

ARTIGO

Inteligência artificial, desempenho ESG e criação de valor: um modelo teórico baseado em capacidades organizacionais

Artificial intelligence, ESG performance, and value creation: a theoretical model based on organizational capabilities

Inteligencia artificial, desempeño ESG y creación de valor: un modelo teórico basado en capacidades organizacionales

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
02/04/2024	18/10/2024	06/12/2024	em processamento

RESUMO

A inteligência artificial pode ampliar a capacidade de previsão, monitoramento e eficiência das organizações, mas sua adoção não garante, por si só, melhoria do desempenho ambiental, social e de governança nem criação sustentável de valor. Este artigo tem como objetivo desenvolver um modelo teórico que explique como capacidades organizacionais condicionam essa relação. Adota-se uma abordagem teórico-integrativa que articula a teoria das capacidades dinâmicas, a literatura sobre governança de inteligência artificial e ética algorítmica, e os estudos consolidados sobre desempenho ESG e criação de valor. Como resultado, propõem-se quatro capacidades mediadoras: integração e qualidade dos dados, aprendizagem organizacional, governança responsável e conversão de informações em decisões estratégicas sustentáveis. O modelo identifica também cinco mecanismos de destruição de valor associados ao uso da inteligência artificial: opacidade decisória, vieses e discriminação sistêmica, consumo intensivo de recursos computacionais, baixa responsabilização e desalinhamento estratégico. Discutem-se ainda a orquestração e a complementaridade entre as capacidades, trajetórias de maturidade organizacional, riscos ligados à dependência de fornecedores tecnológicos e

implicações para a governança corporativa. Conclui-se que os benefícios da inteligência artificial para a agenda ESG dependem da capacidade organizacional de combinar eficiência, legitimidade, controle de riscos e resultados socioambientais, o que distingue a simples aquisição de tecnologia da competência efetiva para utilizá-la de maneira responsável e sustentável.

Palavras-chave: inteligência artificial; ESG; capacidades organizacionais; criação de valor; governança.

ABSTRACT

Artificial intelligence can expand forecasting, monitoring, and efficiency capabilities, but its adoption alone does not guarantee improved environmental, social, and governance performance or sustainable value creation. This article develops a theoretical model explaining how organizational capabilities condition this relationship. It adopts a theoretical-integrative approach that combines dynamic capabilities theory, the literature on artificial intelligence governance and algorithmic ethics, and established research on ESG performance and value creation. Four mediating capabilities are proposed: data integration and quality, organizational learning, responsible governance, and the conversion of information into sustainable strategic decisions. The model also identifies five mechanisms of value destruction associated with artificial intelligence: decision opacity, bias and systemic discrimination, intensive computational resource consumption, weak accountability, and strategic misalignment. The article further discusses the orchestration and complementarity among capabilities, organizational maturity trajectories, risks associated with dependence on technology vendors, and implications for corporate governance. It concludes that the benefits of artificial intelligence for the ESG agenda depend on an organization's ability to combine efficiency, legitimacy, risk control, and social and environmental outcomes, distinguishing mere technology acquisition from the effective capability to use it responsibly and sustainably.

Keywords: artificial intelligence; ESG; organizational capabilities; value creation; governance.

RESUMEN

La inteligencia artificial puede ampliar la capacidad de previsión, seguimiento y eficiencia de las organizaciones, pero su adopción no garantiza por sí sola la mejora del desempeño ambiental, social y de gobernanza ni la creación sostenible de valor. Este artículo tiene como objetivo desarrollar un modelo teórico que explique cómo las capacidades

organizacionales condicionan esta relación. Se adopta un enfoque teórico-integrador que articula la teoría de las capacidades dinámicas, la literatura sobre gobernanza de la inteligencia artificial y ética algorítmica, y los estudios consolidados sobre desempeño ESG y creación de valor. Se proponen cuatro capacidades mediadoras: integración y calidad de los datos, aprendizaje organizacional, gobernanza responsable y conversión de información en decisiones estratégicas sostenibles. El modelo identifica también cinco mecanismos de destrucción de valor asociados al uso de la inteligencia artificial: opacidad decisoria, sesgos y discriminación sistémica, consumo intensivo de recursos computacionales, baja responsabilidad y desalineación estratégica. Se discuten además la orquestación y la complementariedad entre las capacidades, las trayectorias de madurez organizacional, los riesgos vinculados a la dependencia de proveedores tecnológicos y las implicaciones para la gobernanza corporativa. Se concluye que los beneficios de la inteligencia artificial para la agenda ESG dependen de la capacidad organizacional de combinar eficiencia, legitimidad, control de riesgos y resultados socioambientales, lo que distingue la simple adquisición de tecnología de la competencia efectiva para utilizarla de manera responsable y sostenible.

Palabras clave: inteligencia artificial; ESG; capacidades organizacionales; creación de valor; gobernanza.

INTRODUÇÃO

A adoção corporativa de tecnologias baseadas em inteligência artificial tem alterado, em ritmo acelerado, os fundamentos da estratégia competitiva, da arquitetura organizacional e dos modelos de negócio em escala global. Ao mesmo tempo, pressões institucionais, regulatórias e de mercado associadas à agenda ESG (Environmental, Social, and Governance) exigem que as empresas demonstrem compromisso efetivo com a sustentabilidade ambiental, a equidade social e a integridade de governança, para além do discurso declarativo presente em relatórios institucionais. Esse duplo movimento, marcado pela digitalização intensa das operações e pela cobrança crescente por resultados socioambientais mensuráveis, tem levado gestores e pesquisadores a supor, muitas vezes de forma pouco crítica, que a incorporação de ferramentas algorítmicas avançadas produz, quase automaticamente, ganhos de desempenho ESG e criação de valor corporativo.

A evidência disponível não sustenta essa suposição de maneira simples. A aquisição de sistemas de inteligência artificial, a contratação de consultorias

especializadas ou a implementação pontual de modelos de aprendizado de máquina não asseguram, isoladamente, melhorias consistentes nos indicadores ambientais, sociais e de governança das organizações. Ao contrário, parte considerável da literatura crítica sobre tecnologia e sociedade documenta externalidades relevantes associadas a esses sistemas: consumo elevado de energia em centros de dados (Strubell et al., 2019), reprodução de vieses discriminatórios em decisões automatizadas (Buolamwini & Gebu, 2018), opacidade nos processos decisórios e desgaste da legitimidade organizacional perante públicos de interesse diversos (Suchman, 1995; Whittaker, 2021). A relação entre adoção de inteligência artificial e criação de valor sustentável, portanto, não é direta nem automática. Ela é mediada por rotinas gerenciais, ativos intangíveis e capacidades organizacionais específicas, que determinam se a tecnologia será convertida em vantagem competitiva legítima ou em fonte adicional de risco.

A questão que orienta este ensaio teórico pode ser formulada da seguinte maneira: quais capacidades organizacionais medeiam a relação entre a adoção de inteligência artificial e o desempenho ESG, condicionando a criação ou a destruição líquida de valor corporativo? Defende-se, ao longo do texto, que a conversão do potencial técnico da inteligência artificial em resultados sustentáveis depende do desenvolvimento deliberado de competências internas voltadas à integração e à qualidade dos dados, à aprendizagem organizacional contínua, à governança algorítmica responsável e à capacidade de transformar análises técnicas em decisões estratégicas orientadas simultaneamente ao valor econômico e ao valor socioambiental.

