REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

GOVERNANÇA ALGORÍTMICA E TOMADA DE DECISÃO ESTRATÉGICA: IMPACTOS DA INTELIGÊNCIA ARTIFICIAL NA ADMINISTRAÇÃO CONTEMPORÂNEA

ALGORITHMIC GOVERNANCE AND STRATEGIC DECISION-MAKING: IMPACTS OF ARTIFICIAL INTELLIGENCE ON CONTEMPORARY MANAGEMENT

Ciências Humanas, Engenharias, Ciências Sociais Aplicadas •

05/11/2026

REGISTRO DOI: 10.70773/revistatopicos/793607968

Eduardo Silva Vasconcelos[1]

Cleber Barbosa Iack[2]

Fabio Antonio da Silva Arruda[3]

Ary Luiz de Oliveira Peter Filho[4]

Cibely Maria Ferreira de Abreu[5]

Murillo Nazareno Cavalcante Aguiar[6]

Fábio Junio Olimpio Cezário[7]

Magno Oliveira Ramos[8]

Daniel Souza Tabosa[9]

Agnaldo Braga Lima[10]

RESUMO

A inteligência artificial passou a mediar atividades de previsão, recomendação e coordenação em organizações públicas e privadas, alterando o modo como informações são convertidas em decisões estratégicas. Este artigo analisa as implicações da governança algorítmica para a administração contemporânea, especialmente os limites de delegação, a qualidade dos dados, a supervisão humana, os vieses e a responsabilização institucional. Desenvolveu-se revisão sistematizada qualitativa com síntese narrativa de trabalhos acadêmicos e documentos normativos e institucionais verificáveis, sem produção de dados primários ou metanálise. A discussão integra estudos sobre arquiteturas decisórias humano–IA, paradoxos entre automação e ampliação de capacidades, controle organizacional, confiança e resultados de experimentos recentes. Examina também a LGPD, o debate legislativo brasileiro, a regulação europeia e referenciais do NIST, da OCDE e da ISO. Como contribuição teórico-propositiva, apresenta-se o Modelo Integrado de Governança Algorítmica Estratégica (MIGAE), estruturado em seis dimensões interdependentes. A síntese indica que previsões tecnicamente consistentes não garantem escolhas legítimas ou adequadas a contextos incertos. Conclui-se que a contribuição estratégica da IA depende de finalidade justificada, dados rastreáveis, arquitetura decisória proporcional ao risco, supervisão efetiva, mecanismos de contestação e aprendizagem contínua, cuja eficácia organizacional ainda requer validação empírica.

Palavras-chave: Governança algorítmica; Inteligência artificial; Tomada de decisão estratégica; Administração; Responsabilização; Gestão de riscos.

ABSTRACT

Artificial intelligence increasingly mediates forecasting,

recommendation, and coordination in public and private organizations, changing how information is converted into strategic decisions. This article examines the implications of algorithmic governance for contemporary management, focusing on limits to delegation, data quality, human oversight, bias, and institutional accountability. A qualitative systematized review with narrative synthesis was conducted using verifiable scholarly publications and institutional or regulatory documents, without primary data collection or meta-analysis. The discussion brings together research on human–AI decision structures, the automation–augmentation paradox, organizational control, trust, and recent field experiments. It also examines Brazil's data protection law and legislative debate, European regulation, and governance frameworks from NIST, OECD, and ISO. As a conceptual contribution, the paper proposes an Integrated Strategic Algorithmic Governance Model (MIGAE), organized around six interdependent dimensions. The review suggests that technically consistent predictions do not, by themselves, guarantee legitimate or contextually appropriate decisions. Strategic benefits depend on justified purposes, traceable data, risk-proportionate decision structures, meaningful human oversight, contestation mechanisms, and continuous organizational learning; the proposed model still requires empirical validation.
Keywords: Algorithmic governance; Artificial intelligence; Strategic decision-making; Management; Accountability; Risk management.

1. INTRODUÇÃO

A difusão de sistemas de inteligência artificial nas organizações desloca a discussão administrativa da simples informatização de rotinas para a redefinição das condições de produção, interpretação e uso do conhecimento decisório. Modelos preditivos, mecanismos

de recomendação, ferramentas de linguagem generativa e sistemas de otimização podem participar do planejamento de operações, da avaliação de investimentos, da seleção de pessoas e da análise de cenários competitivos. Entretanto, a disponibilidade de previsões não constitui, por si, uma política de decisão. O administrador precisa estabelecer objetivos, identificar limites de dados, considerar interesses afetados e responder por consequências que não podem ser reduzidas à acurácia de um algoritmo. A governança algorítmica surge, nesse cenário, como campo que articula estruturas de autoridade, mecanismos técnicos de controle e critérios institucionais de legitimidade. Faraj, Pachidi e Sayegh (2018) observam que algoritmos de aprendizagem alteram o trabalho organizacional pela opacidade operacional, digitalização extensiva e quantificação antecipatória; tais características ampliam a necessidade de examinar quem define os problemas e quem pode contestar as respostas.

A decisão estratégica apresenta dificuldades particulares porque envolve incerteza radical, horizontes temporais extensos, objetivos concorrentes e situações nas quais o passado fornece orientação imperfeita para o futuro. Uma previsão de demanda pode ajudar a dimensionar investimentos, mas pode tornar-se inadequada quando ocorrem rupturas regulatórias, alterações geopolíticas, novos comportamentos de consumo ou desastres ambientais. De modo semelhante, uma ferramenta que identifica padrões de desempenho histórico pode reforçar estruturas organizacionais que a estratégia pretende transformar. Shrestha, Ben-Menahem e von Krogh (2019) propõem distinguir estruturas decisórias conforme características da tarefa e da contribuição relativa de humanos e sistemas inteligentes, afastando a suposição de que todas as decisões possam seguir uma única arquitetura. O desafio passa a ser

escolher onde automatizar cálculos, onde combinar julgamentos e onde preservar deliberação humana substantiva.

A discussão também possui dimensão social. Sistemas que recomendam metas e distribuem recursos influenciam oportunidades, remuneração, ambiente de trabalho e acesso de grupos sociais a serviços. Critérios aparentemente neutros podem reproduzir assimetrias registradas nos dados; classificações corretas segundo determinada métrica podem ser inadequadas para uma finalidade diferente. Kellogg, Valentine e Christin (2020) mostram que tecnologias algorítmicas reorganizam formas de controle do trabalho por meio de recomendações, registros, avaliações e recompensas. Portanto, a governança não se esgota na segurança da informação ou na escolha de fornecedores. Ela inclui participação das áreas envolvidas, regras de responsabilização, documentação de escolhas e tratamento de conflitos entre eficiência, direitos e objetivos públicos ou empresariais. Uma organização pode demonstrar desempenho computacional satisfatório e, ainda assim, produzir decisões socialmente questionáveis.

Diante dessas tensões, a questão norteadora do artigo é: de que maneira a governança algorítmica pode qualificar a tomada de decisão estratégica sem dissolver a responsabilidade administrativa, a transparência e a capacidade humana de julgamento? O objetivo geral consiste em analisar criticamente os impactos organizacionais da inteligência artificial sobre decisões estratégicas e propor um referencial de governança aplicável à administração contemporânea. Como objetivos específicos, examinam-se arquiteturas de delegação, qualidade e proveniência dos dados, riscos de viés, supervisão, explicabilidade, competências gerenciais, impactos sobre o trabalho, normas relevantes e métricas de avaliação. A contribuição

pretendida é teórico-propositiva: integrar literatura organizacional, evidências empíricas selecionadas e orientações institucionais em um modelo verificável. Não se pressupõe que uma tecnologia seja intrinsecamente superior ao julgamento humano, nem que a adoção de mecanismos de governança assegure automaticamente desempenho estratégico.

2. METODOLOGIA

O estudo foi delineado como revisão sistematizada qualitativa com síntese narrativa e finalidade analítico-propositiva. A seleção documental contemplou trabalhos acadêmicos sobre decisões humano–IA, controle algorítmico, confiança em previsões, aprendizagem organizacional e impactos da automação, além de fontes institucionais de governança de inteligência artificial. Foram consultadas páginas de editoras e registros oficiais das publicações citadas, bem como documentos disponibilizados por NIST, ISO, OCDE, instituições europeias e órgãos legislativos brasileiros. O recorte temático priorizou publicações capazes de sustentar distinções conceituais ou evidências relevantes para decisões organizacionais, sem pretender cobrir integralmente a produção interdisciplinar sobre inteligência artificial. Expressões como algorithmic governance, strategic decision-making, human-AI collaboration, organizational control, AI risk management, explainability e automated decision-making orientaram as consultas temáticas e a organização dos eixos interpretativos.

