

HUMAN-IN THE-LOOP (HITL): UMA ABORDAGEM QUE INTEGRA A PARTICIPAÇÃO HUMANA EM SISTEMAS DE INTELIGÊNCIA ARTIFICIAL E AUTOMAÇÃO¹

BATISTA, E.S.; Graduando; IFSP; São Paulo-SP; lisa.senra@aluno.ifsp.edu.br;
COSTA, I. C.; Graduando; IFSP; São Paulo-SP; ismael.c@aluno.ifsp.edu.br;
SANTOS, S. B.; Graduando; IFSP; São Paulo-SP; samuel.barauna@aluno.ifsp.edu.br;
TEIXEIRA, V. A.; Graduando; IFSP; São Paulo-SP; vanderley.alves@aluno.ifsp.edu.br;
FERREIRA, S. S.; Mestre; IFSP; São Paulo-SP; sandra.silveira@ifsp.edu.br;
NASCIMENTO, R. J.; Mestre; IFSP; São Paulo-SP; ridnal@ifsp.edu.br.

RESUMO

O modelo *Human-in-the-Loop* (HITL) propõe a inserção do julgamento humano em momentos específicos do funcionamento de sistemas automatizados com inteligência artificial (IA), oferecendo uma camada adicional de controle e segurança. Considerando o aumento da preocupação com a autonomia total dos sistemas inteligentes — ainda limitados na compreensão de contextos sociais e éticos — o HITL surge como alternativa para tornar essas tecnologias mais seguras e confiáveis, sobretudo em decisões sensíveis. Esta pesquisa buscou compreender de que maneira o HITL pode contribuir para uma supervisão mais eficiente e segura de sistemas automatizados. O estudo também buscou identificar as principais limitações, benefícios e aplicações práticas dessa abordagem, na perspectiva de aspectos éticos e de segurança. Trata-se de um estudo exploratório, com abordagem qualitativa. Para a realização de uma busca sistemática da literatura e a seleção de artigos científicos voltados ao aprofundamento da compreensão sobre o papel da interação humana em processos decisórios mediados por tecnologias inteligentes, utilizaram-se as plataformas Google, Google Scholar e o Portal CAFe. Os resultados revelam que o HITL já é aplicado em diversas áreas como moderação em redes sociais, rotulagem de imagens para carros autônomos e apoio a diagnósticos médicos. Em comum esses cenários demandam decisões contextualizadas e, muitas vezes subjetivas. A despeito de suas vantagens, como o aumento da confiabilidade e a presença do controle humano em decisões sensíveis, o estudo identificou que o modelo apresenta limitações quanto à escalabilidade, à privacidade de dados e aos custos elevados de capacitação contínua dos operadores. Concluiu-se que o HITL proporciona um equilíbrio essencial entre a autonomia dos sistemas inteligentes e a intervenção humana.

Palavras-chave: HITL; inteligência artificial; tomada de decisão; automação; ética.

INTRODUÇÃO

À medida que a inteligência artificial avança rapidamente, surgem novas preocupações sobre a possibilidade de as máquinas tomarem decisões sem supervisão humana. Embora os algoritmos consigam executar tarefas complexas com eficiência, eles ainda não compreendem adequadamente contextos sociais, culturais e, especialmente, dilemas éticos. Nesse cenário, surge o conceito de Human-in-the-Loop (HITL), que visa integrar o julgamento humano em momentos-chave da tomada de decisão. Segundo Mosqueira-Rey e Hernández (2022), o HITL “atua como um componente de controle e

¹ Pesquisa realizada na disciplina Processos Decisórios – (SPOPRDE) do curso de Tecnologia em Gestão da Produção Industrial

supervisão, permitindo que humanos corrijam erros ou façam ajustes quando a IA não é capaz de lidar com a complexidade de uma tarefa”. Essa interação cria sistemas mais seguros, transparentes e ajustados à realidade social e cultural em que operam.

OBJETIVOS

Esta pesquisa tem como objetivo principal compreender de que maneira o modelo Human-in-the-Loop (HITL) pode contribuir para uma supervisão mais eficaz e segura de sistemas automatizados. A pesquisa também busca identificar os principais benefícios, limitações e aplicações práticas dessa abordagem, com ênfase em contextos que envolvem aspectos éticos e de segurança.

