

HUMAN-IN-THE-LOOP: TRABALHO, ÉTICA E RESILIÊNCIA NA AUDITORIA ESG AUTOMATIZADA

Heloisa Pimentel de Souza¹, Angela Leite Moreno², Mariane Moreira de Souza²

¹Universidade Federal de Alfenas, Alfenas, Brasil (heloisapins@gmail.com)

^{2,3}Universidade Federal de Alfenas, Alfenas, Brasil

Resumo: Este artigo discute o conceito *Human-in-the-Loop* como estratégia de resiliência profissional na auditoria ESG automatizada. Por revisão bibliográfica narrativa, analisa-se como a aplicação de IA facilita e automatiza a coleta e classificação de dados sobre sustentabilidade, mas carece de julgamento ético e interpretação contextual. Argumenta-se que a supervisão humana estruturada, aliada ao letramento de dados, reduz riscos de *greenwashing* e fortalece a responsabilização na auditoria sustentável.

Palavras-chave: Auditoria ESG; Governança algorítmica; Letramento de dados; Ética da inteligência artificial; *Ai-in-the-Loop*.

INTRODUÇÃO

A evolução digital tem alterado de forma acelerada a natureza do trabalho especializado, deslocando tarefas antes exclusivas de profissionais humanos para sistemas de inteligência artificial (IA) capazes de automatizar tarefas ou ampliar as capacidades humanas (Brynjolfsson, 2022). Susskind e Susskind (2018) argumentam que grande parte do trabalho de especialistas, como advogados, médicos, contadores e auditores, pode ser decomposta em tarefas que, ou já são desempenhadas por não especialistas, ou são repetitivas, manuais e podem ser facilmente automatizadas; os autores chamam de “Falácia da IA” (*AI Fallacy*) a crença dos profissionais de que suas tarefas, por serem complexas quando vistas na totalidade, seriam completamente imunes à automação, quando na verdade a tecnologia já está avançando sobre subtarefas específicas dentro dessas profissões.

Paralelamente, a agenda Ambiental, Social e de Governança (ESG) consolidou-se como um componente central na gestão corporativa e avaliação de risco por investidores, reguladores e sociedade civil (Chopra et al., 2024; Vaihekoski e Yahya, 2025). Esse crescimento multiplicou o volume de relatórios produzidos, indicadores utilizados e exigências para divulgação. A coexistência de diferentes frameworks regulatórios e voluntários (como GRI, SASB, TCFD, ISSB e CSRD) impõem volumosas exigências de divulgação cada vez mais amplas e heterogêneas (Alotaibi e Alwathnani, 2025; Telukdarie, Nyathi e Fabchi, 2026). Esse cenário é descrito como um paradoxo por Chopra et al. (2024): quanto mais a divulgação ESG se expande, maior o risco de que ela se torne um instrumento de *greenwashing*, prática que induz o público ao erro

por meio da representação falsa de uma pessoa, empresa ou produto como ambientalmente responsável (Calamai et al., 2025). Nesse contexto, ao invés de efetivamente endereçar os problemas ambientais, sociais e de governança que deveria reportar, a falta de padronização e transparência durante a propagação acelerada de métricas ligadas à sustentabilidade cria um ambiente perfeito para esse tipo de prática enganosa.

A auditoria ESG, nesse cenário, tornou-se uma atividade fortemente burocratizada e trabalhosa. Prodanova, Tarasova e Sotnikova (2023) descrevem a auditoria ESG como componente crítico da governança corporativa, responsável por verificar a aderência das empresas aos indicadores ambientais, sociais e de governança dispersos em múltiplas fontes e formatos, muitas vezes sem padronização entre jurisdições. Del Giudice e Rigamonti (2020) conseguem mostrar empiricamente que a ausência de auditoria externa está associada a maior incidência de *greenwashing*, e que relatórios de sustentabilidade auditados por seres humanos tendem a produzir métricas ESG mais confiáveis, benefício que, no entanto, tem como custo um processo manual, demorado e fortemente dependente de julgamento especializado, o que diminui a adesão das empresas a essa prática.