O objetivo deste artigo é propor um modelo teórico que explique como essas capacidades organizacionais condicionam a relação entre inteligência artificial, desempenho ESG e criação de valor. Para isso, o texto está estruturado em oito seções, além desta introdução. A segunda seção apresenta o referencial teórico, articulando a teoria das capacidades dinâmicas, a literatura sobre governança de inteligência artificial e os estudos sobre desempenho ESG. A terceira seção descreve o percurso metodológico adotado, de natureza teórico-integrativa. A quarta seção expõe o modelo proposto e suas quatro capacidades mediadoras, discutindo também sua orquestração, a maturidade organizacional necessária para sustentá-las e os riscos associados à dependência de fornecedores tecnológicos. A quinta seção examina os mecanismos de destruição de valor ligados ao uso da inteligência artificial em contextos ESG. A sexta seção discute implicações gerenciais e estratégicas, incluindo arquitetura de

governança e gestão do ciclo de vida dos sistemas. A sétima seção apresenta proposições teóricas e uma agenda de pesquisa. Por fim, a oitava seção traz as considerações finais.

Uma observação inicial é necessária. Este artigo não pretende esgotar a discussão sobre inteligência artificial nas organizações, tampouco propor um modelo definitivo. Seu propósito é mais modesto e, ao mesmo tempo, mais específico: oferecer uma estrutura conceitual que ajude pesquisadores e gestores a distinguir entre o uso instrumental da tecnologia e a construção de competências organizacionais capazes de sustentar, de fato, resultados ambientais, sociais e de governança consistentes ao longo do tempo.

REFERENCIAL TEÓRICO

O referencial teórico deste estudo articula três corpos de literatura que, embora tratados com frequência de maneira isolada, mantêm relações conceituais estreitas quando aplicados ao problema da adoção de inteligência artificial em contextos ESG: a teoria das capacidades dinâmicas, os estudos sobre governança de inteligência artificial e ética algorítmica, e a literatura sobre desempenho ESG e criação de valor. A escolha por essa articulação decorre de uma constatação simples: modelos que tratam a tecnologia como fator exógeno e neutro de produtividade tendem a subestimar o papel das escolhas organizacionais na determinação de seus efeitos. Sistemas de inteligência artificial são artefatos sociotécnicos, cujos resultados dependem tanto de suas propriedades técnicas quanto das rotinas, valores e estruturas de decisão das organizações que os implementam (Orlikowski, 2007). Essa premissa orienta a construção do modelo apresentado na quarta seção.

Capacidades dinâmicas e vantagem competitiva na economia digital

A visão baseada em recursos (Resource-Based View) sustenta que a vantagem competitiva de uma empresa decorre da posse de recursos valiosos, raros, difíceis de imitar e organizacionalmente explorados (Barney, 1991). Essa perspectiva, contudo, apresenta limitações relevantes em ambientes de mudança tecnológica acelerada, nos quais a simples posse de recursos não garante desempenho superior sustentado. A teoria das capacidades dinâmicas, desenvolvida por Teece, Pisano e Shuen (1997) e posteriormente refinada por Teece (2007), responde a essa limitação ao deslocar o foco da posse estática de recursos para a habilidade organizacional de integrá-los, construí-los e reconfigurá-los diante de mudanças no ambiente competitivo. Segundo Teece (2007), essa habilidade

se desdobra em três microfundamentos: a capacidade de percepção (sensing), voltada à identificação de oportunidades e ameaças; a capacidade de apreensão (seizing), relacionada à mobilização de recursos para captar o valor identificado; e a capacidade de transformação (transforming), associada à reconfiguração contínua de ativos, estruturas e rotinas organizacionais.

Eisenhardt e Martin (2000) contribuem para essa discussão ao destacar que capacidades dinâmicas não constituem, elas mesmas, fonte direta de vantagem competitiva sustentável; funcionam antes como processos organizacionais que permitem à empresa reconfigurar sua base de recursos de maneira consistente com as mudanças do ambiente. Essa distinção é relevante para o argumento deste artigo. Recursos tecnológicos avançados, entre eles sistemas de inteligência artificial, representam, em sua origem, ferramentas de eficiência operacional. Sua transformação em diferencial competitivo duradouro depende do desenvolvimento de rotinas organizacionais de ordem superior, capazes de absorver a tecnologia, adaptá-la a processos decisórios específicos e realinhá-la às demandas de reguladores, investidores e demais públicos de interesse.

Ao aplicar essa lógica ao contexto da inteligência artificial, é possível argumentar que a tecnologia, isoladamente, não constitui fonte de vantagem competitiva sustentável, tampouco de desempenho ESG superior. Sistemas de aprendizado de máquina, plataformas analíticas e ferramentas de automação encontram-se hoje amplamente disponíveis no mercado, muitas vezes sob a forma de serviços em nuvem oferecidos por um número reduzido de grandes fornecedores. Essa disponibilidade reduz o potencial de diferenciação associado à mera aquisição da tecnologia (Wamba et al., 2021). O diferencial competitivo desloca-se, assim, para as rotinas organizacionais que permitem à empresa utilizar esses recursos de maneira consistente com sua estratégia e com as expectativas de seus públicos de interesse, argumento que fundamenta a proposição das capacidades mediadoras discutidas na quarta seção deste artigo.

Cabe observar que essa mudança de foco, da posse de recursos para a capacidade de orquestrá-los, tem consequências práticas para a forma como conselhos de administração e investidores avaliam projetos corporativos de inteligência artificial. Anúncios de investimento em tecnologia costumam ser recebidos, no mercado de capitais, como sinal positivo, associado a expectativas de ganhos de eficiência e inovação. A teoria das capacidades dinâmicas sugere cautela em relação a essa leitura: o valor do investimento depende menos do montante alocado em hardware, software ou

licenciamento de modelos, e mais da existência prévia, ou do desenvolvimento paralelo, de rotinas organizacionais capazes de absorver essa tecnologia e reconfigurar processos a partir dela. Investidores e conselheiros que avaliam apenas o volume de recursos destinados à inteligência artificial, sem examinar a maturidade organizacional subjacente, correm o risco de superestimar sistematicamente o potencial de criação de valor desses investimentos.

Governança da inteligência artificial, ética algorítmica e legitimidade organizacional

A governança corporativa tradicional foi concebida, em grande medida, para tratar problemas de agência entre acionistas e executivos, com ênfase em mecanismos de controle financeiro e contábil. A crescente presença de sistemas algorítmicos em processos decisórios corporativos amplia o escopo dessa governança, exigindo mecanismos capazes de supervisionar decisões automatizadas, muitas vezes opacas e de difícil auditoria. A literatura sobre ética e governança de inteligência artificial tem se consolidado como campo relevante para a administração estratégica, reunindo contribuições sobre princípios éticos aplicáveis a sistemas autônomos, explicabilidade técnica e mecanismos de responsabilização (Floridi et al., 2018; Jobin, Ienca & Vayena, 2019).

Jobin, Ienca e Vayena (2019), em revisão de documentos orientadores sobre ética em inteligência artificial publicados por governos, empresas e organizações internacionais, identificam convergência em torno de princípios como transparência, justiça, não maleficência, responsabilização e privacidade, ao mesmo tempo em que apontam divergências relevantes sobre como esses princípios devem ser operacionalizados. Essa constatação é importante para o argumento deste artigo: a existência de princípios declarados não equivale à existência de capacidades organizacionais capazes de aplicá-los de maneira consistente. Floridi et al. (2018) reforçam esse ponto ao propor um arcabouço ético para sistemas de inteligência artificial que articula benefício, não maleficência, autonomia, justiça e explicabilidade, destacando que a efetividade desses princípios depende de sua incorporação a processos organizacionais concretos, e não apenas a declarações de intenção.

