UMA VISÃO ABRANGENTE DA COMPUTAÇÃO – VOLUME 5

Organizadores: Mirian Nunes de Carvalho Nunes, Thiago Santana de Oliveira, Ivone Ascar Sauáia Guimarães, Ewerton Ferreira Bastos, Tayssara Elizavieta Martins Varão, Wagner Elvio de Loiola Costa e Antonio Luís de Souto Filho

Pascal
Editora

MIRIAN NUNES DE CARVALHO NUNES
THIAGO SANTANA DE OLIVEIRA
IVONE ASCAR SAUÁIA GUIMARÃES
EWERTON FERREIRA BASTOS
TAYSSARA ELIZAVIETA MARTINS VARÃO
WAGNER ELVIO DE LOIOLA COSTA
ANTONIO LUÍS DE SOUTO FILHO
(Organizadoras)

UMA VISÃO ABRANGENTE DA COMPUTAÇÃO

VOLUME 5

EDITORA PASCAL
2026

Editor Chefe: Prof. Dr. Patrício Moreira de Araújo Filho

Edição e Diagramação: Romilson Carneiro Rodrigues

Edição de Arte: Romilson Carneiro Rodrigues

Bibliotecária: Rayssa Cristhália Viana da Silva – CRB-13/904

Revisão: Os autores

Dados Internacionais de Catalogação na Publicação (CIP)

N972u

Coletânea Uma visão abrangente da computação / Mirian Nunes de Carvalho Nunes *et al*. (Org). São Luís - Editora Pascal, 2026.

132 f. : il.: (Uma visão abrangente da computação; v. 5)

Formato: PDF

Modo de acesso: World Wide Web

ISBN: 978-65-6068-226-9

D.O.I.: 10.29327/5827156

1. Computação. 2. Inteligência artificial. 3. Computação forense. 4. Proteção de dados. I. Nunes, Mirian Nunes de Carvalho. II. Oliveira, Thiago Santana de. III. Guimarães, Ivone Ascar Sauáia. IV. Bastos, Ewerton Ferreira. V. Varão, Tayssara Elizavieta Martins. VI. Costa, Wagner Elvio de Loiola. VII. Souto Filho, Antonio Luís de. VIII. Título.

CDU: 004::343.98

2026

www.editorapascal.com.br

APRESENTAÇÃO

Esta edição da série “Uma visão abrangente da computação” é o resultado da seleção de vários artigos científicos publicados sobre a temática central da obra. Nesta edição são abordados temas como a influência e aplicações em Inteligência Artificial (IA), o uso da computação forense em crimes cibernéticos, uma abordagem sobre a importância da segurança da informação em banco de dados onde se encontram armazenadas diversas informações e uma análise sobre a implementação da lei geral de proteção de dados pessoais, assunto este bastante atual e em evidência em ambientes empresariais devido a sua grande importância. As áreas abordadas como a IA, segurança da informação e a lei geral de proteção de dados, a LGPD, os organizadores ressaltam a sua importância devido a sua grande relevância para a solução de problemas encontrados atualmente em ambientes empresariais.

Os autores desta série científica confirmam o valor dessas áreas da engenharia e ciência da computação e as soluções encontradas, mas principalmente vem reforçar a importância do tema de vanguarda e sua aplicabilidade, contribuindo para que as empresas e centros de pesquisa possam identificar projetos com o potencial de desenvolvimento de novas tecnologias e inovação para o futuro de novas aplicações e soluções de problemas empresariais.

Wagner Elvio de Loiola Costa

Mestre em engenharia elétrica

ORGANIZADORES

Mirian Nunes de Carvalho Nunes

Pós-Graduada em Gestão Educacional pela Faculdades Integradas Potencial - FIP - Cotias - SP; em Arte, Educação e Tecnologias Contemporâneas pela Universidade de Brasília - UnB e em Docência do Ensino Superior pela Universidade Candido Mendes RJ. Graduada em Desenho Industrial pela Universidade Federal do Maranhão - UFMA. Graduada em Formação Pedagógica de Docentes para as áreas do Ensino Médio e Profissionalizante pela Universidade Estadual do Maranhão - UEMA. Exerço cargo de Professora na Faculdade Anhanguera São Luís - MA, ministrando disciplinas da área de Desenho Técnico em programas computacionais (AutoCAD, Inventor, REVIT). Atuo com apoio à Orientação de TCC. Atuei como Professora EaD pela UEMANET.

Thiago Santana de Oliveira

Bacharel em Engenharia Mecânica pelo Instituto Federal do Maranhão (2004), com mestrado em Engenharia de Materiais (2016), na mesma instituição. Atuou como profissional nas áreas de siderurgia e gerenciamento de frota de veículos e equipamentos a diesel, com bons conhecimentos nas ferramentas de gestão da manutenção e produção. Ministra aulas desde 2005, sendo a experiência inicial no ensino médio e técnico. Atualmente, trabalha no ensino superior, onde possui experiência de 10 anos, e técnico, ocupando os cargos de docente e coordenador de curso. Responsável pela organização de eventos na instituição e gestão de documentação dos cursos que coordeno. Possui importantes publicações na área.

Ivone Ascar Sauáia Guimarães

Possui graduação em Tecnologia em Processamento de Dados pelo Centro Universitário do Maranhão (2000), especialização em Gestão em Tecnologia e Negócios em Telecomunicações pela Universidade Estácio de Sá (2001) e Mestrado em Educação pela Universidade Católica de Brasília (2011), com pesquisa voltada à Educação a Distância como elemento de inclusão social. Atualmente cursa Licenciatura em Letras pelo Centro Universitário Claretiano, com conclusão prevista para 2025. Possui mais de 18 anos de experiência no ensino superior, atuando em diversos cursos de graduação, como Sistemas de Informação, Engenharias, Administração, Pedagogia, Radiologia e Turismo, ministrando disciplinas nas áreas de informática, tecnologia da informação, engenharia de software, metodologia científica e educação a distância. Atuou como docente na Universidade Ceuma (2003–2020) e na Faculdade Devry São Luís (2015–2017). Na Universidade Ceuma, integrou o Núcleo de Pesquisas em Sistemas e Tecnologia da Informação (NusTI) e participou de instâncias acadêmicas como NDE, Conselho de Curso, CEPE e CONSU. Possui experiência também como instrutora no SENAI, nas áreas de redes de computadores e lógica computacional. Atualmente desenvolve atividades de consultoria em normalização acadêmica, mentoria em trabalhos científicos, produção de conteúdo educacional e atuação como conteudista para instituições de ensino. Atua ainda como orientadora de projetos de MBA na XP Educação (IGTI), tutora na Uemanet e no Senar, e docente do curso de Ciência da Computação da Faculdade Anhanguera de São Luís, onde coordena o CodeFlow Collab Hub, iniciativa voltada à integração entre ensino, projetos acadêmicos e estágio na área de computação.

ORGANIZADORES

Ewerton Ferreira Bastos

Possui graduação em SISTEMA DE INFORMAÇÃO pela Universidade Ceuma (2010) e mestrado em Engenharia da Computação pela Universidade Estadual do Maranhão (2021). Atualmente é professor do curso de redes de computadores da Faculdade Laboro. Tem experiência na área de Ciência da Computação, com ênfase em Sistemas de Informação, atuando principalmente nos seguintes temas: saúde, gestão, redes, resíduos. rup, poo, spring, hibernate e mvc. e medicamentos.

Tayssara Elizavieta Martins Varão

Realizei estágio supervisionado na PROCOMP INDUSTRIA ELETRONICA LTDA com atividades destinadas ao auxilio técnico e administrativo. Tenho por área de atuação a área Técnica em Eletrotécnica e Engenharia Elétrica. Desenpenhei trabalho instrutora em escola técnica (LED Cursos técnicos e proficionalizantes) entre os anos de 2011e 2019. Concumitante a isso, exerci empreendedorismo no ramo alimentício até o ano de 2020, quando migrei ao ensino superior na então Faculdade Pitágoras, atual Anhanguera, onde leciono atualmente.

Wagner Elvio de Loiola Costa

Doutorando em Engenharia Elétrica com ênfase na área de Automação e Controle no Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Maranhão (PPGEE/CCET/UFMA). Mestre em Engenharia Elétrica com ênfase em Ciência da Computação pela Universidade Federal do Maranhão. Graduação em Engenharia de Elétrica pela Universidade Federal do Maranhão, graduação em Formação Pedagógica em Matemática pela Faculdade PITÁGORAS-São Luís-MA. Técnico laboratório de redes de computadores da Universidade Federal do Maranhão , professor da Faculdade Anhanguera São Luís-MA. Tem experiência na área de Ciência da Computação, com ênfase em Arquitetura de Sistemas de Computação, atuando principalmente nos seguintes temas: redes de computadores, energia elétrica, iot, zabbix e segurança da informação.

Antonio Luís de Souto Filho

Possui graduação em MATEMÁTICA LICENCIATURA pela Universidade Federal do Maranhão (2002), com monografia na área de Topologia: Um teorema sobre Métricas”; tem Especialização em Álgebra, com monografia em Teoria dos Números: O ANEL DOS INTEIROS GUASSIANOS, também pela Universidade Federal do Maranhão e Mestrado Profissional em Matemática também pela Universidade Federal do Maranhão, cuja Dissertação na área de Teoria dos Números: O TEOREMA CHINÊS DOS RESTOS.

SUMÁRIO

1

CHATBOTS NA SAÚDE: UMA REVISÃO DE LITERATURA SOBRE O USO DA INTELIGÊNCIA ARTIFICIAL NA TRIAGEM MÉDICA

CHATBOTS IN HEALTHCARE: A LITERATURE REVIEW ON THE USE OF ARTIFICIAL INTELLIGENCE IN MEDICAL TRIAGE

Márcio José Martins Câmara

Resumo

A Inteligência Artificial (IA) tem promovido transformações significativas no setor de saúde, impulsionando a busca por soluções para a otimização de processos e a gestão da crescente demanda por atendimento. Nesse contexto, agentes conversacionais, conhecidos como *chatbots*, emergiram como ferramentas promissoras no apoio à triagem médica, fornecendo orientação inicial e auxiliando no encaminhamento de pacientes. Este trabalho objetivou analisar os benefícios, desafios e limitações da utilização de *chatbots* baseados em IA na triagem médica, conforme a literatura científica recente. A pesquisa foi conduzida por meio de uma revisão bibliográfica de natureza qualitativa e descritiva, consultando bases de dados como Google Acadêmico, SciELO, PubMed e IEEE Xplore, com foco em publicações dos últimos dez anos (2014-2024). Os resultados indicaram que os *chatbots* são eficazes na coleta estruturada de sintomas, na redução da sobrecarga hospitalar e na ampliação do acesso à informação de saúde, com acurácia notável em áreas específicas como a oftalmologia e a oncologia. Contudo, a literatura ressaltou desafios importantes, como a variabilidade metodológica dos estudos, a necessidade de regulamentação clara sobre ética e responsabilidade profissional, e a dificuldade de integração com os sistemas de prontuários eletrônicos existentes. Concluiu-se que, apesar do potencial para otimizar a triagem, os *chatbots* devem ser implementados como ferramentas complementares e supervisionadas, dependentes de validação clínica robusta e diretrizes regulatórias para garantir segurança e confiabilidade no cuidado ao paciente.

Palavras-chave: *Chatbot*. Inteligência Artificial. Triagem Médica. Saúde Digital. Revisão de Literatura

Abstract

Artificial Intelligence (AI) has promoted significant transformations in the healthcare sector, driving the search for solutions to optimize processes and manage the growing demand for care. In this context, conversational agents, known as chatbots, have emerged as promising tools to support medical triage, providing initial guidance and assisting in patient referrals. This work aimed to analyze the benefits, challenges, and limitations of using AI-based chatbots in medical triage, according to recent scientific literature. The research was conducted through a qualitative and descriptive literature review, consulting databases such as Google Scholar, SciELO, PubMed, and IEEE Xplore, focusing on publications from the last ten years (2014-2024). The results indicated that chatbots are effective in the structured collection of symptoms, in reducing hospital overload, and in expanding access to health information, with notable accuracy in specific areas such as ophthalmology and oncology. However, the literature highlighted important challenges, such as the methodological variability of the studies, the need for clear regulations on ethics and professional responsibility, and the difficulty of integration with existing electronic health record systems. It was concluded that, despite the potential to optimize triage, chatbots should be implemented as complementary and supervised tools, dependent on robust clinical validation and regulatory guidelines to ensure safety and reliability in patient care.

Keywords: Chatbot. Artificial Intelligence. Medical Triage. Digital Health. Literature Review

1 INTRODUÇÃO

A Inteligência Artificial (IA) havia revolucionado diversos setores, e a área da saúde não foi exceção. Diante da crescente demanda por atendimento e da limitação de recursos humanos e infraestruturais, foram buscadas soluções tecnológicas capazes de otimizar processos assistenciais e administrativos. Nesse contexto, agentes conversacionais popularmente conhecidos como chatbots passaram a ser empregados para apoiar atividades de triagem, fornecendo informações iniciais aos usuários e auxiliando no encaminhamento para o atendimento mais adequado.

Os chatbots baseados em IA foram projetados para interagir por meio de linguagem natural, coletando relatos de sintomas, histórico breve e outras informações relevantes para uma pré-avaliação clínica. Na literatura consultada, esses sistemas mostraram potencial para reduzir a sobrecarga em serviços de saúde, ampliar o acesso informativo e oferecer disponibilidades de atendimento fora do horário tradicional, embora o grau de eficácia e a aplicabilidade clínica tenham sido apontados como dependentes da sofisticação dos modelos de processamento de linguagem natural e da integração com fluxos clínicos existentes.

Todavia, a adoção de chatbots na triagem médica também implicou desafios importantes. Foram identificadas limitações relacionadas à precisão nas respostas, à capacidade de interpretação de nuances comunicativas, à aceitação por parte de pacientes e profissionais e a questões de privacidade e segurança de dados sensíveis. Além disso, a literatura ressaltou a necessidade de regulamentações e protocolos que definam responsabilidades e limites de uso desses sistemas em contextos clínicos. Essas considerações tornaram-se ainda mais evidentes no período recente, em que a pandemia de COVID-19 acelerou a implementação de soluções digitais em saúde.

Diante desse cenário, a pesquisa foi conduzida na forma de revisão bibliográfica, de natureza qualitativa e descritiva, com o objetivo de mapear o estado atual do conhecimento sobre a utilização de chatbots baseados em IA na triagem médica. Foram consultadas bases como Google Acadêmico, SciELO, PubMed e IEEE Xplore, considerando publicações no período de 2014 a 2024, em português e inglês, e empregando descritores relacionados a chatbots na saúde, inteligência artificial e triagem médica. A seleção priorizou trabalhos alinhados ao objeto de estudo e de atualidade, buscando identificar benefícios, limitações e lacunas de pesquisa nessa área.

O problema de pesquisa que orientou este trabalho foi de que maneira a utilização de chatbots baseados em Inteligência Artificial tem contribuído para a triagem médica, considerando seus benefícios e limitações? O objetivo geral foi analisar os benefícios, desafios e limitações do uso de chatbots baseados em Inteligência Artificial na triagem médica, segundo a literatura científica. Como objetivos específicos, buscou-se descrever as principais funcionalidades dos chatbots aplicados à triagem médica; apontar os benefícios e desafios do uso da IA na triagem de pacientes; apresentar as discussões da literatura sobre a eficácia desses sistemas, considerando suas limitações; e compreender como a literatura abordou a aceitação dos chatbots por pacientes e profissionais de saúde.

2 DESENVOLVIMENTO

2.1 Metodologia

A pesquisa foi conduzida por meio de uma revisão bibliográfica, de natureza qualitativa e caráter descritivo, com foco na utilização de *chatbots* baseados em Inteligência Artificial aplicados à triagem médica. Essa abordagem permitiu reunir e analisar produções

científicas que discutiram os benefícios, desafios e limitações desses sistemas.

Foram consultadas as bases Google Acadêmico, SciELO, PubMed e IEEE Xplore, considerando publicações entre 2014 e 2024, em português e inglês. Utilizaram-se como descritores os termos "*chatbot* na saúde", "inteligência artificial em saúde" e "triagem médica com *chatbots*". Foram incluídos estudos que abordaram diretamente a aplicação de *chatbots* na triagem de pacientes e excluídos aqueles que tratavam de outras áreas, como marketing ou atendimento ao cliente.

Os trabalhos selecionados foram avaliados quanto à aderência ao tema e organizados em uma síntese narrativa, que destacou objetivos, métodos, resultados e conclusões apresentados pelos autores, possibilitando uma visão crítica sobre as potencialidades e limitações do uso de *chatbots* na triagem médica.

2.2 Resultados e Discussão

A revisão bibliográfica realizada possibilitou identificar um conjunto expressivo de estudos que investigaram a utilização de *chatbots* baseados em Inteligência Artificial (IA) no apoio à triagem médica, indicando um campo de pesquisa em rápida expansão e com grande relevância clínica e social. Os trabalhos analisados revelaram que esses sistemas desempenharam funções cruciais como a coleta estruturada de sintomas, a classificação preliminar de gravidade e a orientação inicial ao paciente, elementos fundamentais para otimizar os fluxos de atendimento em ambientes de alta demanda. No entanto, a análise crítica da literatura evidenciou que, embora os resultados iniciais tenham sido promissores em termos de acesso e agilidade, também foram observadas limitações metodológicas importantes nos estudos, desafios de integração com fluxos clínicos e preocupações éticas crescentes sobre a segurança dos dados e a responsabilidade profissional em casos de falha no diagnóstico preliminar (FRONZA et al., 2023; ADAMOPOULOU; MOUSSIADES, 2020).

Os primeiros estudos focaram na acurácia e na aplicação em nichos clínicos bem definidos, demonstrando o potencial de precisão da IA. Por exemplo, na área da oftalmologia, Park et al. (2023) compararam o desempenho de *chatbots* com estagiários de medicina em 44 casos simulados de triagem ocular. O estudo revelou uma acurácia superior a 90% na identificação de urgências. Contudo, essa alta precisão veio acompanhada de uma tendência à supertriagem, o que, embora garanta a segurança ao não deixar passar casos graves, poderia gerar encaminhamentos desnecessários e, paradoxalmente, aumentar a sobrecarga em outros pontos da cadeia de atendimento. A conclusão dos pesquisadores foi enfática: tais sistemas apresentaram utilidade significativa, mas ainda precisavam ser utilizados de forma complementar e sob supervisão especializada, e não como substitutos autônomos do julgamento clínico.

A aplicação da tecnologia em cenários de monitoramento contínuo também se destacou. Em oncologia, Bibault et al. (2019) relataram que os *chatbots* foram aplicados no monitoramento de efeitos adversos da quimioterapia, na triagem de sintomas e na orientação de pacientes sobre condutas de autocuidado. Revisões posteriores identificaram que, de 21 *chatbots* analisados nesse contexto, 14 relataram altos níveis de satisfação dos usuários. Os pacientes destacaram como principais vantagens a clareza das informações e o apoio contínuo durante o tratamento. Entretanto, os pesquisadores alertaram que a ausência de integração com equipes multiprofissionais e a falta de capacidade do agente conversacional em lidar com informações complexas poderiam comprometer a segurança e a integralidade do cuidado. A literatura, nesse sentido, reforçou a ideia de que a sofisticação do modelo de linguagem natural (NPL) é diretamente proporcional à sua eficácia

e aceitação (PARK et al., 2020).

No campo da saúde mental, a aceitação foi particularmente favorável em virtude do aspecto de anonimato e da alta acessibilidade. You e Gui (2021) e outros autores apontaram que *chatbots* como Woebot, Tess e Wysa foram utilizados com sucesso na triagem inicial de sintomas depressivos e ansiosos, principalmente entre usuários jovens e adultos. Essa aceitação elevada deveu-se, em parte, à facilidade de acesso 24 horas e ao ambiente de confidencialidade percebido. Apesar disso, a literatura enfatizou uma limitação crítica: em casos graves ou de risco de vida, a intervenção automatizada mostrou-se insuficiente, sendo imprescindível o encaminhamento imediato para atendimento humano especializado. Os resultados sugeriram, portanto, que esses agentes foram úteis como primeira linha de acolhimento e suporte leve, mas não como substitutos do tratamento clínico e psicoterápico formal.

A urgência imposta pela crise sanitária da COVID-19 serviu como um acelerador e um teste de estresse para a tecnologia. Durante a pandemia, os *chatbots* foram amplamente empregados para triagem de sintomas respiratórios e para o encaminhamento de pacientes. Barros et al. (2023) relataram experiências no Brasil e em outros países que demonstraram uma redução significativa de filas em prontos-socorros e maior acesso a orientações confiáveis sobre isolamento e vacinação. Embora a acurácia tenha variado conforme o modelo utilizado, os autores destacaram que esses sistemas foram fundamentais para aliviar a sobrecarga hospitalar em um momento de crise sanitária, provando seu valor em cenários de saúde pública. Além disso, a iniciativa de universidades brasileiras, como a UNICAMP, no desenvolvimento de chatbots específicos para orientação da população reforçou a relevância de soluções locais e adaptáveis em emergências (UNICAMP, 2020).

Observa-se, portanto, um padrão claro nos achados: os *chatbots* demonstram maior eficácia imediata não como substitutos do diagnóstico, mas como ferramentas de otimização de fluxo e ampliação do acesso. A capacidade de oferecer um primeiro atendimento disponível 24 horas, coletar sintomas de forma estruturada e fornecer orientação validada (como no manejo de efeitos adversos da quimioterapia ou no isolamento da COVID-19) são os benefícios mais consistentemente relatados. Em contextos de alta demanda ou estigma, como a saúde mental, os *chatbots* preenchem uma lacuna de acessibilidade, sendo valorizados pelos usuários pela confidencialidade e clareza informativa.

Contudo, essa mesma literatura que aponta os benefícios é unânime em sinalizar que a implementação não é trivial e que o campo ainda carece de amadurecimento metodológico. Uma ressalva recorrente é o descompasso entre o potencial técnico e a validação clínica robusta. Muitos estudos, como apontado por Fronza et al. (2023), ainda possuem caráter exploratório e amostras pequenas, dificultando a generalização dos resultados. Essa fragilidade metodológica é um obstáculo para estabelecer métricas padronizadas que comprovem o impacto real em desfechos clínicos, além da simples satisfação do usuário.

Quadro 1. Principais estudos sobre *chatbots* aplicados à triagem médica

Autor (Ano)	**Área de Aplicação**	**Tipo de Estudo**	**Principais Achados**
Park et al. (2023)	Oftalmologia	Estudo comparativo (*Chatbot* vs. estagiários)	Acurácia de 93% nos diagnósticos; tendência à supertriagem.
You & Gui (2021)	Saúde Mental	Revisão / Relato de experiência	Alta aceitação; útil como primeira linha de atendimento; limitações em casos graves.