O material foi lido segundo quatro categorias de evidência. A primeira reúne proposições conceituais que descrevem possibilidades e paradoxos de organização, sem mensurar necessariamente seus resultados. A segunda abrange estudos

experimentais ou observacionais, cujas conclusões permanecem vinculadas às amostras, tarefas e condições investigadas. A terceira corresponde a normas e instrumentos institucionais: seus requisitos e recomendações podem orientar práticas, mas não constituem demonstração empírica de eficácia gerencial. A quarta categoria é a de inferências elaboradas nesta revisão, incluindo o referencial de governança proposto ao final. Essas categorias são explicitadas ao longo da discussão para evitar a passagem indevida de plausibilidade à causalidade ou a generalização de um experimento delimitado para todas as organizações.

A síntese foi estruturada em dimensões relacionadas: autoridade estratégica e limites da delegação; infraestrutura de dados; qualidade e validade das previsões; controles de vieses e consequências distributivas; arquitetura de supervisão e contestação; confiança e aprendizagem de gestores; conformidade e responsabilização; e avaliação contínua dos resultados. Foram privilegiadas convergências e divergências entre fontes, com atenção especial aos contextos em que a capacidade analítica da IA pode coexistir com fragilidades na definição do problema. Não foram produzidos dados originais de empresas ou órgãos públicos, entrevistas, questionários, experimentos computacionais ou medidas próprias de produtividade. Tampouco houve triagem bibliográfica auditável, extração independente por revisores ou avaliação formal de risco de viés de todos os estudos identificados. Por isso, o trabalho não é caracterizado como revisão sistemática formal e não apresenta contagens de seleção ou fluxogramas PRISMA.

As limitações são relevantes para interpretar o alcance das conclusões. Parte das fontes empíricas foi produzida em contextos

específicos, como consultoria, saúde e equipes de inovação, e pode não representar pequenas empresas, administrações públicas ou ambientes organizacionais latino-americanos. Documentos normativos possuem objetivos distintos de pesquisas acadêmicas: sua existência não comprova adoção efetiva, redução de falhas ou conformidade automática. Ademais, regras de inteligência artificial podem sofrer alterações durante o processo legislativo e a implementação regulatória. As referências jurídicas foram tratadas de acordo com seu caráter vigente, normativo ou propositivo, considerando o estado documental verificável na elaboração. O referencial construído é, portanto, instrumento conceitual para planejamento e pesquisa, e sua validade operacional exige estudos prospectivos, dados organizacionais verificáveis e avaliação dos efeitos sobre pessoas afetadas.

3. FUNDAMENTOS DA GOVERNANÇA ALGORÍTMICA

A expressão governança algorítmica pode designar objetos distintos: a utilização de algoritmos para coordenar atividades, o conjunto de regras que governa os próprios sistemas e a distribuição de poder que emerge da mediação tecnológica. Para fins deste estudo, adota-se uma definição organizacional: governança algorítmica é o arranjo de responsabilidades, políticas, controles técnicos e instâncias de deliberação destinado a orientar o desenvolvimento, a aquisição, o uso e a revisão de sistemas automatizados com impacto decisório. Essa definição distingue o mecanismo utilizado para calcular uma recomendação do processo institucional que autoriza sua transformação em decisão. Governar um sistema significa delimitar finalidades, assegurar rastreabilidade, avaliar desempenho e permitir correção ou suspensão. Uma arquitetura pode ser tecnicamente

sofisticada sem dispor de governança proporcional aos efeitos de suas decisões.

A literatura de administração reconhece que decisões resultam de interação entre informação, objetivos, recursos e relações de autoridade. A introdução de modelos de aprendizagem não elimina essas condições; redistribui etapas entre pessoas, fornecedores, interfaces e procedimentos digitais. Faraj, Pachidi e Sayegh (2018) descrevem a política muitas vezes oculta embutida na digitalização do conhecimento organizacional. Ao selecionar atributos e definir um alvo de previsão, uma organização converte escolhas gerenciais em operações aparentemente técnicas. A definição do que deve ser otimizado, como reduzir custos ou maximizar retenção, carrega premissas sobre valor e prioridades. Um modelo pode cumprir fielmente sua função matemática enquanto negligencia aspectos não capturados por seus indicadores. O controle democrático ou corporativo sobre objetivos precede, assim, qualquer debate sobre transparência do código.

A governança também deve ser diferenciada de uma soma de procedimentos de conformidade. Auditorias periódicas, regras contratuais e testes de segurança são necessários, mas tornam-se insuficientes quando não existe responsável capaz de intervir. O NIST AI Risk Management Framework organiza funções de governar, mapear, medir e gerenciar, indicando que a supervisão do ciclo de vida deve acompanhar as condições concretas de aplicação (Tabassi, 2023). A ISO/IEC 42001:2023 estabelece requisitos para sistemas de gestão de inteligência artificial, oferecendo estrutura organizacional voltada à melhoria contínua. Esses referenciais são complementares, porém não equivalentes: o primeiro é instrumento voluntário de gestão de riscos; o segundo é uma norma de sistema de gestão. Sua

utilização responsável exige adaptação às finalidades, capacidades institucionais e obrigações efetivamente aplicáveis.

A governança é mais exigente quando os modelos integram processos decisórios com múltiplos efeitos indiretos. Uma previsão de rotatividade pode orientar programas de retenção, mas também influenciar promoções, recrutamento e distribuição de oportunidades. Se o sistema for avaliado apenas pela previsão estatística, decisões posteriores poderão reproduzir um círculo de intervenções que altera os próprios dados de treinamento. A gestão precisa registrar quais decisões foram tomadas com base em determinada versão, quais alternativas estavam disponíveis e quais resultados ocorreram após a intervenção. Esse registro não é exigência apenas documental: permite identificar mudanças de comportamento provocadas pela presença do algoritmo, conflitos entre áreas e situações nas quais a recomendação inicial deixou de ser pertinente. Governança algorítmica, nesse sentido, constitui uma modalidade de aprendizagem institucional organizada.

4. INTELIGÊNCIA ARTIFICIAL E A ESTRUTURA DAS DECISÕES ESTRATÉGICAS

Decisões estratégicas diferem das rotineiras pela combinação de baixa repetibilidade, consequências elevadas e dependência de condições externas em transformação. Um sistema treinado com decisões anteriores pode estimar resultados dentro de padrões conhecidos, mas não necessariamente avaliar mudanças de missão, redefinições de mercado ou prioridades de interesse público. Shrestha, Ben-Menahem e von Krogh (2019) identificam fatores como interpretabilidade, amplitude de alternativas, velocidade e replicabilidade ao discutir a combinação de decisões humanas e

computacionais. A utilidade desse referencial está em vincular a arquitetura decisória às características do problema, em lugar de definir previamente a substituição de profissionais como objetivo da transformação digital. Uma decisão estratégica pode exigir modelagem preditiva em uma etapa, interpretação qualitativa em outra e deliberação explícita na aprovação final.

A delegação integral pode ser pertinente a operações bem delimitadas, reversíveis e monitoráveis, mas perde justificativa quando há consequências difíceis de reparar ou dados insuficientes sobre contextos novos. Decisões sequenciais podem assumir duas formas: o sistema produz uma recomendação posteriormente examinada por humanos, ou pessoas definem alternativas que são avaliadas por modelos. Decisões agregadas combinam julgamentos de diferentes atores e fontes, exigindo procedimentos claros para resolver desacordos. Esses desenhos não constituem uma hierarquia universal de qualidade. Seu desempenho depende da distribuição de competências, da clareza da tarefa e dos mecanismos que impedem a mera assinatura humana de uma resposta computacional. A supervisão nominal pode oferecer aparência de controle sem alterar a autoridade substantiva exercida pelo sistema.