A legitimidade organizacional, conforme discutida por Suchman (1995), refere-se à percepção generalizada de que as ações de uma organização são desejáveis, próprias ou apropriadas dentro de um sistema socialmente construído de normas, valores e crenças. Aplicada ao contexto da

inteligência artificial, essa noção sugere que a adoção de sistemas algorítmicos sem alinhamento normativo e ético adequado tende a gerar desconfi  ncia institucional e resist  ncia por parte de p  blicos afetados, comprometendo a legitimidade pragm  tica, moral e cognitiva da organiza  o. Em setores intensamente regulados, como o financeiro, o de sa  de e o de energia, a governan  a respons  vel de sistemas de intelig  ncia artificial funciona como mecanismo de prote  o reputacional e de conformidade com marcos regulat  rios crescentemente rigorosos, entre eles o Regulamento Geral de Prote  o de Dados europeu e, no Brasil, a Lei Geral de Prote  o de Dados.

Desempenho ESG e cria  o de valor multidimensional

O desempenho ESG deixou de ser tratado, na literatura de administra  o estrat  gica, como m  trica perif  rica associada a rela  es p  blicas ou filantropia corporativa, passando a integrar discuss  es centrais sobre gest  o de riscos e cria  o de valor de longo prazo (Eccles, Ioannou & Serafeim, 2014; Serafeim, 2014). Eccles, Ioannou e Serafeim (2014), em estudo comparativo entre empresas de alto e baixo desempenho em sustentabilidade, encontram evid  ncias de que a ado  o de pol  ticas voltadas    sustentabilidade corporativa est   associada a desempenho de mercado e desempenho cont  bil superiores no longo prazo, sobretudo quando essas pol  ticas se tornam parte da cultura organizacional, e n  o apenas de programas isolados de responsabilidade social. Serafeim (2014) argumenta ainda que a integra  o da sustentabilidade aos processos organizacionais afeta a qualidade da tomada de decis  o, ao ampliar o conjunto de informa  es consideradas relevantes para a gest  o de riscos e oportunidades.

A cria  o de valor sustent  vel, nesse sentido, n  o decorre da simples adi  o de pr  ticas socioambientais a modelos de neg  cio preexistentes, mas da concilia  o entre desempenho econ  mico e gera  o de valor para m  ltiplos p  blicos de interesse, incluindo trabalhadores, comunidades, fornecedores e o meio ambiente (Freeman, 1984; Porter & Kramer, 2006). Porter e Kramer (2006) introduzem o conceito de valor compartilhado para descrever estrat  gias que geram, simultaneamente, valor econ  mico para a empresa e valor social para a comunidade em que ela opera, argumentando que essa integra  o representa uma fonte relevante de inova  o e diferencia  o competitiva.

A inser  o da intelig  ncia artificial nesse debate exige cautela conceitual. A tecnologia constitui, primariamente, uma ferramenta de otimiza  o e

processamento de informação. Seus efeitos sobre o desempenho ESG e sobre a criação de valor dependem da finalidade estratégica, dos valores institucionais e das diretrizes normativas estabelecidas pela liderança organizacional. Sem direcionamento explícito voltado à sustentabilidade, sistemas de inteligência artificial tendem a otimizar variáveis de curto prazo, como redução de custos e aumento de produtividade, em detrimento de resultados ambientais e sociais de mais difícil mensuração (Amit & Zott, 2001; Dwivedi et al., 2021). Essa constatação reforça a necessidade de um modelo teórico que explicita as condições organizacionais sob as quais a inteligência artificial contribui, de fato, para o desempenho ESG, em vez de presumir essa contribuição como resultado automático da adoção tecnológica.

PERCURSO METODOLÓGICO

Este estudo caracteriza-se como um ensaio teórico de natureza integrativa (Torraco, 2016). A revisão integrativa distingue-se de revisões sistemáticas tradicionais por permitir a síntese conceitual de literaturas provenientes de campos disciplinares distintos, com o objetivo de propor novas perspectivas teóricas, e não apenas sintetizar achados empíricos previamente estabelecidos. Torraco (2016) argumenta que esse tipo de revisão é particularmente adequado para temas emergentes, nos quais a literatura ainda se encontra fragmentada entre diferentes tradições de pesquisa, situação que descreve com razoável precisão o estado atual dos estudos sobre inteligência artificial, ESG e criação de valor.

O desenvolvimento do modelo teórico seguiu quatro etapas. Na primeira, realizou-se um mapeamento das premissas presentes em estudos que relacionam, de maneira direta, a adoção de tecnologias digitais a resultados de sustentabilidade corporativa, identificando lacunas relacionadas à ausência de mediadores organizacionais nessas formulações. Na segunda etapa, articularam-se os construtos centrais da teoria das capacidades dinâmicas, da literatura sobre governança de inteligência artificial e dos estudos sobre desempenho ESG, buscando pontos de convergência conceitual entre esses campos. Na terceira etapa, identificaram-se e conceituaram-se as capacidades organizacionais que, segundo a literatura consultada, funcionam como elos entre a ferramenta tecnológica e os resultados socioambientais pretendidos pelas organizações. Na quarta etapa, consolidou-se o modelo propositivo, articulando as capacidades mediadoras identificadas aos mecanismos de risco e destruição de valor discutidos na quinta seção.

Como todo ensaio teórico, este artigo apresenta limitações metodológicas relevantes. A ausência de dados empíricos impede a verificação estatística das relações propostas, que devem ser compreendidas como proposições teóricas passíveis de investigação futura, e não como resultados comprovados. Além disso, a seleção da literatura consultada, embora tenha buscado abranger contribuições relevantes de diferentes tradições teóricas, reflete escolhas interpretativas do autor, que podem não esgotar a diversidade de perspectivas existentes sobre o tema. Essas limitações são retomadas na sétima seção, ao se propor uma agenda de pesquisa voltada à validação empírica do modelo.

Vale ainda registrar um cuidado metodológico adicional. Ensaaios teóricos correm o risco de apresentar suas proposições como se já contassem com respaldo empírico consolidado, quando na verdade representam hipóteses de trabalho derivadas de raciocínio dedutivo e de analogias com literaturas correlatas. Ao longo deste artigo, procurou-se manter clara a distinção entre afirmações amparadas em pesquisas empíricas específicas, devidamente referenciadas, e inferências teóricas propostas pelo autor a partir da articulação entre diferentes correntes conceituais. Essa distinção é especialmente relevante em um campo ainda em formação, no qual termos como governança de inteligência artificial e capacidades organizacionais para IA carecem de definições operacionais amplamente aceitas, o que dificulta comparações diretas entre estudos e reforça a necessidade de esforços futuros de padronização conceitual e metodológica.

O MODELO TEÓRICO DE CAPACIDADES ORGANIZACIONAIS MEDIADORAS

A premissa central deste modelo é a de que o impacto da inteligência artificial sobre o desempenho ESG e a criação de valor sustentável não ocorre de forma direta, sendo mediado por quatro capacidades organizacionais interdependentes: integração e qualidade dos dados, aprendizagem organizacional, governança responsável e conversão de informações em decisões estratégicas sustentáveis. A fragilidade de qualquer uma dessas dimensões compromete a cadeia de valor associada ao uso da tecnologia, podendo transformar investimentos expressivos em fontes de passivos operacionais, legais e reputacionais.

Essa proposição baseia-se em um argumento relativamente simples, mas com implicações relevantes para a estratégia corporativa: ativos tecnológicos tornaram-se, em grande medida, imitáveis. Softwares de análise preditiva, plataformas de processamento de linguagem natural e

serviços de aprendizado de máquina encontram-se disponíveis, sob condições contratuais semelhantes, para empresas de portes e setores distintos. A diferenciação estratégica, portanto, não reside no software adquirido, mas nas rotinas organizacionais desenvolvidas para governar, interpretar e aplicar criticamente os resultados gerados pelos sistemas algorítmicos. As subseções seguintes detalham cada uma das quatro capacidades propostas.

