

AUDITAR A INTELIGÊNCIA ARTIFICIAL É AUDITAR O PODER? UMA LEITURA CRÍTICA DE AUDITING AI

Auditing Artificial Intelligence as auditing power? A critical reading of auditing AI

Thiago Perez Bernardes de Moraes¹

RESUMO

A presente resenha crítica examina a obra *Auditing AI*, publicada pela The MIT Press em 2026, com o objetivo de discutir sua contribuição para o Direito Digital, em especial para os debates sobre auditoria algorítmica, accountability, assimetria informacional, prova de discriminação automatizada e governança de sistemas de inteligência artificial. A análise adota abordagem crítico-bibliográfica e sustenta que a principal virtude da obra consiste em retirar a auditoria de IA do confinamento técnico, recolocando-a no campo da evidência pública, da responsabilidade institucional e do controle de arquiteturas decisórias opacas. O limite mais sensível do livro, contudo, aparece quando a auditoria é tomada como se fosse mecanismo suficiente de responsabilização. A auditoria pode detectar padrões, organizar controvérsias e constranger narrativas corporativas, mas não produz, por si, sanção, reparação, suspensão de sistemas ou correção estrutural. Conclui-se que *Auditing AI* é obra oportuna e juridicamente relevante, desde que lida contra sua própria domesti-

ABSTRACT

This critical review examines Auditing AI, published by The MIT Press in 2026, in order to discuss its contribution to Digital Law, especially to debates on algorithmic auditing, accountability, informational asymmetry, evidence of automated discrimination and governance of artificial intelligence systems. The analysis adopts a critical-bibliographical approach and argues that the book's main merit lies in moving AI auditing beyond a narrowly technical field, placing it within public evidence, institutional responsibility and the control of opaque decision-making architectures. Its most sensitive limitation, however, emerges when auditing is treated as if it were a sufficient mechanism of accountability. Auditing can detect patterns, organize controversies and challenge corporate narratives, but it does not, by itself, produce sanctions, reparation, system suspension or structural correction. The review concludes that Auditing AI is timely and legally relevant, provided that it is read against its own possible technocratic domestication: auditing AI systems matters only when the evidence obtained is capable of shifting incentives, powers and duties.

¹ Doutor e pós-doutor em Psicologia Social pela Universidad Argentina John Fitzgerald Kennedy, Argentina. Pós-doutor em Direito Constitucional pela Università degli Studi di Messina, Itália. Professor titular do Centro Universitário Campos de Andrade, Uniandrade, e coordenador de projetos pedagógicos do EDUTEA.

cação tecnocrática: auditar sistemas de IA só importa quando a prova obtida consegue deslocar incentivos, poderes e deveres.

Palavras-chave: Direito Digital; inteligência artificial; auditoria algorítmica; accountability; governança tecnológica.

Keywords: *Digital Law; artificial intelligence; algorithmic auditing; accountability; technological governance.*

SUMÁRIO

1. INTRODUÇÃO; 2. IDENTIFICAÇÃO DA OBRA; 3. AUDITORIA COMO FORMA JURÍDICA DA EVIDÊNCIA; 4. O MÉRITO DA OBRA: OPACIDADE CONVERTIDA EM OBJETO PÚBLICO; 5. O LIMITE DA OBRA: DIAGNÓSTICO SEM CONSEQUÊNCIA; 6. ATUALIDADE REGULATÓRIA E IMPLICAÇÕES PARA O DIREITO DIGITAL; 7. CONSIDERAÇÕES FINAIS; REFERÊNCIAS.

1 INTRODUÇÃO

A inteligência artificial entrou no Direito pela porta menos confortável: a da decisão. Sistemas algorítmicos já filtram oportunidades, ordenam visibilidade, modulam crédito, influenciam policiamento, organizam publicidade, apoiam triagens administrativas e participam de ambientes judiciais ou parajudiciais. A dificuldade jurídica, nesse campo, não está só em saber se a máquina erra. Ela erra, e muitas vezes erra de modo seletivo. O problema mais profundo é saber como provar o erro quando a decisão nasce de uma cadeia sociotécnica pouco acessível, distribuída entre modelo, base de dados, fornecedor, contratante, interface, usuário e regime contratual.