Bibault et al. (2019)	Oncologia	Revisão e estudos clínicos	Alta satisfação do usuário (14 de 21 estudos); suporte eficiente em quimioterapia.
Barros et al. (2023)	Triagem de COVID-19	Revisão integrativa	Redução de filas; relevância durante crises sanitárias; precisão variável conforme o sistema.
Fronza et al. (2023)	Aplicações gerais em saúde	Revisão narrativa	Escassez de ensaios clínicos robustos; predominância de estudos exploratórios.
Adamopoulou & Moussiades (2020)	Ética e Tecnologia	Revisão conceitual	Preocupações com privacidade, segurança de dados e regulamentação de IA.
Revisão Perioperatória (2023)	Cirurgia	Meta-análise (8 ensaios clínicos)	73% de satisfação; aumento de 80% no conhecimento do paciente; necessidade de supervisão profissional.

Fonte: Adaptado de Park et al. (2023), You e Gui (2021), Bibault et al. (2019), Barros et al. (2023), Fronza et al. (2023) e Adamopoulou e Moussiades (2020).

A urgência imposta pela crise sanitária da COVID-19 serviu como um acelerador e um teste de estresse para a tecnologia. Durante a pandemia, os *chatbots* foram amplamente empregados para triagem de sintomas respiratórios e para o encaminhamento de pacientes. Barros et al. (2023) relataram experiências no Brasil e em outros países que demonstraram uma redução significativa de filas em prontos-socorros e maior acesso a orientações confiáveis sobre isolamento e vacinação. Embora a acurácia tenha variado conforme o modelo utilizado, os autores destacaram que esses sistemas foram fundamentais para aliviar a sobrecarga hospitalar em um momento de crise sanitária, provando seu valor em cenários de saúde pública. Além disso, a iniciativa de universidades brasileiras, como a UNICAMP, no desenvolvimento de *chatbots* específicos para orientação da população reforçou a relevância de soluções locais e adaptáveis em emergências (UNICAMP, 2020).

Apesar dos benefícios consistentes, a literatura revelou limitações metodológicas e operacionais que precisam ser endereçadas. A revisão de Fronza et al. (2023) apontou que grande parte das pesquisas ainda era de caráter exploratório, com amostras pequenas e ausência de grupos controle. A escassez de ensaios clínicos robustos dificultou a avaliação do impacto real dos *chatbots* em desfechos clínicos concretos, como a redução de complicações ou a diminuição do tempo de atendimento em casos urgentes. Além disso, observou-se uma variabilidade nos critérios de avaliação da eficácia: enquanto alguns trabalhos priorizaram a acurácia diagnóstica, outros focaram apenas na percepção de satisfação do usuário. Essa diversidade metodológica constitui um obstáculo à comparação direta dos resultados, reforçando a necessidade de padronização de métricas para mensurar com clareza o impacto clínico dos *chatbots* na triagem médica.

Um dos desafios mais urgentes discutidos na academia é a lacuna regulatória e ética. Adamopoulou e Moussiades (2020) destacam que a coleta e o armazenamento de dados sensíveis, sobretudo em saúde mental, podem gerar sérios riscos à confidencialidade e à segurança. A conformidade com a LGPD no Brasil e o GDPR na Europa é essencial, mas persiste a dúvida sobre a responsabilização em caso de erro. A União Europeia avançou com o Artificial Intelligence Act (UNIÃO EUROPEIA, 2023), porém a definição de responsabilidade legal diante de triagens incorretas por IA ainda é incerta no cenário global. Essa falta de diretrizes claras sobre limites de atuação e responsabilização constitui um obstá-

culo importante para a adoção ampla de sistemas de IA em saúde.

A "diversidade metodológica" apontada pela literatura não é apenas uma limitação acadêmica; ela tem consequências práticas diretas. A ausência de "padronização de métricas" e a predominância de "estudos exploratórios" dificultam que gestores de saúde e equipes clínicas comparem soluções de forma objetiva. Fica difícil determinar se a "acurácia" relatada por um sistema é comparável à de outro, ou se a "satisfação do usuário" se traduzirá em "desfechos clínicos concretos". Essa falta de comparabilidade e de evidências robustas de ensaios clínicos gera insegurança e retarda a adoção de ferramentas que poderiam, de fato, ser eficazes.

Além das questões metodológicas, um desafio operacional recorrente é a baixa "interoperabilidade com prontuários eletrônicos". A literatura adverte que chatbots implementados como sistemas isolados, que não se comunicam com os fluxos de trabalho existentes, falham em entregar o benefício da otimização. Em vez de aliviar a carga, eles podem gerar "retrabalho", forçando os profissionais a consultar múltiplos sistemas ou a reinserir manualmente os dados coletados pelo *chatbot*. Como indicado nos estudos em oncologia, a falta de integração com a equipe multiprofissional pode comprometer a continuidade do cuidado, tornando a integração técnica um pré-requisito para o sucesso clínico.

Um ponto de tensão identificado na literatura é o fenômeno da "supertriagem". Embora a precisão em áreas como a oftalmologia seja alta, os sistemas são frequentemente programados para errar por excesso de cautela, garantindo que casos graves não sejam negligenciados. Contudo, essa tendência, embora segura para o paciente individual, pode reverter um dos principais benefícios esperados do *chatbot*: a redução da sobrecarga hospitalar. Ao encaminhar mais pacientes do que o necessário para o atendimento especializado, o sistema pode, paradoxalmente, aumentar a sobrecarga em outros pontos da cadeia de atendimento, exigindo um equilíbrio delicado entre a sensibilidade do algoritmo e a eficiência operacional do serviço de saúde.

Quadro 2. Síntese de benefícios e limitações na aplicação de *Chatbots* na Triagem Médica

Aspectos Positivos	**Limitações Encontradas**
Ampliação do acesso a orientações iniciais	Escassez de ensaios clínicos robustos
Disponibilidade de atendimento 24 horas	Ausência de padronização de métricas de avaliação
Redução da sobrecarga hospitalar em períodos de crise (ex.: COVID-19)	Riscos de supertriagem e respostas imprecisas
Apoio educacional e suporte ao manejo de doenças crônicas	Desafios éticos e questões relacionadas à privacidade de dados (LGPD/GDPR)
Economia de tempo e aumento da eficiência dos profissionais de saúde	Baixa interoperabilidade com prontuários eletrônicos e sistemas de saúde

Fonte: Adaptado de Adamopoulou e Moussiades (2020); Fronza et al. (2023).

A integração tecnológica e a aceitação profissional são determinantes para o sucesso da implementação. Os estudos evidenciaram que a implementação isolada dos *chatbots*, sem interoperabilidade com prontuários eletrônicos e fluxos assistenciais, exigiu retrabalho das equipes médicas, neutralizando os benefícios esperados. Pesquisadores defenderam que a adoção bem-sucedida desses sistemas dependeu não apenas da sofisticação técnica, mas também da capacidade de integração ao ecossistema hospitalar e da capacitação contínua dos profissionais para utilizá-los de forma complementar. Nesse sentido, a aceitação institucional e o treinamento das equipes foram apontados como fatores decisivos para a consolidação do uso seguro e eficaz. A aceitação dos pacientes foi, em geral,

positiva, com usuários relatando que a experiência foi simples e esclarecedora; contudo, a cautela por parte dos profissionais sugere que a confiança plena ainda está em construção e depende de resultados clínicos robustos e validados.

A síntese dos estudos demonstrou que os *chatbots* baseados em IA apresentaram potencial relevante para apoiar a triagem médica em diferentes áreas, desde especialidades clínicas até saúde pública. Os resultados apontaram ganhos em acessibilidade, eficiência e satisfação dos pacientes, além de contribuições em educação em saúde. Contudo, também foram identificadas limitações metodológicas, desafios regulatórios e necessidade de maior integração tecnológica. A análise dos trabalhos também apontou caminhos futuros, como a necessidade de ensaios clínicos randomizados e o desenvolvimento de algoritmos capazes de lidar com a linguagem natural em diferentes contextos culturais e linguísticos. Com o avanço dos grandes modelos de linguagem, vislumbra-se que os *chatbots* podem oferecer interações mais precisas, mas a consolidação do uso clínico seguro exigirá protocolos éticos, interoperabilidade com sistemas hospitalares e contínua supervisão profissional, de modo a garantir que a inovação tecnológica seja convertida em benefícios efetivos para os pacientes e para os serviços de saúde.

A "cautela por parte dos profissionais", em contraste com a aceitação positiva dos pacientes, é um fator determinante para a implementação. A literatura sugere que essa hesitação é justificada por dois desafios operacionais e legais: primeiro, a falta de "interoperabilidade com prontuários eletrônicos", que transforma a ferramenta de um auxílio para um "retrabalho"; e segundo, a "ausência de diretrizes normativas" sobre a "responsabilidade legal". O profissional de saúde sente-se receoso ao ser legalmente responsável por uma recomendação de triagem gerada por um algoritmo, especialmente um cuja lógica interna (a "caixa-preta") nem sempre é transparente.

Em contrapartida à cautela profissional, a "aceitação dos pacientes foi, em geral, positiva". Os estudos indicam que essa aceitação é impulsionada por fatores que vão além da simples acurácia. Em saúde mental, o "anonimato" e a "facilidade de acesso 24 horas" foram decisivos, permitindo que usuários buscassem uma primeira linha de apoio sem o estigma associado. Em oncologia, a "clareza das informações e o apoio contínuo" foram relatados como vantagens-chave. Isso sugere que os pacientes valorizam a conveniência, a privacidade e a resposta imediata que os *chatbots* oferecem, preenchendo lacunas deixadas pelo atendimento tradicional.

3 CONCLUSÃO

O presente artigo teve como objetivo analisar os benefícios, desafios e limitações do uso de *chatbots* baseados em Inteligência Artificial na triagem médica, conforme a literatura científica publicada entre 2014 e 2024. O problema de pesquisa, que questionava de que maneira a utilização de *chatbots* tem contribuído para a triagem médica, foi respondido por meio de uma revisão bibliográfica que mapeou os achados mais recentes sobre o tema.

Os resultados demonstram que os *chatbots* oferecem uma contribuição significativa para a triagem médica, atuando principalmente na otimização do acesso e na eficiência inicial do atendimento. Os objetivos específicos propostos foram amplamente contemplados, descrevendo-se as principais funcionalidades (coleta de sintomas, classificação preliminar), apontando-se benefícios (disponibilidade 24 horas, redução de sobrecarga) e desafiose (supervisão profissional, acurácia variável) e compreendendo-se que a aceitação é alta entre pacientes, mas mais cautelosa entre os profissionais.

No entanto, as limitações do estudo de revisão apontam para a imaturidade metodológica do campo, com predominância de estudos exploratórios e a carência de ensaios clínicos randomizados que possam atestar o impacto em desfechos de saúde a longo prazo. Além disso, a principal barreira para a consolidação segura dos *chatbots* reside nas lacunas regulatórias, especialmente em relação à privacidade de dados sensíveis e à definição clara da responsabilidade legal em casos de erro de triagem. Trabalhos futuros devem focar no desenvolvimento de métricas padronizadas para a avaliação da eficácia clínica e na criação de diretrizes éticas e regulatórias robustas.

REFERÊNCIAS

ADAMOPOULOU, E.; MOUSSIADES, L. An overview of chatbot technology. In: **Artificial Intelligence Applications and Innovations**. Cham: Springer, 2020. p. 373–383. Disponível em: https://doi.org/10.1007/978-3-030-49186-4_31. Acesso em: 25 maio 2025.

BARROS, R. Q. F. et al. Eficácia da utilização de chatbots na triagem de sintomas de COVID-19: uma revisão integrativa. **Revista Multidisciplinar em Saúde**, v. 4, n. 3, p. 812–817, 2023. Disponível em: https://editorainterg rar.com.br/publish/index.php/rems/article/view/4047. Acesso em: 25 maio 2025.

BIBAULT, J.-E. et al. **A chatbot versus physicians to provide information for patients with breast cancer: blind, randomized controlled noninferiority trial**. Journal of Medical Internet Research, v. 21, n. 11, e15787, 2019. DOI: https://doi.org/10.2196/15787. Acesso em: 25 agosto 2025.

FRONZA, R.; ROLIM, F. A.; LIMA, E. **Review of artificial intelligence chatbots in healthcare: opportunities and challenges.** International Journal of Medical Informatics, v. 171, 104997, 2023. DOI: https://doi.org/10.1016/j.ijmedinf.2022.104997. Acesso em: 25 agosto 2025.

PARK, C. J. et al. **Accuracy of ChatGPT and Bing AI chat for ophthalmic triage compared to medical trainees**. Canadian Journal of Ophthalmology, v. 58, n. 3, p. 312–318, 2023. DOI: https://doi.org/10.1016/j.jcjo.2023.05.005. Acesso em: 25 agosto 2025.

PARK, C.-W. et al. Artificial Intelligence in Health Care: Current Applications and Issues. **Journal of Korean Medical Science**, v. 35, n. 42, p. e379, 2020. Disponível em: https://doi.org/10.3346/jkms. 2020.35.e379. Acesso em: 25 maio 2025.

REVISÃO PERIOPERATÓRIA. **Chatbots in perioperative care: a systematic review and meta-analysis**. Annals of Surgery Open, v. 4, n. 2, p. e240, 2023. DOI: https://doi.org/10.1097/AS9.0000000000000240. Acesso em: 25 agosto 2025.

UNIÃO EUROPEIA. ***Artificial Intelligence Act*. Regulamento (UE) 2023/2854 do Parlamento Europeu e do Conselho, de 13 de dezembro de 2023**. Jornal Oficial da União Europeia, Bruxelas, 2023. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32023R2854. Acesso em: 2 out. 2025.

UNICAMP. ***Chatbot auxilia triagem de casos de COVID-19*.** Universidade Estadual de Campinas, 2020. Disponível em: https://www.unicamp.br/unicamp/noticias/2020/04/06/chatbot-auxilia-triagem-de-casos-de-covid-19. Acesso em: 4 setembro 2025.

YOU, S.; GUI, R. **Chatbots in mental health: a review and critical analysis**. Frontiers in Psychology, v. 12, p. 799–812, 2021. DOI: https://doi.org/10.3389/fpsyg.2021.799612. Acesso em: 25 agosto 2025.

2

APLICAÇÕES DE IA NO MONITORAMENTO AMBIENTAL

AI APPLICATIONS IN ENVIRONMENTAL MONITORING

Gabriel Ryan Ferreira da Silva
Ivone Ascar Sauaia Guimarães
Mirian Nunes de Carvalho Nunes

Resumo

A Inteligência Artificial (IA) tem se consolidado como uma ferramenta estratégica para o monitoramento ambiental, permitindo a automatização e o aprimoramento de processos de coleta, análise e interpretação de dados ambientais. Este artigo tem como objetivo analisar as principais aplicações da IA no monitoramento de variáveis ambientais, destacando suas contribuições para a detecção, previsão e mitigação de impactos ecológicos. A pesquisa, de natureza bibliográfica e abordagem qualitativa, baseou-se na revisão de estudos científicos recentes que tratam da integração entre IA, sensoriamento remoto e Internet das Coisas (IoT). Os resultados indicam que a IA tem aprimorado a previsão de eventos extremos, o controle da poluição atmosférica, o monitoramento da qualidade da água e solo e a conservação da biodiversidade. Além disso, observou-se que a explicabilidade dos modelos, a qualidade dos dados e a sustentabilidade computacional são fatores decisivos para a efetividade das aplicações. Conclui-se que, embora a IA amplie significativamente as capacidades técnicas de monitoramento, sua eficácia depende da integração entre tecnologia, governança institucional e políticas públicas voltadas à sustentabilidade ambiental.

Palavras-chave: Sustentabilidade, Sensoriamento Remoto, Internet das Coisas (IoT), Previsão Climática.

Abstract

Artificial Intelligence (AI) has become a strategic tool for environmental monitoring, enabling the automation and improvement of processes for collecting, analyzing, and interpreting environmental data. This article aims to analyze the main applications of AI in monitoring environmental variables, highlighting its contributions to the detection, prediction, and mitigation of ecological impacts. The research, of a bibliographic nature and qualitative approach, was based on a review of recent scientific studies dealing with the integration between AI, remote sensing, and the Internet of Things (IoT). The results indicate that AI has improved the prediction of extreme events, the control of air pollution, the monitoring of water and soil quality, and the conservation of biodiversity. Furthermore, it was observed that the explainability of the models, the quality of the data, and computational sustainability are decisive factors for the effectiveness of the applications. It is concluded that, although AI significantly expands the technical capabilities of monitoring, its effectiveness depends on the integration between technology, institutional governance, and public policies focused on environmental sustainability.

Keywords: Sustainability, Remote Sensing, Internet of Things (IoT), Climate Forecasting.

1 INTRODUÇÃO

O monitoramento ambiental representou, ao longo dos últimos anos, uma ferramenta essencial para a preservação dos ecossistemas e para a redução dos impactos provocados pelas atividades humanas. Nesse cenário, a Inteligência Artificial (IA) se destacou como um recurso inovador capaz de potencializar a análise de dados e ampliar as estratégias de conservação, ao oferecer rapidez e precisão na detecção de mudanças ambientais. A utilização de técnicas avançadas, como aprendizado de máquina, redes neurais e processamento de imagens, possibilitou a identificação de padrões complexos e a previsão de eventos críticos, contribuindo para uma gestão ambiental mais eficiente.

A aplicação da IA no monitoramento ambiental justificou-se pela necessidade de superar as limitações dos métodos tradicionais, que envolviam altos custos, coleta manual de dados e pouca eficiência na análise em tempo real. Ao contrário desses modelos, os sistemas inteligentes demonstraram maior capacidade para integrar informações provenientes de sensores, drones e satélites, permitindo a antecipação de riscos ambientais, como desmatamentos, variações climáticas e níveis elevados de poluição. Dessa forma, a tecnologia ofereceu um caminho viável para ações preventivas e corretivas, colaborando diretamente com a sustentabilidade.

O problema de pesquisa residiu na dificuldade de aplicar métodos convencionais de monitoramento de forma ampla e acessível, considerando os custos, a limitação de infraestrutura e a demora nos diagnósticos. Tornou-se necessário compreender como a Inteligência Artificial poderia aprimorar esse processo e contribuir para a identificação precoce de problemas ecológicos, garantindo maior eficiência na preservação ambiental.

Diante dessa realidade, o objetivo geral da pesquisa consistiu em compreender de que forma a IA poderia ser aplicada no monitoramento ambiental, de modo a detectar precocemente problemas ecológicos e apoiar a gestão sustentável dos ecossistemas. Para alcançar esse propósito, a investigação estabeleceu objetivos específicos, como explorar as principais técnicas de IA aplicadas ao monitoramento ambiental, analisar os benefícios dessa tecnologia na detecção de alterações ambientais, discutir os desafios e limitações de sua implementação, apresentar casos reais de aplicação e refletir sobre sua influência na formulação de políticas públicas sustentáveis.

A relevância do estudo fundamentou-se na possibilidade de contribuir para a construção de novos conhecimentos sobre a interação entre tecnologia e meio ambiente, bem como para subsidiar decisões estratégicas em âmbito científico, político e social. Dessa maneira, a análise das aplicações da Inteligência Artificial no monitoramento ambiental apresentou-se como uma oportunidade de destacar não apenas os avanços, mas também os desafios que permeiam essa área, favorecendo o desenvolvimento de soluções inovadoras voltadas à proteção dos ecossistemas.

2 DESENVOLVIMENTO

2.1 Metodologia

Este estudo foi realizado por meio de uma Revisão Bibliográfica, de caráter qualitativo e descritivo, com o objetivo de compreender como a Inteligência Artificial vem sendo aplicada no monitoramento ambiental. Para isso, foram consultadas fontes de literatura acadêmica, como artigos científicos, livros, dissertações, relatórios técnicos e publicações de especialistas, disponíveis em bases de dados reconhecidas, incluindo IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink e Web of Science. A pesquisa priorizou traba-

lhos publicados nos últimos cinco anos, a fim de garantir um panorama atualizado. Os critérios de inclusão consideraram materiais em português e inglês, com embasamento científico e relevância para o tema. Já os critérios de exclusão abrangeram artigos de opinião, resumos sem conteúdo completo e publicações sem rigor acadêmico. As buscas foram realizadas utilizando palavras-chave como “Inteligência Artificial e meio ambiente”, “monitoramento ambiental com IA”, “aprendizado de máquina na preservação ambiental” e “análise de imagens ambientais com IA”. As buscas resultaram em cerca de 28 publicações relevantes, das quais 12 atenderam integralmente aos critérios de inclusão e foram analisadas de forma detalhada na revisão. Os demais trabalhos foram excluídos por não apresentarem abordagem direta sobre o uso de IA em monitoramento ambiental.

A pesquisa permitiu mapear as principais aplicações da IA na preservação dos ecossistemas, além de identificar desafios e impactos dessa tecnologia. Também foram analisadas abordagens de aprendizado de máquina aplicadas ao monitoramento ambiental, à previsão de riscos e à otimização de ações sustentáveis. A revisão bibliográfica ainda contribuiu para a compreensão das limitações e desafios éticos e técnicos envolvidos na implementação dessas tecnologias. Dessa forma, o estudo reuniu informações que possam embasar discussões futuras sobre o uso de IA na gestão ambiental.

2.2 Resultados e Discussão

Os resultados obtidos a partir da revisão bibliográfica indicaram que a Inteligência Artificial (IA) assumiu, de forma crescente, um papel estratégico no monitoramento ambiental, contribuindo para transformar práticas tradicionalmente manuais em sistemas mais automatizados, contínuos e escaláveis. A análise dos trabalhos mostrou que a combinação entre aprendizado de máquina, sensoriamento remoto e redes de sensores (IoT) promoveu saltos qualitativos na capacidade de detecção, mapeamento e previsão de eventos ambientais (Rolnick et al., 2019; Ribeiro et al., 2021).

Diversos estudos evidenciaram que a IA melhorou a previsão de eventos extremos, como enchentes, secas e queimadas, por meio da assimilação de múltiplas fontes de dados e da utilização de modelos preditivos em diferentes escalas. Zhang et al. (2020) observaram que arquiteturas de deep learning foram eficazes na identificação de padrões em imagens de satélite, reduzindo falsos negativos em detecções de desmatamento e alterações no uso do solo.

De forma complementar, Olawade et al. (2024) destacaram que a integração de modelos preditivos com dados meteorológicos e hidrológicos ampliou a acurácia de alertas antecipados. Isso permitiu que autoridades locais e agências de resposta a emergências adotassem medidas preventivas mais oportunas, reduzindo riscos sociais e ambientais.

No campo da qualidade do ar, os estudos revisados mostraram que modelos de machine learning foram aplicados para estimar concentrações de poluentes a partir de variáveis climáticas, sensores de baixo custo e dados de mobilidade urbana. Garcia et al. (2025) relataram que esses modelos forneceram estimativas mais densas do que as obtidas por estações convencionais, possibilitando a identificação de pontos críticos e apoiando medidas de mitigação quase em tempo real.