O paradoxo automação–ampliação, discutido por Raisch e Krakowski (2021), mostra que substituir tarefas e ampliar capacidades humanas não são processos independentes. Ao automatizar análises operacionais, uma organização pode liberar gestores para deliberar sobre cenários complexos; também pode enfraquecer a experiência necessária para perceber inconsistências, caso decisões rotineiras deixem de ser compreendidas. Em contraste, ferramentas de ampliação podem fortalecer o raciocínio, mas criar dependência de recomendações cuja produção não é visível ao usuário. Jarrahi (2018)

argumenta que a complementaridade entre inteligência humana e artificial se torna relevante em ambientes marcados por incerteza e ambiguidade. Portanto, uma estratégia de adoção deve explicitar quais capacidades pretende desenvolver ou preservar, quais tarefas poderão ser automatizadas e como ocorrerá a manutenção do conhecimento profissional.

O planejamento estratégico apoiado por IA precisa ainda distinguir previsão, prescrição e legitimação. A previsão estima cenários a partir de dados e hipóteses; a prescrição sugere ações a partir de critérios de otimização; a legitimação requer justificar por que determinada escolha é compatível com a missão, os recursos e os direitos envolvidos. Confundir essas funções reduz alternativas e torna o debate estratégico dependente de indicadores cuja origem pode ser desconhecida. Por exemplo, uma recomendação de fechamento de unidades pode projetar redução de custos sem modelar perdas de acesso, efeitos territoriais ou compromissos públicos. Cabe ao processo deliberativo apresentar limites, cenários de sensibilidade e consequências para grupos distintos. A participação da IA deve expandir a qualidade das perguntas, não comprimir a estratégia a uma resposta aparentemente precisa.

5. QUALIDADE DOS DADOS, PROVENIÊNCIA E INFRAESTRUTURA INFORMACIONAL

A decisão algorítmica depende de dados produzidos por processos sociais e administrativos anteriores à modelagem. Registros de clientes, desempenho, custos ou absenteísmo refletem categorias de cadastro, incentivos para preenchimento e diferenças na capacidade de observação das unidades. Uma base extensa não é necessariamente representativa das circunstâncias que orientarão

decisões futuras. Problemas de qualidade podem decorrer de campos ausentes, duplicidades, atualização irregular, seleção de casos ou mudança de significado dos indicadores. A governança deve tratar esses fatores como riscos estratégicos, porque uma previsão sofisticada pode amplificar uma distorção originada na coleta. A avaliação de dados, assim, não se resume a saneamento técnico: exige diálogo com as equipes que conhecem o processo e com pessoas potencialmente afetadas.

A proveniência permite reconstruir de onde vieram as informações, quais transformações sofreram e sob quais condições podem ser utilizadas. Sem rastreabilidade, uma organização pode não conseguir explicar por que duas pessoas receberam classificações distintas nem qual fonte sustentou uma recomendação de investimento. Um inventário de dados deve explicitar responsáveis, finalidades, restrições legais, periodicidade de atualização e limitações conhecidas. O mesmo vale para bases adquiridas de terceiros ou informações obtidas por serviços externos. Transparência sobre origem não significa divulgar dados pessoais ou segredos industriais indiscriminadamente; envolve produzir documentação suficiente para auditoria, contestação e correção. A avaliação de fontes deve observar se os dados são apropriados para a tarefa e se o tratamento posterior respeita o contexto em que foram coletados.

A generalização constitui desafio adicional. Modelos ajustados a uma unidade organizacional podem apresentar degradação de desempenho em setores, regiões e períodos distintos. Mudanças de comportamento dos clientes, novas regulamentações e alterações de infraestrutura criam deslocamentos de distribuição que exigem monitoramento. Métricas globais, como médias de erro, podem

ocultar falhas relevantes em subgrupos ou situações raras de alto impacto. Para decisões estratégicas, convém combinar validação retrospectiva, testes em cenários alternativos e acompanhamento prospectivo de resultados. Uma recomendação não deve receber status de evidência definitiva apenas porque o modelo obteve bom desempenho durante seu desenvolvimento. Sua confiabilidade precisa ser tratada como propriedade dependente do contexto e do tempo, reavaliada após mudanças materiais na operação.

A integração de ferramentas generativas acrescenta exigências próprias. Respostas textuais podem organizar informações e formular hipóteses, mas também misturar documentos válidos com conteúdo não confirmado ou produzir referências inexistentes. O perfil do NIST para inteligência artificial generativa destaca categorias de riscos relacionados à informação, à interação humano–sistema e à governança do ciclo de vida (Autio et al., 2024). Na administração, é necessário separar dados recuperados, inferências do modelo e decisões aprovadas. Relatórios que embasam investimentos ou mudanças de pessoal devem permitir verificação das fontes primárias relevantes, explicitar lacunas e documentar versões utilizadas. Uma redação convincente não substitui análise documental, e um resumo automatizado não pode ser a única base para uma decisão de alta consequência.

6. VIESES ALGORÍTMICOS, EQUIDADE E CONSEQUÊNCIAS DISTRIBUTIVAS

O risco de viés algorítmico não surge apenas de programação inadequada. Ele pode estar presente na definição da variável-alvo, na seleção de exemplos, na escolha de indicadores substitutos e no modo como resultados são aplicados. Uma instituição que utiliza

contratações anteriores como referência de qualidade pode reproduzir preferências históricas que não correspondem às competências necessárias para novas funções. Sistemas de crédito, precificação, seleção de fornecedores ou monitoramento de trabalhadores também podem gerar efeitos diferentes sobre populações que não estavam adequadamente representadas nos dados. A avaliação responsável deve distinguir desempenho estatístico de legitimidade social: uma classificação coerente com regularidades passadas não demonstra que essas regularidades devam orientar escolhas futuras. As preocupações com justiça e não discriminação presentes na Recomendação da OCDE sobre Inteligência Artificial reforçam a necessidade de procedimentos de prevenção, acompanhamento e reparação (OCDE, 2024).

Diferentes critérios de justiça podem produzir resultados que não coincidem perfeitamente. Igualar taxas de erro entre grupos, garantir tratamento consistente de perfis semelhantes ou corrigir diferenças de oportunidade são objetivos que exigem escolhas contextuais. A existência de métricas quantitativas não afasta a necessidade de deliberar sobre quais desigualdades são relevantes e quais informações podem legitimamente influenciar determinada decisão. Em organizações com múltiplas unidades, indicadores agregados podem ocultar situações locais de exclusão. O desenho de auditorias deve, portanto, considerar populações afetadas, hipóteses sobre mecanismos de desvantagem, qualidade dos registros e canais de escuta. Mesmo quando dados individuais não permitirem determinadas comparações, a organização continua responsável por examinar efeitos previsíveis e não utilizar a ausência de informação como prova de ausência de risco.

A identificação de viés também depende de observar decisões após a implantação. Um modelo utilizado para distribuir oportunidades de treinamento pode alterar trajetórias profissionais, de modo que futuras medições reflitam parcialmente escolhas produzidas pela própria ferramenta. Se pessoas classificadas como menos promissoras recebem menos recursos, eventual desempenho inferior posterior não comprova que a previsão original representava uma característica intrínseca. Trata-se de risco de retroalimentação institucional. Para reduzi-lo, é necessário examinar intervenções concomitantes, registrar exceções e testar hipóteses alternativas sobre diferenças observadas. Pesquisas sobre o controle algorítmico no trabalho mostram que sistemas de avaliação e recompensa podem modificar condutas, e não apenas retratá-las (Kellogg; Valentine; Christin, 2020). Essa perspectiva impede interpretar dados administrativos como espelhos neutros do comportamento humano.

A mitigação de danos envolve medidas técnicas, organizacionais e procedimentais. A organização pode rever objetivos de otimização, restringir variáveis problemáticas, ampliar testes, documentar limitações, oferecer revisão de decisões e criar mecanismos de suspensão. Nenhuma dessas medidas opera de modo inteiramente isolado. Um sistema pode apresentar boa paridade estatística e ainda ser inadequado por coletar dados em excesso; outro pode respeitar privacidade, mas operar com proxies de baixo significado para a função. A avaliação multicritério deve registrar conflitos entre valores e justificar decisões de projeto. Em processos de alto impacto, a gestão precisa demonstrar que considerou alternativas menos invasivas ou mais verificáveis, inclusive a não adoção de IA. Governança madura não pressupõe que todo problema de administração seja resolvido pela modelagem preditiva.