Auditing AI, obra coletiva publicada pela The MIT Press em 2026, chega precisamente nesse ponto. Seu objeto imediato é a auditoria de inteligência artificial. Seu tema real, contudo, é mais amplo: como sociedades democráticas podem produzir evidência confiável diante de sistemas cuja operação é opaca, proprietária, instável e frequentemente protegida por discursos de neutralidade técnica. O livro tem tom didático, por integrar a coleção Essential Knowledge, mas seu alcance jurídico é maior do que sua superfície introdutória sugere. Ao definir a auditoria de IA

como avaliação rigorosa e sistemática das promessas, expectativas ou padrões aplicáveis a um sistema, a obra desloca a conversa do entusiasmo tecnológico para o terreno da prova, da responsabilidade e do conflito institucional.

Esta resenha sustenta que a contribuição do livro está menos em oferecer uma teoria fechada da auditoria e mais em expor a sua ambivalência. A auditoria é indispensável porque permite observar padrões de comportamento de sistemas que, em muitos casos, não podem ser inteiramente explicados por seus próprios desenvolvedores. Ao mesmo tempo, ela é insuficiente quando não se conecta a regimes jurídicos capazes de impor acesso a dados, publicidade de resultados, revisão independente, sanção, reparação e modificação material dos sistemas auditados. A pergunta que fica depois da leitura é simples e incômoda: quando se audita uma inteligência artificial, está-se auditando um sistema técnico ou uma forma contemporânea de poder?

2 IDENTIFICAÇÃO DA OBRA

A obra resenhada é *Auditing AI*, assinada por The Marquand House Collective e publicada pela The MIT Press, em Cambridge, Massachusetts, no ano de 2026. Trata-se de livro de 206 páginas, organizado em seis capítulos: Introduction; What Is AI Auditing?; The Steps of an AI Audit; Interpreting Audit Results; After the Audit; e A Healthy Audit Ecosystem. A estrutura confirma a natureza pedagógica da coleção, mas a escolha dos casos confere densidade política ao texto.

O livro apresenta exemplos que já compõem uma gramática pública da governança algorítmica: o sistema LYFT de alocação de rins, o caso SABRE no mercado de reservas aéreas, a segmentação discriminatória de anúncios pelo Facebook, a investigação jornalística sobre o COMPAS, o estudo Gender Shades sobre reconhecimento facial, os problemas do Google Image Search, o recorte automático de imagens no Twitter, a ferramenta de triagem de currículos da Amazon e a auditoria civil encomendada pela Airbnb. A seleção é feliz porque impede uma leitura ingênua

da auditoria como simples procedimento técnico. Em cada caso há uma disputa sobre mercado, corpo, raça, reputação, acesso, oportunidade ou autoridade pública.

O livro pode ser lido como manual introdutório, mas essa seria sua leitura mais pobre. A obra funciona melhor quando tratada como uma cartografia de problemas jurídico-institucionais. Sua matéria não é somente a inteligência artificial. É a produção de confiança em ambientes de baixa verificabilidade social. A inteligência artificial, nesse contexto, aparece como infraestrutura decisória capaz de redistribuir chances de vida sem que a pessoa afetada consiga reconstruir, com razoável precisão, a cadeia que levou ao resultado.

3 AUDITORIA COMO FORMA JURÍDICA DA EVIDÊNCIA

O ponto mais relevante da obra está na genealogia da auditoria. Os autores recuperam práticas históricas de teste antidiscriminatório em moradia e emprego, nas quais agentes ou documentos comparáveis eram apresentados a tomadores de decisão para verificar diferenças de tratamento. Essa linhagem é decisiva porque retira a auditoria algorítmica do imaginário estreitamente computacional. A auditoria de IA herda uma tradição de investigação social aplicada a estruturas que negam, dissimulam ou banalizam a discriminação.

Essa herança tem forte consequência jurídica. A discriminação algorítmica raramente se apresenta como intenção explícita. Ela se manifesta como regularidade estatística, como distribuição desigual de erro, como ranqueamento enviesado, como recomendação opaca, como triagem aparentemente neutra ou como interação entre base de dados histórica e objetivo de otimização. O Direito, acostumado a procurar autoria, intenção e nexos causais em moldes relativamente lineares, é forçado a lidar com uma causalidade probabilística, distribuída e mediada por artefatos técnicos.