Garcia et al. (2025) e Gryech et al. (2024) também apontaram o impacto positivo da integração IA–IoT na densificação da malha de monitoramento urbano. Gryech et al. (2024) ressaltaram que esses sistemas possibilitaram maior resolução espacial em áreas urbanas equipadas com sensores distribuídos, favorecendo análises mais detalhadas da qualidade

do ar e intervenções rápidas em episódios de poluição.

O monitoramento da qualidade da água e do solo também apresentou avanços significativos com IA. Ribeiro et al. (2021) documentaram aplicações em que modelos inteligentes detectaram sinais de contaminação em rios a partir de séries temporais de sensores e variáveis físico-químicas. Os modelos permitiram antecipar picos de poluição antes que os impactos fossem sentidos, fortalecendo a segurança hídrica e alimentar.

Em consonância, Moreira e Campos (2020) observaram que soluções automatizadas ajudaram a suprir lacunas de frequência amostral nos métodos tradicionais, garantindo maior escalabilidade das análises. Assim, tornou-se possível ampliar a cobertura espacial e a rapidez de resposta em situações de risco ambiental, algo fundamental em regiões com alta vulnerabilidade ecológica.

No monitoramento da biodiversidade, técnicas de reconhecimento acústico e visual mostraram-se eficazes. Zhang et al. (2020) destacaram que a combinação de imagens ópticas e microfones ambientais gerou indicadores mais robustos de presença/ausência de espécies. Paralelamente, Olawade et al. (2024) relataram que análises bioacústicas baseadas em IA foram capazes de estimar tendências populacionais em florestas tropicais.

Esses resultados favoreceram a priorização de áreas para conservação e contribuíram para a proteção de espécies ameaçadas. As aplicações indicam que a IA, ao ampliar a capacidade de monitorar grandes extensões de ecossistemas, pode se consolidar como ferramenta essencial no suporte a políticas públicas de preservação da biodiversidade.

A utilização de IA no sensoriamento remoto aplicado à agricultura de precisão e ao controle de desmatamentos também foi amplamente relatada. Modelos de machine learning auxiliaram na detecção precoce de queimadas e alterações no uso do solo, permitindo fiscalizações mais ágeis e manejo agrícola mais sustentável (Olawade et al., 2024).

Além disso, iniciativas apoiadas pela European Commission, como o projeto Enviro-IoT (2025), mostraram que a combinação de sensores de baixo custo com processamento inteligente aumentou a viabilidade econômica de monitoramentos ambientais em regiões com recursos limitados. Esse fator é relevante especialmente em países em desenvolvimento.

Do ponto de vista técnico, não houve um algoritmo único dominante. A escolha metodológica variou conforme o problema e os dados disponíveis. Modelos de regressão e árvores de decisão foram aplicados em previsões numéricas, enquanto técnicas de clustering e redes neurais foram mais usadas em classificação de imagens e sinais (Russell; Norvig, 2016).

A literatura destacou também a importância de pipelines de pré-processamento robustos — como filtragem de ruído, normalização e data augmentation —, principalmente em situações em que os dados eram heterogêneos e provinham de múltiplos sensores (Zhang et al., 2020; Ribeiro et al., 2021).

Um ponto recorrente nas discussões foi a questão da explicabilidade e da confiança nos modelos. Gryech et al. (2024) ressaltaram que, em contextos ambientais que impactam políticas públicas, a interpretabilidade dos algoritmos tornou-se requisito fundamental. Assim, técnicas de explainable AI (XAI) foram necessárias para traduzir recomendações em ações concretas.

Ribeiro et al. (2021) e Batista e Lopes (2021) reforçaram que, sem transparência, mesmo modelos de bom desempenho quantitativo poderiam gerar resistência institucional. Portanto, a confiança social e política nos sistemas de IA depende diretamente da clareza

e da auditabilidade dos processos de decisão algorítmica.

A qualidade e a disponibilidade dos dados também surgiram como limitações. Muitos estudos destacaram lacunas de cobertura temporal e espacial, bem como problemas de calibração de sensores de baixo custo e de heterogeneidade das bases de treinamento. Garcia et al. (2025) observaram que esses sensores exigem calibração contínua e mecanismos de compensação por drift.

Ribeiro et al. (2021) acrescentaram que a interoperabilidade entre bancos administrativos, dados satelitais e sensores locais continua sendo um desafio. Para enfrentá-lo, são necessários investimentos em engenharia de dados e padronização, o que demanda esforço institucional e financeiro considerável.

Outro aspecto relevante foi a governança de dados e as implicações éticas. Batista e Lopes (2021) destacaram riscos relacionados à privacidade, como o monitoramento de atividades humanas por sensores ambientais, além de potenciais impactos sociais de decisões automatizadas. Estruturas de governança, segundo os autores, devem incluir transparência e auditoria.

Nesse sentido, Rolnick et al. (2019) e o relatório UNEP (2022) reforçam que os modelos de IA devem ser avaliados não apenas por critérios técnicos, mas também por sua aderência a princípios de sustentabilidade e justiça ambiental. Essa visão amplia a responsabilidade das tecnologias no contexto das mudanças globais.

Quanto à validação e à avaliação de desempenho, os estudos revisados utilizaram métricas diversas, como RMSE, AUC, f1-score e índices de acurácia espacial, dificultando comparações diretas. Zhang et al. (2020) defenderam protocolos de validação cruzada geoespacial, enquanto Gryech et al. (2024) recomendaram a combinação de validação empírica em campo e validação cruzada.

A necessidade de benchmarks públicos e bases de dados padronizadas foi apontada como essencial para acelerar a adoção das soluções e permitir comparações mais justas entre modelos. Esse esforço é visto como condição básica para garantir robustez e confiabilidade nas aplicações de IA em monitoramento ambiental.

Outra dimensão destacada foi a sustentabilidade computacional das soluções. Rolnick et al. (2019) e Olawade et al. (2024) observaram que modelos de deep learning, embora precisos, demandam alto consumo energético e infraestrutura robusta, limitando sua replicabilidade em regiões de baixa disponibilidade tecnológica.

Como alternativa, a literatura recomendou estratégias de compressão de modelos, uso de aprendizado por transferência e algoritmos leves capazes de operar em dispositivos de borda (Enviro-IoT Project, 2025). Essas práticas buscam equilibrar custo e desempenho sem comprometer a precisão das análises.

Além das lacunas técnicas identificadas, observou-se que a Inteligência Artificial também tem se consolidado como ferramenta estratégica na antecipação de impactos climáticos de longo prazo. De acordo com Rolnick et al. (2019), modelos de aprendizado profundo vêm sendo utilizados para simular cenários de aquecimento global e prever o comportamento de variáveis ambientais complexas, como o aumento do nível do mar e a frequência de eventos extremos. O UNEP (2022) complementa que essas tecnologias são essenciais para subsidiar relatórios globais de risco e apoiar decisões sobre mitigação e adaptação climática em escala internacional. Esse tipo de aplicação demonstra que o papel da IA transcende o monitoramento pontual e passa a contribuir para a formulação de políticas sustentáveis baseadas em evidências.

Outro ponto relevante é a integração entre IA e políticas públicas ambientais, o que reforça a importância de uma governança tecnológica inclusiva. Batista e Lopes (2021) destacam que a utilização responsável da IA depende de marcos éticos claros e de mecanismos institucionais que garantam transparência e accountability. Em consonância, Ribeiro et al. (2021) observam que sistemas de monitoramento automatizados podem fortalecer políticas locais de fiscalização e planejamento territorial, desde que acompanhados de capacitação técnica e controle social. Essa sinergia entre ciência de dados e gestão pública amplia a efetividade das ações ambientais e reduz o hiato entre inovação e implementação prática.

A literatura recente também enfatiza os desafios de padronização e interoperabilidade global. Gryech et al. (2024) ressaltam que diferentes países ainda utilizam formatos de dados e protocolos técnicos incompatíveis, o que dificulta a consolidação de bancos de dados internacionais. Olawade et al. (2024) acrescentam que a ausência de normas unificadas compromete a comparabilidade entre modelos e reduz a capacidade de replicar boas práticas. Assim, iniciativas como as propostas pelo UNEP (2022) buscam fomentar padrões abertos e práticas colaborativas, visando consolidar uma infraestrutura global de monitoramento ambiental inteligente.

Finalmente, reforça-se que a ampliação da cobertura geográfica e a consolidação de parcerias institucionais são fatores determinantes para o sucesso dessas tecnologias. Moreira e Campos (2020) argumentam que a integração entre comunidades locais, universidades e órgãos governamentais é o que garante a continuidade operacional dos sistemas e evita a obsolescência tecnológica. Portanto, ao promover um ecossistema colaborativo e ético, a IA tem o potencial de se tornar não apenas uma ferramenta de análise, mas um pilar estratégico na governança ambiental contemporânea.

Por fim, a revisão mostrou que projetos bem-sucedidos de monitoramento ambiental combinaram tecnologia, arranjos institucionais e participação local. Ribeiro et al. (2021) e Moreira e Campos (2020) observaram que saberes locais e manutenção adequada dos sensores foram fundamentais para garantir a sustentabilidade das iniciativas.

Esse alinhamento multifacetado foi decisivo para que soluções de IA gerassem benefícios concretos e duradouros. Iniciativas isoladas, sem governança e suporte institucional, mostraram dificuldades de continuidade e menor impacto em longo prazo (Moreira; Campos, 2020; Ribeiro et al., 2021).

3 CONCLUSÃO

A realização deste estudo permitiu compreender como a Inteligência Artificial tem se consolidado como uma ferramenta essencial para o monitoramento ambiental, promovendo eficiência, precisão e capacidade preditiva na análise de fenômenos ecológicos.

Entretanto, observou-se que ainda existem lacunas de conhecimento relevantes, especialmente relacionadas à integração de bases de dados heterogêneas, à padronização de metodologias de validação, à transparência dos modelos e ao alto custo computacional de algumas soluções. Essas limitações indicam a necessidade de pesquisas futuras voltadas à criação de sistemas mais acessíveis, éticos e sustentáveis.

Dessa forma, os resultados reforçam que a aplicação da IA no monitoramento ambiental é promissora, mas demanda esforços contínuos de aprimoramento técnico e institucional para garantir sua eficácia e equidade em contextos diversos.

Os objetivos propostos foram alcançados, uma vez que foi possível compreender

como a Inteligência Artificial vem sendo aplicada no monitoramento ambiental, identificar as principais tecnologias envolvidas, analisar seus benefícios e desafios e avaliar sua contribuição para práticas mais sustentáveis. Assim, foi possível reconhecer o uso de redes neurais, sensores inteligentes e sistemas de previsão automatizada como recursos de destaque, além de demonstrar a relevância dessas ferramentas na mitigação de impactos ambientais e no apoio à sustentabilidade global.

A pesquisa respondeu ao problema inicialmente proposto — que consistiu em compreender de que forma a Inteligência Artificial pode ser aplicada no monitoramento ambiental para aprimorar a análise e a tomada de decisões sustentáveis — ao evidenciar que a adoção de soluções baseadas em Inteligência Artificial favorece o controle e a interpretação de grandes volumes de dados ambientais, permitindo decisões mais rápidas e embasadas por órgãos públicos e instituições de pesquisa. Apesar dos avanços observados, constatou-se que desafios ainda persistem, especialmente relacionados à limitação de dados em tempo real, à necessidade de infraestrutura tecnológica adequada e à capacitação de profissionais especializados para operar tais sistemas.

De modo geral, o estudo concluiu que a Inteligência Artificial tem se mostrado uma ferramenta promissora no monitoramento ambiental, permitindo maior precisão na coleta e análise de dados, além de contribuir para a tomada de decisões mais eficientes em políticas de sustentabilidade. Observou-se que a integração entre IA e tecnologias emergentes, como IoT e sensores inteligentes, potencializa o acompanhamento de fenômenos ambientais e favorece respostas rápidas frente a situações críticas.

Como proposta para estudos futuros, recomenda-se o aprofundamento das análises sobre a integração entre Inteligência Artificial, Internet das Coisas e aprendizado profundo, com foco em aplicações voltadas à previsão de eventos climáticos extremos e ao monitoramento da qualidade do ar e da água. Além disso, sugere-se investigar a viabilidade de políticas públicas que incentivem a adoção de tecnologias sustentáveis, ampliando o uso da Inteligência Artificial como aliada na preservação ambiental e no desenvolvimento sustentável.

REFERÊNCIAS

BATISTA, T. G.; LOPES, J. M. **Desafios éticos na aplicação da inteligência artificial no meio ambiente**. Revista Brasileira de Ciência, Tecnologia e Inovação, Brasília, v. 13, n. 2, p. 45–58, 2021. Disponível em: https://www.rbcti.gov.br. Acesso em: 9 abr. 2025.

GARCIA, A. et al. **Advancements in air quality monitoring: a systematic review of IoT-based air quality monitoring and AI technologies.** Artificial Intelligence Review, v. 58, 2025. DOI: 10.1007/s10462-025-11277-9. Disponível em: https://link.springer.com/article/10.1007/s10462-025-11277-9. Acesso em: 29 jul. 2025.

GRYECH, I. et al. **Applications of machine learning & Internet of Things for outdoor air pollution monitoring and prediction: A systematic literature review.** Engineering Applications of Artificial Intelligence, v. 137, 2024. Disponível em: https://arxiv.org/abs/2401.01788. Acesso em: 2 ago. 2025.

MOREIRA, M. A.; CAMPOS, F. L. **Introdução ao monitoramento ambiental.** São Paulo: Oficina de Textos, 2020.

OLAWADE, D. B. et al. **Artificial Intelligence in Environmental Monitoring: Advancements, Challenges, and Future Directions.** Hygiene and Environmental Health Advances, UK, v. 12, 2024. Disponível em: https://www.sciencedirect.com/science/article/pii/S2773049224000278. Acesso em: 15 jul. 2025.

RIBEIRO, C. D. et al. **Integração de sensores ambientais com inteligência artificial: avanços e perspectivas.** Revista Tecnologia & Sociedade, Curitiba, v. 8, n. 1, p. 88–103, 2021. Disponível em: https://periodicos.utfpr.edu.br/rts. Acesso em: 8 abr. 2025.

ROLNICK, D. et al. **Tackling climate change with machine learning.** arXiv, [S.l.], 2019. Disponível em: https://arxiv.org/abs/1906.05433. Acesso em: 5 mar. 2025.

RUSSELL, S.; NORVIG, P. **Inteligência artificial**. 3. ed. São Paulo: Pearson Education do Brasil, 2016.

UNEP (United Nations Environment Programme). **Frontiers 2022: Noise, Blazes and Mismatches – Emerging Issues of Environmental Concern**. Nairobi: UNEP, 2022. Disponível em: https://www.unep.org/resources/frontiers-2022. Acesso em: 9 set. 2025.

ZHANG, C. et al. **Deep learning in remote sensing applications: A meta-review.** ISPRS Journal of Photogrammetry and Remote Sensing, [S.l.], v. 160, p. 296–312, 2020. Disponível em: https://doi.org/10.1016/j.isprsjprs.2020.01.001. Acesso em: 12 abr. 2025.

3

INTELIGÊNCIA ARTIFICIAL NA SEGURANÇA CIBERNÉTICA: UMA REVISÃO DE LITERATURA

ARTIFICIAL INTELLIGENCE IN CYBERSECURITY: A LITERATURE REVIEW

Ivanderson França Santos
Ivone Ascar Sauaia Guimarães

Resumo

O Este trabalho apresenta uma revisão de literatura sobre a aplicação da inteligência artificial (IA) na segurança cibernética, destacando suas tecnologias, benefícios, limitações e implicações éticas. A pesquisa foi conduzida por meio de fontes acadêmicas e científicas dos últimos cinco anos, com foco em soluções baseadas em aprendizado de máquina, análise comportamental, sistemas de detecção de intrusão e resposta automatizada a incidentes. Os resultados indicam que a IA contribui significativamente para a detecção proativa de ameaças, automação de tarefas e aumento da eficiência operacional. No entanto, foram identificados desafios técnicos, como falta de explicabilidade dos modelos e risco de viés algorítmico, além de questões éticas e legais relacionadas à privacidade, responsabilidade e transparência. O estudo concluiu que, embora a IA represente um avanço promissor na proteção digital, sua aplicação deve ser acompanhada de governança ética, regulamentações claras e supervisão humana, garantindo que os sistemas inteligentes operem de forma justa, segura e responsável.

Palavras-chave: Ética, Inteligência Artificial, Segurança Cibernética, Tecnologia.

Abstract

This paper presents a literature review on the application of artificial intelligence (AI) in cybersecurity, highlighting its technologies, benefits, limitations, and ethical implications. The research was conducted using academic and scientific sources from the last five years, focusing on solutions based on machine learning, behavioral analysis, intrusion detection systems, and automated incident response. The results indicate that AI significantly contributes to proactive threat detection, task automation, and increased operational efficiency. However, technical challenges were identified, such as the lack of model explainability and the risk of algorithmic bias, in addition to ethical and legal issues related to privacy, accountability, and transparency. The study concluded that, although AI represents a promising advancement in digital protection, its application must be accompanied by ethical governance, clear regulations, and human oversight, ensuring that intelligent systems operate fairly, securely, and responsibly.

Keywords: Ethics, Artificial Intelligence, Cybersecurity, Technology.

1 INTRODUÇÃO

A transformação digital tem redefinido a forma como indivíduos, empresas e governos operam, tornando os sistemas tecnológicos indispensáveis para a vida cotidiana. Com essa crescente dependência de ambientes digitais, a segurança cibernética passou a ocupar um papel central na proteção de dados, infraestruturas críticas e operações estratégicas. A complexidade dos ataques virtuais, que evoluem em velocidade e sofisticação, exige soluções que vão além das abordagens convencionais. Nesse cenário, a Inteligência Artificial (IA) surge como uma alternativa promissora para enfrentar os desafios da proteção digital.

A aplicação da IA na segurança cibernética representa uma mudança de paradigma. Sistemas inteligentes, capazes de aprender com dados e identificar padrões incomuns, têm potencial para detectar ameaças antes que elas se concretizem. Ferramentas como aprendizado de máquina, redes neurais e análise preditiva permitem que os mecanismos de defesa atuem de forma proativa, antecipando riscos e automatizando respostas. Essa capacidade de adaptação e agilidade torna a IA uma aliada estratégica na construção de ambientes digitais mais seguros e resilientes.

Apesar dos avanços, o uso da IA em segurança cibernética não está isento de controvérsias. A mesma tecnologia que fortalece as defesas pode ser explorada por agentes maliciosos para desenvolver ataques mais sofisticados e difíceis de rastrear. Além disso, surgem preocupações éticas e legais sobre a privacidade dos dados, a transparência das decisões automatizadas e a responsabilidade por ações tomadas por sistemas autônomos. Esses dilemas exigem uma análise crítica sobre os limites e implicações do uso da IA em contextos de segurança.

Diante desse panorama, este estudo buscou responder à seguinte questão: como a inteligência artificial pode contribuir para o aprimoramento da segurança cibernética, considerando suas capacidades técnicas, limitações operacionais e os desafios éticos envolvidos? A investigação dessa problemática é essencial para compreender o papel da IA na proteção digital e para orientar práticas mais conscientes e eficazes em sua aplicação.

A relevância da pesquisa está na necessidade urgente de fortalecer os mecanismos de defesa digital frente ao aumento das ameaças cibernéticas. Ao explorar o potencial da IA, este estudo pretende oferecer subsídios para a construção de estratégias mais robustas e inteligentes de proteção. Além disso, ao abordar os riscos e dilemas éticos, contribui para o desenvolvimento de diretrizes que promovam o uso responsável da tecnologia, equilibrando inovação e segurança.

O objetivo geral deste artigo foi analisar como a inteligência artificial pode ser utilizada para melhorar a segurança cibernética, destacando suas aplicações práticas, benefícios e desafios. Os objetivos específicos incluem: investigar as principais tecnologias de IA aplicadas à segurança digital; avaliar os obstáculos técnicos e os riscos associados ao uso indevido da IA; e discutir as implicações éticas e legais da automação em ambientes de proteção cibernética.

2 DESENVOLVIMENTO

2.1 Metodologia

A pesquisa desenvolvida neste trabalho foi do tipo Revisão Bibliográfica, com enfoque qualitativo e descritivo. Essa abordagem tem como objetivo reunir, analisar e sintetizar o

conhecimento científico já produzido sobre a aplicação da inteligência artificial na segurança cibernética, sem a realização de experimentos ou coleta de dados primários. A revisão bibliográfica permite compreender o estado atual da temática, identificar lacunas e apontar tendências emergentes, contribuindo para uma reflexão crítica e fundamentada sobre o uso da IA em ambientes digitais de proteção.

Para a seleção dos materiais, foram consultadas obras acadêmicas disponíveis em bases de dados reconhecidas, como Google Scholar, SciELO, IEEE Xplore e SpringerLink, além de sites especializados em segurança da informação e inteligência artificial. A busca foi restrita a publicações dos últimos cinco anos, com o intuito de garantir a atualidade das informações e contemplar os avanços mais recentes na área. Foram priorizados estudos que apresentassem rigor metodológico, relevância temática e evidências consistentes sobre os impactos da IA na detecção, prevenção e resposta a ameaças cibernéticas. Foram identificados 28 materiais relacionados ao tema, dos quais 16 foram selecionados para leitura e análise detalhada.

A escolha dos textos seguiu critérios de inclusão que consideraram publicações em português e inglês, com foco direto na aplicação da inteligência artificial em contextos de segurança digital. Foram excluídos artigos de revisão, resumos, trabalhos não originais e publicações sem fundamentação teórica sólida. A análise dos materiais buscou identificar diferentes perspectivas sobre o tema, bem como os desafios técnicos, éticos e legais envolvidos na implementação da IA em sistemas de proteção cibernética.

As palavras-chave utilizadas nas buscas incluíram termos como: “inteligência artificial e segurança cibernética”, “machine learning em defesa digital”, “IA na cibersegurança”, “desafios éticos da IA em segurança” e “privacidade e inteligência artificial”. Esses descritores foram combinados de forma estratégica para ampliar o alcance da pesquisa e garantir a identificação de estudos relevantes em diferentes contextos e abordagens.

O processo de revisão permitiu uma visão abrangente e aprofundada sobre o papel da IA na segurança cibernética, destacando suas aplicações práticas, limitações operacionais e implicações éticas. A análise dos textos selecionados foi conduzida de forma sistemática, buscando compreender como a tecnologia tem sido utilizada para fortalecer as defesas digitais e quais são os principais obstáculos enfrentados pelas organizações na adoção dessas soluções.

Dessa forma, a metodologia adotada neste trabalho oferece uma base teórica sólida para a discussão dos impactos da inteligência artificial na segurança cibernética, contribuindo para o desenvolvimento de estratégias mais eficazes e responsáveis no uso dessa tecnologia.