7. EXPLICABILIDADE, TRANSPARÊNCIA E CONTESTAÇÃO DAS DECISÕES

A transparência algorítmica deve ser compreendida como capacidade de oferecer informações relevantes para diferentes públicos, e não como obrigação indiferenciada de divulgar todos os detalhes de um sistema. Desenvolvedores precisam compreender comportamentos técnicos; gestores necessitam conhecer limites de uso, fontes e critérios de aprovação; auditorias requerem registros reconstituíveis; pessoas afetadas necessitam explicações úteis sobre consequências e possibilidades de revisão. O conteúdo adequado varia conforme riscos, contexto e natureza dos dados. A Recomendação da OCDE, atualizada em 2024, articula transparência, possibilidade de contestação e responsabilização ao longo do ciclo de vida de IA (OCDE, 2024). Nesse enquadramento, produzir uma explicação não equivale a provar que a decisão é justa; a explicação deve permitir investigar o resultado e fundamentar eventual correção.

Há diferenças importantes entre interpretabilidade do modelo, explicabilidade posterior e justificativa administrativa. Um modelo pode ser estruturalmente simples e utilizar variáveis inadequadas; outro pode ser tecnicamente complexo, mas operar em processo com documentação robusta, validação externa e mecanismos efetivos de revisão. Explicações posteriores podem ajudar a identificar fatores associados a uma classificação, porém não devem ser apresentadas como descrição infalível do mecanismo causal. A justificativa administrativa incorpora finalidades, alternativas, restrições e consequências que não se esgotam em contribuições de atributos estatísticos. Uma decisão de investimento, por exemplo, necessita demonstrar coerência com riscos, capacidade de

execução, compromissos contratuais e objetivos sociais, além da qualidade da projeção. Transparência significativa conecta essas camadas em vez de substituir decisão responsável por gráficos atraentes.

A contestação exige procedimento acessível e instância com autoridade para modificar resultados. Não basta oferecer canal para reclamações se as respostas são automáticas, genéricas ou incapazes de examinar o caso individual. Também não basta identificar um humano na cadeia decisória quando esse profissional não tem tempo, competência ou autonomia para divergir. Lebovitz, Lifshitz-Assaf e Levina (2022), em estudo situado no diagnóstico médico, observaram maneiras pelas quais profissionais lidam com opacidade ao utilizar IA em julgamentos críticos. Embora o contexto clínico não represente diretamente a administração empresarial, ele evidencia a importância de compreender o trabalho efetivo de interpretação, de testar recomendações e de lidar com incerteza. A aplicação gerencial dessa lição requer canais de escalonamento, documentação das discordâncias e proteção institucional à revisão fundamentada.

Políticas de transparência também precisam considerar sigilo empresarial, proteção de dados e segurança dos sistemas. O dever de explicar decisões não autoriza divulgação indiscriminada de dados pessoais, segredos ou informações que facilitem ataques. Pode ser necessário oferecer descrição de critérios relevantes, documentação controlada para auditoria, registro de versões e informações específicas para o indivíduo afetado. A escolha entre essas modalidades deve estar vinculada ao risco e às obrigações jurídicas concretas. A qualidade do procedimento pode ser avaliada por indicadores como tempo de resposta, proporção de decisões

revisadas com fundamento e capacidade de corrigir fontes de erro. Tais métricas não são resultados apurados neste estudo; representam propostas para pesquisas futuras. O objetivo da transparência deve ser tornar o poder decisório examinável, não criar ritual documental sem consequências práticas.

8. CONFIANÇA, AVERSÃO E APRECIAÇÃO DAS RECOMENDAÇÕES AUTOMATIZADAS

A confiança em sistemas inteligentes não pode ser reduzida à opinião favorável sobre tecnologia. Ela envolve expectativa de desempenho, compreensão de limites, condições de uso e possibilidade de verificar erros. A literatura experimental registra comportamentos aparentemente contrastantes: Dietvorst, Simmons e Massey (2015) identificaram aversão a algoritmos após participantes presenciarem falhas, mesmo quando a previsão algorítmica era comparativamente vantajosa; Logg, Minson e Moore (2019) encontraram, em determinados experimentos, maior adesão a recomendações apresentadas como algorítmicas do que a conselhos humanos. Esses resultados não se anulam porque foram obtidos em situações, comparações e perfis de participantes diferentes. Em conjunto, alertam contra a ideia de uma atitude humana única diante da IA. A qualidade da colaboração depende de como o sistema é apresentado, do conhecimento profissional e da natureza da decisão.

Organizações podem falhar por excesso ou insuficiência de confiança. Na aceitação automática, gestores utilizam uma recomendação sem examinar dados e premissas; na rejeição sistemática, descartam informações úteis por desconfiança genérica. A calibragem exige mostrar indicadores de desempenho

pertinentes, limites de validade, grau de incerteza e procedimentos para revisão. Treinamentos centrados apenas em comandos de ferramenta são insuficientes. Profissionais precisam compreender quando modelos se tornam menos confiáveis, que sinais exigem investigação e como documentar discordância tecnicamente sustentada. O conhecimento necessário não é idêntico em todas as funções: conselhos de administração precisam avaliar exposição estratégica e responsabilidades; gerentes operacionais devem interpretar métricas específicas; equipes técnicas precisam dominar validação e monitoramento. Um programa de governança deve ajustar esses requisitos a cada papel.

A confiança também é influenciada por incentivos institucionais. Quando cumprir a recomendação de um painel é premiado e divergir exige explicações difíceis, a revisão humana torna-se formalidade. O gestor pode aceitar um resultado inadequado para evitar responsabilidade individual, enquanto a organização passa a atribuir à máquina decisões que foram previamente autorizadas por pessoas. Para enfrentar esse fenômeno, convém registrar não apenas casos de concordância, mas justificativas de alteração, erros identificados e medidas adotadas após incidentes. A cultura de aprendizagem deve valorizar contestações fundamentadas e evitar que relatórios de exceção sejam tratados como sinal de desobediência. A governança, assim, intervém no ambiente social que determina o comportamento diante da tecnologia, não apenas nas propriedades do software.

O desenho da interface exerce papel relevante nesse processo. Recomendações apresentadas como certezas podem desencorajar investigação; sistemas que indicam intervalos, premissas e dados ausentes podem facilitar julgamento crítico, desde que a informação

seja compreensível e não sobrecarregue o usuário. A exposição contínua a alertas irrelevantes pode produzir fadiga e reduzir a atenção aos casos críticos. Em contrapartida, ocultar a incerteza para aumentar a adoção pode comprometer qualidade e responsabilidade. A avaliação de interfaces deveria comparar decisões efetivamente tomadas, erros corrigidos, uso adequado de informações e custo cognitivo, e não apenas satisfação com a ferramenta. O objetivo é alcançar confiança proporcional à confiabilidade demonstrada em tarefas pertinentes, preservando competências profissionais para perceber anomalias e modificar trajetórias de ação.

9. COLABORAÇÃO HUMANO–IA E FRONTEIRAS DO DESEMPENHO

Evidências empíricas recentes reforçam que ganhos de desempenho com IA dependem da tarefa e de como o trabalho é organizado. Dell'Acqua et al. (2026a), ao investigar profissionais de consultoria em experimento de campo, utilizaram a noção de fronteira tecnológica irregular para caracterizar situações nas quais assistência generativa melhora determinados desempenhos e pode prejudicar outros, mesmo em atividades aparentemente próximas. Esse resultado não autoriza a conclusão de que toda decisão estratégica com IA será superior ou inferior. Ele demonstra a necessidade de classificar tarefas, observar limites e medir consequências no próprio contexto de utilização. Uma organização que registra apenas produtividade média pode deixar de reconhecer falhas concentradas em tarefas específicas, precisamente aquelas mais sensíveis para a continuidade do negócio.

Outra pesquisa de Dell'Acqua et al. (2026b), realizada com profissionais em desafios de inovação de produtos, investigou como

o acesso à IA interage com trabalho individual e em equipe. Os resultados sugerem potencial para ampliar o desempenho em tarefas delimitadas e reduzir diferenças de perspectiva funcional entre participantes, sem demonstrar que a IA substitui todas as vantagens da colaboração humana. O desenho experimental fornece evidência útil sobre certas atividades de inovação, mas não constitui validação de sistemas autônomos para direção empresarial, licitação pública ou decisões sobre pessoas. A transferência para novos contextos exige considerar complexidade do problema, consequências dos erros, qualidade dos dados e possibilidade de observação independente dos resultados. O valor da evidência reside também em delimitar o que ela efetivamente testa.

