podem estruturar um marco de governança da inteligência artificial que assegure seu alinhamento efetivo com os princípios ESG e que, ao mesmo tempo, promova a criação de valor sustentável? O objetivo principal do artigo é propor um modelo teórico e normativo que articule a supervisão algorítmica com as três dimensões da sustentabilidade corporativa. De forma mais específica, o trabalho busca: a) sistematizar os principais marcos conceituais sobre governança de inteligência artificial e sua relação com a teoria da agência e a teoria das

partes interessadas; b) identificar os pontos de fricção e de sinergia entre os objetivos da transformação digital e os padrões ESG reconhecidos internacionalmente; e c) formular recomendações práticas para conselhos de administração, comitês de auditoria e áreas de sustentabilidade em organizações que operam na região.

A relevância desta análise para o contexto latino-americano não é pequena. Diferentemente de economias com marcos regulatórios de proteção de dados consolidados, como a União Europeia após a entrada em vigor do Regulamento Geral de Proteção de Dados, a maioria dos países da região conta com legislações incipientes ou fragmentadas em matéria de inteligência artificial (Aguerre, 2021). Essa situação transfere boa parte da responsabilidade regulatória para a autorregulação corporativa, o que torna ainda mais urgente compreender como os conselhos de administração e a alta direção podem institucionalizar mecanismos de controle que supram, ao menos parcialmente, as lacunas normativas existentes.

O texto está organizado em sete seções. A segunda seção desenvolve o marco teórico sobre governança algorítmica, teoria da agência ampliada e critérios ESG. A terceira descreve a abordagem metodológica integrativa empregada. A quarta examina as intersecções críticas entre a inteligência artificial e os três pilares ESG, com ênfase na evidência disponível para a região. A quinta discute os mecanismos institucionais concretos para a sustentabilidade algorítmica. A sexta analisa as implicações gerenciais para a alta direção e os conselhos de administração. Por fim, a sétima seção apresenta as considerações finais, as limitações do estudo e uma agenda de pesquisa futura.

MARCO TEÓRICO: GOVERNANÇA DA INTELIGÊNCIA ARTIFICIAL E CRITÉRIOS ESG

A análise da governança algorítmica no contexto corporativo latino-americano revela várias camadas de complexidade institucional que convém desagregar antes de propor qualquer modelo integrador. Por um lado, a pressão competitiva global empurra as empresas a adotar soluções de inteligência artificial para otimizar cadeias de valor e reduzir custos de transação. Por outro lado, a fragilidade histórica dos marcos regulatórios locais em matéria de proteção de dados pessoais, equidade algorítmica e direitos trabalhistas digitais deixa trabalhadores e consumidores em posição de vulnerabilidade diante de decisões cada vez mais automatizadas. Essa assimetria regulatória aumenta os riscos reputacionais e operacionais para as corporações da região e torna necessária uma autorregulação ética séria,

respaldada pela adoção voluntária de padrões internacionais de diligência devida.

A literatura em gestão estratégica documentou que a implementação de tecnologias disruptivas altera os equilíbrios de poder dentro das organizações. Os departamentos de tecnologia e ciência de dados costumam adquirir protagonismo considerável, e com frequência operam em silos desconectados das áreas de sustentabilidade e conformidade normativa. Essa fragmentação interna gera pontos cegos éticos: os vieses algorítmicos passam despercebidos até se materializarem em crises reputacionais ou em sanções legais. A integração efetiva dos princípios ESG pode atuar como contrapeso a essa miopia corporativa, na medida em que subordina os objetivos de otimização técnica aos valores declarados da organização e à sua responsabilidade social ampliada.

A governança da inteligência artificial vem se consolidando, nos últimos anos, como um campo de estudo dentro da administração estratégica e da teoria da organização. Diferentemente das tecnologias informáticas convencionais, os sistemas de aprendizado de máquina apresentam propriedades de opacidade, autonomia decisória e escalabilidade que desafiam os mecanismos tradicionais de controle diretivo (Floridi et al., 2018). A literatura sobre governança corporativa, que remonta aos trabalhos seminais de Jensen e Meckling (1976) sobre os custos de agência, evoluiu de um enfoque centrado quase exclusivamente no controle financeiro para modelos de gestão de riscos mais amplos, que incorporam a supervisão de ativos intangíveis e de tecnologias complexas cujo funcionamento interno é difícil de auditar mesmo para os próprios diretores.

A economia da previsão e a racionalidade algorítmica

Sob a perspectiva da economia da inteligência artificial, essa tecnologia reduz de forma drástica o custo da previsão (Agrawal et al., 2018). Essa redução transforma os processos de tomada de decisão corporativa, pois permite automatizar decisões operacionais e estratégicas em uma velocidade antes impensável. No entanto, substituir a deliberação humana por previsões algorítmicas introduz vieses cognitivos e sistêmicos difíceis de auditar. A racionalidade algorítmica opera sobre correlações estatísticas derivadas de dados históricos, os quais, com frequência, replicam e amplificam padrões de exclusão social e discriminação que já existiam antes da implementação do sistema. O'Neil (2016) documentou de maneira extensa como modelos aparentemente técnicos e neutros, aplicados à avaliação de crédito, à contratação de pessoal ou à vigilância policial,

terminam reproduzindo desigualdades estruturais sob uma aparência de objetividade matemática que dificulta seu questionamento.

Esse fenômeno adquire relevância particular na América Latina, onde os conjuntos de dados disponíveis para treinar modelos costumam refletir desigualdades históricas de acesso ao crédito, à educação formal e ao emprego estável. Um sistema de pontuação de crédito treinado com dados de uma população que, por décadas, teve acesso restrito ao sistema financeiro formal tenderá a perpetuar essa exclusão, mesmo que seus desenvolvedores não tenham tido qualquer intenção discriminatória. Essa observação se conecta ao conceito de "viés de disponibilidade de dados" (data availability bias) discutido por Barocas e Selbst (2016), que advertem que a qualidade ética de um sistema algorítmico depende, em grande medida, da representatividade histórica dos dados com os quais foi treinado.