É justamente nesse contexto que a inteligência artificial passa a desempenhar um papel crescente na auditoria. Ferramentas baseadas em processamento de linguagem natural (NLP) permitem automatizar atividades tradicionalmente intensivas em trabalho humano, como a localização, extração, classificação e consolidação de informações dispersas em relatórios de sustentabilidade, documentos regulatórios e outras fontes não estruturadas



(Telukdarie, Nyathi e Fabchi, 2026; Calamai et al., 2025). Ao reduzir o tempo dedicado a tarefas operacionais e repetitivas, esses sistemas ampliam a capacidade de análise dos auditores e tornam processos de conformidade mais escaláveis e rápidos.

Entretanto, a mesma literatura que documenta esses ganhos também reconhece limites relevantes. Calamai et al. (2025), em revisão sistemática sobre detecção de *greenwashing* via NLP, apontam que a própria definição de *greenwashing* permanece “vaga e ambígua”, dependente de julgamento sobre diferenças sutis entre o discurso corporativo e o desempenho real da empresa, algo que sistemas automatizados de classificação binária não conseguem resolver sozinhos, e que, segundo os autores, exige uma combinação entre análise automatizada e supervisão humana. Telukdarie, Nyathi e Fabchi (2026) chegam a um resultado empírico particularmente revelador: seu *framework* de IA atinge acurácia de mais de 99% na classificação de indicadores ambientais e de governança, mas apenas 63,6% no pilar social o que demonstra claramente como a IA lida bem com dados estruturados e quantitativos, e mal com julgamento contextual e qualitativo, exatamente onde reside parte relevante do trabalho de auditoria ESG.

É nesse ponto que a IA carece de julgamento ético característico do ser humano, interpretação contextual e compreensão sociopolítica e onde os modelos de supervisão humana estruturada, como o *Human-in-the-Loop* (HITL), tornam-se relevantes não apenas como planejamento técnico, mas como estratégia de resiliência profissional para auditores (Zanzotto, 2019; Natarajan et al., 2025). Este artigo discute o papel do modelo HITL como estratégia de resiliência profissional na auditoria ESG automatizada, analisando os impactos da IA sobre o trabalho humano especializado e os desafios éticos associados à automação da análise de dados de sustentabilidade.

MATERIAL E MÉTODOS

Este estudo adota uma abordagem qualitativa e exploratória, fundamentada em uma revisão bibliográfica narrativa de caráter interdisciplinar. A escolha dessa estratégia metodológica justifica-se pela integração de diferentes campos do conhecimento como auditoria e contabilidade, inteligência artificial aplicada e governança algorítmica, permitindo uma análise teórico-crítica sobre o papel do HITL na auditoria ESG automatizada. Embora não se trate de uma revisão sistemática, a busca e a seleção das referências seguiram critérios explícitos, garantindo transparência e rastreabilidade ao processo.

A busca bibliográfica foi realizada nas bases Google Scholar, Scopus e IEEE Xplore, utilizando os

descritores “*human-in-the-loop*”, “*AI ethics*”, “*ESG auditing*”, “*algorithmic governance*”, “*greenwashing detection*”, “*future of work*”, “*AI literacy*” e “*sustainable AI*”, combinados por operadores booleanos (AND e OR). O recorte temporal compreendeu publicações entre 2018 e 2026, período marcado pela consolidação das aplicações de IA e NLP na auditoria ESG e pelo fortalecimento das discussões sobre supervisão humana e governança algorítmica.

Após a leitura dos títulos e resumos, foram selecionados 18 estudos alinhados aos objetivos da pesquisa, priorizando trabalhos que abordassem pelo menos dois dos três eixos temáticos do artigo: (i) automação do trabalho especializado; (ii) auditoria ESG automatizada; e (iii) supervisão humana e governança algorítmica. Os estudos foram analisados por meio de síntese narrativa e organizados nos quatro eixos de discussão apresentados na seção a seguir.