METODOLOGIA

De acordo com a classificação metodológica proposta por Lakatos e Marconi (2003) e Gil (2008), esta pesquisa configura-se como um estudo exploratório com abordagem qualitativa, fundamentado em revisão bibliográfica, utilizando como base materiais científicos, artigos e publicações técnicas que tratam da aplicação do HITL em contextos variados. Essa metodologia permitiu a construção de um cenário abrangente e atualizado sobre o tema, contribuindo para uma compreensão de suas aplicações teóricas e práticas, alinhada aos ensinamentos de Amershi *et al.* (2019), segundo os quais “a interação eficaz entre humanos e IA requer diretrizes claras para maximizar a utilidade e minimizar falhas operacionais”. Para realizar a busca sistemática de literatura e a coleta de artigos de pesquisa relativos ao tema estudado, utilizou-se as plataformas Google, Google Scholar e o Portal café; na revisão textual também foi utilizado a ferramenta ChatGPT, com o prompt “Revise este texto, verificando erros factuais, precisão gramatical, ortografia, pontuação e erros tipográficos.” para ajustes textuais.

FUNDAMENTAÇÃO TEÓRICA

O termo Humano-In-The-Loop (HITL) refere-se a uma abordagem no desenvolvimento e operação de sistemas de inteligência artificial (IA) e machine Learning (ML) (aprendizado de máquina) que envolve a participação de seres humanos durante diversas fases do ciclo de vida desses sistemas.

Diferentemente de sistemas totalmente automatizados, no HITL o ser humano atua como um colaborador essencial para garantir que os modelos aprendam corretamente, tomem decisões mais seguras e se ajustem a contextos complexos que as máquinas sozinhas ainda não compreendem plenamente (Mosqueira-Rey; Hernández, 2022).

Esse processo é caracterizado como se o ser humano e máquina trabalhassem em conjunto, então a máquina realiza uma tarefa inicial, com o processamento de dados ou a geração de uma previsão, após isso o humano revisa os resultados, fornece correções e feedbacks para melhoria, e com isso a máquina então ajusta seus parâmetros para melhorar seu desempenho e realizar as correções apontadas. Esse loop contínuo permite que o sistema evolua de forma mais precisa e confiável ao longo do tempo.

Nesse sentido, o HITL não é apenas uma técnica de IA, mas uma estrutura que ajuda a aprimorar os processos decisórios organizacionais ao fundir o discernimento humano onde a automação sozinha falharia, garantindo resultados mais confiáveis e alinhados aos valores e objetivos de cada organização.

• Onde o HITL é aplicado?

A abordagem Humana-In-The-Loop (HITL) é amplamente utilizada em contextos nos quais a precisão, a ética ou a interpretação contextual são essenciais. Nesses casos, a

supervisão e a intervenção humana garantem que os sistemas de inteligência artificial operem de maneira segura, eficaz e alinhada aos valores humanos, como exemplo: Na Revisão de Conteúdo

Em Plataformas como YouTube, Facebook e Instagram utilizam algoritmos para detectar conteúdos impróprios ou que violem diretrizes da comunidade. No entanto, esses sistemas frequentemente sinalizam casos ambíguos ou controversos que requerem revisão humana. Moderadores especializados avaliam esses conteúdos para decidir se devem ser removidos ou mantidos, garantindo que decisões não sejam injustas ou enviesadas.

• Vantagens e Desvantagens

A integração do ser humano nos processos de inteligência artificial oferece vantagens significativas, especialmente em áreas onde a ética, o julgamento contextual e a adaptação contínua são cruciais. No entanto, também apresenta desafios e limitações.

Vantagens

Melhoria na Precisão dos Modelos: A intervenção humana permite correções que aperfeiçoam os resultados, melhorando o desempenho dos algoritmos de aprendizado de máquina (GOOGLE CLOUD, s.d.). Exemplo Prático: Em diagnósticos médicos por IA, radiologistas humanos revisam imagens para corrigir erros da máquina, refinando o modelo para maior precisão na detecção de doenças.