O marco do triplo resultado e a integração ESG

A integração dos critérios ESG na estratégia corporativa responde à necessidade de gerenciar externalidades socioambientais e de garantir a viabilidade financeira da empresa no longo prazo (Eccles, Ioannou e Serafeim, 2014). Seguindo o enfoque do triplo resultado, formulado originalmente por Elkington (1997), a criação de valor sustentável exige que as organizações equilibrem o desempenho econômico com a responsabilidade ambiental e social. A intersecção entre a inteligência artificial e o desempenho ESG constitui um domínio de estudo relativamente recente, no qual convergem a ética aplicada, a economia institucional e a teoria das partes interessadas formulada por Freeman (1984).

Vale a pena deter-se em uma precisão conceitual. Os critérios ESG não constituem um marco monolítico nem universalmente aceito; convivem diferentes padrões de reporte, entre os quais se destacam os do Global Reporting Initiative (GRI), os da Sustainability Accounting Standards Board (SASB) e, mais recentemente, os padrões do International Sustainability Standards Board (ISSB), criado em 2021 sob o guarda-chuva da Fundação IFRS. Cada um desses marcos prioriza de maneira distinta a materialidade dos indicadores, o que introduz certa ambiguidade ao comparar o desempenho ESG entre empresas e setores. Aplicado à inteligência artificial, esse problema se agrava, porque nenhum dos padrões mencionados incorpora ainda, de forma sistemática, indicadores específicos sobre governança algorítmica, viés de dados ou auditoria de modelos, ainda que organismos como a Organização para a Cooperação e o Desenvolvimento

Econômico já tenham começado a avançar nessa direção (OECD, 2019, 2024).

A Recomendação do Conselho sobre Inteligência Artificial da OCDE, adotada em 2019 e atualizada em 2024, propõe cinco princípios orientadores para o desenvolvimento de uma inteligência artificial confiável: crescimento inclusivo e bem-estar sustentável; valores centrados nas pessoas e equidade; transparência e explicabilidade; robustez, segurança e proteção; e prestação de contas. Esses princípios, embora não vinculantes, influenciaram marcos regulatórios posteriores, incluindo a Lei de Inteligência Artificial da União Europeia, e oferecem um ponto de referência útil para as corporações latino-americanas que buscam se antecipar a uma regulação ainda em construção na região.

Teoria da agência e explicabilidade algorítmica

A teoria da agência ampliada (Eisenhardt, 1989) sugere que os diretores devem prestar contas não apenas aos acionistas tradicionais, mas a um conjunto mais amplo de partes interessadas afetadas pelas decisões automatizadas. A opacidade dos algoritmos aumenta os custos de transação e a assimetria informacional entre a empresa e seus públicos, o que torna necessária a adoção de padrões rigorosos de explicabilidade e de auditoria externa (Kuziemski e Misuraca, 2020). Essa ideia se conecta ao conceito de "caixa-preta algorítmica" (algorithmic black box), que Burrell (2016) analisa a partir de três fontes distintas de opacidade: a opacidade intencional, quando a empresa oculta deliberadamente o funcionamento do sistema por razões de propriedade intelectual; a opacidade técnica, derivada da complexidade matemática de certos modelos; e a opacidade de analfabetismo, que surge quando quem deve supervisionar o sistema carece das competências técnicas necessárias para compreendê-lo.

Essa última fonte de opacidade é especialmente relevante para os conselhos de administração latino-americanos, nos quais a formação técnica em ciência de dados costuma ser limitada entre os membros não executivos. Wilson e Daugherty (2018) sustentam que as organizações que alcançam os melhores resultados na colaboração entre humanos e inteligência artificial são aquelas que investem na capacitação de seus quadros diretivos, de modo que estes possam formular perguntas pertinentes às equipes técnicas e exercer uma supervisão substantiva, e não meramente formal.

Governança corporativa na América Latina: traços distintivos

Para compreender de que maneira a governança da inteligência artificial se articula com os princípios ESG na região, é necessário partir das particularidades da governança corporativa latino-americana. A literatura comparada sobre sistemas de governança corporativa documentou que, diferentemente dos mercados anglo-saxões, caracterizados por uma propriedade acionária dispersa, as empresas latino-americanas apresentam estruturas de propriedade altamente concentradas, com frequência de caráter familiar, e uma separação limitada entre propriedade e controle (La Porta et al., 2000). Essa concentração acionária tem efeitos ambivalentes sobre a governança tecnológica. Por um lado, pode facilitar decisões mais ágeis e uma visão de longo prazo menos condicionada pela pressão trimestral dos mercados de capitais. Por outro lado, reduz os contrapesos internos e pode favorecer que as decisões sobre adoção de inteligência artificial fiquem concentradas em um número reduzido de diretores, sem a revisão independente que um conselho com maior diversidade de interesses ofereceria.

Husted e Allen (2006) e, mais recentemente, Marano e Kostova (2016), estudaram como as práticas de responsabilidade social corporativa na América Latina se desenvolveram em grande medida como resposta a pressões institucionais externas, entre elas a exigência de investidores internacionais e de organismos multilaterais, mais do que como resultado de uma convicção interna consolidada. Essa observação é relevante para o tema que nos ocupa, pois sugere que a adoção de marcos de governança de inteligência artificial na região poderia seguir um padrão semelhante: uma adoção inicial de caráter simbólico, orientada a satisfazer requisitos de reporte ESG exigidos por fundos de investimento internacionais, que só com o tempo se traduz em mudanças substantivas nas práticas organizacionais. Essa distinção entre adoção simbólica e adoção substantiva de práticas de governança corporativa foi documentada originalmente por Westphal e Zajac (1998) no contexto dos planos de remuneração executiva, e é plenamente aplicável à análise dos comitês de ética de inteligência artificial que várias corporações latino-americanas começaram a instalar.