RESULTADOS E DISCUSSÃO

A Disrupção Tecnológica na Auditoria ESG

A última década foi marcada pelo grande aumento do volume de exigências regulatórias relacionadas às divulgações ESG em diferentes jurisdições. Conforme documentam Liu et al. (2025), países como China, Índia e Brasil, além da União Europeia, ampliaram significativamente seus requisitos de divulgação por meio de novos frameworks e normas de sustentabilidade. Como consequência, empresas que atuam em diferentes mercados passaram a atender simultaneamente a múltiplos padrões de reporte heterogêneos, tornando a produção e a auditoria dessas informações um processo cada vez mais complexo e multifacetado.

Entretanto, o aumento das exigências regulatórias não resultou, necessariamente, em maior transparência. Na verdade, Chopra et al. (2024) argumentam que a coexistência de diferentes frameworks favorece a falta de padronização que leva à fragmentação das divulgações ESG e amplia o risco de *greenwashing*, uma vez que permite que as empresas selecionem quais indicadores reportarão e de que forma serão apresentados. Em direção semelhante, Telukdarie, Nyathi e Fabchi (2026) observam que a ausência de métricas harmonizadas e metodologias verificáveis limita a comparabilidade entre empresas e reduz a confiabilidade das divulgações para fins regulatórios e de *benchmarking* internacional.

Nesse contexto, a inteligência artificial tem sido incorporada à auditoria ESG como alternativa para lidar com o crescente volume de informações. Ferramentas baseadas em processamento de linguagem natural (NLP) automatizam atividades



como localização, extração, classificação e consolidação de dados provenientes de relatórios de sustentabilidade e outros documentos não estruturados, reduzindo o tempo dedicado a tarefas operacionais e aumentando a escalabilidade dos processos da auditoria.

Alotaibi e Alwathnani (2025) demonstram esse potencial com o desenvolvimento de um artefato para auditoria automatizada de emissões de Escopo 2 (emissões indiretas de energia). Em seus experimentos, participantes humanos levaram entre duas e cinco horas para localizar, extrair e reconciliar as divulgações de um único relatório, enquanto o sistema automatizado executou o mesmo procedimento em uma fração desse tempo e ainda identificou inconsistências em relatórios de grandes empresas do setor de tecnologia. Esses resultados evidenciam que tarefas repetitivas, estruturadas e quantitativas apresentam elevado potencial de automação, por meio de tecnologias baseadas em aprendizado de máquina, permitindo que os auditores concentrem seus esforços em atividades de maior complexidade analítica.

Os benefícios da automação, entretanto, não são homogêneos entre os diferentes pilares do ESG. Telukdarie, Nyathi e Fabchi (2026) obtiveram acurácia de 99,1% na classificação de indicadores ambientais e de 99,3% para indicadores de governança, mas apenas 63,6% para o pilar social. Segundo os autores, essa diferença decorre da natureza predominantemente qualitativa das informações sociais, caracterizadas por terminologia ambígua, categorias sobrepostas e grande dependência de contexto. Esses resultados demonstram que o desempenho da IA está diretamente relacionado ao grau de estruturação dos dados analisados, sendo muito menor em tarefas que exigem interpretação contextual e julgamento profissional e ético.

Limitações semelhantes são observadas na literatura sobre detecção automatizada de *greenwashing*. Em revisão sistemática sobre aplicações de NLP nessa área, Calamai et al. (2025) identificam a ausência de bases de dados amplamente validadas contendo casos reais de *greenwashing*, fazendo com que grande parte dos estudos utilize indicadores indiretos, como discrepâncias entre discurso corporativo e desempenho ambiental ou excesso de linguagem positiva. Os autores concluem que a identificação confiável de práticas de *greenwashing* continua exigindo julgamento humano e defendem o desenvolvimento de processos auditáveis que combinem análise automatizada e supervisão humana.