Maior Confiabilidade e Transparência: O envolvimento humano aumenta a transparência nas decisões da IA, pois os humanos podem entender como as decisões são tomadas. Isso constrói confiança entre os usuários e o público. (MOSQUEIRA-REY; HERNÁNDEZ, 2022). Exemplo Prático: Em um sistema de análise de crédito, um agente humano revisa decisões de risco da IA explicando os motivos aos clientes e intervindo se necessário, aumentando a confiabilidade.

Adaptação a Situações Complexas ou Inéditas: O julgamento humano é capaz de lidar com exceções, ambiguidades e contextos imprevisíveis que desafiam a automação pura. Em casos ambíguos ou controversos que os algoritmos sozinhos não conseguem resolver, o HITL permite que humanos forneçam interpretação contextual e discernimento ético. (KLIPPA, s.d.). Exemplo Prático: Na moderação de conteúdo em redes sociais, moderadores humanos avaliam posts que a IA classifica como "incertos" (sarcasmo, ironia), usando seu discernimento para decidir a ação correta.

Redução de Vieses e Aumento da Ética: Modelos de IA podem gerar vieses presentes nos dados de treinamento. A supervisão humana permite identificar, monitorar e mitigar esses vieses, promovendo a justiça e a ética nos sistemas de IA. Isso é crucial para garantir que o sistema não tome decisões tendenciosas ou injustas

Desvantagens

Custo Operacional Elevado: A necessidade de intervenção humana pode ser financeiramente dispendiosa e logisticamente complexa de escalar. Contratar, treinar e manter equipes aumenta os custos operacionais e limita a velocidade de processamento em larga escala. Exemplo Prático: Uma empresa que precisa processar milhões de documentos diariamente para categorização. Usar HITL para revisar cada documento significaria contratar e gerenciar centenas de pessoas, o que eleva muito os custos e torna o processo lento demais para o volume.

Fadiga e Erros Humanos: A revisão manual repetitiva de algumas tarefas HITL pode levar à fadiga e à exaustão mental dos operadores humanos, resultando em erros e

inconsistências na rotulagem ou validação dos dados. Exemplo Prático: Trabalhadores revisando milhares de fragmentos de áudio para transcrever e categorizar ruídos para um assistente de voz. Após horas de trabalho repetitivo, a atenção pode diminuir, levando a transcrições imprecisas ou à perda de detalhes importantes, comprometendo a qualidade do dado de treinamento

Latência nas Respostas: Inclusão de humanos no fluxo de trabalho inevitavelmente diminui a velocidade com que as tarefas podem ser concluídas, introduzindo latência nas respostas, especialmente em comparação com sistemas totalmente automatizados que operam em tempo real. Exemplo Prático: Em um sistema de detecção de fraudes em transações financeiras online, a IA precisa de decisões em milissegundos. Se cada transação suspeita exigir a revisão de um humano antes de ser bloqueada, o concretizam antes da intervenção humana.

LIMITAÇÕES TÉCNICAS, SOCIAIS E ÉTICAS:

Escalabilidade: À medida que os sistemas se tornam mais complexos e os volumes de dados aumentam, a capacidade de envolver humanos de forma eficiente se torna limitada (KLIPPA, s.d.).

Privacidade e Segurança de Dados: A participação humana na análise de dados sensíveis exige cuidados rigorosos com a proteção de informações pessoais. Garantir a conformidade com regulamentações de privacidade (como a LGPD ou GDPR) e prevenir vazamentos ou acessos indevidos é um desafio contínuo que demanda infraestrutura.

Treinamento e Capacitação: Os humanos inseridos no loop precisam ser constantemente treinados e atualizados, não apenas sobre as ferramentas e interfaces, mas também sobre os objetivos específicos e as limitações do sistema de IA. Isso demanda um investimento contínuo em educação e desenvolvimento para garantir que os operadores humanos mantenham a qualidade.

Responsabilidade e Accountability: Nas decisões em saúde e em justiça, a definição clara de responsabilidades entre as ações da máquina e a intervenção humana ainda é um desafio em discussão. Determinar quem é "responsável" por um erro ou um resultado indesejado quando um humano está no loop de decisão pode ser complexo e exige pontos éticos bem estabelecidos e baseados em leis.