2.2 Resultados

A aplicação da inteligência artificial (IA) na segurança cibernética tem revolucionado a forma como organizações enfrentam ameaças digitais. Entre as tecnologias mais utilizadas, destacam-se o aprendizado de máquina (machine learning), a análise comportamental, os sistemas de detecção de intrusão baseados em IA e os mecanismos de resposta automatizada. Essas ferramentas têm se mostrado eficazes na identificação de padrões suspeitos, antecipação de ataques e mitigação de riscos em tempo real (FORTINET, 2025).

O aprendizado de máquina é uma das abordagens mais promissoras. Ele permite que sistemas de segurança aprendam com dados históricos e adaptem suas respostas a novas ameaças. Algoritmos supervisionados e não supervisionados são utilizados para detectar

anomalias no tráfego de rede, reconhecer malwares e prever vulnerabilidades antes que sejam exploradas (HACKONE, 2024). Essa capacidade de adaptação contínua é essencial em um cenário onde os ataques evoluem rapidamente.

A análise comportamental é outra tecnologia amplamente empregada. Por meio da observação de padrões de uso, a IA consegue identificar desvios que podem indicar atividades maliciosas, como acesso não autorizado ou movimentações incomuns de dados. Essa abordagem é especialmente útil na detecção de ameaças internas, que muitas vezes passam despercebidas por sistemas tradicionais (MONFRE; SILVA; VICENTIN, 2023).

Sistemas de detecção de intrusão baseados em IA (IDS inteligentes) analisam o tráfego de rede em tempo real, buscando sinais de comprometimento. Diferente dos métodos baseados em assinaturas, que dependem de atualizações constantes, os IDS inteligentes conseguem identificar ataques desconhecidos, como zero-day exploits, com maior precisão (PONTES; VASCONCELOS; SILVA, 2023). Isso amplia significativamente a capacidade de defesa das organizações.

A resposta automatizada a incidentes é uma funcionalidade que tem ganhado destaque. Quando uma ameaça é detectada, sistemas alimentados por IA podem isolar dispositivos comprometidos, bloquear endereços IP suspeitos e iniciar protocolos de contenção sem intervenção humana. Essa agilidade reduz o tempo de resposta e limita os danos causados por ataques (FORTINET, 2025).

Ferramentas como o IBM QRadar, Darktrace e CrowdStrike Falcon exemplificam o uso prático da IA na segurança cibernética. Essas plataformas combinam análise preditiva, automação e inteligência de ameaças para oferecer proteção avançada. O QRadar, por exemplo, utiliza machine learning para correlacionar eventos e identificar riscos emergentes (HACKONE, 2024).

A tecnologia de User and Entity Behavior Analytics (UEBA) também tem sido amplamente adotada. Ela permite que sistemas de segurança monitorem o comportamento de usuários e dispositivos, detectando atividades incomuns que possam indicar comprometimento. Essa abordagem é eficaz na prevenção de ataques de engenharia social e na proteção contra fraudes (MALWAREBYTES, 2024).

A IA também tem sido aplicada na detecção de phishing, uma das ameaças mais comuns. Algoritmos treinados conseguem identificar padrões linguísticos e estruturais em e-mails e sites fraudulentos, alertando os usuários antes que interajam com conteúdos maliciosos. Essa tecnologia tem reduzido significativamente os casos de roubo de credenciais (PONTES; VASCONCELOS; SILVA, 2023).

Outra aplicação relevante é o gerenciamento automatizado de vulnerabilidades. Sistemas baseados em IA avaliam continuamente a infraestrutura digital, identificando falhas e sugerindo correções. Essa abordagem proativa evita que brechas sejam exploradas por cibercriminosos, fortalecendo a postura de segurança das organizações (HACKONE, 2024).

A integração da IA com tecnologias emergentes, como blockchain e computação quântica, também tem sido explorada. Essa combinação permite a criação de sistemas mais robustos e resilientes, capazes de resistir a ataques sofisticados. O uso de IA para validar transações e monitorar redes distribuídas é um exemplo dessa sinergia (MONFRE; SILVA; VICENTIN, 2023).

A automação de tarefas rotineiras é outro benefício da IA na segurança cibernética. Atividades como análise de logs, verificação de conformidade e geração de relatórios são realizadas por algoritmos, liberando os profissionais de segurança para funções estratégicas. Isso aumenta a eficiência operacional e reduz o risco de erro humano (FORTINET,

2025).

A IA também contribui para a melhoria da autenticação de usuários. Sistemas que analisam impressões digitais, padrões de digitação e voz conseguem validar identidades com maior precisão, dificultando o acesso indevido. Essa abordagem é especialmente útil em ambientes críticos, como saúde e finanças (HACKONE, 2024).

A eficácia da IA na detecção de ameaças tem sido comprovada por estudos recentes. Relatórios da indústria, como o da Fortinet (2025), indicam que sistemas inteligentes têm demonstrado uma alta taxa de precisão na identificação de ataques, superando em muitos casos os métodos tradicionais. Essa capacidade reforça a importância da IA como ferramenta de defesa.

No entanto, é importante destacar que a eficácia da IA depende da qualidade dos dados utilizados no treinamento dos algoritmos. Dados enviesados podem comprometer os resultados, gerando falsos positivos ou negligenciando ameaças reais. Por isso, é essencial que as organizações adotem práticas de governança de dados (HACKONE, 2024).

Além disso, a IA deve ser integrada a uma estratégia ampla de segurança, que inclua políticas claras, capacitação de profissionais e colaboração entre setores. A tecnologia, por si só, não é suficiente para garantir proteção. É necessário um ecossistema que favoreça sua aplicação ética e eficaz (MONFRE; SILVA; VICENTIN, 2023).

Em síntese, as tecnologias de IA têm se mostrado altamente eficazes na detecção e prevenção de ameaças cibernéticas. Elas oferecem agilidade, precisão e capacidade de adaptação, elementos essenciais para enfrentar os desafios da segurança digital contemporânea. Com o avanço contínuo dessas ferramentas, espera-se que a IA desempenhe um papel cada vez mais central na proteção de dados e sistemas.

Apesar dos avanços significativos da inteligência artificial (IA) na segurança cibernética, sua implementação enfrenta uma série de desafios técnicos, operacionais e éticos. Um dos principais obstáculos é a complexidade dos modelos utilizados, que muitas vezes operam como “caixas-pretas”, dificultando a interpretação das decisões tomadas pelos sistemas automatizados (BAVARESCO; WEBBER, 2024). Essa falta de transparência pode comprometer a confiança dos usuários e limitar a adoção da tecnologia em ambientes críticos.

A IA explicável (XAI) tem sido proposta como solução para esse problema, permitindo que os sistemas justifiquem suas ações de forma compreensível para humanos. No entanto, ainda há limitações na aplicação prática dessas abordagens, especialmente em cenários de alta complexidade, como ataques em tempo real (BAVARESCO; WEBBER, 2024). A ausência de explicações claras pode gerar resistência por parte de gestores e operadores de segurança, que precisam entender os critérios utilizados para bloquear acessos ou sinalizar ameaças.

Outro desafio relevante é o viés algorítmico. Quando os dados utilizados para treinar os modelos de IA são incompletos ou tendenciosos, os sistemas podem reproduzir essas distorções, afetando negativamente a eficácia das defesas digitais **(**HEGGLER; SZMOSKI; MIQUELIN, 2025). Isso é especialmente preocupante em ambientes corporativos, onde decisões automatizadas podem impactar diretamente a reputação e a operação das empresas.

Além disso, a dependência excessiva da IA pode reduzir a intervenção humana em momentos críticos. Embora a automação seja desejável para agilizar respostas, ela também pode comprometer a tomada de decisões estratégicas, especialmente em situações que exigem julgamento contextual ou sensibilidade ética (SAMPAIO; SABBATINI; LIMON-

GI, 2024). A combinação entre inteligência artificial e supervisão humana continua sendo essencial para garantir segurança e responsabilidade.

A evolução dos ataques cibernéticos também representa um desafio crescente. Cibercriminosos têm utilizado IA para desenvolver malwares adaptativos, deepfakes e campanhas de phishing altamente personalizadas, dificultando a detecção por sistemas convencionais (MALWAREBYTES, 2024). Essa "corrida armamentista digital" exige que as defesas baseadas em IA evoluam continuamente para acompanhar as novas táticas de ataque.

O uso indevido da IA por agentes mal-intencionados é uma preocupação crítica. Ferramentas generativas podem ser exploradas para criar códigos maliciosos, manipular saídas de sistemas ou enganar usuários com conteúdos falsificados (DATACAMP, 2024). A facilidade de acesso a modelos avançados torna essas ameaças mais acessíveis, ampliando a superfície de ataque e exigindo novas estratégias de contenção.

A segurança dos próprios sistemas de IA também precisa ser considerada. Algoritmos vulneráveis podem ser manipulados por atacantes para gerar respostas incorretas ou ignorar ameaças reais. A proteção da infraestrutura que sustenta a IA é tão importante quanto a defesa dos dados que ela analisa (DATACAMP, 2024). Sem medidas robustas de segurança, a IA pode se tornar um vetor de ataque em vez de uma barreira.

A governança da IA é outro ponto sensível. Muitas organizações ainda não possuem diretrizes claras sobre o uso ético e seguro da tecnologia, o que pode levar à implementação inadequada e ao surgimento de riscos legais e operacionais (ASSOCIAÇÃO..., 2024). A criação de políticas internas de compliance e auditoria é fundamental para garantir que a IA seja utilizada de forma responsável.

A privacidade dos dados é uma das principais preocupações associadas à IA na segurança cibernética. A coleta e o processamento de grandes volumes de informações podem violar direitos fundamentais se não forem acompanhados de práticas rigorosas de anonimização e consentimento (SENADO APROVA..., 2024). O cumprimento da Lei Geral de Proteção de Dados (LGPD) e de regulamentações internacionais é indispensável nesse contexto.

A falta de capacitação técnica também limita a eficácia da IA. Profissionais de segurança precisam entender como os algoritmos funcionam, como interpretar suas saídas e como ajustar os parâmetros para melhorar a performance. Sem essa qualificação, há risco de uso inadequado ou subutilização das ferramentas disponíveis (SAMPAIO; SABBATINI; LIMONGI, 2024).

A interoperabilidade entre sistemas de IA e outras tecnologias de segurança é outro desafio. Muitas soluções operam de forma isolada, dificultando a integração e a troca de informações entre plataformas. A criação de padrões técnicos e protocolos de comunicação pode facilitar essa integração e ampliar a eficácia das defesas (ASSOCIAÇÃO..., 2024).

A escalabilidade das soluções de IA também deve ser considerada. Sistemas que funcionam bem em ambientes controlados podem apresentar falhas quando aplicados em larga escala, especialmente em redes distribuídas ou com alto volume de tráfego. Testes rigorosos e simulações são necessários para validar a robustez das ferramentas (PONTES; VASCONCELOS; SILVA, 2023).

A atualização constante dos modelos é essencial para manter a eficácia da IA. Novas ameaças exigem que os algoritmos sejam treinados com dados recentes e relevantes. A falta de atualização pode tornar os sistemas obsoletos e vulneráveis, comprometendo a segurança da organização (DATACAMP, 2024).

A colaboração entre setores é fundamental para enfrentar os desafios da IA na segurança cibernética. Equipes técnicas, jurídicas e de governança devem atuar de forma integrada para garantir que as soluções adotadas estejam alinhadas com os objetivos estratégicos e com as normas vigentes (SAMPAIO; SABBATINI; LIMONGI, 2024). Essa abordagem multidisciplinar fortalece a resiliência organizacional.

Dessa forma, é necessário promover uma cultura de segurança que valorize a ética, a transparência e a responsabilidade no uso da IA. A tecnologia deve ser vista como uma aliada, mas também como uma ferramenta que exige vigilância constante. A construção de ambientes digitais seguros depende da combinação entre inovação e prudência, entre automação e supervisão humana.

A crescente incorporação da inteligência artificial (IA) em sistemas de segurança cibernética tem gerado debates relevantes sobre os limites éticos e legais dessa tecnologia. Embora a IA ofereça avanços significativos na proteção de dados e na resposta a ameaças digitais, sua aplicação levanta preocupações sobre privacidade, transparência, responsabilidade e justiça. Esses aspectos tornam-se ainda mais críticos quando decisões automatizadas afetam diretamente indivíduos e organizações (LAMB, 2024).

A privacidade dos dados é uma das principais questões éticas envolvidas. Sistemas de IA operam com grandes volumes de informações, muitas vezes sensíveis, para identificar padrões e antecipar riscos. Sem políticas claras de consentimento e anonimização, há risco de violação de direitos fundamentais, especialmente em ambientes corporativos e governamentais (RODOLFO, 2025). A proteção da privacidade deve ser um princípio estruturante na concepção de soluções baseadas em IA.

Outro ponto de tensão é a responsabilidade pelas decisões automatizadas. Em sistemas de segurança, algoritmos podem bloquear acessos, sinalizar comportamentos suspeitos ou isolar dispositivos. Quando essas ações afetam usuários legítimos ou causam prejuízos, surge a dúvida sobre quem deve ser responsabilizado: o desenvolvedor, o operador ou o próprio sistema? A ausência de marcos legais específicos dificulta a atribuição de culpa e a reparação de danos (SILVA, 2023).

A transparência dos algoritmos é essencial para garantir confiança e auditabilidade. Muitos modelos de IA funcionam como “caixas-pretas”, dificultando a compreensão das decisões tomadas. Essa opacidade compromete a supervisão e a responsabilização, especialmente em contextos críticos como segurança nacional ou proteção de infraestruturas essenciais. A explicabilidade da IA — ou seja, sua capacidade de justificar decisões — é um requisito ético fundamental (LAMB, 2024).

O viés algorítmico também representa um risco ético significativo. Dados utilizados para treinar modelos de IA podem refletir desigualdades sociais, culturais ou econômicas, levando a decisões discriminatórias ou injustas. Na segurança cibernética, isso pode significar bloqueios indevidos, priorização incorreta de ameaças ou exclusão de grupos vulneráveis. A equidade deve ser um valor central na construção de sistemas inteligentes (RODOLFO, 2025).

Do ponto de vista legal, a regulamentação da IA ainda é incipiente em muitos países, incluindo o Brasil. A Lei Geral de Proteção de Dados (LGPD) estabelece diretrizes importantes, mas não aborda de forma específica os desafios da IA em segurança cibernética. A ausência de normas detalhadas sobre responsabilidade, transparência e uso ético da IA cria um cenário de insegurança jurídica para empresas e usuários (SILVA, 2023).

A governança da IA exige a criação de políticas internas que orientem seu uso de forma ética e legal. Empresas devem estabelecer comitês de ética, protocolos de auditoria

e mecanismos de supervisão para garantir que os sistemas operem dentro dos limites aceitáveis. A autorregulação pode ser um caminho viável enquanto não há legislação específica, mas deve ser acompanhada por fiscalização externa e participação da sociedade civil (INTELIGÊNCIA..., 2024).

A formação ética dos profissionais envolvidos no desenvolvimento e operação de sistemas de IA é igualmente essencial. A capacitação técnica deve ser acompanhada por disciplinas que abordem filosofia, direitos humanos e ética aplicada, promovendo uma visão mais ampla e responsável da tecnologia. A educação ética contribui para decisões mais conscientes e alinhadas com os valores sociais (SICHMAN, 2021).

O impacto social da IA na segurança cibernética também merece atenção. A automação de processos pode gerar desemprego, marginalização de grupos vulneráveis e ampliação de desigualdades. É necessário refletir sobre os efeitos da tecnologia na sociedade e buscar soluções que promovam inclusão e justiça, evitando que a IA se torne um instrumento de exclusão (INTELIGÊNCIA..., 2024).

A cooperação internacional é fundamental para enfrentar os desafios éticos e legais da IA. A tecnologia transcende fronteiras e exige esforços coordenados entre países para estabelecer padrões comuns. Organizações multilaterais têm promovido debates sobre governança global da IA, buscando harmonizar princípios e práticas que garantam segurança e respeito aos direitos humanos (SICHMAN, 2021).

Casos recentes de uso indevido da IA em segurança digital reforçam a necessidade de regulamentação. Deepfakes, manipulação de dados e ataques automatizados são exemplos de como a tecnologia pode ser explorada de forma maliciosa. A legislação deve prever sanções para abusos e mecanismos de prevenção que protejam os usuários e a integridade dos sistemas (SILVA, 2023).

A participação da sociedade civil é essencial na construção de uma IA ética. Usuários, pesquisadores e organizações não governamentais devem ser incluídos nos processos de decisão, garantindo pluralidade de perspectivas e defesa dos direitos fundamentais. A transparência nos processos de desenvolvimento e implementação é chave para essa inclusão (RODOLFO, 2025).

A ética na IA não se limita à intenção dos desenvolvedores, mas envolve todo o ciclo de vida da tecnologia: desde a coleta de dados até a tomada de decisões. Cada etapa deve ser guiada por princípios claros, como beneficência, justiça, responsabilidade e respeito à dignidade humana (LAMB, 2024).

A criação de certificações éticas para sistemas de IA pode ser uma alternativa para promover boas práticas. Essas certificações avaliariam critérios como transparência, segurança, privacidade e impacto social, oferecendo aos usuários uma garantia mínima de conformidade e estimulando a concorrência ética entre fornecedores (INTELIGÊNCIA...., 2024).

Portanto, é necessário reconhecer que a ética na IA é um processo contínuo. À medida que a tecnologia evolui, novos dilemas surgem, exigindo atualização constante das normas, práticas e valores. A construção de uma IA segura e justa depende do compromisso coletivo com a responsabilidade, a transparência e o respeito aos direitos humanos (SICHMAN, 2021).

3 CONCLUSÃO

A presente revisão de literatura teve como objetivo analisar como a inteligência artificial (IA) pode ser utilizada para aprimorar a segurança cibernética, considerando suas aplicações práticas, benefícios, limitações e implicações éticas. Ao longo do estudo, foi possível observar que os objetivos propostos foram amplamente atendidos, com a identificação das principais tecnologias de IA empregadas na proteção digital, como aprendizado de máquina, análise comportamental, sistemas de detecção de intrusão e resposta automatizada a incidentes.

A investigação permitiu responder à questão central do trabalho, demonstrando que a IA contribui significativamente para a detecção proativa de ameaças, automação de tarefas e aumento da eficiência operacional. No entanto, também foram evidenciadas limitações importantes, como a falta de explicabilidade dos modelos, o risco de viés algorítmico e os desafios relacionados à governança ética e à privacidade dos dados. Esses aspectos revelam que, embora a IA represente um avanço promissor, sua aplicação exige cautela, supervisão humana e regulamentações claras.

Entre as dificuldades enfrentadas durante a pesquisa, destaca-se a escassez de estudos que abordem de forma integrada os aspectos técnicos e éticos da IA na segurança cibernética. A fragmentação das abordagens limita a compreensão holística do tema e reforça a necessidade de investigações multidisciplinares. Além disso, a rápida evolução tecnológica impõe desafios à atualização constante dos modelos e à capacitação dos profissionais envolvidos.

Como recomendação, sugere-se que futuros trabalhos aprofundem a análise sobre a IA explicável (XAI), explorando soluções que tornem os sistemas mais transparentes e compreensíveis. Também é relevante investigar estratégias de mitigação de viés algorítmico e formas de integração entre IA e políticas de compliance. A colaboração entre áreas técnicas, jurídicas e éticas pode fortalecer a governança da IA e promover sua aplicação responsável.

Por fim, este estudo contribui para o avanço do conhecimento ao oferecer uma visão crítica e atualizada sobre o papel da inteligência artificial na segurança cibernética. Ao destacar tanto os benefícios quanto os riscos envolvidos, promove uma reflexão necessária sobre o uso consciente da tecnologia, reforçando a importância de equilibrar inovação e responsabilidade na construção de ambientes digitais mais seguros e justos.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO/IEC 42001:2023: Tecnologia da informação — Inteligência artificial — Sistema de gestão**. Rio de Janeiro: ABNT, 2024.

BAVARESCO, Maurício Zalamena; WEBBER, Carine Geltrudes. IA explicável para reduzir a assimetria de informação no consumo. **Revista de Informática Aplicada (Redin), v. 20, n. 2, 2024**. Disponível em: https://seer.faccat.br/index.php/redin/article/view/3538. Acesso em: 10 jul. 2025.

DATACAMP. O que é machine learning adversarial? Tipos de ataques e defesas. **DataCamp Blog**, 29 jul. 2024. Disponível em: https://www.datacamp.com/pt/blog/adversarial-machine-learning. Acesso em: 12 jul. 2025.

FORTINET. **2025 Global Threat Landscape Report**. Sunnyvale: FortiGuard Labs, 2025. Disponível em: https://www.fortinet.com/resources/reports/threat-landscape-report. Acesso em: 15 jul. 2025.

HACKONE. Machine Learning e IA na Segurança Cibernética: Integração no auxílio à detecção e prevenção de ameaças cibernéticas. **Blog - Hackone**, 28 maio 2024. Disponível em: https://blog.hackone.com.br/2024/05/28/machine-learning-e-ia-na-seguranca-cibernetica-integracao-no-auxilio-a-deteccao-e-prevencao-de-ameacas-cibernetica/. Acesso em: 20 jul. 2025.

HEGGLER, J. M.; SZMOSKI, R. M.; MIQUELIN, A. F. As dualidades entre o uso da Inteligência Artificial na educação e os riscos de vieses algorítmicos. **Educação & Sociedade**, v. 46, e289323, 2025. Disponível em: https://doi.org/10.1590/ES.289323. Acesso em: 25 jul. 2025.

INTELIGÊNCIA artificial: desafios éticos e oportunidades. **Atlas Governance**, 23 ago. 2024. Disponível em: https://welcome.atlasgov.com/blog/mercado/inteligencia-artificial-desafios-eticos-e-oportunidades. Acesso em: 02 ago. 2025.

LAMB, Luís C. Inteligência artificial, ética e o futuro. **Jornal da USP,** 2024. Disponível em: https://jornal.usp.br/wp-content/uploads/2024/05/8-Luis-Lamb-certo.pdf Acesso em: 14 ago. 2025.

MALWAREBYTES. Riscos da IA e da segurança cibernética. **Malwarebytes**, 2024. Disponível em: https://www.malwarebytes.com/pt-br/cybersecurity/basics/risks-of-ai-in-cyber-security. Acesso em: 05 ago. 2025.

MONFRE, G. A.; SILVA, F. G.; VICENTIN, A. C. Inteligência Artificial Aplicada à Cibersegurança: Soluções Estratégicas para um Ambiente Digital Resiliente. **Revista Matiz Online,** 13. ed., set. 2023. Disponível em: https://immes.edu.br/wp-content/uploads/2025/04/Artigo_MATIZ_2023_Ciberseguranca.pdf. Acesso em: 10 ago. 2025.