Integração e qualidade dos dados

A primeira capacidade mediadora refere-se à habilidade organizacional de coletar, tratar, estruturar e integrar bases de dados heterogêneas, incluindo tanto informações financeiras e operacionais tradicionais quanto métricas socioambientais. Sistemas de inteligência artificial dependem, de maneira essencial, de dados íntegros, acurados e representativos; a máxima consolidada na ciência de dados segundo a qual dados de baixa qualidade produzem resultados de baixa qualidade aplica-se de forma direta a esse contexto. A ausência de curadoria rigorosa compromete modelos preditivos e distorce diagnósticos estratégicos, com consequências que se estendem à própria mensuração do desempenho ESG.

Organizações que mantêm dados dispersos em silos departamentais enfrentam dificuldades sistemáticas para monitorar emissões indiretas de gases de efeito estufa associadas à cadeia de valor (escopo 3), auditar fornecedores quanto a práticas de trabalho análogo ao escravo ou mensurar indicadores internos de diversidade e inclusão. Sem integração de dados, modelos preditivos tornam-se pouco confiáveis ou produzem diagnósticos distorcidos sobre a sustentabilidade da organização. O desenvolvimento dessa capacidade requer investimento continuado em arquiteturas de dados, governança de metadados e conformidade com marcos regulatórios de proteção de dados, entre eles a Lei Geral de Proteção de Dados brasileira e o Regulamento Geral de Proteção de Dados europeu.

Há um aspecto adicional que merece destaque: a integração de dados financeiros e socioambientais em uma mesma plataforma analítica permite identificar correlações entre ecoeficiência e rentabilidade que, de outro modo, permaneceriam ocultas em análises setorializadas. Essa integração contribui para superar a dicotomia, ainda presente em parte da prática gerencial, entre desempenho financeiro e responsabilidade socioambiental, tratando-os como dimensões complementares de uma mesma estratégia de criação de valor. A qualidade dos dados funciona, assim, como alicerce sobre

o qual as demais capacidades mediadoras se apoiam; sem esse alicerce, as capacidades subsequentes operam sobre uma base informacional frágil.

Aprendizagem organizacional e capacidade absorativa

A tecnologia de inteligência artificial evolui em ciclos de inovação acelerados, o que exige das organizações o desenvolvimento contínuo de capacidade absorativa, entendida por Cohen e Levinthal (1990) como a habilidade de reconhecer o valor de informações novas e externas, assimilá-las e aplicá-las a fins comerciais. Essa capacidade não se resume à aquisição de conhecimento técnico especializado; envolve também o cultivo de competências analíticas e de pensamento crítico entre colaboradores, especialistas setoriais e gestores em diferentes níveis hierárquicos.

A aprendizagem organizacional assegura que a empresa não permaneça na posição de usuária passiva de soluções comerciais, mas se torne capaz de interpretar com profundidade analítica os resultados gerados por modelos preditivos, compreender suas margens de erro, identificar limitações metodológicas e ajustar processos operacionais e estratégicos com base nas lições extraídas da operação dos sistemas. Cohen e Levinthal (1990) destacam que a capacidade absorativa depende de conhecimento prévio relacionado, o que sugere que organizações com histórico de investimento em capacitação técnica e analítica tendem a apresentar maior facilidade para absorver e aplicar corretamente novas tecnologias de inteligência artificial. Essa constatação tem implicações relevantes para empresas de menor porte ou com estruturas de capacitação menos desenvolvidas, que podem enfrentar dificuldades adicionais para converter a adoção tecnológica em aprendizado organizacional efetivo.

A aprendizagem organizacional também cumpre função de retroalimentação em relação às demais capacidades do modelo. Auditorias de vieses revelam necessidades de treinamento específico; falhas operacionais identificadas na operação de sistemas motivam revisão de políticas de governança; e a participação de diferentes áreas funcionais na avaliação dos sistemas amplia a capacidade organizacional de identificar riscos que equipes técnicas, isoladamente, poderiam não perceber. Esse processo aproxima a gestão de sistemas de inteligência artificial de um processo de aprendizagem institucional contínuo, no qual métricas, decisões e consequências são periodicamente comparadas e revisadas.

Governança responsável da inteligência artificial

A terceira capacidade mediadora corresponde à habilidade organizacional de estabelecer diretrizes normativas de conformidade, explicabilidade técnica, mitigação de vieses algorítmicos e responsabilização institucional. Essa capacidade assegura que os modelos utilizados pela organização respeitem direitos fundamentais, evitem a reprodução de discriminações sistêmicas em processos sensíveis, como triagem de currículos, concessão de crédito ou alocação de recursos, e mantenham trilhas de auditoria acessíveis a órgãos reguladores e ao público externo.

A governança responsável distingue-se de declarações formais de princípios éticos por sua dimensão operacional. Não basta que a organização publique diretrizes gerais sobre uso ético de inteligência artificial; é necessário que essas diretrizes se traduzam em processos concretos de avaliação de risco, documentação de sistemas, testes de equidade e mecanismos efetivos de contestação por parte de indivíduos afetados por decisões automatizadas. A ausência dessa tradução operacional explica, em boa medida, por que organizações com códigos de ética robustos ainda registram incidentes graves relacionados a vieses algorítmicos e falhas de responsabilização (Jobin et al., 2019).

Conversão de informações em decisões estratégicas sustentáveis

A quarta capacidade mediadora representa a etapa final e, em certo sentido, mais decisiva do modelo: a habilidade gerencial de traduzir análises algorítmicas em decisões táticas e estratégicas orientadas à criação de valor compartilhado. Dispor de painéis sofisticados de previsão de riscos climáticos, projeções de eficiência energética ou modelos de precificação de risco social não é suficiente. A organização precisa de agilidade decisória, capital político interno e flexibilidade estrutural para realocar recursos financeiros, redesenhar portfólios de produtos sob critérios de economia circular e ajustar suas relações com públicos de interesse relevantes.

Essa capacidade é, com frequência, a mais negligenciada nos investimentos corporativos em inteligência artificial, que tendem a concentrar recursos na aquisição de tecnologia e na formação de equipes técnicas, sem investimento equivalente em processos decisórios capazes de absorver e aplicar os resultados gerados por essa tecnologia. O resultado observado em muitos casos é a produção de análises tecnicamente robustas que permanecem dissociadas da estratégia organizacional, sem qualquer efeito prático sobre decisões de investimento, alocação de recursos ou relacionamento com públicos de interesse.

Orquestração das capacidades e complementaridades

As quatro capacidades descritas não produzem resultados de forma isolada. A qualidade dos dados pode aumentar a precisão de modelos analíticos, mas não assegura que suas recomendações sejam convertidas em decisões coerentes com compromissos ambientais, sociais e de governança. Da mesma forma, uma estrutura formal de governança algorítmica pode se tornar meramente protocolar quando a organização não dispõe de profissionais capazes de interpretar limitações estatísticas, questionar premissas e traduzir evidências em escolhas estratégicas. O modelo pressupõe, portanto, complementaridade entre as quatro dimensões: o valor emerge da combinação entre infraestrutura informacional, aprendizagem, controle institucional e capacidade decisória, e não da presença isolada de qualquer uma dessas dimensões.

Essa complementaridade ajuda a explicar por que organizações que utilizam tecnologias semelhantes, muitas vezes adquiridas dos mesmos fornecedores, apresentam resultados distintos em termos de desempenho ESG. Recursos tecnológicos podem ser comprados no mercado; as rotinas que permitem integrá-los à estratégia, ao sistema de incentivos e aos mecanismos de responsabilização dependem de aprendizagem acumulada ao longo do tempo, o que confere a essas rotinas um caráter tácito e de difícil imitação, consistente com os argumentos clássicos da teoria das capacidades dinâmicas (Teece, 2007).