Outras literaturas de auditoria reforçam essa conclusão. Del Giudice e Rigamonti (2020)

demonstram que empresas submetidas à auditoria externa apresentam menor incidência de *greenwashing* e maior confiabilidade em seus ratings ESG. De forma semelhante, Vaihekoski e Yahya (2025) verificam que auditorias de maior qualidade reduzem a assimetria de informação entre o desempenho real das organizações e suas divulgações de sustentabilidade, fortalecendo a confiança dos consumidores e investidores. Em conjunto, esses estudos indicam que, embora a IA represente um avanço importante para automatizar tarefas operacionais da auditoria ESG, ela ainda não substitui o julgamento humano em atividades que envolvem interpretação, avaliação ética e compreensão contextual.

Assim, a literatura converge para um ponto central: a automação da auditoria ESG já se mostra tecnicamente viável para atividades estruturadas, como extração de informações, classificação de indicadores e verificação de consistência dos dados reportados. Entretanto, sua eficácia diminui à medida que aumentam as demandas por interpretação contextual, avaliação qualitativa e julgamento ético. Essa lacuna desloca o debate da viabilidade técnica da automação para a necessidade de mecanismos de supervisão humana capazes de assegurar transparência, confiabilidade e responsabilização nas decisões apoiadas por inteligência artificial, discussão que fundamenta a seção seguinte.

IA e Automação do Trabalho Especializado

A distinção entre automação (*automation*) e aumento de capacidades (*augmentation*) é fundamental para compreender os impactos da inteligência artificial sobre o trabalho especializado na contemporaneidade. Brynjolfsson (2022) argumenta que a IA pode seguir dois caminhos distintos, ou ela substituirá tarefas anteriormente executadas por humanos, ou ampliará suas capacidades, aumentando produtividade e geração de valor. Segundo o autor, enquanto a automação comum tende a reduzir o valor econômico do trabalho humano e seu poder de barganha, tecnologias voltadas ao aumento de capacidades do profissional favorecem o surgimento de novos produtos, serviços e atividades, preservando o protagonismo dos seres humanos. Apesar disso, o autor observa que organizações frequentemente priorizam estratégias de automação por apresentarem menor custo de implementação e retorno econômico mais imediato.

Essa discussão é aprofundada por Susskind e Susskind (2018), que a aplicam diretamente às profissões especializadas. Para os autores, atividades desempenhadas por especialistas podem ser decompostas em tarefas menores, muitas das quais são passíveis de automação, ainda que o trabalho completo permaneça dependente de julgamento



humano. Essa percepção contraria o que denominam como “Falácia da IA”, isto é, a crença de que tarefas baseadas em experiência e conhecimento tácito seriam, por natureza, totalmente imunes à automação. Na prática, a decomposição das atividades revela componentes repetitivos e estruturados de forma que são elegíveis a serem executados por sistemas inteligentes sem eliminar a necessidade de um profissional, apenas realocando esforços para outras etapas do fluxo de trabalho da área.

Marguerit (2025) reforça essa distinção ao analisar dados do mercado de trabalho norte-americano entre 2015 e 2022, identificando efeitos distintos para tecnologias voltadas à automação e ao aumento de capacidades. Enquanto a exposição à automação está associada à redução dos salários e à diminuição da criação de empregos em ocupações de baixa qualificação, tecnologias orientadas ao aumento de capacidades favorecem o surgimento de novas atividades, profissões e ganhos salariais, especialmente em ocupações intensivas em conhecimento. Esses resultados sugerem que cargos altamente especializados, como a auditoria, tendem a ser mais beneficiados quando a tecnologia é utilizada como ferramenta de apoio ao trabalho humano, e não como mecanismo de substituição.

Assim, a forma como a IA é incorporada à auditoria ESG torna-se um fator decisivo. Bankins, Hu e Yuan (2024) identificam dois mecanismos complementares pelos quais a IA transforma o trabalho: um voltado à automação de tarefas repetitivas e outro baseado na complementaridade entre capacidades humanas e sistemas inteligentes. Os autores destacam ainda que a IA generativa cria novas demandas profissionais, como competências para interpretar, validar e comunicar resultados produzidos por modelos de IA, ampliando o escopo do conjunto de habilidades exigidas dos trabalhadores em vez de simplesmente reduzir sua participação no processo produtivo.