A administração estratégica envolve dimensões que escapam à produção de respostas plausíveis. Negociar interesses conflitantes, interpretar relações políticas, compreender compromissos sociais e deliberar sobre riscos não são tarefas redutíveis à geração de opções textuais. A IA pode ampliar o repertório de cenários e apoiar análises de consistência; a decisão final requer identificar quem arcará com perdas e quem terá legitimidade para assumir obrigações. Essa divisão não decorre de uma suposta superioridade humana em todas as atividades, mas da distinção entre capacidades analíticas e estruturas de responsabilidade. Um modelo não possui deveres organizacionais comparáveis aos de administradores legalmente investidos de poder. A arquitetura de colaboração deve traduzir esse fato em fluxos de aprovação, evidências recuperáveis e responsabilidades nominadas.

A experimentação organizacional constitui estratégia apropriada quando se pretende explorar novos usos. Projetos-piloto devem definir previamente objetivos, linhas de comparação, critérios de

interrupção e registro de incidentes. Em atividades críticas, testes podem começar em ambiente sem efeitos sobre pessoas ou serviços reais, utilizando dados protegidos e revisão independente. A avaliação precisa captar benefícios e custos: tempo de elaboração de alternativas, qualidade dos resultados, esforço de verificação, regressões, riscos de confidencialidade e impactos sobre capacidades humanas. A recomendação de ampliar o uso somente após evidência relevante não significa rejeitar inovação. Significa tratar a adoção como hipótese gerencial sujeita à refutação e à correção. Uma organização que não consegue dizer em quais tarefas o sistema melhora resultados ainda não dispõe de base suficiente para generalizar sua utilização.

10. TRABALHO, PODER ORGANIZACIONAL E RESPONSABILIDADE GERENCIAL

A incorporação de decisões algorítmicas modifica estruturas de trabalho porque transforma a produção de registros em instrumento de coordenação. Sistemas que distribuem tarefas, classificam desempenho e produzem recomendações sobre trabalhadores podem reduzir tempo administrativo, mas também intensificar monitoramento e alterar a autonomia profissional. Kellogg, Valentine e Christin (2020) descrevem múltiplos mecanismos de controle por algoritmos, evidenciando que a mediação digital não é neutra quanto ao poder. Em processos estratégicos, essas informações podem aparecer como dados agregados de produtividade ou eficiência. O gestor precisa examinar se as métricas representam adequadamente a atividade e se incentivos associados a elas alteram comportamentos de maneira indesejada. Melhorar um indicador pode não significar melhorar o serviço, a aprendizagem ou as condições de trabalho.

Decisões de seleção, promoção, remuneração ou desligamento exigem cautela reforçada porque afetam trajetórias individuais e podem reproduzir desigualdades. Um modelo pode identificar correlações históricas entre atributos e resultados sem distinguir circunstâncias de oportunidade, disponibilidade de treinamento ou barreiras institucionais. A gestão de pessoas não deve transformar predições probabilísticas em diagnósticos definitivos sobre capacidade profissional. Procedimentos adequados incluem revisão dos critérios de cargo, avaliações contextualizadas, documentação de dados e possibilidade de contestação. Também é necessário examinar grupos potencialmente prejudicados e assegurar que instrumentos utilizados sejam apropriados à finalidade declarada. A presença de supervisão humana não resolve automaticamente o problema quando os responsáveis recebem apenas uma classificação e não possuem meios de investigar suas limitações.

As competências da liderança passam a incluir alfabetização em dados e interpretação de resultados, mas permanecem vinculadas à escuta, negociação e responsabilidade ética. Um gestor que conhece a taxa média de acerto de um modelo não está necessariamente preparado para avaliar seu impacto sobre uma equipe ou comunidade. Formação relevante combina noções de risco, proteção de dados, leitura de métricas, avaliação de decisões, aspectos jurídicos e capacidade de dialogar com especialistas. A distribuição dessas competências deve acompanhar as decisões efetivamente desempenhadas por cada função. Em organizações pequenas, a ausência de equipes técnicas extensas recomenda reduzir a complexidade das aplicações ou buscar avaliação externa; não justifica delegação irrestrita a fornecedores. Responsabilidade gerencial inclui reconhecer limites de capacidade institucional e adequar a tecnologia ao contexto disponível.

A participação dos trabalhadores constitui fonte de informação para detectar efeitos não previstos. Quem executa as tarefas pode identificar discrepâncias entre o indicador e o trabalho real, formas de adaptação oportunista às métricas ou consequências de decisões aparentemente eficientes. Canais de consulta e revisão devem existir antes e depois da implantação, especialmente quando há mudanças expressivas de monitoramento e distribuição de atividades. A liderança precisa demonstrar como contribuições recebidas foram consideradas e que procedimentos podem ser modificados. Sem retorno verificável, participação torna-se prática meramente simbólica. O desenho de governança também deve preservar privacidade e evitar vigilância desproporcional, reconhecendo que registros digitais produzidos no trabalho não autorizam qualquer finalidade posterior. O sucesso da transformação algorítmica depende tanto da qualidade técnica das ferramentas quanto da legitimidade dos processos nos quais operam.

11. MARCOS REGULATÓRIOS E O CONTEXTO BRASILEIRO

A regulamentação da inteligência artificial pode influenciar escolhas estratégicas antes mesmo de estabelecer proibições específicas. Exigências de transparência, proteção de dados, documentação e responsabilização alteram custos de implantação, critérios de contratação e avaliação de riscos reputacionais. No Brasil, a Lei nº 13.709/2018, Lei Geral de Proteção de Dados Pessoais, estabelece parâmetros para o tratamento de dados pessoais e prevê, em seu artigo 20, o direito de solicitar revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem interesses do titular. O dispositivo abrange perfis pessoais, profissionais, de consumo e de crédito, além de

prever informações sobre critérios e procedimentos, observados limites legais relativos a segredos comercial e industrial (Brasil, 2018). É fundamental não afirmar que esse artigo assegura, em sua redação vigente, revisão necessariamente efetuada por pessoa natural, exigência que não consta do texto atual.

A LGPD não se confunde com lei geral sobre todos os impactos da IA. Um algoritmo pode produzir efeitos sobre logística, consumo energético ou planejamento de investimentos sem necessariamente realizar tratamento de dados pessoais em todas as operações. Quando existem dados pessoais, a organização precisa examinar bases legais, finalidade, adequação, necessidade, segurança e direitos dos titulares, sem presumir que o emprego de IA cria uma hipótese autônoma de autorização. No plano administrativo, a gestão de risco deve contemplar fornecedores que coletam, armazenam ou reutilizam informações submetidas a ferramentas externas. Contratos, políticas de acesso e rotinas de retenção precisam ser compatíveis com as finalidades declaradas. A avaliação jurídica, portanto, deve ocorrer na concepção do projeto e nas modificações posteriores, em vez de ser solicitada apenas após a implantação.

O debate legislativo brasileiro também exige precisão temporal. O Projeto de Lei nº 2.338/2023, destinado a disciplinar o uso da inteligência artificial, foi aprovado pelo Senado Federal e remetido à Câmara dos Deputados, onde permanecia em tramitação na verificação realizada em outubro de 2026 (Brasil, 2023). Suas proposições podem informar discussões sobre governança e risco, mas não devem ser apresentadas como obrigações legais já vigentes. Da mesma forma, recomendações internacionais e normas voluntárias não substituem o ordenamento nacional. Confundir

projeto legislativo, norma técnica e obrigação legal constitui erro de gestão que pode induzir investimentos mal direcionados ou criar falsa sensação de conformidade. Uma política responsável identifica a natureza e o âmbito de aplicação de cada instrumento, bem como a necessidade de acompanhamento de mudanças.