A interdependência entre capacidades também produz efeitos de retroalimentação relevantes para a governança organizacional. Decisões sustentáveis bem documentadas melhoram a qualidade dos dados disponíveis para ciclos futuros de análise; incidentes identificados em auditorias revelam necessidades específicas de capacitação; e a participação de áreas funcionais distintas, como jurídico, sustentabilidade e operações, amplia a capacidade de identificar riscos que equipes técnicas poderiam não perceber isoladamente. A organização deixa de avaliar apenas a acurácia técnica dos modelos e passa a examinar também a qualidade social, ambiental e econômica das decisões apoiadas por eles.

Maturidade organizacional e trajetória de desenvolvimento

As capacidades mediadoras podem ser compreendidas como dimensões de maturidade organizacional, que evoluem ao longo do tempo em resposta a experiências acumuladas, mudanças regulatórias e decisões de investimento. Em estágios iniciais, a organização tende a concentrar-se na aquisição de ferramentas e na execução de projetos-piloto, com integração limitada entre áreas e forte dependência de fornecedores externos. Os dados

permanecem dispersos, a avaliação de riscos ocorre de maneira reativa e as decisões são orientadas principalmente por indicadores operacionais de curto prazo. Nesse estágio, a inteligência artificial pode gerar ganhos localizados de eficiência, mas dificilmente altera de forma consistente o desempenho ESG ou o processo de criação de valor da organização.

Em um estágio intermediário, a organização começa a formalizar políticas, definir responsabilidades específicas e ampliar o letramento analítico de seus gestores. Projetos passam a ser avaliados segundo critérios de relevância estratégica, e a empresa desenvolve mecanismos para documentar modelos, acompanhar resultados e revisar decisões periodicamente. A aproximação entre as áreas de tecnologia, sustentabilidade, jurídico e negócio reduz a fragmentação característica do estágio anterior, embora ainda possam persistir iniciativas paralelas, métricas incompatíveis entre áreas e disputas relacionadas à propriedade dos dados. O principal desafio nesse estágio consiste em transformar estruturas formais de governança em rotinas efetivamente incorporadas ao cotidiano organizacional.

A maturidade avançada caracteriza-se pela incorporação da governança de inteligência artificial ao sistema regular de gestão da organização. A alta direção acompanha riscos e oportunidades de maneira sistemática, os modelos são monitorados ao longo de todo o seu ciclo de vida e os resultados são avaliados segundo dimensões econômicas, ambientais e sociais de forma integrada. A organização desenvolve capacidade para comparar alternativas tecnológicas, interromper sistemas inadequados quando necessário e aprender com incidentes ocorridos. Nesse estágio, a inteligência artificial deixa de ser tratada como iniciativa excepcional e passa a integrar a arquitetura decisória da empresa, submetida aos mesmos princípios de responsabilidade, controle e criação de valor aplicados a outros ativos estratégicos.

Essa evolução, contudo, não segue trajetória necessariamente linear. Mudanças regulatórias, substituição de lideranças, processos de fusão e aquisição ou incidentes graves podem acelerar ou interromper o desenvolvimento das capacidades descritas. Além disso, uma mesma organização pode apresentar maturidade elevada em determinadas áreas e fragilidade acentuada em outras. Instituições financeiras, por exemplo, costumam desenvolver controles robustos para modelos de concessão de crédito, área submetida a regulação específica e escrutínio contínuo, ao mesmo tempo em que utilizam sistemas pouco supervisionados em áreas como recursos humanos ou marketing, nas quais a pressão regulatória é

historicamente menor. A avaliação da maturidade organizacional deve, portanto, considerar processos e aplicações específicas, evitando classificações genéricas baseadas apenas na existência de políticas corporativas declaradas.

A trajetória de desenvolvimento das capacidades também precisa ser compatível com a materialidade dos riscos envolvidos em cada aplicação. Não é necessário que todas as organizações construam estruturas de governança idênticas, mas é indispensável que a intensidade dos controles acompanhe o impacto potencial das decisões automatizadas. Empresas de menor porte podem adotar mecanismos simplificados de governança, desde que preservem documentação mínima, responsabilização clara e possibilidade concreta de revisão das decisões tomadas com apoio de sistemas algorítmicos. O objetivo não é reproduzir estruturas burocráticas complexas em qualquer contexto, mas assegurar que a velocidade da inovação tecnológica não ultrapasse a capacidade institucional de compreender e administrar suas consequências.

Relações com fornecedores e dependência tecnológica

Grande parte das organizações adota inteligência artificial por meio de plataformas, serviços em nuvem e modelos desenvolvidos por terceiros, e não por meio de desenvolvimento interno integral. Essa estrutura amplia o acesso a tecnologias avançadas, inclusive para empresas de menor porte, mas cria relações de dependência que afetam diretamente a governança, a segurança da informação e a capacidade de aprendizagem organizacional. Quando contratos de fornecimento não garantem documentação adequada, portabilidade de dados ou informação transparente sobre atualizações dos sistemas, a organização contratante perde autonomia para avaliar criticamente a ferramenta utilizada e corrigir problemas identificados. A responsabilidade pelos impactos gerados, no entanto, permanece com a organização que utiliza a tecnologia em seus processos internos, independentemente de quem a desenvolveu.

A seleção de fornecedores de tecnologia deveria incluir, de maneira mais sistemática do que ocorre atualmente na prática corporativa, critérios técnicos e critérios ESG combinados. Além de desempenho técnico, custo e disponibilidade, é necessário avaliar práticas de proteção de dados, consumo de recursos computacionais e energéticos, gestão de vieses nos modelos oferecidos, transparência quanto a limitações conhecidas e condições de trabalho ao longo da cadeia tecnológica do fornecedor. Contratos precisam definir obrigações claras de comunicação de incidentes, de alterações

relevantes nos sistemas e de limitações técnicas conhecidas, além de assegurar acesso a evidências suficientes para fins de auditoria, respeitados os direitos de propriedade intelectual do fornecedor.

A dependência tecnológica excessiva pode ser reduzida por meio de estratégias de arquitetura modular, adoção de padrões interoperáveis, capacitação interna contínua e manutenção de conhecimento técnico mínimo sobre o funcionamento dos processos automatizados utilizados pela organização. Não é necessário que a empresa desenvolva internamente todos os componentes tecnológicos que utiliza, mas é recomendável que conserve competência suficiente para compreender o funcionamento geral das soluções adotadas, interpretar criticamente seus resultados e substituir fornecedores quando necessário, sem comprometer a continuidade de suas operações. Essa capacidade protege a continuidade operacional da organização e reduz o risco de que decisões estratégicas fiquem condicionadas, de maneira excessiva, a interesses comerciais de fornecedores externos.

MECANISMOS DE DESTRUIÇÃO DE VALOR E RISCOS DA IA EM CONTEXTOS ESG

A ausência, a atrofia ou o enfraquecimento das capacidades organizacionais mediadoras propostas neste modelo pode transformar a adoção de inteligência artificial em vetor de destruição de valor corporativo e de erosão do desempenho ESG. A literatura consultada permite identificar cinco mecanismos principais associados a esse risco: opacidade decisória, vieses e discriminação sistêmica, consumo intensivo de recursos computacionais, baixa responsabilização e desalinhamento estratégico. As subseções seguintes examinam cada um desses mecanismos com maior profundidade.