PONTES, Gabriel Rodrigues; VASCONCELOS, Cláudio Nunes; SILVA, Flávio de Pilla. Sistema de detecção de intrusão utilizando métodos de aprendizagem de máquina em redes de computadores. **Revista de Ciência e Inovação,** v. 4, n. 1, p. 74-87, 2023. Disponível em: https://periodicos.iffarroupilha.edu.br/index.php/cienciai-novacao/article/view/388. Acesso em: 15 ago. 2025.

RODOLFO, B. Implicações Éticas das Tecnologias de Inteligência Artificial: Direitos Autorais, Privacidade, Segurança e Regulação. **Comunicação e Sociedade**, v. 47, e025005, 2025. Disponível em: https://doi.org/10.17231/comsoc.47(2025).6088. Acesso em: 20 ago. 2025.

SAMPAIO, Rafael Cardoso; SABBATINI, Marcelo; LIMONGI, Ricardo. **Diretrizes para o uso ético e responsável da Inteligência Artificial Generativa: um guia prático para pesquisadores**. São Paulo: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação - Intercom, 2024.

SENADO aprova marco regulatório de inteligência artificial. **Correio do Povo**, 10 dez. 2024. Disponível em: https://www.correiodopovo.com.br/not%C3%ADcias/pol%C3%ADtica/senado-aprova-marco-regulat%C3%B3rio-de-intelig%C3%AAncia-artificial-1.1560503. Acesso em: 01 set. 2025.

SICHMAN, Jaime S. Inteligência Artificial e sociedade: avanços e riscos. **Estudos Avançados**, v. 35, n. 101, p. 37-50, jan./abr. 2021. Disponível em: https://doi.org/10.1590/s0103-4014.2021.35101.004. Acesso em: 05 set. 2025.

SILVA, Francisco Alves da. Responsabilidade civil e inteligência artificial: explorando soluções e desafios da era digital. **RECIMA21 - Revista Científica Multidisciplinar,** v. 4, n. 11, e4114434, 2023. Disponível em: https://recima21.com.br/recima21/article/view/4434. Acesso em: 10 set. 2025.

4

DESAFIOS E SOLUÇÕES NA IMPLEMENTAÇÃO DE REDES DE FIBRA ÓPTICA FTTH

CHALLENGES AND SOLUTIONS IN THE IMPLEMENTATION OF FTTH FIBER OPTIC NETWORKS

Kaio Costa Cavalcanti
Ivone Ascar Sauaia Guimarães
Mirian Nunes de Carvalho Nunes

Resumo

A massificação das redes de fibra óptica Fiber to the Home (FTTH) consolidou-se como a principal infraestrutura de conectividade de alta velocidade, mas também introduziu novos desafios críticos de segurança da informação. Este artigo tem como objetivo analisar as principais vulnerabilidades em redes FTTH e as soluções técnicas e de gestão essenciais para a mitigação de riscos. A pesquisa, de natureza bibliográfica e abordagem qualitativa, baseou-se na revisão de estudos científicos recentes, normas técnicas e legislações vigentes, como a LGPD e a GDPR. Os resultados indicam que a segurança em redes ópticas exige uma abordagem multicamadas, superando a simples proteção física. Ameaças como o grampeamento óptico são mitigadas por contramedidas lógicas, como a criptografia AES no padrão GPON, e por sistemas de monitoramento proativo baseados em Inteligência Artificial. Além disso, observou-se que a complexidade operacional, os custos de implementação e a conformidade legal são fatores decisivos para a efetividade das soluções. Conclui-se que, embora a tecnologia FTTH ofereça robustez, sua segurança eficaz depende da integração entre defesas tecnológicas avançadas, processos operacionais maduros e uma sólida governança de dados alinhada às exigências legais.

Palavras-chave: Segurança de Redes. Fibra Óptica. FTTH. LGPD. Criptografia.

Abstract

Artificial The widespread adoption of Fiber to the Home (FTTH) optical fiber networks has established itself as the primary high-speed connectivity infrastructure, but it has also introduced new critical information security challenges. This article aims to analyze the main vulnerabilities in FTTH networks and the essential technical and management solutions for risk mitigation. The research, bibliographic in nature with a qualitative approach, was based on a review of recent scientific studies, technical standards, and current legislation, such as the LGPD and GDPR. The results indicate that security in optical networks requires a multi-layered approach, going beyond simple physical protection. Threats such as optical tapping are mitigated by logical countermeasures, like AES encryption in the GPON standard, and by proactive monitoring systems based on Artificial Intelligence. Furthermore, it was observed that operational complexity, implementation costs, and legal compliance are decisive factors for the effectiveness of the solutions. It is concluded that, although FTTH technology offers robustness, its effective security depends on the integration of advanced technological defenses, mature operational processes, and solid data governance aligned with legal requirements.

Keywords: Network Security. Optical Fiber. FTTH. LGPD. Encryption.

1 INTRODUÇÃO

A expansão das redes de fibra óptica, impulsionada pela tecnologia Fiber to the Home (FTTH), transformou fundamentalmente o acesso à internet em escala global, inaugurando uma era de conectividade com alta velocidade e maior estabilidade. Essa infraestrutura tornou-se a espinha dorsal para a transmissão de um volume crescente de informações, suportando desde o entretenimento digital até operações críticas de empresas e governos. A fibra óptica, por sua natureza dielétrica, foi por muito tempo percebida como um meio de transmissão intrinsecamente seguro, imune a interferências eletromagnéticas que afetam os cabos metálicos.

Contudo, essa percepção de segurança absoluta foi desafiada à medida que a dependência dessa tecnologia aumentou. A crescente transmissão de dados sensíveis, tanto de natureza pessoal quanto corporativa, evidenciou que a robustez do meio físico não elimina por completo as vulnerabilidades. A possibilidade de ameaças, que vão desde a interceptação física do sinal luminoso até ataques lógicos mais complexos, expôs a necessidade de uma análise mais crítica sobre a proteção dessas redes, tornando a segurança da informação um pilar tão essencial quanto o próprio desempenho da conexão.

A relevância deste estudo foi amplificada pelo novo cenário regulatório global, marcado pela promulgação de leis rigorosas de proteção de dados, como a General Data Protection Regulation (GDPR) na Europa e a Lei Geral de Proteção de Dados Pessoais (LGPD) no Brasil. Tais legislações impuseram aos provedores de serviço a responsabilidade direta pela segurança das informações dos usuários. Nesse contexto, a segurança em redes FTTH deixou de ser um mero diferencial técnico para se tornar uma obrigação legal e um fator determinante para a confiança do consumidor, tornando indispensável a investigação sobre como alinhar a infraestrutura tecnológica às exigências de conformidade.

Diante do exposto, a pesquisa foi orientada pela seguinte problemática: Quais são as principais vulnerabilidades de segurança em redes de fibra óptica FTTH e quais contramedidas técnicas e de gestão são essenciais para mitigar os riscos e assegurar a conformidade com as leis de proteção de dados? Desta forma, o objetivo geral deste trabalho consistiu em analisar as vulnerabilidades e as principais estratégias de segurança aplicáveis a redes FTTH, integrando as dimensões técnica e regulatória do problema.

Para alcançar o objetivo proposto, foram traçados os seguintes objetivos específicos: primeiramente, foi realizado um mapeamento das principais ameaças à segurança em redes ópticas, com especial atenção ao grampeamento de fibra. Em seguida, foram discutidas as contramedidas técnicas mais eficazes, como a criptografia e o monitoramento inteligente da rede. Por fim, analisou-se a intersecção entre a segurança da infraestrutura FTTH e as diretrizes estabelecidas pela LGPD e GDPR, consolidando a compreensão sobre os desafios e soluções no cenário atual.

2 DESENVOLVIMENTO

2.1 Metodologia

O presente estudo foi conduzido por meio de uma revisão bibliográfica de natureza qualitativa e descritiva, com o objetivo de analisar a produção científica sobre segurança em redes de fibra óptica FTTH. Para o levantamento do referencial teórico, realizou-se uma busca por publicações nas bases de dados acadêmicas Google Scholar, IEEE Xplore, ScienceDirect e Scopus. O recorte temporal da pesquisa compreendeu trabalhos publicados nos últimos cinco anos, abrangendo o período entre 2020 e 2025, a fim de garantir a análi-

se dos materiais mais recentes e estritamente relevantes para o cenário tecnológico atual.

Os descritores utilizados para a busca, nos idiomas português e inglês, foram: "segurança em redes FTTH" (FTTH *network security*), "proteção de dados em fibra óptica" (*data protection in fiber optics*), "criptografia em redes ópticas" (*cryptography in optical networks*) e "vulnerabilidades em redes ópticas" (*optical network vulnerabilities*). Como critérios de inclusão, foram selecionados artigos científicos, livros e capítulos de livros que abordassem diretamente as ameaças, contramedidas e a gestão de segurança em redes FTTH. Foram adotados como critérios de exclusão os trabalhos que não apresentavam aderência direta ao tema, bem como resumos, artigos de opinião e trabalhos não disponíveis na íntegra para consulta. A aplicação desses filtros resultou em uma seleção inicial de 28 trabalhos. Após a leitura analítica e a aplicação final dos critérios, 10 publicações foram selecionadas como o material fundamental para esta pesquisa, por sua relevância direta aos objetivos do estudo. Todas as 10 obras selecionadas estão compreendidas no recorte temporal estipulado de 2020 a 2025. A análise crítica deste material permitiu a fundamentação das discussões apresentadas neste artigo.

2.2 Resultados e Discussão

Os resultados obtidos a partir da revisão bibliográfica indicam que a segurança em redes de fibra óptica FTTH é um domínio que passou por uma reavaliação crítica, transitando de uma percepção de robustez inerente para a necessidade de uma estratégia de defesa multicamadas, proativa e alinhada a imperativos legais. A análise dos trabalhos mostrou que, embora o meio físico ofereça vantagens de segurança sobre tecnologias legadas, a crescente criticidade e o valor econômico dos dados transmitidos expôs vulnerabilidades significativas que demandam uma combinação de contramedidas físicas, lógicas e de gestão (Santos; Moreira, 2022).

A vulnerabilidade física mais proeminente, amplamente discutida na literatura, é o grampeamento óptico (*optical tapping*). Kashyap, Joshi e Singh (2020) detalham que essa ameaça se materializa por meio de técnicas que não exigem o rompimento da fibra, como a indução de macrocurvaturas (*macrobending*). Essa ação provoca o escape de uma fração mínima do sinal luminoso, que pode ser capturada e analisada por um atacante sem causar uma interrupção de serviço que alerte o provedor. A eficácia do ataque reside na sua sutileza, comprometendo a confidencialidade dos dados de forma silenciosa. É crucial notar que o modelo de ameaça não se restringe a atores externos; a ameaça interna, representada por funcionários ou técnicos terceirizados com acesso legítimo à infraestrutura física (postes, caixas de emenda, armários de distribuição), representa um risco igualmente elevado, pois esses indivíduos possuem o conhecimento e a oportunidade para instalar dispositivos de interceptação com baixa probabilidade de detecção imediata.

A eficácia do monitoramento passivo da potência do sinal tem se mostrado limitada na prática. Segundo Kashyap, Joshi e Singh (2020), invasores munidos de equipamentos sofisticados conseguem atuar camuflados nas oscilações naturais da rede, o que invalida a detecção baseada apenas em níveis simples de atenuação. Para superar essa lacuna técnica, ganham destaque soluções mais precisas, como os Reflectômetros Ópticos no Domínio do Tempo (OTDRs), que mapeiam anomalias físicas na linha. O grande desafio, contudo, é financeiro: manter OTDRs monitorando a rede de acesso em tempo real exige um investimento alto, o que leva muitas operadoras a restringir essa proteção ao backbone, deixando a 'última milha' — justamente o trecho mais exposto fisicamente — vulnerável.

Diante da impossibilidade de garantir uma proteção física infalível, a discussão con-

verge para a camada lógica como a principal linha de defesa. A literatura recente, como Santos e Moreira (2022), reforça o princípio fundamental de que a confidencialidade deve ser inerente ao dado, e não apenas ao meio físico. Nesse contexto, a criptografia se consolida como a ferramenta essencial, conforme apontado por estudos sobre a segurança da camada física em redes ópticas (Chen; Liu, 2020). A sua aplicação é particularmente crítica na arquitetura Ponto-Multiponto (P2MP) das redes GPON, onde o sinal downstream é difundido para todos os usuários de um mesmo splitter. O próprio padrão GPON, ciente desse risco, tornou obrigatório o uso do algoritmo AES para cifrar o tráfego, garantindo que cada equipamento terminal (ONT) decifre apenas os dados que lhe são destinados.

A eficácia do AES, contudo, depende inteiramente da segurança do processo de gerenciamento de chaves criptográficas. O padrão GPON utiliza o protocolo OMCI (*ONT Management and Control Interface*) para a troca de chaves entre a OLT e cada ONT. Se um atacante conseguir explorar uma vulnerabilidade nesse protocolo ou de alguma forma capturar as chaves durante o processo de autenticação de uma nova ONT na rede, a criptografia do enlace pode ser comprometida. Portanto, a robustez da segurança não está apenas no algoritmo em si, mas em toda a cadeia de processos que gerencia o ciclo de vida das chaves, um ponto que Stallings (2017) enfatiza como frequentemente sendo o elo mais fraco em sistemas criptográficos.

A segurança da rede, contudo, não se encerra na proteção física da fibra. Barros e Sampaio (2023) alertam que é indispensável uma visão ponta a ponta, focada especialmente nos equipamentos terminais (ONTs). Por estarem na casa do cliente, esses dispositivos tornam-se vetores frequentes de ataque, seja por configurações inseguras, firmware obsoleto ou falhas em protocolos de aplicação. Há ainda um risco estrutural difícil de mitigar: a cadeia de suprimentos. Backdoors inseridos durante a fabricação das ONTs representam uma ameaça silenciosa que pode comprometer milhares de usuários simultaneamente, fugindo completamente ao controle operacional do provedor.

Diante da insuficiência das defesas estáticas tradicionais, a tendência atual caminha para a gestão inteligente. Zhang et al. (2021) propõem o uso de Inteligência Artificial e Machine Learning para romper com a lógica reativa baseada apenas em assinaturas de ataques conhecidos. A proposta é treinar algoritmos para assimilar o 'DNA' do funcionamento normal da rede, monitorando em tempo real variáveis como latência, jitter, padrões de tráfego e até a potência do sinal. Ao dominar esse padrão, a IA consegue detectar desvios sutis que passariam despercebidos, como um upload atípico na madrugada (sugerindo uma botnet) ou uma queda de performance localizada que denuncie uma tentativa de grampo físico.

A implementação de tais sistemas de IA gera uma sinergia valiosa, convergindo as operações de segurança (SecOps) com as operações de rede (NetOps). As mesmas plataformas que detectam anomalias de segurança podem ser utilizadas para prever falhas de hardware, otimizar o roteamento de tráfego e melhorar a Qualidade de Experiência (QoE) do usuário final. Essa dupla funcionalidade fortalece o argumento de negócio para o investimento em plataformas inteligentes, pois o retorno sobre o investimento (ROI) não se limita à mitigação de riscos, mas se estende à otimização da performance e à redução de custos operacionais, conforme a visão de gestão de redes proposta por Zhang et al. (2021).

Em consonância, a proposta de Zhang et al. (2021) também aborda a questão da auditabilidade e da integridade dos registros de segurança por meio do blockchain. Em ataques sofisticados, um dos primeiros alvos são os logs do sistema, que podem ser alterados ou apagados para ocultar os rastros do invasor. O uso de um blockchain para registrar eventos de segurança cria um livro-razão distribuído e imutável, fornecendo uma trilha de

auditoria confiável, essencial para investigações forenses e para a comprovação de conformidade legal. Essa imutabilidade é particularmente valiosa no contexto da LGPD, que exige que as empresas sejam capazes de demonstrar as medidas de segurança adotadas e de fornecer relatórios precisos em caso de um incidente de segurança.

Paralelamente à evolução técnica, a dimensão regulatória redefiniu as obrigações dos provedores de serviço A promulgação de leis como a GDPR na Europa e a LGPD no Brasil transformou a proteção de dados em um requisito legal mandatório. A análise do impacto direto dessas leis na operação dos provedores é um tema central na literatura recente (Gomes; Almeida, 2022). Essa mudança legislativa introduziu um cálculo econômico inegável: o custo de um incidente de segurança não se resume mais a perdas técnicas, mas inclui multas regulatórias severas, despesas com remediação, ações judiciais e, talvez o mais danoso, a perda de reputação e a consequente evasão de clientes. Nesse cenário, o investimento em segurança deixa de ser um custo operacional para se tornar uma estratégia de mitigação de risco financeiro e de preservação do valor da marca.

O princípio de "Privacidade desde a Concepção" (*Privacy by Design*), central nessas legislações, implica que a segurança não pode ser um adendo, mas deve ser uma consideração fundamental no próprio planejamento e arquitetura da rede FTTH. Isso se traduz em decisões de engenharia, como a segmentação lógica da rede para isolar diferentes tipos de tráfego e a escolha de fornecedores de equipamentos que demonstrem um forte compromisso com práticas seguras de desenvolvimento de software. A análise dos trabalhos permitiu traçar um paralelo direto entre os requisitos legais e as contramedidas técnicas. A criptografia AES, o monitoramento com IA e os logs em blockchain são exemplos concretos das "medidas técnicas" exigidas pelo Artigo 46 da LGPD, enquanto as "medidas administrativas" incluem a elaboração de um plano de resposta a incidentes e a realização de treinamentos de conscientização para funcionários.

Um ponto recorrente nas discussões foi a questão dos desafios e limitações. A complexidade operacional aumenta exponencialmente em ambientes multivendor, uma realidade na maioria dos provedores, que utilizam equipamentos de diferentes fabricantes. A heterogeneidade de OLTs e ONTs dificulta a implementação de uma política de segurança unificada, complica os ciclos de atualização de firmware e pode criar lacunas de segurança na interface entre sistemas distintos, desafiando a visão de uma gestão centralizada e inteligente. O custo de tecnologias avançadas, como plataformas de IA, e a necessidade de mão de obra qualificada para operá-las, também foram apontados como barreiras significativas.

A questão da explicabilidade dos modelos de IA, conhecida como XAI (*Explainable AI*), também emergiu como um fator crítico. Para que uma equipe de operações de rede confie em um alerta gerado por IA, é fundamental que o sistema possa fornecer uma justificativa compreensível para sua decisão. Modelos "caixa-preta", mesmo que precisos, podem gerar resistência. Adicionalmente, o uso de IA para monitorar padrões de tráfego, embora essencial para a segurança, tangencia questões éticas de privacidade. Os provedores devem navegar com cuidado, garantindo que o monitoramento se restrinja a metadados e análises de comportamento anonimizadas, em conformidade com os princípios da LGPD, para não cruzar a linha entre a proteção da rede e a vigilância dos usuários.

Olhando para o futuro, a literatura sinaliza a necessidade de evoluir para arquiteturas ainda mais resilientes. A ascensão da computação quântica, por exemplo, representa uma ameaça a longo prazo para os algoritmos criptográficos atuais, como o AES. Ataques do tipo "coletar agora, decifrar depois" (*harvest now, decrypt later*) podem estar ocorrendo, onde dados cifrados são armazenados hoje para serem quebrados por computadores

quânticos no futuro. Isso impulsiona a pesquisa em Criptografia Pós-Quântica (PQC), um campo que busca desenvolver algoritmos resistentes a esse novo paradigma computacional, sendo uma consideração estratégica para a segurança de dados de longo prazo, como apontam estudos recentes sobre a teoria e prática da segurança em redes (Chen; Liu, 2020).

Por fim, a evolução das estratégias de defesa aponta para uma mudança estrutural e definitiva: a adoção da Arquitetura de Confiança Zero (*Zero Trust Architecture* - ZTA). Esse modelo representa um rompimento drástico com a velha ideia de 'perímetro seguro' (o conceito de castelo e fosso), partindo da premissa realista de que a ameaça pode não apenas vir de fora, mas já estar operando dentro da infraestrutura. No cenário específico das redes FTTH, isso implica uma postura de desconfiança padrão: nenhum componente — seja a OLT central no provedor ou a ONT na residência do cliente — recebe credibilidade automática apenas por estar conectado fisicamente à rede. Sob a ótica do Zero Trust, cada solicitação de acesso ou troca de dados deve ser rigorosamente verificada, autenticada e autorizada em tempo real. Na prática, a ZTA funciona como o alicerce integrador que valida todas as outras camadas de defesa: ela torna indispensável o monitoramento inteligente e contínuo detalhado por Zhang et al. (2021), exige a blindagem física e lógica dos dispositivos terminais descrita por Silva e Costa (2021) e impõe a aplicação rigorosa da criptografia ponta a ponta defendida por Santos e Moreira (2022), garantindo que a segurança persista mesmo se o meio físico for comprometido.

3 CONCLUSÃO

A realização deste estudo permitiu compreender como a segurança em redes FTTH evoluiu de uma percepção de robustez física para uma disciplina complexa e multicamadas, essencial para a sustentabilidade dos serviços de telecomunicações. Os objetivos propostos foram alcançados, uma vez que foi possível mapear as principais vulnerabilidades, como o grampeamento óptico e as falhas em nível de protocolo, além de discutir a relevância de soluções como a criptografia ponta a ponta, o monitoramento inteligente e a adequação às novas leis de proteção de dados.

A pesquisa respondeu ao problema inicialmente proposto ao evidenciar que as contramedidas técnicas e de gestão são interdependentes, e que a segurança eficaz só é alcançada por meio de uma estratégia integrada que combina proteção física, lógica e administrativa. Apesar dos avanços tecnológicos observados, constatou-se que desafios ainda persistem, especialmente relacionados aos custos de implementação de soluções avançadas (como IA e OTDRs contínuos) para pequenos provedores, à complexidade de gerenciar ambientes multivendor e à ameaça constante representada pelo fator humano, seja por erro ou ação maliciosa interna.

Como proposta para estudos futuros, recomenda-se o aprofundamento das análises sobre a aplicação de algoritmos de Machine Learning na detecção de anomalias em tempo real em redes PON. Além disso, sugere-se investigar a viabilidade técnica e econômica da implementação de Arquiteturas de Confiança Zero (Zero Trust) no contexto específico das redes de acesso FTTH e a preparação da infraestrutura para os desafios futuros da Criptografia Pós-Quântica (PQC), garantindo a segurança dos dados a longo prazo.

REFERÊNCIAS

BARROS, L. F.; SAMPAIO, K. M. **Autenticação e segurança de dispositivos em redes FTTH: uma análise de vetores de ataque em ONTs**. Revista Brasileira de Segurança Cibernética, v. 11, n. 1, p. 29-41, 2023.