Ética da inteligência artificial: do princípio à prática

A proliferação de códigos de ética da inteligência artificial no setor corporativo tem sido notável nos últimos anos. Jobin, Ienca e Vayena (2019) analisaram 84 documentos de diretrizes éticas publicados por governos, empresas e organizações internacionais, e encontraram uma convergência relativamente alta em torno de cinco princípios: transparência, justiça e equidade, não maleficência, responsabilidade e privacidade. No entanto, os

mesmos autores advertem que essa convergência discursiva contrasta com uma divergência substancial na maneira como cada princípio é interpretado e implementado na prática, o que abre espaço para o que a literatura tem chamado de "ethics washing", ou seja, o uso de uma linguagem ética como cobertura reputacional sem mudanças operacionais correspondentes (Wagner, 2018).

Rahwan (2018) propõe o conceito de "sociedade no circuito" (society-in-the-loop) para descrever um modelo de governança em que os sistemas algorítmicos não são apenas auditados por especialistas técnicos, mas incorporam mecanismos de participação das comunidades afetadas por suas decisões. Essa perspectiva é particularmente pertinente para o contexto latino-americano, onde a desconfiança histórica em relação às instituições econômicas faz com que a legitimidade de um sistema de governança dependa, em boa medida, de que as partes interessadas percebam que têm algum grau de voz real em seu desenho e em sua supervisão, e não apenas na comunicação posterior de seus resultados.

O panorama regulatório emergente na América Latina

Ainda que a região careça de uma legislação integral sobre inteligência artificial comparável à Lei de Inteligência Artificial da União Europeia, nos últimos anos vários países começaram a avançar nessa direção, com resultados desiguais. O Brasil discute desde 2021 o Projeto de Lei 2338, que propõe um marco de classificação de sistemas de inteligência artificial segundo seu nível de risco, inspirado de forma explícita no modelo europeu, e que inclui disposições específicas sobre auditoria algorítmica e sobre responsabilidade civil dos fornecedores de tecnologia. O Chile aprovou em 2024 uma política nacional de inteligência artificial que, embora não tenha caráter vinculante, estabelece princípios orientadores semelhantes aos da OCDE. O Peru promulgou em 2023 uma lei que declara de interesse nacional o desenvolvimento da inteligência artificial, ainda que sua regulamentação continue pendente em vários aspectos centrais. O México, por sua vez, ainda não conta com uma lei federal específica, e o debate regulatório tem se concentrado principalmente no âmbito da proteção de dados pessoais.

Essa heterogeneidade regulatória gera um problema de coordenação para as corporações multilatinas, isto é, aquelas empresas de origem latino-americana que operam em vários países da região simultaneamente. Uma mesma prática de uso de inteligência artificial pode estar sujeita a obrigações de transparência distintas conforme o país onde é aplicada, o

que complica o desenho de políticas corporativas unificadas. Diante dessa fragmentação, algumas empresas optaram por adotar como padrão interno o marco regulatório mais exigente entre os países onde operam, uma estratégia conhecida na literatura de conformidade normativa como "nivelamento por cima" (leveling up), que reduz o risco de descumprimento em qualquer das jurisdições e, de quebra, simplifica a gestão interna da política de governança tecnológica.

Essa lacuna regulatória parcial reforça o argumento central deste trabalho: na ausência de uma obrigação legal homogênea e exigível em toda a região, a responsabilidade de estabelecer padrões rigorosos de governança algorítmica recai, em grande medida, sobre os próprios conselhos de administração e sobre a disciplina que os mercados de capitais internacionais possam impor por meio de seus critérios de investimento responsável.

Em síntese, o marco teórico que sustenta este trabalho integra cinco corpos de literatura: a economia da previsão e a racionalidade algorítmica, o marco do triplo resultado e os padrões ESG, a teoria da agência aplicada à explicabilidade algorítmica, a literatura sobre governança corporativa e ética da inteligência artificial em economias emergentes, e o panorama regulatório emergente próprio da região. A partir dessa integração, constrói-se, nas seções seguintes, um modelo que articula essas dimensões para a análise do caso latino-americano.

METODOLOGIA

Este estudo adota uma metodologia de revisão teórica de caráter integrativo (Torraco, 2016), orientada a sintetizar literatura científica interdisciplinar sobre governança algorítmica, ética corporativa e desempenho ESG. Optou-se por esse desenho, e não por uma revisão sistemática em sentido estrito, porque o objetivo do trabalho não é quantificar a frequência de determinados achados empíricos, mas construir um marco conceitual capaz de articular corpos de literatura que, até agora, se desenvolveram de maneira relativamente independente: a governança corporativa, a ética da inteligência artificial e os estudos de sustentabilidade empresarial em economias emergentes.

O processo metodológico foi estruturado em quatro fases, planejadas para garantir a rastreabilidade analítica do trabalho.

Na primeira fase, realizou-se um levantamento bibliográfico centrado em marcos regulatórios, diretrizes éticas e modelos de governança de

inteligência artificial, tanto em economias emergentes quanto em economias desenvolvidas, utilizando bases de dados acadêmicas internacionais, entre elas Scopus, Web of Science e Google Scholar. Priorizaram-se artigos publicados em revistas indexadas de administração, ética aplicada e economia institucional, entre elas Journal of Business Ethics, Academy of Management Review, Strategic Management Journal, Corporate Governance: An International Review, Business Ethics Quarterly e Minds and Machines, além de documentos técnicos de organismos multilaterais como a OCDE e a UNESCO.

Na segunda fase, contrastaram-se os objetivos declarados da transformação digital corporativa com os requisitos normativos dos principais padrões internacionais de reporte ESG, entre eles o GRI, o SASB e o ISSB. Esse contraste permitiu identificar as áreas em que os marcos de sustentabilidade existentes já contemplam, ainda que de maneira incipiente, critérios aplicáveis à governança tecnológica, e aquelas em que persiste uma lacuna normativa evidente.