Viés Humano: Assim como os algoritmos podem herdar vieses dos dados de treinamento, os próprios humanos, mesmo com as melhores intenções, podem introduzir julgamentos subjetivos e enviesados em suas opiniões ou anotações. Esse viés humano pode, por sua vez, afetar a neutralidade e a imparcialidade do sistema de IA contradizendo um dos benefícios potenciais do HITL (GILLESPIE, 2018).

RESULTADOS E DISCUSSÃO

O HITL vem sendo aplicado em diversos contextos do cotidiano. Um exemplo prático é a moderação de conteúdo em redes sociais como YouTube, Instagram e TikTok. Nessas redes, mesmo com o uso de algoritmos, a intervenção humana ainda é essencial para julgar conteúdos ambíguos ou sensíveis (GILLESPIE, 2018). outro uso notável ocorre na indústria automotiva, em especial na rotulagem de imagens para veículos autônomos, onde humanos auxiliam na interpretação correta de cenários urbanos complexos. Já no setor da saúde, o HITL é adotado em diagnósticos médicos assistidos por IA, nos quais profissionais avaliam os resultados sugeridos pelas máquinas antes de qualquer decisão clínica.

Por outro lado, depender da ação humana traz desafios, como dificuldades para escalar o sistema, a exigência de capacitação constante e o aumento nos custos. Além disso, há o risco de viés humano influenciar os resultados. A segurança da informação e a

privacidade dos dados também exigem atenção especial, principalmente em ambientes regulados, como saúde e finanças. A seguir, na figura Vantagens e Desvantagens do HITL, apresenta-se uma síntese das principais vantagens e desvantagens do uso do modelo Human-in-the-Loop, com base nos estudos analisados:

Figura1: Vantagens e Desvantagens do HITL

Vantagens do HITL	Desvantagens do HITL
Aumento da precisão em tarefas críticas	Maior custo operacional
Supervisão ética e social em decisões sensíveis	Necessidade de treinamento constante
Possibilidade de correção de erros da IA	Escalabilidade limitada em ambientes com grande volume de dados
Adaptação a contextos ambíguos e não previstos	Risco de introduzir viés humano nos processos
Reforçar e corrigir o aprendizado das máquinas	Depende da disponibilidade e atenção dos operadores

Fonte: Mosqueira-Rey e Hernández (2022); Amershi et al. (2019); Google Cloud (s.d.).

CONSIDERAÇÕES FINAIS/CONCLUSÃO

A adoção do modelo Human-in-the-Loop representa um avanço significativo para a confiabilidade dos sistemas baseados em inteligência artificial. Inserir o julgamento humano nos pontos-chave do processo contribui para decisões mais éticas, mais flexíveis e mais alinhadas com valores humanos. Embora ainda existam limitações operacionais e desafios estruturais, sua aplicação tem se mostrado essencial em cenários onde erros não são toleráveis. Dessa forma, o HITL não apenas melhora a performance dos sistemas automatizados, mas também reforça a importância do fator humano na era digital.

REFERÊNCIAS

AMERSHI, Saleema et al. **Guidelines for Human-AI Interaction**. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019. Disponível em: <https://dl.acm.org/doi/10.1145/3290605.3300233>. Acesso em: 06 ago. 2025.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GILLESPIE, Tarleton. **Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media**. New Haven: Yale University Press, 2018.

GOOGLE CLOUD. **What is Human-in-the-Loop (HITL) in AI & ML?** [S. l.]: Google Cloud, [s.d.]. Disponível em: <https://cloud.google.com/discover/human-in-the-loop?hl=pt-BR>. Acesso em: 06 ago. 2025.

KLIPPA. **Guia Human-in-the-Loop**. Klippa, s.d. Disponível em: <https://www.klippa.com/pt/blog/informacao/guia-human-in-the-loop/>. Acesso em: 06 ago. 2025.

LAKATOS, E. M.; MARCONI, M. A. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003]

MOSQUEIRA-REY, M.; HERNÁNDEZ, D. **Human-in-the-loop machine learning: a state of the art**. Artificial Intelligence Review, 2022. Disponível em: <https://link.springer.com/article/10.1007/s10462-022-10246-w>. Acesso em: 06 ago. 2025.