Opacidade, explicabilidade e risco decisório

A opacidade algorítmica não se resume à dificuldade técnica de compreender modelos complexos, especialmente aqueles baseados em redes neurais profundas. Ela decorre também de decisões organizacionais relacionadas à documentação, ao acesso à informação e à distribuição de responsabilidades dentro da empresa. Um sistema pode ser tecnicamente explicável, no sentido de que suas técnicas permitem interpretar como determinada decisão foi produzida, e ainda assim permanecer institucionalmente opaco, quando seus critérios não são comunicados aos públicos afetados ou quando não existe instância organizacional responsável

por revisar seus resultados. A explicabilidade relevante para a governança corporativa precisa, portanto, combinar compreensão técnica, justificativa gerencial e possibilidade concreta de contestação por parte de indivíduos afetados pelas decisões (Floridi et al., 2018).

Em decisões de alto impacto, é recomendável que a organização registre a finalidade do sistema utilizado, os dados empregados em seu treinamento, as métricas de desempenho adotadas, os limites conhecidos do modelo e as condições que exigem intervenção humana direta. Esse tipo de registro protege a organização contra a falsa impressão de neutralidade algorítmica, com frequência utilizada como justificativa para decisões que, na prática, refletem escolhas técnicas e institucionais específicas. O registro também permite reconstruir o processo decisório em auditorias internas e externas, algo especialmente relevante quando decisões automatizadas afetam acesso a crédito, oportunidades de emprego, acesso a serviços essenciais ou exposição a riscos ambientais.

Vieses, representatividade e impactos sociais

Os vieses observados em sistemas de inteligência artificial não resultam apenas de erros de programação, embora esses erros também possam ocorrer. Eles refletem, com frequência, desigualdades históricas presentes nos dados utilizados para treinar os modelos, escolhas inadequadas de variáveis, amostras pouco representativas de determinados grupos populacionais ou objetivos de otimização que ignoram consequências distributivas relevantes (Buolamwini & Gebru, 2018). Uma organização pode melhorar a eficiência média de determinado processo e, simultaneamente, piorar seus resultados para grupos específicos, situação que passa despercebida quando a avaliação de desempenho considera apenas médias agregadas.

Por essa razão, a avaliação de desempenho de sistemas de inteligência artificial deveria ser desagregada sempre que exista risco relevante de tratamento desigual, permitindo comparar taxas de erro, benefícios e custos entre diferentes públicos afetados. A governança responsável exige que a organização defina previamente quais diferenças de tratamento são aceitáveis, quais exigem correção imediata e quais tornam determinado sistema incompatível com seus compromissos ESG declarados. Esse julgamento envolve valores, direitos e prioridades estratégicas, e por isso não deveria ser delegado exclusivamente à equipe técnica responsável pelo desenvolvimento dos modelos. A participação de profissionais de conformidade, sustentabilidade, recursos humanos e áreas jurídicas reduz a

probabilidade de que decisões tecnicamente eficientes produzam efeitos socialmente injustos, tornando o controle de vieses uma responsabilidade organizacional compartilhada, e não uma atribuição isolada da área de tecnologia.

Pegada ambiental e eficiência computacional

O potencial ambiental positivo da inteligência artificial costuma ser associado, na literatura aplicada e na prática corporativa, à otimização energética, à manutenção preditiva de equipamentos e ao monitoramento de emissões. Esses benefícios, contudo, precisam ser comparados de forma sistemática com os recursos consumidos pelo próprio sistema utilizado. Modelos de grande escala demandam capacidade computacional significativa, energia elétrica, água para resfriamento de centros de dados e renovação frequente de equipamentos, com impactos ambientais que, em muitos casos, permanecem pouco visíveis para a organização usuária do sistema (Strubell, Ganesh & McCallum, 2019). A avaliação ambiental de sistemas de inteligência artificial não deveria considerar apenas a etapa de treinamento dos modelos, mas todo o seu ciclo de vida, incluindo armazenamento de dados, execução de inferências, atualizações periódicas e descarte de hardware.

A capacidade de integração de dados discutida anteriormente pode contribuir para esse tipo de controle, ao relacionar indicadores de desempenho técnico com dados de consumo energético e resultados operacionais. Em vez de adotar automaticamente o modelo mais complexo disponível no mercado, a organização deveria selecionar a solução proporcional ao problema que pretende resolver, avaliando se ganhos marginais de acurácia justificam custos ambientais e financeiros adicionais associados a modelos de maior escala. Delmas e Burbano (2011) alertam, em contexto mais amplo, para o risco de que iniciativas apresentadas como sustentáveis produzam, na prática, impactos ocultos incompatíveis com metas climáticas corporativas, fenômeno conhecido como esverdeamento corporativo (greenwashing). Esse risco aplica-se de maneira direta a projetos de inteligência artificial promovidos sob argumentos de sustentabilidade sem avaliação criteriosa de seus custos ambientais reais.

Responsabilização e supervisão humana

A delegação de tarefas decisórias a sistemas automatizados não elimina a responsabilidade humana pelos resultados dessas decisões. Ao contrário, exige que a organização defina com maior precisão quem autoriza o uso de determinado sistema, quem acompanha seu desempenho ao longo do tempo,

quem possui autoridade para interrompê-lo e quem responde, formalmente, por eventuais danos causados por seu funcionamento. Expressões genéricas como "decisão do algoritmo", ainda comuns em justificativas organizacionais, tendem a diluir a responsabilidade e a enfraquecer a governança, ao sugerir uma autonomia técnica que, na prática, não existe de forma isolada dos processos organizacionais que definiram e implementaram o sistema. A tecnologia deve permanecer inserida em uma cadeia decisória identificável, com autoridades, limites e procedimentos claros de escalonamento em caso de falhas.

A supervisão humana efetiva não se resume à inclusão formal de uma pessoa na etapa final do processo decisório. Quando o profissional responsável não possui informação suficiente, tempo adequado ou autonomia real para contrariar a recomendação gerada automaticamente, sua participação torna-se meramente formal, sem capacidade efetiva de correção. A supervisão precisa ser desenhada de modo que o responsável compreenda os fundamentos da recomendação apresentada, reconheça situações excepcionais que exijam análise adicional e possa suspender o processo decisório sem sofrer incentivos organizacionais contrários a essa decisão. Essa condição depende de investimento em treinamento, de desenho adequado de cargos e responsabilidades, e de uma cultura organizacional que favoreça o questionamento de recomendações automatizadas, em vez de tratá-las como decisões definitivas e incontestáveis.

Desalinhamento e destruição estratégica de valor

Projetos de inteligência artificial podem destruir valor mesmo quando funcionam corretamente do ponto de vista técnico. Isso ocorre, por exemplo, quando o sistema resolve problemas de baixa relevância estratégica, automatiza processos que deveriam ser reformulados em vez de apenas acelerados, ou desloca recursos financeiros e gerenciais de iniciativas mais importantes para a organização. O entusiasmo tecnológico, presente com frequência em discussões corporativas sobre inteligência artificial, pode levar a organização a definir a solução tecnológica antes de compreender adequadamente a necessidade que pretende atender, produzindo investimentos de difícil integração e resultados pouco relacionados à estratégia geral da empresa. A análise de materialidade ESG, entendida como a identificação sistemática de temas relevantes para determinado setor e para os públicos de interesse da organização, oferece um filtro útil para priorizar aplicações de inteligência artificial capazes de responder a riscos e oportunidades efetivamente significativos, em vez de projetos motivados principalmente por modismos de mercado.