O mesmo padrão é observado em escala internacional. Segundo o relatório da Organização Internacional do Trabalho (OIT, 2025), mais de 30% do emprego global apresenta potencial de exposição à IA, seja por automação, aumento de capacidades ou combinação de ambos. Entretanto, o número de ocupações com potencial predominante de aumento é aproximadamente seis vezes superior ao de ocupações suscetíveis à automação completa, ou seja, no estágio atual de desenvolvimento da IA generativa, a lógica de complementaridade entre humanos e tecnologia ainda prevalece, com destaque em ocupações de maior qualificação.

Em conjunto, as evidências teóricas (Brynjolfsson, 2022; Susskind e Susskind, 2018) e empíricas (Marguerit, 2025; OIT, 2025; Bankins, Hu e Yuan, 2024) convergem para um mesmo entendimento de

que a principal questão não é se a inteligência artificial será incorporada no trabalho especializado, mas sim de que forma isso ocorrerá. No contexto da auditoria ESG, a adoção de sistemas voltados ao aumento de capacidades e produtividade tende a fortalecer a atuação dos auditores, enquanto estratégias baseadas na substituição integral do trabalho humano ampliam grandemente riscos técnicos, organizacionais e profissionais. Essa distinção fundamenta a discussão, na seção seguinte, sobre o HITL como mecanismo capaz de integrar eficiência tecnológica, supervisão humana e resiliência profissional.

Human-in-the-Loop como Modelo de Resiliência

O conceito de HITL representa mais do que um mecanismo genérico de supervisão humana em sistemas baseados em inteligência artificial. Conforme proposto por Zanzotto (2019), ele é um paradigma de desenvolvimento de IA que reconhece a importância do papel ativo dos seres humanos na produção, validação e utilização do conhecimento incorporado aos modelos computacionais. Para o autor, manter o humano “*no loop*” não constitui apenas uma exigência técnica para aumentar a qualidade dos sistemas, mas também uma questão de ética e justiça, ao reconhecer que o conhecimento utilizado para treinar e aperfeiçoar algoritmos foi previamente produzido por trabalhadores humanos.

Essa concepção também implica compreender que diferentes formas de interação entre humanos e IA produzem resultados distintos. Natarajan et al. (2025) distinguem sistemas *Human-in-the-Loop* (HIL), nos quais a IA executa a maior parte das tarefas e o humano atua predominantemente como revisor ou validador, de sistemas *AI-in-the-Loop* (A²L), caracterizados por uma colaboração contínua entre humanos e inteligência artificial ao longo de todo o processo decisório. Segundo os autores, avaliar essas arquiteturas apenas por métricas tradicionais de desempenho, como precisão ou recall, é insuficiente, pois tais indicadores não revelam o grau real de participação humana e nem a distribuição das responsabilidades e tarefas durante a tomada de decisão.

Essa distinção assume grande importância na auditoria ESG. Um sistema que apenas produz relatórios para aprovação formal do auditor preserva a responsabilidade jurídica do profissional apenas de maneira aparente. Em contrapartida, modelos concebidos para apoiar continuamente o processo decisório permitem que o auditor participe da interpretação das evidências, da validação dos resultados e da construção do parecer final, preservando seu julgamento profissional e humano mesmo diante da automação crescente no fluxo de trabalho cotidiano.



A necessidade dessa participação efetiva está diretamente relacionada ao conceito de *accountability*. Drori e Teeni (2024), ao investigarem o uso de IA generativa na revisão de artigos científicos, argumentam que a responsabilização humana somente é legítima quando o responsável realmente compreende e supervisiona os processos automatizados que fundamentam a sua decisão. Caso contrário, a assinatura humana transforma-se em um procedimento meramente formal, incapaz de garantir responsabilidade real sobre o resultado produzido pela IA. A interpretação crítica das evidências é resultado direto da experiência desses profissionais, sendo essa percepção o ponto mais importante da auditoria ESG.