No plano internacional, o Regulamento (UE) 2024/1689 estabelece disciplina europeia para sistemas de IA e classifica determinadas utilizações em emprego, gestão de trabalhadores e recrutamento como aplicações de alto risco, conforme seus critérios e condições (União Europeia, 2024). A Recomendação da OCDE, revisada em 2024, fornece princípios sobre desenvolvimento confiável, transparência, robustez e prestação de contas (OCDE, 2024). A ISO/IEC 42001:2023, por sua vez, estabelece requisitos para sistema organizacional de gestão de IA, enquanto o NIST AI RMF 1.0 apresenta estrutura voluntária para avaliação de riscos (ISO; IEC, 2023; Tabassi, 2023). O diálogo entre esses referenciais pode apoiar empresas multinacionais e instituições brasileiras, desde que se observem diferenças de jurisdição, vigência, objetivo e natureza normativa. A convergência de valores não significa identidade de requisitos.

12. CONTROLES OPERACIONAIS, CIBERSEGURANÇA E RISCOS DE TERCEIROS

A governança algorítmica depende de controles que atravessam o ciclo de vida do sistema: definição de finalidade, aquisição, desenvolvimento, treinamento, validação, operação, monitoramento e desativação. Em organizações complexas, a responsabilidade costuma ser distribuída entre áreas de negócio, tecnologia, segurança, jurídico, auditoria e gestão de pessoas. Essa distribuição

exige coordenação para evitar lacunas nas quais todos utilizam os resultados, mas nenhum setor responde pelas premissas. Um inventário de aplicações deve informar responsáveis, finalidade, criticidade, fontes de dados, fornecedores e mecanismos de intervenção. A gestão de mudanças precisa registrar atualizações de modelos, alterações de parâmetros e substituições de serviços, pois um mesmo produto comercial pode apresentar comportamento distinto após mudanças técnicas. A governança se enfraquece quando o inventário formal não corresponde aos sistemas efetivamente utilizados em departamentos e equipes.

Riscos de cibersegurança também possuem dimensão estratégica. Ferramentas generativas conectadas a documentos internos e sistemas operacionais podem expor informações confidenciais, processar instruções adversariais ou executar ações além do escopo previsto. O perfil de IA generativa do NIST recomenda tratar riscos específicos desses sistemas mediante controles vinculados à finalidade e ao ambiente de uso (Autio et al., 2024). A organização deve estabelecer fronteiras de permissões, registrar operações sensíveis, controlar integrações e definir situações nas quais o sistema apenas propõe ações sem executá-las. Em atividades de alto impacto, a mesma ferramenta não deveria ser a única responsável por produzir recomendações, validar seu próprio resultado e autorizar sua implementação. O princípio de separação de funções continua pertinente, mesmo quando parte do trabalho é automatizada.

A contratação de soluções de terceiros pode criar dependências informacionais difíceis de perceber na fase comercial. Documentos de apresentação enfatizam capacidades, mas a organização necessita examinar limitações de serviço, condições de tratamento

de dados, possibilidade de auditoria, disponibilidade de registros e procedimentos diante de incidentes. Também convém avaliar portabilidade, custos de saída e viabilidade de manter operações caso o fornecedor altere preços, modelos ou políticas. Uma ferramenta incorporada ao planejamento estratégico pode tornar-se infraestrutura crítica sem que a organização disponha de mecanismos para verificar suas premissas. A governança de fornecedores deve, portanto, alinhar requisitos de segurança, documentação, continuidade e responsabilização antes da expansão do uso. O risco não decorre somente de ataques; pode resultar de indisponibilidade, degradação silenciosa de qualidade ou perda da capacidade interna de avaliação.

O monitoramento pós-implantação deve combinar indicadores técnicos e organizacionais. Taxas de erro, deriva de dados, falhas de integração e incidentes de confidencialidade são importantes, mas não descrevem integralmente os efeitos sobre clientes, trabalhadores e estratégias. Relatórios periódicos podem examinar reclamações, revisões de decisões, dependência de fornecedores, resultados não previstos e permanência da justificativa de adoção. Os responsáveis precisam ter autoridade para reduzir escopo, alterar parâmetros ou suspender ferramentas quando riscos superam benefícios demonstráveis. O registro de incidentes deve alimentar revisão de processo, e não somente correção pontual do software. Sistemas confiáveis não são aqueles apresentados como incapazes de falhar; são aqueles cuja falha pode ser percebida, contida, examinada e utilizada para melhorar a organização.

13. PROPOSTA DE MODELO INTEGRADO DE GOVERNANÇA ESTRATÉGICA

Com base na literatura e nos instrumentos analisados, propõe-se um Modelo Integrado de Governança Algorítmica Estratégica, denominado nesta revisão MIGAE, como construção conceitual não validada empiricamente. Sua finalidade é tornar observável a ligação entre decisão estratégica, qualidade dos dados, responsabilidades e consequências. O modelo compreende seis dimensões articuladas: legitimidade da finalidade; qualidade e rastreabilidade informacional; adequação da arquitetura de decisão; supervisão e contestação; gestão de riscos e direitos; e aprendizagem institucional contínua. As dimensões não configuram escala estatística pronta, índice de maturidade universal ou certificação de conformidade. Trata-se de referência para estruturar perguntas, distribuir funções e planejar verificações antes e depois da incorporação de IA. Seu valor potencial depende de adaptação ao setor, à organização e à criticidade de cada aplicação.

A primeira dimensão demanda explicitar o problema a ser resolvido e justificar por que uma ferramenta algorítmica é apropriada. Perguntas relevantes incluem quais resultados a organização deseja melhorar, que alternativas não algorítmicas foram consideradas e quem poderá sofrer consequências da escolha. A segunda dimensão concentra-se nas bases informacionais: origem, representatividade, qualidade, limitações, proteção de dados e mecanismos de atualização. Sem avaliação das duas primeiras dimensões, indicadores de desempenho do modelo podem oferecer precisão aparente para uma decisão mal formulada. Por essa razão, um projeto não deveria avançar apenas porque demonstra capacidade tecnológica. A autorização estratégica pressupõe evidência de pertinência, custo proporcional e condições mínimas de controle.

A terceira dimensão define a arquitetura decisória de acordo com risco, reversibilidade e natureza das evidências. Atividades rotineiras com consequências limitadas podem admitir automação maior; tarefas estratégicas ambíguas requerem deliberação compartilhada e documentação das premissas; decisões que afetam direitos, renda ou acesso a oportunidades exigem salvaguardas proporcionais. A quarta dimensão estabelece supervisão significativa e mecanismos de contestação: responsáveis devem dispor de competência, tempo e autoridade para revisar recomendações, enquanto indivíduos afetados precisam de canais adequados de esclarecimento e revisão. Esses requisitos não são satisfeitos pela presença simbólica de um humano na interface. A governança deve verificar se a organização consegue efetivamente alterar decisões e corrigir problemas de dados ou de objetivos.

A quinta dimensão incorpora riscos técnicos, jurídicos, sociais e reputacionais, articulando práticas de segurança, privacidade, avaliação de vieses, continuidade operacional e gestão de fornecedores. Sua implementação pode dialogar com as funções do NIST AI RMF e com sistemas de gestão baseados na ISO/IEC 42001, sem presumir equivalência entre referenciais (Tabassi, 2023; ISO; IEC, 2023). A sexta dimensão exige aprendizagem: cada uso deve gerar evidências de resultados, custos, reclamações, falhas e mudanças de contexto. O modelo proposto prevê ciclos de revisão nos quais uma aplicação pode ser mantida, modificada, restrita ou descontinuada. A existência de metas quantitativas não dispensa apreciação de resultados qualitativos, sobretudo quando a finalidade inclui justiça, confiança e condições de trabalho.

Uma futura validação do MIGAE poderia utilizar estudos de caso longitudinais, comparação entre unidades organizacionais e análise

de documentos decisórios. Antes de implantar a ferramenta, seriam definidos indicadores de qualidade, tempo, custo, incidentes, contestabilidade e efeitos distributivos. Durante a operação, registrar-se-iam mudanças de processo, intervenções humanas e resultados efetivos, incluindo falhas e situações de recusa. A validação deveria examinar se as seis dimensões são compreensíveis para gestores, se geram decisões mais justificáveis e se possuem aplicabilidade em organizações de portes distintos. A existência de uma estrutura conceitual coerente não comprova melhoria causal; essa demonstração dependeria de estratégias de pesquisa adequadas e de dados coletados de forma transparente. A proposta permanece, portanto, aberta a revisão crítica e aperfeiçoamento.