É nesse ponto que a resenha dialoga com a discussão já desenvolvida por Freitas (2025) na própria Revista Rede de Direito Digital, Intelec-

tual & Sociedade. Ao tratar do enviesamento em sistemas de inteligência artificial, a autora mostra que o problema envolve fontes variadas de viés, qualidade da base de dados, desenho do modelo, estereótipos, decisões passadas, explicabilidade, auditabilidade, accountability e direitos fundamentais. A contribuição de Auditing AI conversa diretamente com essa chave de leitura: auditar não significa apenas perguntar se o sistema é tecnicamente eficiente, mas se seus efeitos são juridicamente toleráveis dentro de um ambiente atravessado por assimetrias de poder, informação e vulnerabilidade.

A auditoria, portanto, atua como forma jurídica da evidência. Ela não substitui a decisão judicial, a fiscalização administrativa ou a deliberação legislativa. Seu papel é outro: criar condições mínimas para que o conflito saia da intuição e se converta em objeto demonstrável. Quando uma plataforma afirma que seu sistema é neutro, seguro ou eficiente, a auditoria pergunta: neutro em relação a quem? Seguro para quem? Eficiente a que custo? A obra ganha força quando insiste que a boa auditoria nasce de uma disposição cética, voltada a testar promessas e não a confirmá-las.

4 O MÉRITO DA OBRA: OPACIDADE CONVERTIDA EM OBJETO PÚBLICO

Auditing AI tem um mérito raro em livros introdutórios sobre tecnologia: ele não confunde explicação com controle. A obra reconhece que muitos sistemas funcionam como caixas-pretas, inclusive para parte dos agentes que os desenvolvem, operam ou contratam. Também evita reduzir o problema ao acesso ao código-fonte. O código pode ser relevante, sem dúvida, mas não esgota o comportamento social do sistema. Uma arquitetura algorítmica pode apresentar documentação técnica coerente e produzir efeitos normativamente inaceitáveis. O inverso também é possível: um sistema pouco elegante, do ponto de vista computacional, pode ser mais controlável quando há logs, rastreabilidade, supervisão humana efetiva e canais de contestação.

O livro acerta ao enfatizar que auditorias podem trabalhar com entradas e saídas, inspeção de interfaces, testes por usuários, bases públicas, experimentos de correspondência, análise de resultados e, em alguns casos, acesso interno. Essa pluralidade impede que a auditoria se transforme em uma liturgia única. A depender do sistema, o melhor desenho pode ser estatístico, jurídico, jornalístico, comunitário, pericial ou híbrido. A consequência é importante para o Direito Digital: a qualidade de uma auditoria depende da pergunta que ela formula, do padrão que adota, do grupo que protege, dos dados a que tem acesso e da consequência que consegue acionar.

A obra também é útil por discutir a independência do auditor. A auditoria conduzida pelo próprio agente interessado pode ter mais dados, mas sofre com captura, delimitação estratégica de escopo e seleção reputacional de resultados. A auditoria externa pode ter mais liberdade, mas frequentemente opera com acesso imperfeito. O livro não resolve a tensão, e talvez não pudesse resolvê-la. Ainda assim, nomeia o dilema com clareza suficiente para deslocar o debate do slogan para a engenharia institucional. A pergunta deixa de ser “houve auditoria?” e passa a ser: quem pagou, quem escolheu as métricas, quem delimitou a amostra, quais dados foram negados, que achados não foram publicados, que padrão jurídico foi utilizado e que efeitos se seguiram ao relatório?

Essa mudança de pergunta é a contribuição mais fértil da obra para a Revista Rede de Direito Digital, Intelectual & Sociedade. A auditoria de IA não deve ser vista como apêndice da conformidade, mas como instrumento de disputa sobre a inteligibilidade de sistemas que decidem. Ao transformar opacidade em objeto público, a auditoria incomoda a neutralidade performada por empresas e governos. Ela obriga o sistema a se apresentar ao mundo por seus efeitos, e não apenas por suas promessas.