O alinhamento estratégico entre projetos de inteligência artificial e a agenda ESG requer indicadores que ultrapassem medidas convencionais de produtividade e redução de custos. Um projeto voltado à gestão da cadeia de suprimentos, por exemplo, deveria ser avaliado também por sua contribuição efetiva à rastreabilidade de produtos, à melhoria das condições de trabalho ao longo da cadeia e à redução de impactos ambientais associados à produção e à logística. Um sistema utilizado na gestão de pessoas precisa considerar, além de indicadores de eficiência, a qualidade das decisões produzidas, seus efeitos sobre a diversidade da força de trabalho, a confiança dos colaboradores e o bem-estar organizacional. A criação de valor sustentável depende da incorporação dessas dimensões ao desenho inicial do projeto de inteligência artificial, e não de sua adição posterior, como item complementar em relatórios de sustentabilidade elaborados após a implementação do sistema.

IMPLICAÇÕES GERENCIAIS E ESTRATÉGICAS

As reflexões desenvolvidas ao longo deste artigo oferecem implicações relevantes para executivos, membros de conselhos de administração e gestores de sustentabilidade. Em primeiro lugar, evidenciam-se os limites de abordagens tecnocêntricas, que tratam a inteligência artificial como solução pronta e descolada do contexto institucional em que será utilizada. Investir em tecnologia sem investir, de maneira equivalente, no fortalecimento das capacidades organizacionais mediadoras discutidas neste artigo equivale, na prática, a acelerar processos que podem ser ineficientes ou insustentáveis, apenas com maior velocidade e menor possibilidade de correção tempestiva. Em segundo lugar, o modelo sugere que a gestão da sustentabilidade na era digital demanda a criação de instâncias interdisciplinares de governança algorítmica, reunindo especialistas em dados, profissionais de conformidade ESG e lideranças estratégicas, capazes de atuar na curadoria contínua dos dados utilizados, na auditoria periódica de vieses e na validação ética das decisões apoiadas por sistemas automatizados.

Arquitetura de governança e responsabilidades

A institucionalização das capacidades propostas neste modelo exige uma arquitetura de governança compatível com o porte, o setor de atuação e o nível de risco de cada organização. O conselho de administração deveria estabelecer princípios gerais e supervisionar riscos relevantes associados ao uso de inteligência artificial, enquanto a alta direção traduz essas orientações em políticas concretas, alocação de recursos e definição de prioridades. Comitês interdisciplinares podem contribuir significativamente

para a avaliação de projetos específicos, mas não deveriam substituir a responsabilidade das áreas proprietárias dos processos em que os sistemas são utilizados. Cada sistema relevante precisa ter um responsável de negócio identificável, um responsável técnico e uma função independente de controle, capaz de questionar decisões tomadas pelas duas primeiras instâncias quando necessário.

Uma política corporativa de inteligência artificial deveria estabelecer critérios claros para classificação de risco, documentação mínima exigida, procedimentos para aquisição de soluções externas, proteção de dados, testes prévios à implementação, supervisão humana contínua e critérios objetivos para descontinuação de sistemas. Sistemas de baixo impacto podem seguir procedimentos simplificados de governança, enquanto aplicações que afetam direitos fundamentais, segurança física ou decisões financeiras relevantes exigem avaliação mais aprofundada. Essa proporcionalidade evita burocratizar excessivamente iniciativas de baixo risco, sem reduzir o controle necessário sobre usos potencialmente danosos da tecnologia.

Ciclo de vida dos sistemas de inteligência artificial

A governança responsável precisa acompanhar todo o ciclo de vida do sistema, e não apenas sua fase inicial de implementação. Na fase de concepção, devem ser definidos com clareza o problema que o sistema pretende resolver, sua finalidade específica, os públicos potencialmente afetados e os critérios objetivos de sucesso a serem utilizados em sua avaliação. Durante o desenvolvimento, são necessários testes sistemáticos de qualidade, robustez técnica, segurança da informação e equidade entre diferentes grupos afetados. Antes da implantação em produção, a organização deveria avaliar impactos potenciais e preparar procedimentos de contingência para cenários de falha. Após a entrada em operação, o desempenho do sistema precisa ser monitorado continuamente, de modo a identificar mudanças na distribuição dos dados, deterioração progressiva da acurácia ou efeitos não antecipados durante as fases anteriores.

A descontinuação de sistemas também integra, de maneira relevante, esse ciclo de vida. Sistemas que perderam sua finalidade original, que utilizam tecnologias obsoletas ou que apresentam riscos desproporcionais em relação aos benefícios gerados não deveriam permanecer em operação apenas em razão do investimento já realizado em sua implementação, raciocínio que constitui uma forma reconhecida de viés de custo irreversível. A capacidade organizacional de encerrar ou substituir uma solução

tecnológica demonstra maturidade estratégica e reduz a tendência de perpetuar decisões anteriores independentemente de sua adequação atual. Registros detalhados de versões, alterações realizadas e incidentes ocorridos permitem que o aprendizado obtido em determinado projeto seja transferido para iniciativas futuras da organização.

Indicadores para avaliação de valor sustentável

A mensuração dos resultados obtidos com sistemas de inteligência artificial deveria combinar indicadores técnicos, econômicos e ESG, evitando a limitação a métricas puramente operacionais. Indicadores de acurácia, tempo de processamento e disponibilidade do sistema são necessários, mas insuficientes para avaliar seu valor real para a organização e para seus públicos de interesse. É recomendável que a organização acompanhe também consumo de recursos computacionais e energéticos, distribuição de erros entre diferentes grupos afetados, reclamações recebidas, decisões revertidas após contestação, percepção de confiança por parte de colaboradores e clientes, e efeitos observados sobre públicos relevantes específicos. Esse conjunto ampliado de indicadores permite identificar situações em que ganhos operacionais aparentes ocultam deterioração social ou ambiental não capturada por métricas técnicas convencionais.

O valor efetivamente criado por determinado sistema deveria ser avaliado em relação a uma linha de base clara e a alternativas plausíveis de solução para o mesmo problema. Nem toda melhoria observada decorre necessariamente da inteligência artificial adotada, e nem toda tarefa organizacional exige, de fato, automação para ser executada com qualidade adequada. Comparações sistemáticas com processos anteriores, projetos-piloto bem documentados e avaliações periódicas reduzem o risco de atribuição indevida de resultados à tecnologia, quando outros fatores organizacionais poderiam explicar as melhorias observadas. A qualidade dessa mensuração também depende do reconhecimento de efeitos de longo prazo, como dependência excessiva de fornecedores tecnológicos específicos, perda de conhecimento técnico interno decorrente da terceirização de funções críticas, ou alteração relevante nas relações da organização com trabalhadores e consumidores.

PROPOSIÇÕES TEÓRICAS E AGENDA DE PESQUISA

O modelo desenvolvido ao longo deste artigo permite formular proposições teóricas capazes de orientar estudos empíricos futuros. A primeira proposição estabelece que a relação entre adoção de inteligência artificial e

desempenho ESG tende a ser positiva apenas quando a organização apresenta níveis adequados de integração e qualidade dos dados utilizados pelos sistemas. A segunda proposição sugere que a aprendizagem organizacional amplia a capacidade da empresa de transformar resultados analíticos em mudanças concretas de processos e decisões estratégicas. A terceira proposição sustenta que a governança responsável reduz a incidência dos mecanismos de destruição de valor discutidos na quinta seção, associados a opacidade, vieses e baixa responsabilização institucional.

A quarta proposição afirma que a capacidade de conversão de informações em decisões sustentáveis medeia a relação entre capacidade analítica organizacional e criação de valor. Uma empresa pode produzir previsões tecnicamente robustas sem, no entanto, alterar seus investimentos ou práticas de gestão em razão delas; nesse caso, a inteligência gerada pelo sistema permanece dissociada da estratégia organizacional, apesar de sua qualidade técnica. A quinta proposição indica que as quatro capacidades apresentam efeito complementar entre si: a deficiência acentuada em qualquer uma delas tende a limitar os benefícios produzidos pelas demais, mesmo quando essas outras capacidades se encontram bem desenvolvidas. Essa hipótese de complementaridade pode ser examinada, em pesquisas futuras, por meio de modelos configuracionais ou de análises comparativas qualitativas entre organizações de diferentes setores e portes.