Na terceira fase, identificaram-se os pontos de fricção, de sinergia e os paradoxos estratégicos entre ambos os domínios, por meio de um exercício de codificação temática da literatura revisada. Esse exercício se apoiou na técnica de análise de conteúdo qualitativo descrita por Elo e Kyngäs (2008), que permite organizar categorias conceituais emergentes a partir da leitura sistemática de um corpus documental.

Por fim, na quarta fase, sintetizaram-se os achados para estruturar um modelo integrador de governança corporativa aplicável ao contexto latino-americano, apresentado e discutido nas seções quarta, quinta e sexta deste artigo. Como toda revisão integrativa, este trabalho não pretende esgotar a literatura disponível sobre o tema, mas oferecer uma síntese argumentada capaz de orientar tanto a reflexão acadêmica quanto a prática diretiva.

INTERSECÇÕES CRÍTICAS ENTRE A INTELIGÊNCIA ARTIFICIAL E OS PILARES ESG

A análise integrativa realizada neste estudo mostra que a inteligência artificial impacta cada um dos três pilares ESG de maneira diferenciada, o que exige respostas de governança específicas e, ao mesmo tempo, coordenadas entre si. A seguir, examina-se cada dimensão separadamente, sem perder de vista que, na prática organizacional, essas três dimensões interagem de forma constante.

Dimensão ambiental (E) e sustentabilidade operacional

A inteligência artificial oferece ferramentas de considerável potencial para a modelagem climática, a otimização de redes elétricas, a gestão eficiente de recursos hídricos e a redução de emissões industriais. Empresas de setores intensivos em energia começaram a usar modelos preditivos para otimizar o consumo energético de suas plantas e reduzir sua pegada de carbono. No entanto, essa mesma tecnologia tem uma pegada ecológica operacional considerável, associada ao consumo intensivo de energia e de água nos centros de dados de grande escala necessários para treinar e operar os modelos de aprendizado profundo. Strubell, Ganesh e McCallum (2019) documentaram que o treinamento de um único modelo de processamento de linguagem natural de grande porte pode gerar emissões de dióxido de carbono comparáveis às de vários automóveis ao longo de toda a sua vida útil, dado que tem sido citado com frequência na literatura posterior sobre sustentabilidade digital.

Essa contradição ambiental, entre o potencial da inteligência artificial para mitigar as mudanças climáticas e seu próprio custo energético, obriga as corporações a auditar de maneira ativa todo o ciclo de vida de suas infraestruturas de computação, e não apenas o efeito ambiental das aplicações finais da tecnologia. Na América Latina, onde vários países ainda dependem em grande medida de fontes de energia não renovável para sua matriz elétrica, a instalação de novos centros de dados apresenta um dilema adicional: a localização geográfica dessas infraestruturas pode determinar se sua pegada de carbono é ou não compensada por energias limpas disponíveis localmente. Algumas corporações da região começaram a exigir de seus fornecedores de serviços em nuvem informações desagregadas sobre a origem da energia utilizada, prática que ainda não é generalizada, mas que aponta na direção correta.

Dimensão social (S) e equidade algorítmica

Os sistemas automatizados de tomada de decisão aplicados à gestão de recursos humanos, à avaliação de risco de crédito ou ao atendimento ao cliente podem incorporar vieses históricos discriminatórios presentes nos dados com os quais foram treinados, o que afeta negativamente a equidade social e a inclusão de populações vulneráveis na América Latina. Esse risco não é meramente hipotético. Estudos realizados nos Estados Unidos sobre sistemas de pontuação de crédito e de reconhecimento facial demonstraram taxas de erro significativamente mais altas para pessoas de pele escura e para mulheres, em comparação com homens de pele clara (Buolamwini e Gebru, 2018). Embora a evidência empírica específica para a região latino-americana ainda seja escassa, há razões para pensar que fenômenos

semelhantes possam ocorrer, dada a sobreposição histórica entre desigualdade racial, desigualdade de gênero e desigualdade de acesso a serviços financeiros formais em vários países da região.

Um âmbito particularmente sensível é o da seleção de pessoal por meio de sistemas automatizados. Empresas de diferentes setores começaram a usar algoritmos para filtrar currículos, analisar entrevistas em vídeo ou avaliar o desempenho de seus funcionários. Quando esses sistemas são treinados com dados históricos de contratação que refletem padrões anteriores de discriminação, por idade, por gênero ou por origem socioeconômica, correm o risco de perpetuar esses mesmos padrões sob uma aparência de neutralidade técnica. Raji e Buolamwini (2019) documentaram que a publicação de auditorias externas sobre o desempenho diferencial desses sistemas, desagregado por grupos demográficos, pode incentivar as empresas desenvolvedoras a corrigir seus modelos, o que sugere que a transparência, mais do que a simples proibição regulatória, pode ser uma alavanca de mudança efetiva em contextos onde a capacidade de fiscalização estatal é limitada, como ocorre em boa parte da América Latina.

Dimensão de governança (G) e transparência institucional

A opacidade dos algoritmos de aprendizado de máquina enfraquece os mecanismos tradicionais de prestação de contas perante acionistas, reguladores e sociedade civil, o que exige novas formas de auditoria algorítmica independente e de explicabilidade técnica (Raji et al., 2020). Essa dimensão se conecta diretamente com a literatura clássica de governança corporativa, na medida em que a introdução de sistemas de decisão automatizada acrescenta uma camada adicional de complexidade aos já conhecidos problemas de agência entre diretores, acionistas e outras partes interessadas.

Um problema particular da dimensão de governança na região é a presença escassa de perfis técnicos especializados em tecnologia dentro dos conselhos de administração. Diferentemente do que ocorre em mercados mais desenvolvidos, onde algumas empresas já contam com comitês específicos de tecnologia ou com conselheiros independentes com formação em ciência de dados, na América Latina essa prática ainda é pouco frequente. Essa carência limita a capacidade efetiva de supervisão do conselho sobre as decisões relacionadas à inteligência artificial, e transfere, na prática, boa parte da responsabilidade de governança para as próprias equipes técnicas que desenvolvem ou adquirem os sistemas, o que reproduz

precisamente o problema da opacidade de analfabetismo descrito por Burrell (2016) e mencionado na seção anterior.