CHEN, W.; LIU, Y. **Physical layer security in passive optical networks: from theory to practice**. IEEE Communications Magazine, v. 58, n. 3, p. 45-51, 2020.

GOMES, P. R.; ALMEIDA, R. S. **Governança de segurança e conformidade com a LGPD em provedores de internet**. Revista de Direito, Tecnologia e Inovação, v. 7, n. 1, p. 112-128, 2022.

KASHYAP, A.; JOSHI, S.; SINGH, M. **Optical fiber tapping: vulnerabilities and countermeasures**. Journal of Optical Communications, v. 41, n. 2, p. 123–130, 2020.

RIBEIRO, A. C. **Criptografia Pós-Quântica (PQC) e seus impactos na segurança de infraestruturas de longa duração**. Anais do Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg), p. 201-214, 2022.

SANTOS, R. L.; MOREIRA, E. M. **Segurança em redes ópticas: técnicas e desafios**. Revista Brasileira de Segurança da Informação, v. 4, n. 2, p. 78–89, 2022.

SILVA, L. M.; COSTA, J. F. **Vulnerabilidades em redes GPON: ataques de negação de serviço e rogue ONTs**. Journal of Network and Computer Applications, v. 125, p. 78-89, 2021.

VIANA, J. P. **O impacto da GDPR e LGPD na arquitetura de redes de telecomunicações: da teoria à prática**. International Journal of Data Privacy Law, v. 3, n. 2, p. 145-160, 2021.

ZHANG, Y.; CHEN, X.; WU, D. **Intelligent security management in fiber-optic networks using AI and blockchain**. Computer Networks, v. 196, p. 108173, 2021.

5

A DEPENDÊNCIA DA INTELIGÊNCIA ARTIFICIAL: NA TOMADA DE DECISÃO CLÍNICA, ANÁLISE DOS IMPACTOS SOCIAIS, PSICOLÓGICOS E ÉTICOS

THE DEPENDENCE ON ARTIFICIAL INTELLIGENCE: IN CLINICAL DECISION-MAKING, ANALYSIS OF SOCIAL, PSYCHOLOGICAL, AND ETHICAL IMPACTS

Thiago Kauã Costa de Abreu
Geovana da Conceição Avelar Ribeiro
Ivone Ascar Sauáia Guimarães
Mirian Nunes de Carvalho Nunes

Resumo

O presente estudo abordou a dependência da inteligência artificial (IA) e seus efeitos na autonomia humana, na tomada de decisão crítica e na responsabilidade ética em um futuro crescentemente automatizado. O objetivo geral foi mostrar o impacto causado pelo crescente uso da IA, que fornece respostas e soluções prontas, incluindo seus efeitos psicológicos, sociais e éticos. A metodologia utilizada foi a revisão de literatura, com pesquisa de livros, dissertações e artigos científicos nas bases de dados SciELO e Google Acadêmico, abrangendo trabalhos publicados nos últimos 8 anos. Os resultados destacaram o impacto transformador da IA na tomada de decisão clínica, ampliando as capacidades diagnósticas e terapêuticas, como a identificação de tumores malignos. Contudo, o estudo revela que o uso da IA pode gerar um risco de dependência, acarretando a redução de habilidades e competências importantes, como a diminuição da criatividade, capacidade de pensamento crítico e intuição. O uso excessivo de tecnologia também pode impactar negativamente a saúde mental, levando a um aumento do isolamento social, ansiedade e tristeza. Por fim, a pesquisa aponta que é fundamental o desenvolvimento ético da IA para evitar a ampliação de desigualdades e a geração de novos perigos, como a discriminação algorítmica e a quebra de privacidade, sendo urgente a criação de uma base ética sólida. É essencial garantir que a tecnologia seja usada de forma ética e responsável, preservando a autonomia e o pensamento crítico humanos.

Palavras-chave: Tecnologia. Dependência. IA.

Abstract

The present study addressed the dependence on artificial intelligence (AI) and its effects on human autonomy, critical decision-making, and ethical responsibility in an increasingly automated future. The overall objective was to demonstrate the impact caused by the growing use of AI, which provides ready-made answers and solutions, including its psychological, social, and ethical effects. The methodology used was a literature review, drawing on books, dissertations, and scientific articles from the SciELO and Google Scholar databases, covering works published in the last eight years. The results highlighted the transformative impact of AI on clinical decision-making, expanding diagnostic and therapeutic capabilities, such as the identification of malignant tumors. However, the study reveals that the use of AI may lead to a risk of dependency, resulting in the reduction of important skills and competencies, such as decreased creativity, critical thinking, and intuition. Excessive use of technology may also negatively affect mental health, leading to increased social isolation, anxiety, and sadness. Finally, the research indicates that the ethical development of AI is essential to prevent the widening of inequalities and the emergence of new risks, such as algorithmic discrimination and breaches of privacy, making the creation of a solid ethical foundation urgent. It is essential to ensure that technology is used ethically and responsibly, preserving human autonomy and critical thinking.

Keywords: Technology. Dependence. AI.

1 INTRODUÇÃO

Este estudo se trata da dependência da inteligência artificial primando mostrar qual é o efeito causado pela dependência da IA em relação a autonomia humana, a capacidade de tomada de decisão crítica e a responsabilidade ética em um futuro cada vez mais automatizado. Tendo como objetivo mostrar o impacto causado pelo crescente uso da Inteligência Artificial ao fornecer respostas rápidas e soluções prontas, incluindo seus efeitos psicológicos, sociais e éticos.

O estudo mostra os riscos da dependência e ajuda a identificar como a IA afeta o pensamento crítico, a perda da capacidade de buscar e analisar informações de forma independente, levando a uma homogeneização do pensamento, diminuição da pluralidade de ideias e a um futuro composto apenas por indivíduos receptores de informações pré-selecionadas por máquinas.

As aplicações da IA são vastas e abrangem diversos setores, tendo como exemplo na área da saúde e na ética. Na medicina, a IA auxilia no diagnóstico de doenças, no desenvolvimento de tratamentos personalizados e na descoberta de novos medicamentos. Na ética, por meio da análise de grandes volumes de dados e da detecção de padrões, a IA ajuda profissionais e instituições a identificar riscos, reduzir erros e garantir mais transparência nos processos decisórios.

No entanto, o rápido avanço da IA também levanta questões importantes sobre seu impacto na sociedade. A automação impulsionada pela IA pode levar à perda de empregos em alguns setores, exigindo a requalificação da força de trabalho. Além disso, a IA leva também a redução da autonomia humana e levanta questões éticas relacionadas à privacidade, segurança e viés algorítmico.

Este estudo explorou os impactos sociais da IA na tomada de decisão clínica, os efeitos psicológicos da dependência da IA e as implicações éticas na responsabilidade humana. Teve como relevância garantir que a tecnologia seja usada de forma ética e responsável, preservando a autonomia, as capacidades humanas, o pensamento crítico e evitar que a sociedade dependa de algoritmos para tomadas de decisões, diminuindo a capacidade de pensar e agir de forma independente.

2 DESENVOLVIMENTO

2.1 Metodologia

O tipo de pesquisa realizada foi uma revisão de literatura, onde foram pesquisados livros, dissertações e artigos científicos selecionados através de busca nas seguintes bases de dados: SciELO e Google Acadêmico. O período dos artigos pesquisados foi dos trabalhos publicados nos últimos 8 anos. As palavras-chave utilizadas na busca foram: “Dependência”, “Tecnologia” e “Inteligência Artificial. Os principais autores, foram: Dantas et al. (2024), Bernal (2017), Sousa et al. (2025), EUA (2017), Santos, Simões e Neves (2023) e Lucas e Santos (2021).

2.2 Resultados e Discussão

Os resultados apresentados destacam o impacto transformador da inteligência artificial (IA) na tomada de decisão clínica, ampliando as capacidades diagnósticas e terapêuticas. Estes sistemas utilizam algoritmos e aprendizado de máquina para dar aos

profissionais de saúde ferramentas que podem prever com alta precisão os resultados de tratamentos e o progresso de doenças (Dantas *et al*., 2024).

A Inteligência Artificial já apresenta desempenho superior ao dos seres humanos em diversas atividades na área da saúde, como a identificação de tumores malignos e a realização de pesquisas clínicas. No entanto, sua implementação em larga escala ainda enfrenta obstáculos relacionados a fatores técnicos, culturais e organizacionais. A inteligência artificial está revolucionando a área da saúde, trazendo muitos avanços importantes (Dantas *et al*., 2024).

Contudo, este estudo também revela que o uso da IA pode gerar um risco de dependência, sobretudo quando tarefas simples e rápidas passam a ser automatizadas ou quando atividades que deveriam ser realizadas por pessoas, com o objetivo de estimular o aprendizado, são substituídas. Isso pode acarretar a redução de habilidades e competências importantes. Além disso, o uso excessivo da tecnologia impacta negativamente a saúde mental, já que a diminuição da interação humana real pode comprometer a empatia e o vínculo emocional (Santos; Simões, Neves 2023).

No passado, quando alguém precisava obter informações sobre fatos ou adquirir conhecimentos em geral, os livros e enciclopédias eram praticamente a única fonte disponível. Isso significava que a pessoa interessada precisava percorrer sumários e capítulos de várias obras em busca do conteúdo desejado, sem nenhuma garantia de que realmente o encontraria. Com o avanço da tecnologia, esse processo passou a ser muito mais ágil e preciso (Santos; Simões, Neves 2023).

A inteligência artificial está em constante desenvolvimento e tende a estar cada vez mais presente no cotidiano. Contudo, sua utilização também envolve riscos, como a criação de dependência, a redução da busca por conhecimento e a substituição de atividades humanas fundamentais, como calcular, ler e pesquisar. Por isso, é essencial considerar não apenas os benefícios que a IA oferece, mas também os seus possíveis prejuízos, já que ambos merecem ser analisados e debatidos com a mesma atenção (Santos; Simões, Neves 2023).

A pesquisa aponta que desenvolver uma IA de forma ética é fundamental, pois uma aplicação inadequada pode ampliar desigualdades e gerar novos perigos, como a discriminação nos algoritmos e a quebra da privacidade. Com o avanço da IA em diferentes áreas, surgem dúvidas sobre sua real eficácia na proteção dos direitos humanos e na promoção da equidade social (Bernal, 2017).

O avanço da Inteligência Artificial nas instituições sociais torna ainda mais urgente a criação de uma base ética sólida. Com a expansão dessa tecnologia em diversos setores, as formas de interação humana também se transformam. Os obstáculos a enfrentar não se limitam a questões técnicas, mas exigem um compromisso ético central. A conexão entre ética e tecnologia se expressa em ações colaborativas que buscam assegurar que as inovações sejam direcionadas para o benefício coletivo (Sousa *et al*., 2025).

Em seguida o quadro 1, mostra a relação entre os impactos sociais, psicológicos e éticos da Inteligência Artificial.

Quadro 1. Impactos sociais, psicológicos e éticos da IA

AUTOR	TÓPICO	DESCRIÇÃO	EXEMPLO PRÁTICO
Dantas, *et al.*	Impactos sociais da IA na tomada de decisão clínica.	A inteligência artificial na decisão clínica é uma ferramenta inovadora com potencial para reduzir desigualdades em saúde, influenciadas por fatores socioeconômicos, culturais e ambientais.	a IA pode identificar padrões ocultos, detecção de tumores malignos, condução de pesquisas clínicas, fornecer melhores diagnósticos e prognósticos.
Santos; Simões, Neves.	Efeitos psicológicos da dependência da IA.	A alta dependência de sistemas de inteligência artificial pode resultar na diminuição da criatividade, capacidade de pensamento crítico e intuição. É essencial encontrar um ponto de equilíbrio, usando somente a IA para auxiliar nas decisões.	A redução da busca por conhecimento e a substituição de atividades humanas fundamentais, como calcular, ler e pesquisar.
Sousa, *et al.*	Implicações éticas da IA na responsabilidade humana.	As implicações éticas e os desafios operacionais são significativos, levantando questões importantes sobre a confiabilidade, privacidade dos dados.	Ao examinarmos as práticas de vigilância, percebemos que o uso de algoritmos sem diretrizes éticas pode intensificar a discriminação, especialmente se sua aplicação não for supervisionada.

Fonte: Adaptado de Lauar et al (2024).

O **Quadro 1** mostra de forma resumida os principais impactos da inteligência artificial, destacando aspectos sociais, psicológicos e éticos. Nele, são destacados autores e suas contribuições, descrevendo como a IA influencia a tomada de decisão clínica, os efeitos da dependência tecnológica no comportamento humano e as implicações éticas relacionadas à responsabilidade e ao uso adequado da tecnologia.

A utilização da inteligência artificial nos sistemas de saúde tem aperfeiçoado a administração dos cuidados médicos e aprimorado o método de tomada de decisões clínicas. Essa tecnologia é reconhecida por sua habilidade de aprender, se ajustar e prever consequências, tem a habilidade de redefinir os métodos tradicionais usados na medicina. No entanto, essa mudança destaca uma dualidade entre ganhos de eficiência e desafios éticos (EUA, 2017).

Ao analisar dados de forma objetiva e consistente, a IA pode oferecer uma análise imparcial e baseada em indícios. Alguns sistemas de IA estão sendo envolvidos na área da saúde como no diagnóstico médico, ajudando os profissionais a compreender dados complexos e melhorando a precisão das análises. A qualidade das decisões também é melhorada pela IA ao mitigar vieses cognitivos e emotivos que frequentemente influenciam decisões humanas (Lauar *et al*., 2024).

Ao incorporar algoritmos avançados e técnicas de aprendizado de máquina, esses sistemas oferecem aos profissionais de saúde maneiras capazes de prever, com maior precisão, os resultados dos tratamentos e a evolução das doenças, melhorando a tomada de decisões clínicas e o uso eficiente dos recursos (EUA, 2022).

O uso da inteligência artificial aumenta preocupações éticas, como a responsabilidade por erros e o risco de preconceitos nos algoritmos. Conforme esses sistemas passam a desempenhar funções importantes, torna-se necessário definir claramente quem será

responsabilizado por eventuais falhas ou problemas (Silva; Nogaroli, 2020).

Além disso, o sistema judiciário brasileiro ainda não prevê abertamente a responsabilidade de quaisquer atos praticados por sistemas de Inteligência Artificial. Isso gera uma brecha na preservação dos pacientes e impossibilita o ajuste de possíveis prejuízos (Lucas; Santos, 2021).

Outro ponto é as mudanças sociais resultantes das interações entre pessoas e máquinas. Evidencia-se que a relação entre tecnologia, princípios e sociedade exige um cuidado, garantindo que isso não se transforme em um problema ao desenvolvimento humano (Bernal, 2017).

O uso da internet se torna um vício quando a pessoa começa a usar sem um propósito e passa a se preocupar com isso, deixando que a internet controle sua vida. Assim como outras formas de dependência, o vício em IA pode provocar prejuízos físicos e mentais. É fundamental que os sinais demonstrados pela pessoa prejudicada sejam analisados (Lerner *et al*., 2024).

A dependência excessiva de dispositivos e apps digitais levam a uma instabilidade entre o mundo real e virtual, e afeta as próprias emoções. Isso pode levar a um aumento do isolamento social, ansiedade e tristeza (Sergipe, 2022).

Essa atitude comportamental se espalhou por várias áreas, segurando a ideia de que as ações das pessoas são norteadas, em grande parte, não por decisões racionais, mas por tendências cognitivas, fazendo com que elas se comportem de maneira "previsivelmente irracional". (Lerner *et al*., 2024).

Atualmente, um dos cuidados também é garantir que os avanços em IA sigam normas éticas essenciais e forneçam justiça social de maneira responsável. Isso envolve transparência nos algoritmos, proteção à privacidade dos dados e mitigação de vieses (Lucas; Santos, 2021).

Segundo Biondi e Cernev (2023, p.3) "a ética digital deve estar alinhada às soluções de IA para enfrentar os desafios do mundo real". Nesse sentido é essencial manter um diálogo consciente e responsável nessa área, permitindo que o progresso tecnológico aconteça de forma organizada aos valores sociais. Assim, a integração entre inovação tecnológica e princípios éticos torna-se indispensável para evitar riscos e desigualdades. No entanto, o excesso da utilização de sistemas automatizados traz preocupações éticas relacionadas ao uso correto nas operações e ao risco de preconceitos presentes nos algoritmos usados (Silva; Nogaroli, 2020).

Os efeitos sociais da IA demandam uma avaliação cuidadosa da conexão entre tecnologia e os direitos individuais, pois, obter dados sem a permissão das pessoas gera dilemas éticos, mostrando a importância de discutir os limites da intervenção tecnológica no dia a dia. Assim, promovendo a privacidade em um contexto onde a coleta de dados é frequente (Sousa *et al*., 2025).

As normas para desenvolver inteligências artificiais devem ser baseadas em responsabilidade social, a ética é fundamental. É importante assegurar que os direitos humanos e as liberdades dos cidadãos sejam protegidos, mesmo com o desenvolvimento da tecnologia (Sousa, 2025).

É essencial que as regras morais estejam em primeiro lugar ao usar a inteligência artificial. Em vez de destacar apenas a eficiência, o principal propósito deve ser garantir justiça e igualdade. Isso significa priorizar a transparência nos processos e proteger os direitos individuais (Lucas; Santos, 2021).

Novas leis estão começando a debater como os produtos digitais são feitos, apontando um problema que antes era ignorado: o uso de truques de design para influenciar os usuários e fazê-los passar mais tempo nas plataformas. Com a inteligência artificial, essas práticas se tornaram ainda mais potentes (Pizarro, 2024).

Sendo assim, as regras morais na inteligência artificial não devem ser vistas como obstáculo para o progresso tecnológico, mas como um instrumento para garantir um futuro mais justo e sustentável. Manter uma avaliação moral contínua e promover conversas abertas são essenciais para fortalecer a confiança das pessoas nas novas tecnologias. Investir em educação é fundamental para garantir que a tecnologia tenha um impacto assertivo na sociedade (Sousa, 2025).

O trabalho ressalta que para garantir que os benefícios da IA na assistência ao paciente sejam alcançados de forma ética e segura, é crucial que haja regulamentações fortes e uma fiscalização rigorosa. Isso é necessário para evitar o aumento das desigualdades e a criação de novos riscos, garantindo que a tecnologia seja usada de maneira responsável (Dantas *et al*., 2024).

3 CONCLUSÃO

A análise realizada neste trabalho atingiu o objetivo proposto de mostrar o impacto causado pelo crescente uso da Inteligência Artificial (IA) ao fornecer respostas rápidas e soluções prontas, destacando seus efeitos psicológicos, sociais e éticos. A pesquisa buscou examinar a dependência da IA em relação à autonomia humana, à capacidade de tomada de decisão crítica e à responsabilidade ética em um cenário cada vez mais automatizado. Por meio da revisão de literatura, foi possível identificar uma dualidade nos avanços da IA: por um lado, a tecnologia aperfeiçoa a administração dos cuidados médicos e melhora a tomada de decisões clínicas, por exemplo, ao auxiliar no diagnóstico de doenças, por outro, ela introduz riscos significativos de dependência, que comprometem a saúde mental e podem levar à redução de habilidades e competências humanas essenciais, como o pensamento crítico e a busca independente por conhecimento. Dessa forma, o estudo confirmou a natureza multifacetada do problema, respondendo à questão central sobre os efeitos da dependência da IA na autonomia e ética humanas.

As considerações finais do estudo reforçam a necessidade de um equilíbrio consciente na utilização da IA, para que ela atue como uma ferramenta de auxílio e não como um substituto total das capacidades humanas. Os resultados e a discussão apontaram que a alta dependência de sistemas automatizados pode resultar na diminuição da criatividade, da capacidade de pensamento crítico e da intuição. Além dos impactos psicológicos, a pesquisa destacou as implicações éticas urgentes, como a responsabilidade por eventuais erros dos sistemas e o risco de preconceitos presentes nos algoritmos, que podem intensificar a discriminação. Portanto, a principal limitação do estudo reside na rápida evolução da própria tecnologia, exigindo uma reavaliação contínua de seus impactos e das regulamentações necessárias. A recomendação fundamental é que o avanço da IA seja sempre pautado por uma base ética sólida e transparente, protegendo os direitos humanos e a privacidade dos dados.

Para trabalhos futuros, sugere-se a realização de estudos de campo, a fim de quantificar e analisar o nível real de dependência da IA em setores específicos, como na saúde e na ética. Outra proposta é aprofundar a pesquisa sobre o esquema legal brasileiro e internacional para a responsabilização jurídica em casos de falhas algorítmicas (um ponto ainda em desenvolvimento e com brechas). É imprescindível que o diálogo consciente e

responsável sobre a integração entre inovação tecnológica e princípios éticos seja mantido, garantindo que o progresso da IA contribua para um futuro mais justo e sustentável, no qual a tecnologia seja um instrumento de aprimoramento e não de limitação da capacidade humana. A vigilância ética e a educação sobre o uso adequado dessas ferramentas são cruciais para que a sociedade não se torne meramente receptora de informações pré-selecionadas por máquinas.

REFERÊNCIAS

ALVES, Pedro André Brites. **A Dependência da Internet:** Efeitos na saúde. 2014. Dissertação (Mestrado em Sistemas e Tecnologias da Informação para a Saúde) Instituto Superior de Engenharia, Coimbra, 2014. Disponível em: https://core.ac.uk/download/pdf/62708011.pdf. Acesso em: 2 abr. 2025.

BERNAL, Juan Guillermo Díaz. **Encontros da tecnologia e sociedade da informação:** perspectivas da filosofia da educação no século XXI. Uberlândia: [s.n.]. 2017.

BRUNO, Fernanda; FALTAY, Paulo; LERNER, Alice; STRECKER, Helena. IA emocional e design capcioso: a questão da soberania para a subjetividade. **Liinc em Revista**, [S. l.], v. 20, n. 2, 2024. DOI: 10.18617/liinc.v20i2.7311. Disponível em: https://revista.ibict.br/liinc/article/view/7311. Acesso em: 2 abr. 2025.

DANTAS, Daniela Alves. et al. **Inteligência artificial na tomada de decisão clínica:** Impactos, ética e eficiência. [S.l.]: Científica Digital, 2024.

FALL, Mikal J; ECKHARDT, Christopher D; McMULLEN, Patricia A; FARMER, Steven J. Effect of cognitive impairment on the ability to learn fall prevention strategies in older adults. **Physical Therapy**, v. 98, n. 4, p. 293-302, 2018. DOI: 10.1093/ptj/pzx117. Disponível em: https://pubmed.ncbi.nlm.nih.gov/29507784/. Acesso em: 2 abr. 2025

FONCECA, Marcelo Pires. et al. **Inteligência Artificial na tomada de decisão: ameaça ou oportunidade para gestores? São José dos Pinhais:** Las Ciencias Sociales, 2024.