Pesquisas futuras poderiam desenvolver escalas específicas para mensurar cada uma das quatro capacidades propostas e testar empiricamente o modelo em diferentes contextos setoriais. Instituições financeiras, indústrias intensivas em recursos naturais, empresas de tecnologia e organizações do setor de saúde apresentam perfis de risco e estruturas regulatórias suficientemente distintos para justificar investigações específicas em cada um desses contextos. Estudos longitudinais permitiriam observar como as capacidades organizacionais evoluem ao longo do tempo, em resposta a incidentes relevantes, mudanças regulatórias ou processos de substituição tecnológica. Estudos de caso aprofundados poderiam, por sua vez, detalhar processos internos de aprendizagem organizacional e conflitos frequentemente observados entre áreas técnicas, áreas de sustentabilidade e a alta direção das empresas.

Uma frente adicional de pesquisa consiste em investigar resultados distributivos associados à adoção de inteligência artificial em contextos ESG. É possível que a criação de valor ocorra predominantemente para acionistas, enquanto custos relevantes são transferidos para trabalhadores,

consumidores ou comunidades afetadas pelas operações da organização. A avaliação empírica desse fenômeno deveria identificar, de maneira explícita, quem recebe os benefícios gerados pela tecnologia, quem assume os riscos associados a seu uso e como os mecanismos de governança discutidos neste artigo alteram essa distribuição ao longo do tempo. Essa abordagem amplia a análise para além do desempenho agregado da organização e aproxima a pesquisa sobre inteligência artificial e ESG de debates mais amplos sobre legitimidade organizacional, justiça distributiva e responsabilidade corporativa, temas centrais para a administração estratégica contemporânea.

CONSIDERAÇÕES FINAIS

A articulação entre inteligência artificial, desempenho ESG e criação de valor corporativo constitui um dos temas mais relevantes para a teoria e a prática da administração estratégica contemporânea. Como procurou-se demonstrar ao longo deste artigo, a implantação de ativos tecnológicos avançados não constitui, isoladamente, garantia de sustentabilidade corporativa ou de vantagem competitiva duradoura. O valor econômico e socioambiental gerado pela aplicação da inteligência artificial em contextos ESG é condicionado, de forma bastante direta, pela robustez de quatro capacidades organizacionais mediadoras: a integração rigorosa de dados, a aprendizagem institucional contínua, a governança algorítmica responsável e a capacidade gerencial de converter análises técnicas em decisões estratégicas sustentáveis.

Quando essas competências estão ausentes ou pouco desenvolvidas na organização, a tecnologia tende a atuar como aceleradora de riscos preexistentes, gerando opacidade decisória, vieses discriminatórios e destruição de valor em vez de sua criação. Esse achado tem implicações práticas relevantes para gestores: sugere que investimentos em inteligência artificial deveriam ser avaliados não apenas por seu potencial técnico, mas também pela capacidade organizacional disponível para absorvê-los de maneira responsável. A distinção entre adquirir tecnologia e desenvolver competência efetiva para utilizá-la de forma sustentável, central ao argumento apresentado neste artigo, parece particularmente relevante em um momento no qual muitas organizações enfrentam pressão para adotar soluções de inteligência artificial em ritmo acelerado, muitas vezes sem tempo adequado para desenvolver as capacidades necessárias para fazê-lo de maneira responsável.

O modelo proposto não pretende oferecer respostas definitivas sobre como cada organização deve estruturar sua governança de inteligência artificial, tampouco estabelecer um conjunto único de práticas aplicáveis indistintamente a diferentes setores e contextos. Seu propósito é mais modesto: oferecer uma estrutura conceitual capaz de orientar tanto a reflexão acadêmica quanto a prática gerencial sobre as condições organizacionais necessárias para que a inteligência artificial contribua, de fato, para o desempenho ESG e para a criação de valor sustentável, em vez de presumir essa contribuição como resultado automático da adoção tecnológica.

Como agenda para investigações futuras, recomenda-se a validação empírica deste modelo por meio de abordagens metodológicas variadas, incluindo modelagem de equações estruturais para testar as relações propostas entre as capacidades mediadoras e os resultados organizacionais, além de estudos de caso comparados entre diferentes setores industriais, capazes de aprofundar a compreensão sobre nuances setoriais na governança algorítmica voltada à sustentabilidade. Espera-se que este ensaio teórico contribua para deslocar o debate sobre inteligência artificial e sustentabilidade corporativa de uma perspectiva centrada exclusivamente na tecnologia para uma perspectiva centrada nas capacidades organizacionais necessárias para utilizá-la de maneira responsável, distinção que parece cada vez mais relevante à medida que essas tecnologias se tornam parte estrutural da estratégia corporativa contemporânea.

REFERÊNCIAS

- Amit, R., & Zott, C. (2001). Value creation in e-business. *Strategic Management Journal*, 22(6-7), 493–520. <https://doi.org/10.1002/smj.187>
- Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120. <https://doi.org/10.1177/014920639101700108>
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77–91.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128–152. <https://doi.org/10.2307/2393553>
- Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. *California Management Review*, 54(1), 64–87. <https://doi.org/10.1525/cm.2011.54.1.64>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI):

Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>

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., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, 21(10-11), 1105–1121. [https://doi.org/10.1002/1097-0266\(200010/11\)21:10/11<1105::AID-SMJ133>3.0.CO;2-E](https://doi.org/10.1002/1097-0266(200010/11)21:10/11<1105::AID-SMJ133>3.0.CO;2-E)

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.

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>

Orlikowski, W. J. (2007). Sociomaterial practices: Exploring technology at work. *Organization Studies*, 28(9), 1435–1448. <https://doi.org/10.1177/0170840607081138>

Porter, M. E., & Kramer, M. R. (2006). Strategy and society: The link between competitive advantage and corporate social responsibility. *Harvard Business Review*, 84(12), 78–92.

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>

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

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571–610. <https://doi.org/10.2307/258788>

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>

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509–533. [https://doi.org/10.1002/\(SICI\)1097-0266\(199708\)18:7<509::AID-SMJ882>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z)

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>

Wamba, S. F., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, 164, 120482. <https://doi.org/10.1016/j.techfore.2020.120482>

Whittaker, M. (2021). The steep cost of capture. *Interactions*, 28(6), 50–55. <https://doi.org/10.1145/3488666>

NOTA BIOGRÁFICA

CONTRIBUIÇÃO DO AUTOR

Lincon Lopes: conceituação, metodologia, investigação, análise formal, redação do manuscrito original, revisão e edição, e supervisão. Curadoria de dados e obtenção de financiamento não se aplicam a este estudo teórico-integrativo, que não utilizou conjunto de dados primários nem recebeu financiamento externo.

DECLARAÇÕES EDITORIAIS

Financiamento: o estudo não recebeu financiamento externo.

Conflito de interesses: o autor declara não haver conflito de interesse.

Uso de inteligência artificial: não declarado pelo autor.

APROVAÇÃO ÉTICA

Não se aplica. O estudo é de natureza teórico-integrativa e não envolveu participantes humanos, dados pessoais identificáveis ou experimentação animal.

DISPONIBILIDADE DOS DADOS

Não se aplica. O artigo não produziu conjunto de dados primários.

© 2026 O autor. Publicado pela Noemica sob a licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0).