Interações entre pilares: sinergias e tensões

Convém destacar que as três dimensões ESG não operam de maneira isolada diante da inteligência artificial, mas interagem entre si, ora de forma sinérgica, ora de forma conflitante. Um exemplo de sinergia é o uso de modelos preditivos para otimizar rotas logísticas, o que pode reduzir simultaneamente o consumo de combustível, melhorar as condições de trabalho dos motoristas ao diminuir jornadas excessivas, e facilitar o reporte transparente de indicadores de eficiência para os investidores. Um exemplo de tensão, por sua vez, é o uso de sistemas de vigilância algorítmica do desempenho no trabalho que, embora possam melhorar indicadores de produtividade reportáveis sob critérios de governança, geram ao mesmo tempo efeitos negativos sobre o bem-estar e a privacidade dos trabalhadores, ou seja, sobre a dimensão social.

Essa observação é coerente com a literatura sobre trade-offs entre dimensões ESG, que aponta que a maximização simultânea das três dimensões nem sempre é possível, e que as organizações precisam, com frequência, tomar decisões que privilegiam uma dimensão sobre outra em função de seu contexto específico (Hahn, Figge, Pinkse e Preuss, 2018). Aplicado à governança da inteligência artificial, esse marco sugere que as corporações latino-americanas precisam de mecanismos explícitos de deliberação interna que permitam identificar e resolver, de maneira consciente e documentada, essas tensões, em vez de deixar que se resolvam de fato segundo a lógica de eficiência que costuma predominar nas equipes técnicas.

MECANISMOS INSTITUCIONAIS PARA A SUSTENTABILIDADE ALGORÍTMICA

Para mitigar os riscos descritos na seção anterior, as corporações precisam institucionalizar mecanismos internos de controle que assegurem o alinhamento entre as implementações tecnológicas e seus compromissos ESG. A partir da revisão de literatura realizada, identificam-se cinco mecanismos que, em conjunto, constituem a base de um modelo de governança algorítmica orientado à sustentabilidade.

Comitês éticos de inteligência artificial com poder de veto

O primeiro mecanismo é a criação de comitês éticos de inteligência artificial dotados de autoridade real, incluindo o poder de veto sobre projetos automatizados de alto risco. A literatura sobre comitês de ética corporativa adverte, no entanto, que sua eficácia depende criticamente de sua composição e de sua posição hierárquica dentro da organização. Um comitê composto exclusivamente por pessoal da área de tecnologia, ou que se reporta diretamente ao diretor de tecnologia, dificilmente exercerá uma supervisão independente. Pelo contrário, os comitês mais eficazes costumam incluir representantes das áreas de conformidade normativa, de recursos humanos, de sustentabilidade e, quando possível, membros externos com formação em ética aplicada ou em direitos digitais, e se reportam diretamente ao conselho de administração ou a um comitê do conselho (Rességuier e Rodrigues, 2020).

Auditorias periódicas de equidade e viés

O segundo mecanismo consiste na realização periódica de auditorias de equidade e de viés sobre os sistemas algorítmicos em produção, e não apenas na fase de desenho. Raji et al. (2020) propõem um marco de auditoria interna de ponta a ponta que documenta cada etapa do ciclo de vida de um modelo, desde a coleta de dados até o monitoramento posterior à sua implementação, com o objetivo de identificar pontos de fricção ética antes que se traduzam em danos concretos. Esse tipo de auditoria exige, no entanto, capacidades técnicas especializadas que muitas corporações latino-americanas de médio porte ainda não possuem internamente, o que abre uma oportunidade de mercado para empresas de auditoria algorítmica independente, um setor ainda incipiente na região, mas em expansão.

Relatórios de transparência algorítmica

O terceiro mecanismo é a publicação de relatórios de transparência algorítmica dirigidos às partes interessadas, que documentem de maneira acessível, e não meramente técnica, como a organização utiliza sistemas de inteligência artificial em suas decisões mais sensíveis. A literatura sobre comunicação corporativa de sustentabilidade adverte que esses relatórios correm o risco de se tornar exercícios de relações públicas se não vierem acompanhados de indicadores verificáveis externamente (Cho, Laine, Roberts e Rodrigue, 2015). Por essa razão, recomenda-se que os relatórios de transparência algorítmica sejam integrados aos reportes ESG já existentes, e que sejam submetidos aos mesmos processos de verificação externa aplicados, por exemplo, aos indicadores de emissões de carbono.

Capacitação e alfabetização algorítmica do conselho de administração

O quarto mecanismo, com frequência subestimado na literatura, é a capacitação específica dos membros do conselho de administração e da alta direção em conceitos básicos de inteligência artificial e de ética de dados. Como discutido na seção teórica, a opacidade de analfabetismo constitui uma barreira significativa para a supervisão efetiva. Programas de capacitação contínua, complementados pela incorporação gradual de conselheiros independentes com formação técnica, podem reduzir essa lacuna e fortalecer a capacidade do conselho de formular perguntas pertinentes e de exigir explicações compreensíveis das equipes técnicas.

Mecanismos de participação de partes interessadas externas

O quinto mecanismo, inspirado na proposta de sociedade no circuito de Rahwan (2018), consiste em incorporar canais formais de consulta com as comunidades e os grupos afetados pelos sistemas algorítmicos da empresa, particularmente em setores como serviços financeiros, saúde ou emprego. Esses canais podem assumir a forma de painéis consultivos externos, pesquisas periódicas ou mesas de diálogo com organizações da sociedade civil especializadas em direitos digitais. Embora essa prática ainda seja pouco comum entre as corporações latino-americanas, algumas experiências no setor financeiro da região mostram que a incorporação desse tipo de mecanismo pode antecipar conflitos regulatórios e melhorar a legitimidade social dos projetos de automação antes de sua implementação completa.