IPESAÚDE (Sergipe). **Especialista do IPESaúde adverte sobre impacto do excesso do uso de tecnologia na saúde mental.** Disponível em: https://ipesaude.se.gov.br/especialista-do-ipesaude-adverte-sobre-impacto-do excesso-do-uso-de-tecnologia-na-saude-mental/. Acesso em: 2 abr. 2025.

LAUAR, José Amin De Gusmão. et al. **Inteligência Artificial (IA) E Sua Aplicação Na Saúde.** [S.l.]: IOSR Journal Of Humanities And Social Science, 2024.

MARR, Bernard. **Os 15 maiores riscos da inteligência artificial.** Forbes Brasil, 5 jun. 2023. Disponível em: https://forbes.com.br/forbes-tech/2023/06/os-15-maiores riscos-da-inteligencia-artificial/. Acesso em: 12 maio 2025.

MENDES, Claudia, **Proteção de dados pessoais:** privacidade versus avanço tecnológico. Rio de Janeiro: Anja Czymmeck, 2019.

NAVARIN, Nicolò; DONADINI, Irene; DE RIZ, Laura. Artificial Intelligence in the Workplace. **Diagnostics** (Basel), v. 12, n. 3, p. 567, 2022. DOI: 10.3390/diagnostics12030567. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC8959062/. Acesso em: 2 abr. 2025.

PIZARRO, Carolina V. Projetar eticamente na era da IA: o papel do designer enquanto agente humanizador de tecnologias. In: CONFERÊNCIA LATINO AMERICANA DE ÉTICA EM INTELIGÊNCIA ARTIFICIAL, 1, 2024, Niteroi. **Anais** [...]. Porto Alegre: Sociedade Brasileira de Computação, 2024. p. 1-4. DOI: https://doi.org/10.5753/laai-ethics.2024.32437.

SANTOS, Alanis Soares dos; SIMÕES, Lucas de Brito; NEVES, João Emmanuel D Alkmin. **Inteligência Artificial e suas Dependências na Vida Humana**. Campinas: Revista Brasileira em Tecnologia da Informação, 2023.

SANTOS, Denise Oliveira dos; LUCAS, Luciana Berbigier. **Considerações sobre os desafios jurídicos do uso da Inteligência Artificial na medicina.** Porto Alegre: Revista da Faculdade de Direito da UFRGS, 2021.

SILVA, Rodrigo da Guia; NOGAROLI, Rafaella**. Inteligência artificial na análise diagnóstica:** benefícios, riscos e responsabilidade do médico. Rio de Janeiro: Thomson Reuters Brasil, 2020.

SOUSA, Elineuda do Socorro Santos Picanço; PEREIRA, Karin Regina de Bem; VARGAS, Sandra Regina de; FONSECA, Luciana da; RONCATO, Lúcia Lopes Borges. a ética da inteligência artificial: implicações sociais e responsabilidade. **MISSIONEIRA**, Santo Ângelo. v. 27, n. 2, p. 3-13, mai. 2025. Disponível em: https://cemipa.com.br/revistas/index.php/missioneira/article/view/23/12. Acesso em: 2 abr. 2025.

6

A INTELIGÊNCIA ARTIFICIAL NA SAÚDE: PRECISÃO E CONFIABILIDADE NO DIAGNÓSTICO MÉDICO

ARTIFICIAL INTELLIGENCE IN HEALTHCARE: PRECISION AND RELIABILITY IN MEDICAL DIAGNOSIS

Ana Gabrielle Silva de Souza
Mirian Nunes de Carvalho
Ivone Ascar Sauáia Guimarães

Resumo

A Inteligência Artificial tem ganhado destaque na área da saúde por sua capacidade de auxiliar na identificação de doenças e na análise de exames médico, utilizando algoritmos e técnicas de aprendizado de máquina capazes de reconhecer padrões em imagens, resultados laboratoriais e históricos clínicos, o que possibilita diagnósticos mais rápidos, muitas vezes, com alta taxa de acerto, representando um apoio significativo para os profissionais de saúde na tomada de decisões e na melhoria do atendimento aos pacientes. O objetivo deste estudo é analisar a assertividade dos diagnósticos gerados por IA, buscando responder se eles podem ser considerados precisos e confiáveis, com base em uma revisão bibliográfica que reuniu estudos recentes sobre o uso da IA em diagnósticos médicos. Apesar dos avanços, a tecnologia enfrenta desafios, como a dependência da qualidade dos dados utilizados no treinamento dos sistemas, que pode comprometer os resultados quando esses dados são incompletos ou apresentam vieses, além da necessidade de supervisão constante por parte de médicos, já que a IA não substitui a experiência clínica, e da exigência de maior transparência nos algoritmos, permitindo que os profissionais compreendam de que forma os diagnósticos foram obtidos. Diante disso, as considerações finais apontam que a inteligência artificial, possui grande potencial para transformar o diagnóstico médico, oferecendo mais rapidez, precisão e eficiência, mas sua eficácia deve ser avaliada de maneira crítica e seus resultados entendidos como um recurso de apoio complementar ao trabalho humano, e não como substitutos da análise clínica realizada por especialistas.

Palavras-chave: Exames clínicos. Aprendizado de máquina. Eficiência.

Abstract

Artificial Intelligence (AI) has gained prominence in the healthcare field for its ability to assist in identifying diseases and analyzing medical exams, using algorithms and machine learning techniques capable of recognizing patterns in images, laboratory results, and clinical histories. This enables faster diagnoses, often with a high accuracy rate, representing significant support for healthcare professionals in decision-making and improving patient care. The objective of this study is to analyze the accuracy of diagnoses generated by AI, seeking to determine if they can be considered precise and reliable, based on a literature review that compiled recent studies on the use of AI in medical diagnoses. Despite the advances, the technology faces challenges, such as the dependence on the quality of the data used in training the systems, which can compromise the results when this data is incomplete or biased, as well as the need for constant supervision by physicians, since AI does not replace clinical experience, and the requirement for greater transparency in the algorithms, allowing professionals to understand how the diagnoses were obtained. Therefore, the final considerations indicate that artificial intelligence has great potential to transform medical diagnosis, offering greater speed, accuracy, and efficiency, but its effectiveness must be critically evaluated and its results understood as a complementary support resource to human work, and not as a substitute for clinical analysis performed by specialists.

Keywords: Clinical examinations. Machine learning. Efficiency.

1 INTRODUÇÃO

A inteligência artificial, desenvolvida no campo da ciência da computação, passou a ter grande importância para o avanço de várias áreas do conhecimento. Na saúde, seu uso trouxe mudanças relevantes, especialmente na forma de realizar diagnósticos médicos, garantindo mais precisão e confiança nos resultados.

O desenvolvimento de tecnologias baseadas em algoritmos de aprendizado de máquina e redes neurais possibilitou a redução de erros, o aprimoramento da tomada de decisões e a otimização do tempo destinado à investigação de enfermidades. Assim, constatou-se que a aplicação da inteligência artificial no contexto médico representou um marco importante para melhoria da qualidade e da segurança nos serviços prestados à população.

Justificou-se a realização desse trabalho pela importância de analisar a confiabilidade dos diagnósticos médicos realizados com o apoio da inteligência artificial. Essa análise permitiu compreender se a IA poderia ter desempenho igual ou melhor que os métodos convencionais, sendo esse o objetivo geral do estudo.

Foram considerados aspectos como a exatidão de respostas, a estabilidade dos resultados em diferentes situações e a capacidade de reduzir erros, ampliando a segurança no processo diagnóstico. Buscou-se compreender de que maneira essas ferramentas organizaram, analisaram e transformaram grandes volumes de dados em resultados úteis e confiáveis, além de identificar seus desafios e limitações, sendo esse o objetivo específico.

2 DESENVOLVIMENTO

2.1 Metodologia

A pesquisa desenvolvida adotou como método uma revisão bibliográfica, com o objetivo de reunir e analisar estudos que abordam a utilização da Inteligência artificial na área da saúde, enfatizando sua precisão e confiabilidade no diagnóstico médico. Para a coleta de informações, foram consultadas publicações disponíveis no Google Acadêmico, priorizando trabalhos publicados nos últimos cinco anos, a fim de contemplar avanços recentes sobre o tema.

Foram definidos como critérios de inclusão os artigos escritos no idioma português, inglês e espanhol, que apresentassem resultados e análises relevantes sobre a aplicação da IA em diagnósticos médicos, destacando aspectos de desempenho, confiabilidade e acurácia dos sistemas. Quanto aos critérios de exclusão, foram descartados textos que se limitassem a apresentar resumos ou primeiras impressões, não atendendo aos objetivos propostos.

Foram identificadas 19 publicações, entre artigos científicos, dissertações e livros. Após a leitura dos resumos e aplicação dos critérios de inclusão e exclusão, 12 estudos foram selecionados para análise integral, refletindo os avanços recentes sobre a precisão e a confiabilidade da Inteligência Artificial no diagnóstico médico. As palavras-chave utilizadas foram “inteligência artificial”, “medicina” e “diagnóstico médico”.

2.2 Resultados e Discussão

A inteligência artificial tem se tornado cada vez mais importante na medicina, ajudando a melhorar diagnósticos, tratamentos e o gerenciamento de informações dos pacientes. Segundo Magalhães (2024), a incorporação da IA na medicina não representa apenas um avanço tecnológico, mas sim uma transformação que redefine os limites da prática médica.

Conforme De Morais (2023), a tecnologia tem demonstrado utilidade na detecção precoce de diferentes patologias, incluindo casos oncológicos e cardíacos. Sendo assim, a IA tem se mostrado uma ferramenta essencial para obtenção de diagnósticos precisos e confiáveis na medicina.

Por exemplo, em radiologia, a IA pode ajudar a identificar e delinear automaticamente estruturas anatômicas, permitindo a avaliação detalhada de órgãos e tecidos em diferentes planos, isso é particularmente valioso para planejamento cirúrgico (De Morais, 2023). Além disso, a integração entre IA e sistemas hospitalares possibilita a criação de banco de dados cada vez mais robustos, que auxiliam na pesquisa científica.

A Inteligência Artificial é capaz de identificar padrões indicativos de doenças em estágios iniciais, muitas vezes imperceptíveis ao olho humano. Isso é particularmente evidente na detecção precoce de cânceres, onde a IA pode destacar microcalcificações, massas ou outras alterações sutis, permitindo intervenções precoces e tratamentos mais eficazes (De Morais, 2023). Um diagnóstico rápido é essencial para iniciar o tratamento precocemente.

Esses sistemas são capazes de analisar grandes volumes de dados clínicos, identificar padrões complexos e comparar informações médicas com bases de conhecimentos já existentes. A IA tem demonstrado impressionante precisão na detecção de doenças, como câncer de pele e pneumonia, através da análise de imagens médicas (De MORAIS, 2023).

Com a possibilidade de identificar precocemente anomalias ou indícios de patologias em estágios iniciais, aumentando significativamente as chances de recuperação do paciente, cada vez mais médicos vêm incorporando a IA em sua rotina de trabalho. A IA é cada vez mais empregada na área da saúde como um segundo cérebro, um ser pensante, utilizada como auxílio no diagnóstico de doenças e no atendimento a pacientes (NUNES, GUIMARÃES e DADALTO, 2022).

O Machine Learning, ou aprendizado de máquina, é um tipo de algoritmo de IA que permite aos sistemas aprenderem a partir de dados históricos, sem a necessidade de programação manual para cada situação. Pode-se reconhecer que o fornecimento de um diagnóstico rápido por um software com inteligência artificial pode ser, muitas vezes, fator crucial para o imediato início do tratamento (NOGAROLI E SILVA, 2020).

Em emblemático exemplo do desenvolvimento da Inteligência Artificial aplicada à seara de diagnósticos médicos, pesquisadores da Universidade de Oxford (Inglaterra) desenvolveram, no Hospital John Radcliffe, um aparelho inteligente, que, por meio de machine learning, propõe o diagnóstico de doenças cardíacas (NOGAROLI E SILVA, 2020). Dessa forma, o Machine Learning torna-se uma ferramenta que contribui para diagnósticos mais rápidos e precisos.

O machine learning surge como uma solução eficiente para transformar dados brutos em informações estratégicas, capazes de apoiar decisões clínicas, otimizar processos internos e melhorar a qualidade do cuidado ao paciente. Com a utilização de tecnologias de computação cognitiva e machine learning na análise dos bancos de dados de hospitais, aumenta-se a eficiência do atendimento por meio do monitoramento de dados vitais (NOGAROLI E SILVA, 2020).

Apesar dos avanços significativos da IA na área da saúde, persistem questionamentos acerca de sua precisão. Uma das principais discussões refere-se à possibilidade de a IA substituir o papel do médico no processo diagnóstico, conforme Soares (2023), embora os benefícios da IA sejam evidentes, essas tecnologias não substituem o papel do médico – elas servem como ferramentas para aprimorar decisões e fortalecer a relação médico-paciente.

Simões, Mattar e Morato (2024), afirmam que o médico continua sendo essencial no atendimento, pois é ele quem realiza o raciocínio clínico necessário para compreender a origem da patologia e, a partir disso, solicitar os exames complementares quando necessários. Essa atuação humana contribui para fortalecer o relacionamento médico e paciente.

Para Raulin e Angel (2025), a confiança, empatia e comunicação devem permanecer sendo os pilares dessa relação, com a IA sendo utilizada como uma ferramenta que enriquece, e não compromete a prática médica. Servindo, portanto, como um apoio, a IA contribui para o aprimoramento e melhoria da qualidade dos cuidados prestados aos pacientes.

A assertividade dos diagnósticos feitos por IA comparado com aqueles produzidos por médicos especialistas gera divergências de opiniões. Em um trabalho realizado com o tema “Algoritmo versus humanos”, Pereira (2021) afirma que, a racionalidade total só poderia ser alcançada por fórmulas matemáticas ou programas de computador. Por outro lado, o uso da intuição é um aspecto importante do processo de decisão.

Pereira (2021) afirmou que, as ferramentas computacionais são limitadas e o que algoritmo utilizado para a decisão racional apresentou comportamento mais similar a um especialista intuitivo humano, portanto, este comportamento mais similar a um especialista intuitivo e a possibilidade de processamento de diversos formatos de dados podem afetar a racionalidade das decisões.

Dessa forma, ele demonstra uma posição equilibrada, reconhecendo o valor das ferramentas computacionais, mas destacando que a intuição humana continua sendo um elemento essencial no processo decisório. Por sua vez, Madriz (2024) diz que um diagnóstico preciso e oportuno é a base de um atendimento de saúde eficaz. No entanto, os métodos tradicionais enfrentam limitações devido a subjetividade humana e ao excesso de dados disponíveis.

Conforme Nogaroli e Silva (2020), não se trata de pugnar por uma substituição dos profissionais da saúde por sistema de IA, mas tão somente de reconhecer os potenciais benefícios dessa nova tecnologia no que tange, sobretudo, ao auxílio dos profissionais na tomada de decisão. Complementando a ideia de que a IA atua como suporte.

O aumento da quantidade de dados disponíveis favorece a precisão dos sistemas de apoio à decisão médica, o que contribui para melhor compreensão dos diagnósticos e tratamento de diversas condições de saúde (NOGAROLI E SILVA, 2020). A tabela a seguir apresenta prós e contras da utilização da IA na tomada de decisões relacionadas ao diagnóstico.

Tabela 1. Prós e contras da Inteligência Artificial no diagnóstico

Autor	Posição	Argumento
Nogaroli e Silva (2020)	A favor	Os sistemas de IA oferecem suporte crucial à decisão clínica por sua capacidade de processar e analisar grandes volumes de dados de forma rápida e eficiente.
Nunes, Guimarães e Dadalto (2022)	Contra	A Inteligência Artificial, se utilizada como sistema de tomada de decisão, pode levar a erros sobre a conduta médica a ser seguida, dependendo da capacidade do sistema para identificar o problema que acomete o paciente, o que eliminaria uma das vantagens de seu uso.
De Morais (2023)	A favor	A tecnologia de inteligência artificial está revolucionando a medicina diagnóstica, proporcionando avanços significativos na precisão, eficiência e personalização dos cuidados de saúde. Embora haja desafios a serem superados, a colaboração multidisciplinar e o desenvolvimento de regulamentações adequadas podem garantir que a IA seja uma ferramenta valiosa para melhorar a saúde e o bem-estar dos pacientes.
Simões, Mattar e Morato (2024)	Contra	Um dos principais riscos associados ao uso da IA na medicina reside na possibilidade de diagnósticos equivocados ou tardios, logo, podem resultar em tratamentos inadequados ou na ausência de intervenção médica oportuna, comprometendo seriamente a saúde e o bem-estar do paciente.
Lanzagorta-Ortega, CarrilloPérez e Carrillo-Esper (2022)	A favor	A IA também reduzirá a frequência de erros médicos e aumentará a precisão do diagnóstico por meio da integração, análise e interpretação de informações por algoritmos e softwares

Fonte: Adaptado pelo autor com base em Nogaroli e Silva (2020). Nunes, Guimarães e Dadalto (2022). De Morais (2023). Simões, Mattar e Morato (2024). Lanzagorta-Ortega, Carrillo-Pérez e Carrillo-Esper (2022)

Percebe-se que a Inteligência Artificial é capaz de processar grandes volumes de dados clínicos de forma rápida e estruturada. De acordo com Lanzagorta-Ortega, Carrillo-Pérez e Carrillo-Esper (2022), a IA avança no desenvolvimento de algoritmos capazes de realizar triagens diagnósticas e de tratamento para doenças comuns, utilizando informações de histórico médico e sinais clínicos dos pacientes.

Após analisar a tabela, os prós e contras levantam vários questionamentos relacionados a aplicação da IA, apesar de ser vista para alguns de maneira positiva, para outros sua rápida implementação coloca desafios substanciais à formação médica e à prática da medicina (NOGUEIRA, 2020).

Há numerosos desafios e limitações no que diz respeito à aplicação da Inteligência Artificial. Um desses obstáculos à implementação da IA na medicina é a desconfiança quanto à objetividade e segurança dessas tecnologias, por parte dos médicos e dos cidadãos (NOGUEIRA, 2020).

Dentre esses obstáculos, sobressaem-se as questões de caráter ético. Para garantir uma implementação responsável da IA na Medicina, é crucial estabelecer diretrizes éticas, legais e sociais que preservem os direitos humanos e assegurem o uso seguro e justo dessas tecnologias (Saad, 2024). A tabela 2 apresenta os desafios éticos apontados por dois autores em relação ao uso da IA.

Tabela 2. Desafios éticos relacionados à Inteligência Artificial

Autor	Ano	Desafios éticos apontados
Saad, Maria Auxiliadora	2024	Ausência de regulamentação e discos locais; Falta de uma política governamental sobre o tema; Inexistência de comitês de IA em instituições de saúde; Robotização na relação com o paciente; Falhas na privacidade de dados; Dificuldade de escolha das tecnologias de IA mais adequadas; Problemas relacionados à autonomia do profissional de saúde; Condutas imorais; Erros de diagnósticos.
Nunes, Guimarães e Dadalto	2022	Responsabilidade civil da equipe de saúde devido à decisão apoiada em IA; Formação deficitária das equipes de saúde para lidar adequadamente com o manejo desse sistema; Confiança dos profissionais da saúde nesse sistema; integridade científica no processo de construção da IA; Proteção e compartilhamento dos dados sensíveis captados para alimentar o sistema.

Fontes: Adaptado pelo autor com base em Saad (2024). Nunes, Guimarães e Dadalto (2022)

Todavia para Nunes, Guimarães e Dadalto (2022), no que tange às novas tecnologias, tão importante quanto se preocupar com a ética nos testes é atentar-se à integridade científica das pesquisas, à veracidade dos dados, ao consentimento do titular das informações e, principalmente, à lisura e ao rigor científico do projeto apresentado.

Complementando essa discussão, Nogueira (2024) enfatiza que, após a identificação dos desafios, é fundamental propor e aplicar estratégias que visem à mitigação de riscos e à promoção de um uso responsável e seguro da tecnologia.

É indiscutível que mesmo com os inúmeros desafios relacionados a Inteligência Artificial tem desempenhado um papel fundamental no aprimoramento de processos e contribuindo de forma significativa. Apesar dos desafios, a IA oferece oportunidades substanciais para aprimorar a detecção precoce, personalização do tratamento e otimização dos recursos de saúde (MAGALHÃES, 2024).

Além disso, muitos profissionais da área médica têm demonstrado receptividade quanto ao uso da IA, reconhecendo seu potencial. Esse reconhecimento evidencia que, quando utilizada de forma ética e responsável, pode ser vista como uma aliada indispensável. De qualquer forma, não há dúvidas de que a Inteligência Artificial continuará a se consolidar na medicina e a promover avanços significativos no futuro (SOARES, 2023).

3 CONCLUSÃO

O presente estudo teve como objetivo analisar a assertividade dos diagnósticos realizados por Inteligência Artificial e verificar se esses sistemas podem ser precisos e confiáveis no contexto médico. A pesquisa demonstrou que a IA tem se mostrado uma ferramenta promissora no auxílio à identificação de doenças e na interpretação de exames clínicos, oferecendo resultados mais rápidos e contribuindo para decisões médicas mais seguras.

Assim, considera-se que o objetivo proposto foi alcançado, uma vez que ficou evidenciado o potencial da IA como suporte ao trabalho dos profissionais de saúde. Entretanto, observou-se que ainda existem limitações dessas tecnologias a prática clínica, tais fatores podem comprometer a precisão dos diagnósticos e a confiança dos profissionais na utilização desses sistemas, reforçando que a experiência humana permanece indispensável.

Diante desses resultados recomenda-se que futuras pesquisas aprofundem a discussão sobre diretrizes éticas, legais e técnicas que orientem o uso responsável da Inteligência Artificial na medicina. Concluo, portanto, que a IA é um grande aliado dos médicos quando usada de forma responsável, contribuindo significativamente para a evolução da medicina e a melhoria do cuidado ao paciente.

REFERÊNCIAS

DE MORAIS, J. J. *et al.* **Impacto da tecnologia de Inteligência Artificial na medicina diagnóstica.** *Revista Ibero-Americana de Humanoides, Ciências e Educação*, v. 9, n. 7, p. 1303-1214, 2023. Disponível em: https://periodicorease.pro.br/rease/article/view/10699. Acesso em: 27 set. 2025. (Modelo revista).

LANZAGORTA-ORTEGA, D.; CARRILLO-PÉREZ, D. L.; CARRILLO-ESPER, R. **Inteligência Artificial em medicina: presente y futuro.** México: *Gaceta Médica de México*, v. 158,p.17-21,2022. Disponível em: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0016-38132022001100017&lng=es&nrm=iso. Acesso em: 27 set. 2025. (Modelo revista).

MADRIZ, L. J. S. *et al*. **Inteligência Artificial aplicada al diagnóstico médico: uma revisión actual.** *Revista Científica de Salud y Desarrollo Humano*, v. 5, n. 2, p. 274-288, 2024. Disponível em: https://revistavitalia.org/index.php/vitalia/article/view/183. Acesso em: 27 set. 2025. (Modelo revista).