Essa discussão é aprofundada por Liu e Xu (2025) por meio do conceito de *Meaningful Human Control* (MHC). Segundo os autores, o controle humano significativo depende de duas dimensões complementares, a capacidade da intervenção humana sobre o funcionamento da IA e a existência de mecanismos claros de *accountability*. Para que decisões automatizadas sejam socialmente legítimas, seu funcionamento deve permanecer rastreável aos agentes humanos responsáveis. Os autores alertam, entretanto, que a simples presença de um profissional supervisionando o sistema não elimina riscos como o *automation bias*, fenômeno no qual operadores passam a aceitar automaticamente recomendações produzidas pela IA, reduzindo seu próprio julgamento crítico. Assim, manter um auditor “*no loop*” somente produz benefícios se a sua participação permanece efetivamente ativa e crítica durante todo o processo decisório.

Sob a perspectiva da governança, Heilingner (2022) reforça essa interpretação ao argumentar que grande parte das discussões sobre ética da inteligência artificial permanece excessivamente abstrata e que o foco fica principalmente em temas generalistas em vez de mecanismos concretos de implementação. Para o autor, princípios éticos somente produzem efeitos quando traduzidos em estruturas institucionais capazes de distribuir responsabilidades, criar mecanismos de supervisão e garantir transparência durante as decisões automatizadas. Nesse contexto, o HITL deixa de ser apenas uma recomendação ética e passa a fazer parte de um modelo concreto de governança algorítmica.

Essa compreensão converge com a própria literatura de auditoria ESG, na qual Alotaibi e Alwathnani (2025) apresentam o sistema automatizado para auditoria de emissões de Escopo 2 e concluem que a principal contribuição da IA consiste em complementar, e não substituir, o julgamento profissional dos auditores. Os autores defendem explicitamente que a adoção dessas tecnologias deve fortalecer o papel do auditor na interpretação e

validação das informações produzidas pelos sistemas automatizados e não de substituí-lo.

Em conjunto, essas evidências permitem compreender o HITL como um modelo de resiliência profissional para a auditoria ESG. Mais do que preservar a presença formal do auditor durante o processo, o HITL busca garantir que a inteligência artificial amplie suas capacidades analíticas sem substituir sua autonomia decisória. Assim, a resiliência profissional não decorre da resistência à automação, mas da construção de ferramentas e modelos de colaboração que combinem eficiência tecnológica, julgamento humano, *accountability* e governança algorítmica.

Alfabetização de Dados e Novas Competências

Já que modelos como o HITL pressupõem um profissional capaz de exercer julgamento crítico sobre os resultados produzidos por uma ferramenta de IA, o desenvolvimento dessas competências na formação dos profissionais especialistas deixa de ser um aspecto operacional e passa a constituir necessidade para que o modelo seja aplicado corretamente. Liu e Xu (2025) denominam essa capacidade de *oversight literacy* (letramento de supervisão), entendida como um dos pilares do controle humano significativo (*Meaningful Human Control*), ao lado de mecanismos de transparência, explicabilidade e estruturas institucionais de governança. Sem esse conjunto mínimo de competências, a supervisão humana tende a tornar-se apenas formal, aumentando a vulnerabilidade ao viés de automação.

Essa perspectiva é aprofundada por Bankins, Hu e Yuan (2024), que utilizam o conceito de *AI literacy* (letramento em IA) para descrever as competências necessárias dentro da interação crítica com sistemas inteligentes. Os autores organizam a base letramento em três dimensões: os conhecimentos técnicos básicos sobre o funcionamento da IA, competências relacionadas ao trabalho, como por exemplo decidir o momento correto para delegar tarefas ao sistema e quando deve-se manter a decisão sob responsabilidade humana, e habilidades interpessoais e socioemocionais, que são cada vez mais valorizadas justamente por sua menor suscetibilidade à automação. Desse modo, trabalhadores com maior letramento em IA apresentam melhor capacidade de adaptação, maior confiança calibrada na tecnologia e utilização mais eficaz dos sistemas inteligentes.