A combinação desses cinco mecanismos, comitê ético com poder de veto, auditorias periódicas de equidade, relatórios de transparência verificáveis, capacitação do conselho e participação de partes interessadas externas, constitui o que este trabalho denomina um modelo de governança algorítmica orientado à sustentabilidade. Trata-se de um modelo exigente, que requer investimento de tempo e de recursos, mas que é coerente com a evidência disponível sobre os fatores que distinguem a adoção simbólica da adoção substantiva de práticas de governança corporativa (Westphal e Zajac, 1998).

O caso do setor financeiro como referência regional

O setor financeiro latino-americano oferece um bom ponto de referência para observar, de maneira concreta, como essas tensões se manifestam na prática. Os bancos e as plataformas de crédito digital da região adotaram rapidamente modelos de pontuação de crédito baseados em aprendizado de máquina, que incorporam variáveis alternativas às tradicionais, como o comportamento em redes sociais, os padrões de uso do telefone celular ou o histórico de pagamento de serviços básicos. Essa inovação permitiu, em

vários casos documentados, ampliar o acesso ao crédito formal para populações historicamente excluídas do sistema financeiro, o que representa uma contribuição positiva para a dimensão social da sustentabilidade corporativa.

No entanto, a mesma inovação introduz riscos que nem sempre são evidentes à primeira vista. O uso de variáveis alternativas, como o consumo de dados móveis ou a geolocalização, pode acabar operando como um indicador indireto (proxy) de variáveis socioeconômicas que a regulação de proteção de dados proíbe utilizar diretamente, como o nível de renda ou o local de residência. Esse fenômeno, conhecido na literatura como discriminação por proxy, foi documentado em mercados desenvolvidos e há razões sólidas para pensar que também ocorra na América Latina, embora a evidência empírica específica ainda seja escassa. A regulação financeira da região, historicamente centrada no risco de solvência e na lavagem de ativos, nem sempre conta ainda com os instrumentos técnicos necessários para fiscalizar esse tipo de risco algorítmico, o que reforça, mais uma vez, a importância da autorregulação corporativa discutida nas seções anteriores.

Superintendências financeiras de vários países da região, entre elas as da Colômbia e do México, começaram a emitir orientações sobre o uso responsável da inteligência artificial na avaliação de risco de crédito, ainda que seu caráter continue sendo, em sua maioria, recomendatório e não vinculante. Esse exemplo setorial ilustra de maneira clara o argumento geral deste trabalho: a tecnologia pode contribuir de forma genuína para os objetivos ESG de uma organização, mas apenas se sua implementação vier acompanhada de mecanismos de auditoria e de prestação de contas capazes de detectar e corrigir os efeitos indesejados que, do contrário, permaneceriam invisíveis até se tornarem um problema reputacional ou regulatório de grande escala.

IMPLICAÇÕES ESTRATÉGICAS PARA A ALTA DIREÇÃO

Os resultados deste ensaio sugerem que a alta direção corporativa na América Latina precisa superar a visão puramente instrumental da tecnologia, segundo a qual a inteligência artificial é avaliada exclusivamente em função de seus ganhos de eficiência e de seu retorno financeiro imediato. A governança da inteligência artificial, integrada de maneira efetiva com os princípios ESG, representa um fator de diferenciação competitiva e de legitimidade institucional em mercados globais cada vez mais regulados e mais atentos aos riscos reputacionais associados à tecnologia.

Em primeiro lugar, a evidência revisada sugere que as empresas que se antecipam a padrões regulatórios internacionais, em vez de esperar que a legislação local os torne obrigatórios, obtêm vantagens de reputação diante de investidores institucionais que já incorporam critérios de governança tecnológica em seus processos de diligência devida. Fundos de investimento de grande porte começaram a incluir perguntas específicas sobre governança de inteligência artificial em seus questionários de avaliação ESG, o que transforma esse tema, para as empresas que buscam financiamento internacional, em uma questão de acesso ao capital, e não apenas de ética corporativa.

Em segundo lugar, a integração da governança de inteligência artificial com os princípios ESG exige uma redefinição das responsabilidades dentro da estrutura diretiva. Em muitas organizações latino-americanas, a responsabilidade sobre a inteligência artificial recai exclusivamente na área de tecnologia, enquanto a responsabilidade sobre ESG recai em uma área de sustentabilidade separada, com pouca interação entre ambas. Este trabalho sugere que a criação de instâncias de coordenação formal, como o comitê ético de inteligência artificial descrito na seção anterior, é indispensável para superar essa fragmentação funcional e evitar que os pontos cegos éticos já mencionados se traduzam em crises reputacionais evitáveis.

Em terceiro lugar, a alta direção deve considerar que a gestão de riscos algorítmicos não pode ser improvisada depois que um sistema já está em produção. A literatura sobre gestão de riscos de tecnologias emergentes recomenda incorporar avaliações de impacto ético desde as etapas iniciais de desenho de qualquer projeto de inteligência artificial, de maneira análoga a como se realizam as avaliações de impacto ambiental em projetos de infraestrutura física (Reisman, Schultz, Crawford e Whittaker, 2018). Essa prática, ainda pouco difundida na região, permitiria identificar riscos de viés, de privacidade e de exclusão social antes que um sistema alcance sua fase de implementação em massa, quando corrigi-lo se torna consideravelmente mais custoso.

Por fim, convém assinalar que o investimento em governança algorítmica não deve ser entendido como um custo puramente defensivo. Empresas que desenvolveram capacidades sólidas de auditoria e de explicabilidade de seus sistemas costumam estar mais bem posicionadas para escalar novas aplicações de inteligência artificial com maior velocidade, porque contam com protocolos já testados para avaliar riscos e obter a aprovação interna necessária. Nesse sentido, a governança rigorosa pode funcionar, mais do

que como um freio, como uma capacidade organizacional que permite uma inovação mais rápida e, ao mesmo tempo, mais responsável.