5 O LIMITE DA OBRA: DIAGNÓSTICO SEM CONSEQUÊNCIA

A principal fragilidade de Auditing AI aparece quando a auditoria é tratada como se seu valor institucional decorresse quase automaticamen-

te da qualidade do diagnóstico. O livro reconhece que auditorias identificam problemas sem necessariamente produzir responsabilização. A crítica jurídica precisa insistir nesse ponto. Um relatório bem desenhado pode se tornar prova. Pode também virar peça de governança cosmética. Entre uma coisa e outra existe uma diferença institucional enorme.

O risco de audit-washing atravessa toda a obra, embora pudesse receber formulação mais dura. Empresas, agências públicas e fornecedores tecnológicos aprendem rapidamente a converter linguagem de governança em amortecedor reputacional. Um relatório, um selo, uma matriz de risco, uma política de ética, um painel de transparência e um comitê consultivo podem funcionar como instrumentos reais de controle. Podem igualmente servir para estabilizar a confiança social sem alterar a estrutura decisória que produz o dano. A auditoria, nesse segundo cenário, não controla o poder. Ela administra a percepção pública do poder.

Esse limite é especialmente sensível no campo jurídico porque a auditoria opera na fronteira entre conhecimento e coerção. Conhecer o dano não equivale a repará-lo. Medir disparidade não equivale a impedir sua reprodução. Demonstrar enviesamento não equivale a redistribuir oportunidades. A obra ficaria mais forte se enfrentasse de modo mais direto a passagem do achado técnico para a consequência jurídica. Uma auditoria robusta precisa estar acoplada a deveres de documentação, preservação de registros, acesso regulatório, obrigação de resposta, contraditório técnico, canais de contestação, sanções proporcionais e mecanismos de reparação coletiva quando houver dano difuso.

Há ainda um problema de temporalidade. Sistemas de IA mudam. Modelos são atualizados, bases são substituídas, parâmetros são recalibrados, fornecedores são trocados, interfaces são redesenhadas, regras de moderação variam. Uma auditoria pontual pode capturar um momento e perder o fenômeno. Esse aspecto exige pensar auditoria como regime, e não como evento. O livro avança nessa direção ao defender ecossistemas saudáveis de auditoria, mas a resposta regulatória precisa ser mais incisiva: sistemas de alto impacto exigem monitoramento recorrente, padrões mínimos de rastreabilidade e obrigação de reavaliação quando houver mudança relevante de finalidade, desenho ou contexto de uso.

6 ATUALIDADE REGULATÓRIA E IMPLICAÇÕES PARA O DIREITO DIGITAL

A atualidade do livro fica evidente diante do Regulamento (UE) 2024/1689, conhecido como Artificial Intelligence Act, e do Projeto de Lei n. 2.338/2023, em tramitação no Brasil. A arquitetura regulatória contemporânea aproxima inteligência artificial, gestão de risco, documentação, transparência, supervisão humana, avaliação de impacto e prestação de contas. Nesse cenário, a auditoria deixa de ser técnica periférica e passa a integrar a infraestrutura prática da regulação.

O diálogo com Guadamuz (2025), também publicado na Revista Rede de Direito Digital, Intellectual & Sociedade, reforça esse ponto. Ao examinar a Lei de Inteligência Artificial da União Europeia, o autor enfatiza a centralidade da documentação, da transparência e das obrigações incidentes sobre provedores, inclusive em relação a modelos de uso geral, dados de treinamento e riscos associados à criação de conteúdo digital. Auditing AI oferece uma peça complementar a esse debate: a regulação pode exigir documentação, mas alguém precisa testar a suficiência, a consistência e os efeitos reais daquilo que foi documentado. Sem auditoria, a transparência corre o risco de virar depósito de informação incompreensível. Sem consequência jurídica, a auditoria corre o risco de virar protocolo de aparência.

No contexto brasileiro, a questão assume contornos próprios. Grande parte dos sistemas utilizados por empresas, plataformas e órgãos públicos é desenvolvida ou controlada por agentes externos ao espaço regulatório nacional. Isso coloca problemas de soberania informacional, jurisdição, contratação pública, dependência tecnológica e capacidade estatal. Importar modelos de auditoria sem estrutura institucional para exigir dados, preservar evidências e impor consequências produziria uma regulação decorativa. O Brasil não precisa somente de boas palavras sobre IA confiável. Precisa de arranjos capazes de transformar desconfiança social em procedimento de verificação e procedimento de verificação em dever jurídico.

