



HUMAN-IN-THE-LOOP (HITL): Uma abordagem que integra a participação humana em sistemas de inteligência artificial e automação

HUMAN-IN-THE-LOOP (HITL): An approach that integrates human participation into artificial intelligence and automation systems

Elisa de Souza Senra Lopes Batista

Graduanda em Gestão da Produção Industrial
IFSP / Campus São Paulo
<https://orcid.org/0009-0001-3168-7007>
elisa.senra@aluno.ifsp.edu.br

Ismael Correia Costa

Graduando em Gestão da Produção Industrial
IFSP / Campus São Paulo
<https://orcid.org/0009-0007-1068-5081>
ismael.c@aluno.ifsp.edu.br

Ridnal João do Nascimento

Mestre em Tecnologia Nuclear
Docente do IFSP / Campus São Paulo
<https://orcid.org/0000-0002-8649-0164>
ridnal@ifsp.edu.br

Samuel Barauna Santos

Graduando em Gestão da Produção Industrial
IFSP / Campus São Paulo
<https://orcid.org/0009-0008-3579-4441>
samuel.barauna@aluno.ifsp.edu.br

Sandra Silveira Ferreira

Mestra em Engenharia Mecânica
Docente do IFSP / Campus São Paulo
<https://orcid.org/0009-0002-6259-1590>
sandra.silveira@ifsp.edu.br

Vanderley Alves

Graduando em Gestão da Produção Industrial
IFSP / Campus São Paulo
<https://orcid.org/0009-0002-2399-6915>
vanderley.alves@aluno.ifsp.edu.br

EDIÇÃO ESPECIAL - X ENCONTRO DE INICIAÇÃO CIENTÍFICA E PÓS-GRADUAÇÃO DO IFSP - CAMPUS SÃO PAULO (EICPOG)

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 de controle e segurança. Diante da preocupação com a autonomia total desses sistemas — ainda limitados na compreensão de contextos sociais e éticos — o HITL surge como alternativa para torná-los seguros e confiáveis em decisões sensíveis. Esta pesquisa buscou compreender de que maneira o HITL pode contribuir para uma supervisão

também eficiente e segura de sistemas automatizados, além de identificar limitações, benefícios e aplicações práticas dessa abordagem sob a perspectiva ética e de segurança. Trata-se de um estudo exploratório, com abordagem qualitativa. Realizou-se uma busca sistemática da literatura, selecionando artigos nas plataformas Google, Google Scholar e Portal CAFE, com foco na interação humana em processos decisórios mediados por tecnologias inteligentes. Os resultados revelam que o HITL já é aplicado em áreas como moderação em redes sociais, rotulagem de imagens para carros autônomos e apoio a diagnósticos médicos, cenários que exigem decisões subjetivas. Identificaram-se vantagens, como maior confiabilidade, mas também limitações ligadas à escalabilidade, privacidade de dados e custos de capacitação contínua. Conclui-se que o HITL oferece um equilíbrio entre autonomia tecnológica e intervenção humana.

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

Abstract

The Human-in-the-Loop (HITL) model proposes the insertion of human judgment at specific moments in the operation of automated systems with artificial intelligence (AI), offering a layer of control and security. Given the concern about the total autonomy of these systems—still limited in their understanding of social and ethical contexts—HITL emerges as an alternative to make them safe and reliable in sensitive decisions. This research sought to understand how HITL can contribute to an equally efficient and safe supervision of automated systems, as well as to identify limitations, benefits, and practical applications of this approach from an ethical and security perspective. This is an exploratory study with a qualitative approach. A systematic literature search was conducted, selecting articles from the Google, Google Scholar, and CAFE Portal platforms, focusing on human interaction in decision-making processes mediated by intelligent technologies. The results reveal that HITL is already applied in areas such as social media moderation, image labeling for autonomous cars, and support for medical diagnoses, scenarios that require subjective decisions. Advantages were identified, such as greater reliability, but also limitations related to scalability, data privacy, and ongoing training costs. It is concluded that HITL offers a balance between technological autonomy and human intervention.

Keywords: HITL; artificial intelligence; decision making; automation; ethics.

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 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.

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. Em sistemas integrados de gerenciamento no agendamento de serviços, como microempresas de trabalho artístico, o sistema coleta os dados do cliente por aplicativos de mensagens, o humano revisa e faz o orçamento e a IA atua na geração de insights de mercado para o desenvolvimento de estratégias promocionais personalizadas e na tomada de decisões (Lima, Campanharo & Freitas, 2025).

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 esse conteúdo 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.). Como 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). Assim como, 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.). Como 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 concretizem 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 Lei Geral de Proteção de Dados Pessoais LGPD no Brasil) 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).

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.

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

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

Considerações Finais

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-chaves 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, S., Weld, D., Vorvoreanu, M., Fournay, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). *Guidelines for human-AI interaction*. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). <https://doi.org/10.1145/3290605.3300233>
- Gil, A. C. (2008). *Métodos e técnicas de pesquisa social* (6ª ed.). Atlas.
- Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- Google Cloud. (s.d.). *What is Human-in-the-Loop (HITL) in AI & ML?* <https://cloud.google.com/discover/human-in-the-loop?hl=pt-BR>
- Klippa. (s.d.). *Guia Human-in-the-Loop*. <https://www.klippa.com/pt/blog/informacao/guia-human-in-the-loop/>
- Lakatos, E. M., & Marconi, M. A. (2003). *Fundamentos de metodologia científica* (5ª ed.). Atlas.
- Lima, C., Campanharo, B., & Freitas, R. (2025). Sistema integrado de agendamentos (SIA) e uso de inteligência artificial (IA) voltada para a geração de insights de mercado. *Revista Para Graduandos – REGRASP*, 10(3), 1–17. <https://doi.org/10.47734/regrasp.v10.03.p01-17>
- Mosqueira-Rey, M., & Hernández, D. (2022). Human-in-the-loop machine learning: A state of the art. *Artificial Intelligence Review*. <https://link.springer.com/article/10.1007/s10462-022-10246-w>