Um elemento adicional que merece atenção da alta direção é a relação entre governança da inteligência artificial e a retenção de talento especializado. Os profissionais jovens de ciência de dados e de engenharia de software, em particular aqueles formados em universidades com programas de ética tecnológica, mostram uma sensibilidade crescente ao uso responsável da tecnologia que desenvolvem. Pesquisas realizadas entre profissionais do setor em diferentes países mostraram que uma parcela significativa desses trabalhadores considera a ausência de políticas claras de ética de dados um fator relevante na hora de avaliar uma oferta de emprego ou de permanecer em uma organização. Para as corporações latino-americanas, que já enfrentam uma concorrência intensa por talento técnico qualificado diante de empresas globais que oferecem condições de trabalho remoto e salários em moeda estrangeira, contar com uma política de governança algorítmica crível pode se tornar um argumento adicional de atração e de retenção de pessoal, um benefício raramente contabilizado nas análises de retorno sobre investimento desses programas, mas consistente com a evidência disponível sobre o vínculo entre propósito organizacional e comprometimento com o trabalho.

Por último, a alta direção deve considerar o papel dos órgãos de controle interno e de auditoria nesse processo. A função de auditoria interna, tradicionalmente voltada ao controle financeiro e à conformidade normativa, precisa ampliar seu alcance para incorporar a revisão periódica dos sistemas algorítmicos de maior risco. Isso implica, entre outras coisas, capacitar as equipes de auditoria em conceitos básicos de aprendizado de máquina e estabelecer protocolos de trabalho conjunto com as áreas de tecnologia, de modo que a auditoria não dependa exclusivamente das informações que as próprias equipes técnicas decidem compartilhar. Sem essa ampliação de alcance, o risco de que a governança algorítmica fique reduzida a uma declaração de princípios sem verificação efetiva continua sendo alto.

CONSIDERAÇÕES FINAIS

A governança da inteligência artificial e sua articulação com os princípios ESG configuram, hoje, um tema central para as corporações latino-americanas. Como foi argumentado ao longo deste trabalho, a tecnologia por si só não garante sustentabilidade alguma; ela requer um arcabouço institucional de governança rigorosa que subordine a eficiência algorítmica ao bem-estar social e ambiental das comunidades onde a empresa opera. O

modelo proposto, articulado em torno de cinco mecanismos institucionais, oferece um ponto de partida para que conselhos de administração e equipes diretivas possam avançar nessa direção de maneira estruturada, e não meramente reativa diante de pressões regulatórias ou de crises reputacionais pontuais.

Ao longo do artigo, mostrou-se que a inteligência artificial impacta de maneira diferenciada cada um dos três pilares ESG. Na dimensão ambiental, o potencial da tecnologia para otimizar processos convive com uma pegada energética considerável que as empresas precisam auditar de forma ativa. Na dimensão social, os sistemas automatizados podem perpetuar desigualdades históricas se não forem submetidos a auditorias de equidade rigorosas e periódicas. Na dimensão de governança, a opacidade dos algoritmos exige novas capacidades de supervisão que boa parte dos conselhos de administração da região ainda não desenvolveu. Diante desse panorama, a proposta central do trabalho é que nenhuma dessas três dimensões pode ser gerenciada de forma isolada, e que sua integração efetiva depende de mecanismos de coordenação interna explícitos, respaldados por autoridade real dentro da estrutura organizacional.

Limitações do estudo

Apesar da consistência conceitual do modelo proposto, este estudo apresenta limitações que convém reconhecer com clareza. Por se tratar de uma revisão teórico-integrativa, o trabalho não incorpora evidência empírica quantitativa nem estudos de caso setoriais específicos sobre corporações latino-americanas, de modo que o marco proposto requer validação em ambientes corporativos reais da região. Além disso, a literatura revisada provém, em sua maioria, de contextos institucionais diferentes do latino-americano, principalmente dos Estados Unidos e da Europa, o que exige certa cautela ao extrapolar achados originados nesses contextos para realidades regulatórias, culturais e econômicas diferentes. Por fim, dada a velocidade com que evoluem tanto a tecnologia de inteligência artificial quanto a regulação a ela associada, alguns dos marcos normativos discutidos neste trabalho, em particular os relacionados a padrões de reporte ESG, podem se modificar de maneira significativa nos próximos anos.

Agenda de pesquisas futuras

Como orientação para futuras agendas de pesquisa, recomenda-se o desenvolvimento de estudos empíricos quantitativos, por meio de modelagem de equações estruturais ou outras técnicas equivalentes, que

permitam medir o efeito mediador da governança de inteligência artificial sobre o desempenho ESG em setores industriais intensivos em dados, como o financeiro, o de telecomunicações e o de varejo. Também seria valioso realizar estudos de caso comparados entre corporações latino-americanas que implementaram comitês éticos de inteligência artificial e aquelas que não o fizeram, com o objetivo de identificar de maneira mais precisa os fatores organizacionais que facilitam, ou que dificultam, a adoção substantiva, e não meramente simbólica, desses mecanismos. Por fim, dada a escassez de evidência empírica específica sobre viés algorítmico em populações latino-americanas, recomenda-se estimular pesquisas que documentem de maneira sistemática o desempenho diferencial de sistemas de pontuação de crédito, de seleção de pessoal e de avaliação de risco aplicados na região, seguindo metodologias semelhantes às empregadas por Buolamwini e Gebru (2018) no contexto estadunidense.

Posicionamento autoral

Do ponto de vista autoral, sustenta-se que a adoção acrítica da inteligência artificial, sob a retórica genérica da modernização corporativa, representa uma ameaça à coesão social na América Latina se não vier acompanhada de mecanismos de governança sólidos e verificáveis. A tecnologia não é um fator exógeno nem neutro: é o resultado de decisões institucionais e políticas concretas, que devem ser governadas de maneira transparente e, na medida do possível, com participação efetiva das comunidades afetadas por suas consequências. O desafio para as corporações da região não consiste em escolher entre inovação tecnológica e responsabilidade social, mas em construir as capacidades institucionais necessárias para que ambas avancem em conjunto.