A leitura de Auditing AI também ajuda a sofisticar a discussão sobre prova em litígios algorítmicos. Em muitos casos, a pessoa afetada por

decisão automatizada terá acesso apenas ao resultado final. A empresa ou o Estado deterá os registros, a documentação, os critérios de classificação, a base de treinamento e o conhecimento operacional. A auditoria pode funcionar como ponte entre suspeita individual e evidência coletiva. Esse ponto é decisivo para ações civis públicas, tutela de direitos fundamentais, proteção de consumidores, proteção de dados pessoais e controle de discriminação indireta. O Direito precisará lidar com padrões, e não somente com episódios.

7 CONSIDERAÇÕES FINAIS

Auditing AI é uma obra relevante porque apresenta a auditoria de inteligência artificial como prática de conhecimento situada entre tecnologia, Direito e política. Seu maior valor está em mostrar que sistemas algorítmicos devem ser julgados por seus efeitos observáveis, por suas promessas testáveis e pelos padrões jurídicos ou sociais que alegam cumprir. Essa escolha é correta. A retórica da inovação costuma pedir confiança antes da prova. A auditoria inverte a ordem: exige prova antes da confiança.

A crítica central é que a auditoria não pode ser celebrada como solução autônoma. Ela é uma etapa de inteligibilidade. Pode abrir caminho para controle, sanção e reparação, mas também pode ser absorvida como rotina de compliance sem potência transformadora. O livro percebe esse risco, embora ainda confie bastante na formação de um ecossistema saudável de auditoria. A experiência jurídica recomenda prudência. Ecossistemas saudáveis raramente nascem sem incentivos adequados, obrigações claras e custos reais para quem conserva sistemas danosos.

A pergunta final permanece aberta. Quem audita os auditores? Quem define o padrão de aceitabilidade? Quem responde quando o relatório identifica dano e nada muda? Quem garante acesso a dados quando o segredo empresarial bloqueia a prova? Quem protege pesquisadores, jornalistas e organizações civis que testam sistemas poderosos? Auditing AI não responde integralmente a essas questões. Sua importância está em torná-las inevitáveis.

Para o Direito Digital, a obra deve ser lida como convite à institucionalização crítica da auditoria algorítmica. Uma auditoria fraca confirma a autoridade do sistema. Uma auditoria forte desloca o ônus da justificação. Se sistemas de inteligência artificial já decidem, classificam, filtram e distribuem oportunidades, auditar esses sistemas significa disputar a forma jurídica da visibilidade. Nesse sentido, auditar a inteligência artificial é, no limite, auditar o poder.

REFERÊNCIAS

BRASIL. Senado Federal. **Projeto de Lei n. 2.338, de 2023**. Dispõe sobre o uso da inteligência artificial. Brasília, DF: Senado Federal, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 30 abr. 2026.

FREITAS, Cinthia Obladen de Almendra. Enviesamento em sistemas de inteligência artificial e seus reflexos no Direito. **Revista Rede de Direito Digital, Intelectual & Sociedade**, Curitiba, v. 5, n. 9, p. 173-198, 2025. DOI: 10.5380/rrddis.v5i9.100107.

GUADAMUZ, Andrés. Os direitos autorais na Lei de Inteligência Artificial da União Europeia. **Revista Rede de Direito Digital, Intelectual & Sociedade**, Curitiba, v. 5, n. 9, p. 19-38, 2025. DOI: 10.5380/rrddis.v5i9.100121.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. **Artificial Intelligence Risk Management Framework (AI RMF 1.0)**. Gaithersburg: NIST, 2023. DOI: 10.6028/NIST.AI.100-1.

THE MARQUAND HOUSE COLLECTIVE. **Auditing AI**. Cambridge, Massachusetts: The MIT Press, 2026.

UNIÃO EUROPEIA. **Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho**, de 13 de junho de 2024. Estabelece regras harmonizadas em matéria de inteligência artificial e altera atos legislativos da União. Jornal Oficial da União Europeia, L 2024/1689, 12 jul. 2024. Disponível em: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>. Acesso em: 30 abr. 2026.

Recebido em 30 de abril de 2026.

Aprovado em 09 junho de 2026.