Evidências do mercado de trabalho reforçam essa interpretação. O relatório da Organização Internacional do Trabalho (OIT, 2025) mostra que habilidades específicas relacionadas à IA raramente são demandadas de forma isolada, aparecendo, na maioria das ocupações, combinadas a diversas competências cognitivas, gerenciais e



socioemocionais. O estudo conclui que estratégias de qualificação voltadas exclusivamente ao domínio técnico da IA são insuficientes, sendo mais efetivas estratégias multifacetadas que agregam competências digitais, pensamento crítico e capacidade de resolução de problemas complexos.

Esses resultados dialogam diretamente com as evidências apresentadas anteriormente, nas quais percebemos que os benefícios da IA tendem a concentrar-se em ocupações de alta qualificação quando utilizada para ampliar capacidades humanas, e não para substituí-las. No contexto da auditoria ESG, isso significa que o potencial da IA depende da capacidade do auditor de interpretar, questionar e validar criticamente os resultados produzidos pelos sistemas automatizados. Sem esse letramento, o profissional deixa de atuar como agente de supervisão e passa apenas a confirmar decisões tomadas pela IA, reproduzindo exatamente o cenário de *automation bias* descrito por Liu e Xu (2025).

Dessa forma, a resiliência profissional do auditor ESG não depende apenas da adoção formal de modelos HITL, mas da preparação dos profissionais para exercer uma supervisão efetivamente qualificada. Isso desloca parte da responsabilidade para os processos de formação e capacitação profissional, que passam a demandar a incorporação do letramento de dados, tecnologia e de inteligência artificial como competências centrais da auditoria contemporânea. Essa necessidade é reconhecida pela própria literatura de auditoria ESG automatizada, na qual Alotaibi e Alwathnani (2025) defendem que ferramentas de IA devem complementar o julgamento profissional dos auditores, reforçando que a eficácia desses sistemas depende, em última instância, da capacidade humana de interpretar, validar e contextualizar suas recomendações.

CONCLUSÃO

A inteligência artificial já demonstra capacidade concreta de transformar a auditoria ESG ao automatizar tarefas de coleta, extração, classificação e consolidação de dados antes inteiramente manuais, produzindo ganhos de eficiência amplamente documentados na literatura (Alotaibi e Alwathnani, 2025; Telukdarie, Nyathi e Fabchi, 2026; Liu et al., 2025). Ao mesmo tempo, os estudos analisados evidenciam limites importantes dessa automação, especialmente em atividades que envolvem interpretação contextual, julgamento ético e avaliação qualitativa. A menor acurácia na análise de indicadores sociais, a dificuldade de identificar práticas de *greenwashing* e a ausência de mecanismos próprios de responsabilização demonstram que a IA, isoladamente, ainda não é capaz de substituir o trabalho especializado do auditor.

Nesse contexto, a auditoria ESG tende a evoluir para um modelo híbrido, no qual sistemas inteligentes assumem atividades estruturadas e repetitivas, enquanto profissionais concentram-se em funções relacionadas ao julgamento crítico, à interpretação contextual e à tomada de decisão. Essa perspectiva está alinhada à literatura sobre automação do trabalho especializado, que diferencia aplicações voltadas à substituição daquelas orientadas ao aumento das capacidades humanas (Brynjolfsson, 2022; Susskind e Susskind, 2018; Marguerit, 2025).

Os resultados desta revisão indicam que o HITL representa um modelo consistente para operacionalizar essa integração entre inteligência artificial e trabalho humano. Mais do que inserir um profissional como etapa final de validação, o HITL pressupõe participação ativa durante o processo decisório, mecanismos claros de *accountability* e preservação do julgamento humano, conforme discutido por Zanzotto (2019), Natarajan et al. (2025), Drori e Teeni (2024) e Liu e Xu (2025).