14. RESULTADOS DA SÍNTESE E DISCUSSÃO

A síntese identifica convergência na percepção de que a IA pode ampliar recursos de análise, mas não elimina os problemas de definição de objetivos e atribuição de responsabilidades. Shrestha, Ben-Menahem e von Krogh (2019) oferecem uma estrutura para escolher combinações entre decisões humanas e automatizadas, enquanto Raisch e Krakowski (2021) mostram como automação e ampliação se condicionam mutuamente. A contribuição de Faraj, Pachidi e Sayegh (2018) acrescenta que os sistemas de aprendizagem reconfiguram relações de conhecimento e poder. Consideradas conjuntamente, essas perspectivas recomendam analisar simultaneamente tarefa, arquitetura decisória e contexto institucional. O ganho estratégico não pode ser deduzido apenas da velocidade de processamento nem da quantidade de informações disponíveis. Ele exige avaliação da pertinência dos problemas e da capacidade de responder por consequências.

A literatura experimental também impõe limites à universalização. Estudos sobre aversão e apreciação algorítmica demonstram comportamentos distintos diante de conselhos automatizados, sugerindo que confiança depende das condições de comparação, da expertise e da experiência com falhas (Dietvorst; Simmons; Massey, 2015; Logg; Minson; Moore, 2019). Pesquisas de campo de 2026 sobre tarefas de consultoria e inovação indicam possibilidades de ganho, mas também limites heterogêneos de desempenho conforme a natureza das atividades (Dell'Acqua et al., 2026a; 2026b). Esses resultados não devem ser somados como se avaliassem uma mesma população ou intervenção. A síntese narrativa reconhece a especificidade dos desenhos e utiliza as diferenças para formular hipóteses organizacionais, sem criar medida combinada de efeito que não foi calculada.

Outra convergência diz respeito ao caráter contínuo da governança. A ISO/IEC 42001, o NIST AI RMF e a Recomendação da OCDE organizam exigências ou princípios de documentação, controle, supervisão e responsabilização, ainda que possuam natureza jurídica e institucional distinta (ISO; IEC, 2023; Tabassi, 2023; OCDE, 2024). O reconhecimento desses pontos comuns sustenta a construção de rotinas de inventário, avaliação, monitoramento e revisão. Contudo, sua adoção formal não garante impacto positivo. Uma organização pode possuir políticas documentadas e, ao mesmo tempo, permitir que decisões automatizadas sejam incontestáveis na prática. A lacuna entre procedimento escrito e comportamento gerencial constitui questão central para pesquisas futuras e para auditorias organizacionais.

O contexto regulatório reforça a necessidade de distinguir obrigações vigentes e instrumentos em discussão. A LGPD já

estabelece direitos e deveres relativos a dados pessoais, enquanto o PL nº 2.338/2023 encontrava-se em tramitação no Legislativo brasileiro em outubro de 2026 (Brasil, 2018; 2023). O Regulamento Europeu de IA constitui referência normativa específica de outra jurisdição e pode ter efeitos conforme as condições de aplicação, não sendo automaticamente idêntico às exigências brasileiras (União Europeia, 2024). A comparação é útil para antecipar questões gerenciais e organizar processos compatíveis com distintas obrigações, sem apresentar harmonização completa como fato consumado. O ambiente normativo demanda acompanhamento regular e interpretação jurídica contextualizada.

A principal contribuição interpretativa reside em deslocar o foco da adoção tecnológica para a qualidade da decisão institucionalmente situada. A pergunta relevante deixa de ser apenas quanto trabalho pode ser automatizado e passa a incluir quais tarefas justificam delegação, que evidências sustentam as recomendações, quem arca com riscos e como resultados podem ser questionados. O MIGAE responde a essa agenda ao integrar propósito, dados, arquitetura, supervisão, risco e aprendizagem. Sua limitação central decorre da natureza propositiva e documental: não foram medidos custos de implantação, efeitos sobre desempenho ou índices de maturidade em organizações reais. Ainda assim, o modelo oferece categorias verificáveis para futuras comparações e ajuda a impedir que sofisticação computacional seja confundida com legitimidade administrativa.

15. CONSIDERAÇÕES FINAIS

A investigação demonstra, em termos conceituais e documentais, que a governança algorítmica constitui elemento decisivo para

compreender os impactos da inteligência artificial na administração contemporânea. A incorporação de sistemas de previsão, recomendação e geração de conteúdo amplia possibilidades de análise, mas transforma também a organização da autoridade e a atribuição de responsabilidade. A qualidade da decisão estratégica não decorre diretamente de maior volume de dados ou da adoção de modelos avançados. Ela exige que a instituição defina seus objetivos, conheça as limitações das ferramentas e preserve capacidade de deliberar sobre consequências. O problema central não consiste em escolher abstratamente entre decisões humanas e decisões algorítmicas, mas em estabelecer condições sob as quais cada contribuição pode ser utilizada de maneira verificável e legítima. A administração permanece responsável por suas escolhas, inclusive quando operações intermediárias são executadas por sistemas automatizados.

A literatura analisada sugere que a divisão de funções entre pessoas e máquinas deve acompanhar as características da tarefa. Processos repetitivos, claramente delimitados e com consequências reversíveis podem admitir maior automação, enquanto situações de alta incerteza e impacto exigem interpretação contextual, análise de alternativas e supervisão efetiva. O paradoxo entre automação e ampliação evidencia que ganhos de eficiência podem coexistir com perda de conhecimento ou dependência tecnológica. Por isso, a implantação precisa ser acompanhada de investimento em competências gerenciais e técnicas, assim como em práticas de revisão que não penalizem discordâncias fundamentadas. A presença de profissionais na etapa final é insuficiente quando eles não dispõem de informação e autoridade para contrariar recomendações. Liderança responsável exige construir as condições reais do julgamento, e não apenas declarar sua importância.

A governança dos dados foi identificada como condição anterior à confiabilidade dos modelos. Informações incompletas, registros produzidos sob incentivos inadequados ou variáveis sem pertinência podem conduzir a decisões problemáticas mesmo quando algoritmos apresentam bom desempenho estatístico. Proveniência, qualidade, atualização e documentação precisam integrar o processo estratégico desde a definição do problema. Também se tornam essenciais mecanismos que revelem alterações de comportamento, degradação de desempenho e efeitos produzidos pelas próprias recomendações. A interpretação de resultados deve respeitar diferenças entre previsão, explicação e causalidade, evitando apresentar como destino inevitável uma regularidade encontrada em dados históricos. A organização deve manter meios de reconstruir decisões e corrigir suas premissas quando o contexto se modifica.

Os riscos distributivos revelam que métricas de desempenho não substituem critérios de justiça e legitimidade. Decisões que afetam trabalhadores, consumidores ou comunidades exigem atenção a impactos diferenciados, direitos de revisão e limites do tratamento de dados. A identificação de vieses não se encerra no cálculo de indicadores, pois envolve escolhas sobre finalidades, oportunidades e consequências. A transparência deve oferecer informações úteis a gestores, auditores e pessoas afetadas, respeitando restrições legítimas de privacidade e segurança. Igualmente, a contestação precisa produzir possibilidades efetivas de correção. Sistemas formalmente explicáveis podem permanecer injustos se não existem canais de escuta ou capacidade institucional para alterar resultados. A governança, portanto, combina técnicas de avaliação com mecanismos procedimentais de responsabilização.

A análise regulatória confirma a importância de distinguir fontes com naturezas distintas. A legislação brasileira de proteção de dados estabelece direitos e obrigações efetivos; propostas legislativas sobre IA precisam ser acompanhadas sem serem tratadas prematuramente como leis; normas técnicas e recomendações internacionais podem orientar boas práticas, mas não substituem análise do caso concreto. Essa diferenciação é especialmente relevante em instituições que utilizam aplicações importadas ou contratam fornecedores que operam em múltiplos países. A administração deve construir inventários de sistemas, registrar responsabilidades, revisar contratos e prever procedimentos de resposta a incidentes. O objetivo da conformidade não é apenas evitar sanções, mas fortalecer a consistência das decisões e a capacidade de demonstrar por que determinada tecnologia foi considerada adequada aos interesses legítimos da organização.