REFERÊNCIAS

Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction machines: The simple economics of artificial intelligence. Harvard Business Press.

Aguerre, C. (2021). Género y gobernanza de la inteligencia artificial en América Latina. Revista CIDOB d'Afers Internacionals, (127), 87-108.

Argyris, C., & Schön, D. A. (1978). Organizational learning: A theory of action perspective. Addison-Wesley.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104(3), 671-732. https://doi.org/10.15779/Z38BG31

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of Machine Learning Research, 81, 77-91.

Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. Big Data & Society, 3(1), 1-12. https://doi.org/10.1177/2053951715622512

Cascio, W. F., & Boudreau, J. W. (2016). The search for global competence: From international HR to strategic agility. Human Resource Management Review, 26(1), 5-15. https://doi.org/10.1016/j.hrmr.2015.09.002

Cho, C. H., Laine, M., Roberts, R. W., & Rodrigue, M. (2015). Organized hypocrisy, organizational façades, and sustainability reporting. Accounting, Organizations and Society, 40, 78-94. https://doi.org/10.1016/j.aos.2014.12.003

Cihon, P., Maas, M. M., & Kemp, L. (2020). Should artificial intelligence governance be centralised? Design lessons from history. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (pp. 228-234). ACM. https://doi.org/10.1145/3375627.3375857

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management Science, 60(11), 2835-2857. https://doi.org/10.1287/mnsc.2014.1984

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. Academy of Management Review, 14(1), 57-74. https://doi.org/10.5465/amr.1989.4279003

Elkington, J. (1997). Cannibals with forks: The triple bottom line of 21st century business. Capstone.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. Journal of Advanced Nursing, 62(1), 107-115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5

Freeman, R. E. (1984). Strategic management: A stakeholder approach. Pitman.

Gasser, U., & Almeida, V. A. (2017). Dialogues on artificial intelligence: Governing through code and beyond. IEEE Internet Computing, 21(4), 4-11. https://doi.org/10.1109/MIC.2017.2911432

Hahn, T., Figge, F., Pinkse, J., & Preuss, L. (2018). A paradox perspective on corporate sustainability: Descriptive, instrumental, and normative aspects. Journal of Business Ethics, 148(2), 235-248. https://doi.org/10.1007/s10551-017-3587-2

Husted, B. W., & Allen, D. B. (2006). Corporate social responsibility in the multinational enterprise: Strategic and institutional approaches. Journal of International Business Studies, 37(6), 838-849. https://doi.org/10.1057/palgrave.jibs.8400227

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(4), 305-360. https://doi.org/10.1016/0304-405X(76)90026-X

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making. Telecommunications Policy, 44(6), 101976. https://doi.org/10.1016/j.telpol.2020.101976

La Porta, R., López-de-Silanes, F., Shleifer, A., & Vishny, R. W. (2000). Investor protection and corporate governance. Journal of Financial Economics, 58(1-2), 3-27. https://doi.org/10.1016/S0304-405X(00)00065-9

Marano, V., & Kostova, T. (2016). Unpacking the institutional complexity in adoption of CSR practices in multinational enterprises. Journal of Management Studies, 53(1), 28-54. https://doi.org/10.1111/joms.12124

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 1-21. https://doi.org/10.1177/2053951716679679

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.

Organisation for Economic Co-operation and Development. (2019). Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449.

Organisation for Economic Co-operation and Development. (2024). Recommendation of the Council on Artificial Intelligence (amended). OECD/LEGAL/0449.

Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. Ethics and Information Technology, 20(1), 5-14. https://doi.org/10.1007/s10676-017-9430-8

Raji, I. D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 429-435). ACM. https://doi.org/10.1145/3306618.3314244

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In Proceedings of the 2020 ACM Conference on Fairness, Accountability, and Transparency (pp. 33-44). ACM. https://doi.org/10.1145/3351095.3372842

Reisman, D., Schultz, J., Crawford, K., & Whittaker, M. (2018). Algorithmic impact assessments: A practical framework for public agency accountability. AI Now Institute.

Rességuier, A., & Rodrigues, R. (2020). AI ethics should not remain toothless! A call to bring back the teeth of ethics. Big Data & Society, 7(2), 1-5. https://doi.org/10.1177/2053951720942541

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 3645-3650). ACL. https://doi.org/10.18653/v1/P19-1355

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal, 28(13), 1319-1350. https://doi.org/10.1002/smj.640

Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. Human Resource Development Review, 15(4), 404-428. https://doi.org/10.1177/1534484316671606

Wagner, B. (2018). Ethics as an escape from regulation: From ethics-washing to ethics-shopping? In E. Bayamlioğlu, I. Baraliuc, L. Janssens, & M. Hildebrandt (Eds.), Being profiled: Cogitas ergo sum (pp. 84-89). Amsterdam University Press.

Westphal, J. D., & Zajac, E. J. (1998). The symbolic management of stockholders: Corporate governance reforms and shareholder reactions. Administrative Science Quarterly, 43(1), 127-153. https://doi.org/10.2307/2393593

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics: Towards a focus on tensions. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 195-200). ACM. https://doi.org/10.1145/3306618.3314289

Wilson, H. J., & Daugherty, P. R. (2018). Collaborative intelligence: Humans and AI are joining forces. Harvard Business Review, 96(4), 114-123.

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector: Applications and challenges. International Journal of Public Administration, 42(7), 596-615. https://doi.org/10.1080/01900692.2018.1498103

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.

DECLARAÇÕES EDITORIAIS

CONTRIBUIÇÃO DO AUTOR: CONCEITUALIZAÇÃO, METODOLOGIA, INVESTIGAÇÃO, ANÁLISE FORMAL, CURADORIA DE DADOS, REDAÇÃO DO MANUSCRITO ORIGINAL E REVISÃO E EDIÇÃO.

Financiamento: O estudo não recebeu financiamento externo.

Conflito de interesses: O autor declara não ter conflito de interesses.

Uso de inteligência artificial: O uso de ferramentas de inteligência artificial não foi declarado pelo autor.