Entretanto, a efetividade desse modelo depende da qualificação dos próprios profissionais. A supervisão humana somente produz benefícios quando acompanhada de letramento em dados e em inteligência artificial suficiente para interpretar, questionar e validar criticamente os resultados produzidos pelos sistemas automatizados. Dessa forma, a resiliência profissional dos auditores não decorre da resistência à adoção da IA, mas da capacidade de utilizá-la como ferramenta de ampliação das competências humanas, preservando a responsabilidade e a confiabilidade do processo de auditoria.

Por fim, este trabalho contribui ao aproximar três campos frequentemente discutidos de forma isolada: a automação do trabalho especializado, a auditoria ESG e a governança da inteligência artificial, propondo o HITL como um elo conceitual entre eles. Como limitação, trata-se de um estudo baseado em revisão bibliográfica, não contemplando validação empírica em ambientes reais de auditoria. Pesquisas futuras podem investigar a implementação prática de modelos HITL em processos de auditoria ESG, avaliando seus impactos na qualidade das decisões, eficiência operacional e confiabilidade.

REFERÊNCIAS

- Alotaibi, E. M.; Alwathnani, A. M. Ai-Enabled ESG Compliance Audit for Stakeholders. Sustainability, 2025.
- Bankins, S.; Hu, X.; Yuan, Y. Artificial intelligence, workers, and future of work skills. Current Opinion in Psychology, v. 58, 101828, 2024.



- Brynjolfsson, E. The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence. *Daedalus*, v. 151, n. 2, p. 272-287, 2022.
- Calamai, T. et al. Detecting Greenwashing: A Natural Language Processing Literature Survey. *arXiv preprint arXiv:2502.07541*, 2025.
- Chopra, S. S. et al. Navigating the Challenges of Environmental, Social, and Governance (ESG) Reporting: The Path to Broader Sustainable Development. *Sustainability*, v. 16, 606, 2024.
- Del Giudice, A.; Rigamonti, S. Does Audit Improve the Quality of ESG Scores? Evidence from Corporate Misconduct. *Sustainability*, v. 12, 5670, 2020.
- Drori, I.; Te'eni, D. Human-in-the-Loop AI Reviewing: Feasibility, Opportunities, and Risks. *Journal of the Association for Information Systems*, 2024.
- Heilinger, J.-C. The Ethics of AI Ethics. A Constructive Critique. *Philosophy & Technology*, v. 35, 61, 2022.
- International Labour Organization (OIT). Artificial intelligence adoption and its impact on jobs. G20 Framework Working Group, 2025.
- Liu, C.; Xu, W. Human-controllable AI: Meaningful Human Control. Preprint (capítulo em preparação), *arXiv:2512.04334*, 2025.
- Liu, Y.; Song, J.; Zhou, B.; Liu, J. Artificial intelligence applications and corporate ESG performance. *International Review of Economics and Finance*, v. 104, 104559, 2025.
- Marguerit, D. Augmenting or Automating Labor? The Effect of AI Development on New Work, Employment, and Wages. *LISER*, *arXiv:2503.19159*, 2025.
- Natarajan, S. et al. Human-in-the-loop or AI-in-the-loop? Automate or Collaborate? Proceedings of the AAAI Conference on Artificial Intelligence (AAAI-25), 2025.
- Prodanova, N.; Tarasova, O.; Sotnikova, L. ESG audit and its fundamental concepts. *E3S Web of Conferences*, v. 402, 13025, 2023.
- Susskind, D.; Susskind, R. The Future of the Professions. *Proceedings of the American Philosophical Society*, v. 162, n. 2, jun. 2018.
- Telukdarie, A.; Nyathi, M. H. L.; Fabchi, R. J. Environmental, social, and governance sustainability: an AI-centric approach driving data standardization and automation. *Frontiers in Sustainability*, v. 7, 1793155, 2026.
- Vaihekoski, M.; Yahya, H. Environmental, Social, and Governance (ESG) and Firm Valuation: The Moderating Role of Audit Quality. *Journal of Risk and Financial Management*, v. 18, 148, 2025.
- Zanzotto, F. M. Viewpoint: Human-in-the-loop Artificial Intelligence. *Journal of Artificial Intelligence Research*, v. 64, p. 243-252, 2019.