O Modelo Integrado de Governança Algorítmica Estratégica apresentado neste artigo sintetiza essas exigências em seis dimensões articuladas: finalidade, dados, arquitetura decisória, supervisão, gestão de riscos e aprendizagem. Trata-se de proposta conceitual que ainda necessita validação em instituições reais. Sua adoção não deve ser confundida com certificação ou garantia de desempenho. Pesquisas futuras podem comparar organizações que implementam práticas distintas, examinar custos e benefícios ao longo do tempo e investigar consequências sobre qualidade decisória, confiança, autonomia e justiça. A avaliação deverá incluir resultados desfavoráveis e contextos nos quais o uso de IA não produza vantagens, evitando um viés de confirmação tecnológica. A construção de evidência robusta dependerá de protocolos transparentes, dados observados, análises de sensibilidade e explicitação das limitações de generalização.

A contribuição mais ampla deste estudo consiste em reconhecer que a transformação algorítmica da administração não elimina a centralidade da escolha humana; modifica os meios pelos quais essa escolha é informada, executada e justificada. Estratégias organizacionais continuam expressando prioridades, disputas de recursos e compromissos com pessoas e instituições. A IA pode aumentar capacidade de percepção e elaboração de cenários, mas sua utilidade deve ser demonstrada em tarefas específicas e confrontada com custos, riscos e alternativas. O avanço responsável requer organizações capazes de aprender com erros, preservar a possibilidade de contestação e interromper aplicações quando seus efeitos forem inadequados. Governança algorítmica, assim, representa menos uma promessa de decisões perfeitas e mais um compromisso permanente com decisões fundamentadas, revisáveis e compatíveis com a responsabilidade administrativa contemporânea.

REFERÊNCIAS BIBLIOGRÁFICAS

AUTIO, Chloe et al. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. Gaithersburg: National Institute of Standards and Technology, 2024. NIST AI 600-1. DOI: https://doi.org/10.6028/NIST.AI.600-1.

BRASIL. Congresso Nacional. Projeto de Lei nº 2.338, de 2023. Dispõe sobre o uso da inteligência artificial. Brasília, DF: Congresso Nacional, 2023. Disponível em: https://www.congressonacional.leg.br/materias/materias-bicamerais/-/ver/pl-2338-2023. Acesso em: 28 out. 2026.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 28 out. 2026.

DELL'ACQUA, Fabrizio et al. Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality. Organization Science, 2026a. DOI: https://doi.org/10.1287/orsc.2025.21838.

DELL'ACQUA, Fabrizio et al. The Cybernetic Teammate: A Field Experiment on Generative AI and Teamwork. Organization Science, v. 37, n. 4, p. 1217–1242, 2026b. DOI: https://doi.org/10.1287/orsc.2025.20702.

DIETVORST, Berkeley J.; SIMMONS, Joseph P.; MASSEY, Cade. Algorithm aversion: people erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, v. 144, n. 1, p. 114–126, 2015. DOI: https://doi.org/10.1037/xge0000033.

FARAJ, Samer; PACHIDI, Stella; SAYEGH, Karla. Working and organizing in the age of the learning algorithm. Information and Organization, v. 28, n. 1, p. 62–70, 2018. DOI: https://doi.org/10.1016/j.infoandorg.2018.02.005.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO); INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC). ISO/IEC 42001:2023: Information technology — Artificial intelligence — Management system. Geneva: ISO, 2023. Disponível em: https://www.iso.org/standard/42001. Acesso em: 28 out. 2026.

JARRAHI, Mohammad Hossein. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business Horizons, v. 61, n. 4, p. 577–586, 2018. DOI: https://doi.org/10.1016/j.bushor.2018.03.007.

KELLOGG, Katherine C.; VALENTINE, Melissa A.; CHRISTIN, Angèle. Algorithms at work: the new contested terrain of control. Academy of Management Annals, v. 14, n. 1, p. 366–410, 2020. DOI: https://doi.org/10.5465/annals.2018.0174.

LEBOVITZ, Sarah; LIFSHITZ-ASSAF, Hila; LEVINA, Natalia. To engage or not to engage with AI for critical judgments: how professionals deal with opacity when using AI for medical diagnosis. Organization Science, v. 33, n. 1, p. 126–148, 2022. DOI: https://doi.org/10.1287/orsc.2021.1549.

LOGG, Jennifer M.; MINSON, Julia A.; MOORE, Don A. Algorithm appreciation: people prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, v. 151, p. 90–103, 2019. DOI: https://doi.org/10.1016/j.obhdp.2018.12.005.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO (OCDE). Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449). Paris: OECD, 2024. Adotada em 2019, alterada em 2024. Disponível em: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. Acesso em: 28 out. 2026.

RAISCH, Sebastian; KRAKOWSKI, Sebastian. Artificial intelligence and management: the automation–augmentation paradox. Academy of Management Review, v. 46, n. 1, p. 192–210, 2021. DOI: https://doi.org/10.5465/amr.2018.0072.

SHRESTHA, Yash Raj; BEN-MENAHEM, Shiko M.; VON KROGH, Georg. Organizational decision-making structures in the age of artificial intelligence. California Management Review, v. 61, n. 4, p. 66–83, 2019. DOI: https://doi.org/10.1177/0008125619862257.

TABASSI, Elham. Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg: National Institute of Standards and Technology, 2023. NIST AI 100-1. DOI: https://doi.org/10.6028/NIST.AI.100-1.

UNIÃO EUROPEIA. Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que estabelece regras harmonizadas em matéria de inteligência artificial. Jornal Oficial da União Europeia, 2024. Disponível em: https://eur-lex.europa.eu/eli/reg/2024/1689/oj. Acesso em: 28 out. 2026.

---

[1] Doutor em Ciências – Engenharia Elétrica. Instituto Federal Goiano – IF Goiano. E-mail: acesse o artigo original para visualizar o e-mail.

[2] Doutorado em Estatística e Investigação Operacional. Universidade de Lisboa –ULisboa. E-mail: acesse o artigo original para visualizar o e-mail. ORCID: https://orcid.org/0000-0001-6236-6989.

[3] Pós-doutorado em Psicologia Organizacional; Doutor e Mestre em Administração. Instituto Rui Barbosa de Pós Graduação. E-mail: acesse o artigo original para visualizar o e-mail.

[4] Doutor em Administração de Empresas. Universidad TrêsFronteras (UNINTER). E-mail: acesse o artigo original para visualizar o e-mail.

[5] Mestre em Economia (UFC – CAEN); Especialista emNeuromarketing; Especialista em Marketing Digital, Global e de Relacionamento; Especialista em Planejamento Estratégico; Especialista em Comportamento Organizacional. Economia e Administração; Universidade Estadual do Ceará – UECE. E-mail: acesse o artigo original para visualizar o e-mail.

[6] Mestrando em Ciências e Meio Ambiente pela Universidade Federal do Pará – UFPA. Gestão Pública, História e Pedagogia. E-mail: acesse o artigo original para visualizar o e-mail.

[7] Doutorando e Mestrando em Ciências da Educação – Christian BusinessSchool. Pós-graduado em Psicologia Cognitivo Comportamental – FACULESTE; Pós-graduado em Análise do Comportamento Aplicada – UNIFAHE; Pós-graduado em Saúde Coletiva – Faculdade IGUAÇU. Graduado em Pedagogia – UNIP; Graduado em Psicologia – Estácio de Sá; Graduado em Educação Especial – UNICV. E-mail: acesse o artigo original para visualizar o e-mail. ORCID: https://orcid.org/0009-0005-1557-4159.

[8] Doutorando em Difusão do Conhecimento – IFBA; Mestre em Ciências Contábeis – UFBA; Mestre em Ciências Farmacêuticas – UNEB. Ciências Contábeis – UNEB; Farmácia – UFBA. E-mail: acesse o artigo original para visualizar o e-mail.

[9] Mestrado em Direito - Centro UniversitárioUninta. Pós em Administração Pública e Gestão Estratégica - Faculdade FACUMINAS. Pós em Direito Administrativo - FAMART. Direito - Universidade Estadual Vale do Acaraú – UVA. História - Universidade Estadual Vale do Acaraú – UVA. LATTES: https://lattes.cnpq.br/2194894682016709. ORCID:

https://orcid.org/0009-0000-9577-4014. E-mail: acesse o artigo original para visualizar o e-mail.

[10] Doutor e Mestre em Ciências e Meio Ambiente. Universidade Federal do Pará – UFPA; Matemático e Pedagogo. E-mail: acesse o artigo original para visualizar o e-mail.