MAGALHÃES, M. I. S. *et al.* **Impacto da Inteligência Artificial no diagnóstico médico: desafios e oportunidades**. *Revista Ibero-Americana de Humanoides, Ciências e Educação*, v. 10,n.1,p.1477-1485,2024. Disponível em: https://periodicorease.pro.br/rease/article/view/12819. Acesso em: 27 set. 2025. (Modelo revista).

NOGUEIRA, C. A. V**. Desafios na implementação da Inteligência Artificial na medicina: uma revisão narrativa.** 2024. Disponível em: http://hdl.handle.net/10400.5/100449. Acesso em: 27 set. 2025. (Modelo trabalho).

NUNES, H. C.; GUIMARÃES, R. M. C.; DADALTO, L. **Desafios bioéticos do uso da inteligência artificial em hospitais.** *Revista Bioética*, 2022. Disponível em: https://doi.org/10.1590/19838042022301509PT. Acesso em: 27 set. 2025. (Modelo revista).

NOGAROLI, R.; SILVA, R. G. **Inteligência Artificial na análise diagnóstica: benefícios, riscos e responsabilidade do médico.** In: *Debates Contemporâneos em Direito Médico e da Saúde*. São Paulo: Thomson Reuters Brazil, 2020. p. 69-91. Disponível em: https://www.researchgate.net/profile/Rodrigo-Silva-47/publication/343881300. Acesso em: 27 set. 2025. (Modelo trabalho).

PEREIRA, D. V. C. **Algoritmo versus humanos: a assertividade e comportamento da racionalidade e intuição no processo de tomada de decisão**. Brasília, 2022. Disponível em: http://repositorio.unb.br/handle/10482/44651. Acesso em: 27 set. 2025. (Modelo trabalho).

RAULIN, M. L. F.; ANGEL, D. J**. Inteligência Artificial na medicina: impactos e desafios.** *Revista Ibero-Americana de Humanoides, Ciências e Educação*, v. 11, n. 1, p. 2801-2814, 2025. Disponível em: https://periodicorease.pro.br/rease/article/view/18024. Acesso em: 27 set. 2025. (Modelo revista).

SAAD, M. A. N. *et al.* **Inteligência Artificial na medicina: desafios éticos e a ausência de regulamentações e diretrizes locais.** Niterói: Conferência Latino-Americana de Ética em Inteligência Artificial. SBC, p. 97-100, 2024. Disponível em: https://doi.org/10.5753/laai-ethics.2024.32461. Acesso em: 27 set. 2025. (Modelo trabalho).

SIMÕES, S. C. A.; MATTAR, J. G.; MORATO, B. B**. Inteligência Artificial aplicada à saúde: riscos, benefícios e limites.** Belo Horizonte: Anais do Simpósio Brasileiro de Inteligência Artificial, p. 13, 2023. Disponível em: https://editorapascal.com.br/wpcontent/uploads/2024/09/SIMPOSIO-DE-IA.pdf. Acesso em: 27 set. 2025. (Modelo trabalho).

SOARES, R. A. *et al.* **O uso da Inteligência Artificial na medicina: aplicações e benefícios.** *Research, Society and Development*, v. 12, n. 4, p. e5012440856-e5012440856, 2023. Disponível em: https://rsdjournal.org/rsd/article/view/40856. Acesso em: 27 set. 2025. (Modelo revista).

7

A DIFERENÇA DAS ESTRATÉGIAS DE DESENVOLVIMENTO PARA ENGENHARIA DE SOFTWARE

DIFFERENCES BETWEEN SOFTWARE DEVELOPMENT STRATEGIES

Willame Silva Oliveira Filho
Tayssara Elizavieta Martins Varão
Mirian Nunes de Carvalho Nunes

Resumo

Este trabalho aborda o problema da crescente necessidade de metodologias ágeis ao invés da abordagem tradicional no desenvolvimento de um software, discutindo em quais contextos cada abordagem pode ser mais adequada. O objetivo geral é descrever as diferenças entre as metodologias ágeis e a tradicional na engenharia de software, analisando suas limitações e aplicabilidades em diferentes contextos. A pesquisa foi realizada por meio de uma revisão de literatura de natureza qualitativa e descritiva, com base em livros, dissertações e artigos científicos, consultados em bases de dados como Google Acadêmico e SciELO. As considerações apontam que não há uma metodologia plenamente superior; O estudo indica que a escolha de abordagem pode depender de fatores relacionados às características do projeto, algo aprofundado ao longo do trabalho.

Palavras-chave: Metodologias ágeis. Modelo cascata. Processos de software. Gerenciamento de projetos. Scrum.

Abstract

This study addresses the problem of the growing need for agile methodologies instead of the traditional approach in software development, discussing the contexts in which each approach may be more appropriate. The overall objective is to describe the differences between agile and traditional methodologies in software engineering, analyzing their limitations and applicability in different contexts. The research was conducted through a qualitative and descriptive literature review, based on books, dissertations, and scientific articles consulted in databases such as Google Scholar and SciELO. The findings indicate that there is no fully superior methodology; the study suggests that the choice of approach may depend on factors related to the project's characteristics, an aspect explored throughout the work.

Keywords: Agile methodologies. Waterfall model. Software processes. Project management. Scrum.

1 INTRODUÇÃO

A engenharia de software é um campo dinâmico que busca aplicar princípios de engenharia como sistematização, padronização, controle de qualidade e gerenciamento de riscos, para criar sistemas de software de alta qualidade. Ao realizar qualquer tarefa, independentemente da maneira na qual essa seja realizada, é necessário seguir um passo a passo, um roteiro com sequência de ações Da mesma forma, o desenvolvimento de software exige planejamento prévio, determinando finalidades, funcionalidades e ferramentas. Segundo Pressman (2016), esse processo envolve um conjunto estruturado de atividades que orienta a construção do sistema.

Com base nessa informação, compreende-se que é necessário planejamento para o desenvolvimento de um software. Nesse contexto, torna-se importante ressaltar a existência de duas abordagens principais na elaboração deste planejamento: a metodologia tradicional e as metodologias ágeis. A abordagem tradicional caracteriza-se por planejamento e documentação rígidas, enquanto as metodologias ágeis partem de um pensamento mais flexível, lidando melhor com imprevistos e interação contínua com o cliente.

Atualmente, o desenvolvimento de software progride constantemente, e um dos fatores principais para o sucesso do projeto é a escolha e a prática de determinada metodologia. A relevância do estudo justificou-se pela necessidade de compreender de maneira comparativa as duas abordagens, evidenciando suas principais diferenças, casos de uso e limitações diante diferentes projetos de software, impactando diretamente prazos, custos, qualidade e satisfação do cliente.

Diante dessas descrições, percebe-se que não basta apenas compreender as diferenças entre as metodologias; é necessário também identificar em quais cenários cada abordagem tende a ser mais eficaz. Nesse sentido, define-se o problema central desta pesquisa: em quais situações a metodologia tradicional e as metodologias ágeis se mostram mais adequadas para o desenvolvimento de software, considerando seus impactos nos processos e nos resultados finais dos projetos?

O objetivo geral do trabalho consistiu em descrever as diferenças entre as metodologias ágeis e a metodologia tradicional na engenharia de software, incluindo suas limitações e viabilidades. Já os objetivos específicos foram identificar as principais divergências práticas e conceituais entre as duas metodologias e avaliar em quais situações cada metodologia é mais favorável.

2 DESENVOLVIMENTO

2.1. Metodologia

A pesquisa foi conduzida por meio de uma revisão bibliográfica, de caráter qualitativo e descritivo, método amplamente utilizado para reunir, organizar e analisar conhecimentos já produzidos sobre um tema (MARCONI; LAKATOS, 2017). As buscas conduzidas foram realizadas nas seguintes bases de dados: Google Acadêmico e SciELO, IEEE Xplore, ACM Digital Library, ScienceDirect, bem como livros e dissertações disponíveis em meio digital. As obras utilizadas foram publicadas nos últimos dez anos. A busca considerou como descritores: “metodologias ágeis”, “engenharia de software” e “metodologia tradicional”.

Os critérios de inclusão abrangeram publicações completas, que apresentaram análises comparativas entre metodologias ágeis e a tradicional, enquanto os critérios de exclusão descartaram artigos de opinião, resumos, documentos duplicados ou que não

apresentaram relevância direta ao tema. Esse método de revisão bibliográfica possibilitou identificar, organizar e analisar os estudos já publicados sobre o tema, sem recorrer a experimentações ou estudos de caso.

2.2. Resultados e Discussão

O presente estudo buscou analisar as diferenças entre as metodologias ágeis e a tradicional no desenvolvimento de software, avaliando suas viabilidades em distintos contextos. A pesquisa bibliográfica revelou dados qualitativos e descritivos que permitem compreender com maior profundidade as características e particularidades de cada abordagem.

Conforme o estudo de Mishra e Alzoubi (2023), *Structured software development versus agile software development: a comparative analysis*, a metodologia conhecida como modelo cascata, ou *waterfall*, é um modelo de processo prescritivo, onde a fase atual precisa ser completada antes do início da próxima. Segundo Thesing, Feldmann e Burchardt (2021),o modelo cascata é uma abordagem bem inflexível, adequada para projetos que precisam de planejamento estável e antecipado, com perspectiva de longo prazo.

A metodologia tradicional segue uma sequência de fases, evidenciadas na Figura 1 a seguir.

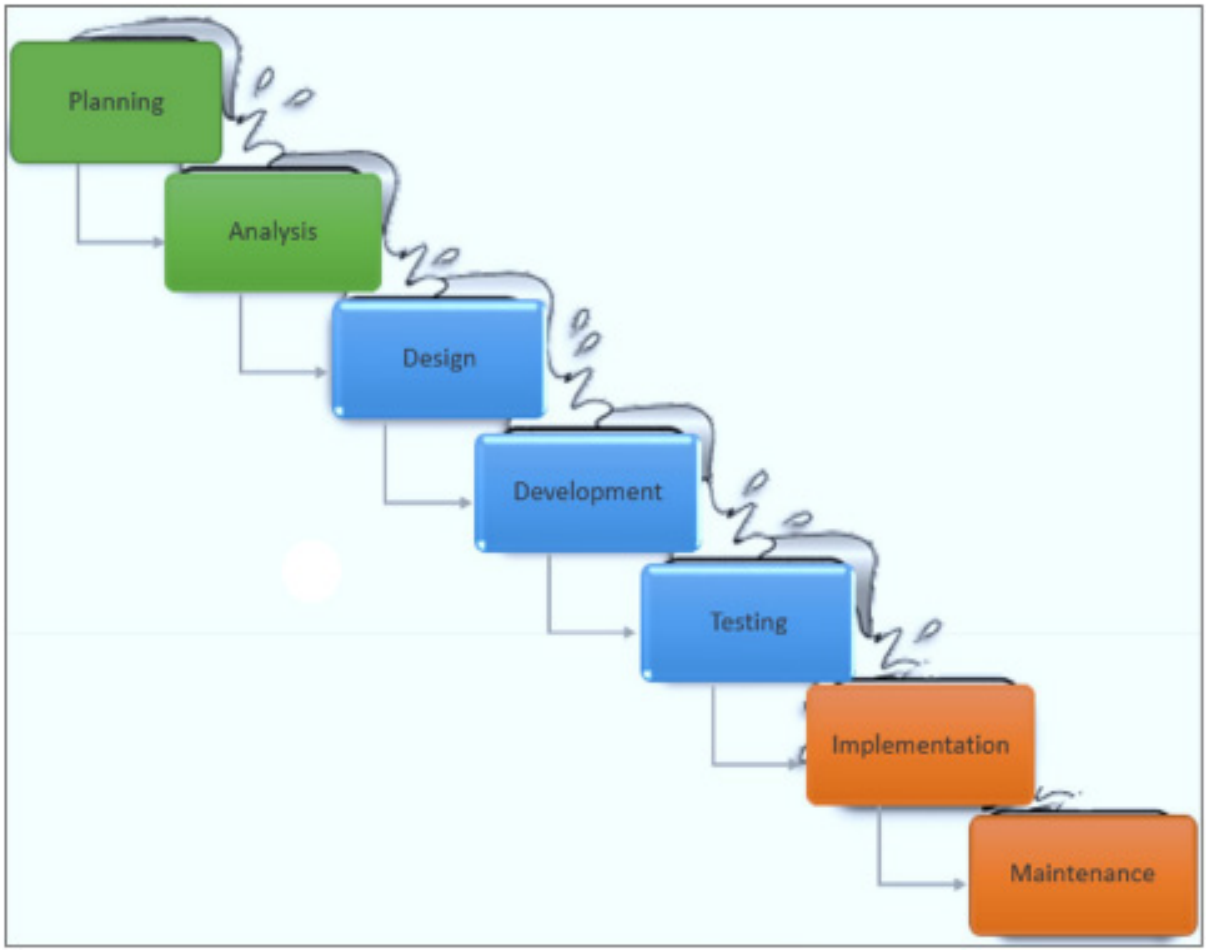

Figura 1. Modelo de desenvolvimento em cascata

Fonte: Adaptado de MISHRA e ALZOUBI (2023)

A imagem demonstra que cada fase depende da conclusão da fase anterior, começando pelo planejamento e análise de requisitos, seguido do projeto do sistema, desen-

volvimento, testes, implementação, e por fim, manutenção. Essa metodologia é indicada para projetos com requisitos estáveis e bem compreendidos, situação que pode ocorrer quando adaptações ou aperfeiçoamentos bem definidos precisam ser feitos em um sistema existente, como aponta Pressman (2016).

Em contrapartida, as metodologias ágeis, como *Scrum, Kanban* e *Extreme Programming* (XP), propõem uma abordagem mais baseada em *feedbacks* e flexibilidade. O *Modern Agile*, proposto por Kerievsky (2016), reforça os princípios de entrega contínua de valor, valorização das pessoas e maximização da sua produtividade. Segundo Indriasari et al. (2022), as metodologias ágeis têm como objetivo gerar produtos de software com qualidade de forma econômica por meio de uma série de ciclos curtos de desenvolvimento, acelerando o seu desenvolvimento em pequenas versões. Dessa forma, nota-se que a agilidade não se limita unicamente à adoção de um conjunto de práticas, mas representa uma mudança cultural com foco em colaboração, foco no cliente e à flexibilidade, diferenciando-se dos modelos tradicionais.

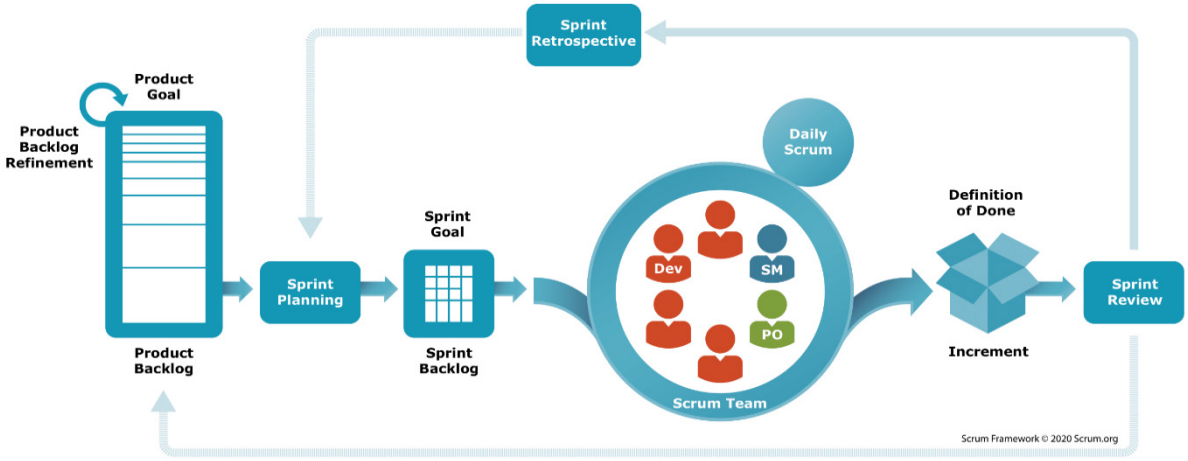

Figura 2. Representação do ciclo de Scrum.

Fonte: Adaptado de SCRUM.ORG (2023).

Como demonstrado na Figura 2, o *Scrum*, um framework de desenvolvimento ágil é caracterizado por ciclos curtos e iterativos chamados Sprints, nos quais a equipe define junto com o cliente, as funcionalidades prioritárias que serão entregues. O processo envolve reuniões diariamente para alinhamento, desenvolvimento incremental com testes automatizados, revisão ao final de cada sprint com apresentação ao cliente e uma retrospectiva para identificar pontos de melhoria. Essa abordagem favorece entregas frequentes, alta adaptabilidade e constante evolução do produto.

Um dos achados mais surpreendentes ao comparar métodos ágeis com o modelo cascata é relacionado às taxas de sucesso de projetos. Informações compiladas de diversos estudos, como o conhecido CHAOS Report (STANDISH GROUP, 2015), revelam uma diferença evidente. Projetos ágeis tiveram 39% de sucesso, enquanto projetos cascata apenas 11%.

Segundo o STANDISH GROUP (2015), projetos em modelo cascata apresentaram 29% de falha e 60% com desempenho insatisfatório, enquanto projetos ágeis tiveram 9% de falha e 52% com desempenho insatisfatório. Essa diferença evidencia que a natureza iterativa do método ágil contribui para maior taxa de sucesso e melhor entrega de valor.

O alto índice preocupante de 29% de fracasso em projetos Cascata pode ser melhor compreendido quando o fracasso não é visto como um incidente isolado, mas sim como um procedimento contínuo. Nizam (2022), ao examinar 22 estudos de caso, apresenta um

modelo processual do fracasso que frequentemente começa com a equipe.

Esses problemas iniciais, muitas vezes percebidos como pequenos alertas sobre falhas ou inconsistências, tendem a ser ignorados em metodologias com fases longas e *feedback* tardio. Pequenos desvios não corrigidos desde o princípio acabam se acumulando ao longo do tempo resultando em complicações críticas durante as etapas finais, uma situação que contribui para explicar o alto índice de fracasso desses projetos (NIZAM, 2022).

O debate detalhado dessas informações nos leva diretamente à questão da confiabilidade da metodologia convencional. Conforme Thesing, Feldmann e Burchardt (2021), suposições equivocadas, requisitos iniciais abstratos e mal interpretados podem ter grande impacto no projeto. O método sequencial que requer a conclusão total de cada fase antes do início da próxima acaba resultando em custos elevados para modificações, uma vez que a rigidez do método dificulta a adaptação a mudanças.

Essa inflexibilidade cria um cenário favorável para a etapa seguinte do processo de falha descrito por Nizam (2022), denominada “Recusa de Desvios”. Nessa fase, a equipe reconhece os problemas, mas fica paralisada pelo alto custo da mudança. Eles tendem a ocultar as más notícias na esperança de resolver a situação sem o conhecimento da gestão, um fenômeno conhecido como *“Mum Effect”*.

Se durante os testes for encontrado um erro de requisito, isso pode se tornar uma correção complexa e custosa, por outro lado, os métodos ágeis com ciclos breves e *feedback* contínuos possibilitam à equipe detectar prontamente erros para que possam ser corrigidos precocemente. Essa característica está alinhada com o *Modern Agile*, que tem como uma das prioridades “identificar problemas antecipadamente”, conforme proposto por Kerievsky (2016).

Um aspecto crucial da pesquisa está ligado à felicidade do cliente e à entrega de valor, elementos recorrentes em relatórios anuais como o *“State of Agile”*. Segundo o relatório de 2023, 52% das organizações afirmam que a principal razão para adotar métodos ágeis é a capacidade de responder rapidamente a mudanças, enquanto 44% destacam a melhoria na satisfação do cliente (DIGITAL.AI, 2023). Esses dados evidenciam que a satisfação do cliente está diretamente relacionada à entrega contínua de resultados funcionais e ao ajuste constante das prioridades ao longo do projeto.

Essa perspectiva sustenta a visão apresentada por Alami, Zahedi e Krancher (2024), que destaca a relevância dos fatores humanos, como a confiança e comunicação aberta, na interação entre cliente e equipe nas abordagens das metodologias ágeis. A manutenção da participação do cliente durante todo o ciclo de desenvolvimento representa um aspecto característico das metodologias ágeis, assegurando que o produto final atenda com maior proximidade às suas expectativas desde as fases inaugurais da entrega.

Os pontos mencionados destacam a relação direta com os objetivos específicos deste artigo. Ao analisar os efeitos das duas estratégias utilizadas observamos que no modelo tradicional enfatiza-se o cumprimento rigoroso de um plano inicial; já no ágil a qualidade é desenvolvida de maneira progressiva e validada de forma contínua (PARGAONKAR, 2023).

A comunicação entre as equipes e o nível de contentamento dos clientes tendem a ser mais elevados em ambientes que adotam as metodologias ágeis, devido aos feedbacks regulares e à colaboração intensiva. Isso está relacionado diretamente ao objetivo deste estudo de avaliar como as duas estratégias influenciam a qualidade do software, a comunicação interna e o contentamento dos clientes.

Embora tenha suas limitações em ambientes dinâmicos, o método *Waterfall* demonstra a suas qualidades em contextos específicos. Uma abordagem sequencial e line-

ar torna o processo de desenvolvimento mais previsível, possibilitando um planejamento claro, com marcos bem estabelecidos e um cronograma direto (PARGAONKAR, 2023). Essa abordagem estruturada é fortalecida pela ênfase na documentação detalhada em cada etapa, que funciona como uma referência importante para futuras manutenções e assegura uma compreensão abrangente dos requisitos do projeto.

Para os administradores de projeto, essa transparência facilitará o controle, possibilitando uma alocação de recursos mais eficiente, um seguimento de progressão mais preciso também é viabilizado. Por esses motivos, esse método convencional é considerado especialmente adequado para projetos mais simples, onde os requisitos são claros desde o princípio (PARGAONKAR, 2023).

Por outro lado, a flexibilidade das metodologias ágeis apresenta seus próprios obstáculos. A necessidade de colaboração contínua e participação ativa de todos os membros da equipe e partes interessadas tornam o processo intensivo em recursos, podendo afetar outras atividades em andamento. Em projetos de grande porte e complexidade, a gestão ágil pode se tornar um obstáculo, já que as múltiplas etapas iterativas contínuas podem dificultar a coordenação global e o cumprimento dos prazos.

A agilidade também é alcançada em parte pela redução da ênfase na documentação detalhada. No entanto, tal flexibilidade pode resultar em registros incompletos sobre algumas decisões de design ao longo do projeto (PARGAONKAR, 2023). No contexto dessa pesquisa, essa característica demonstra a vulnerabilidade das metodologias ágeis em depender fortemente da comunicação contínua. Assim, quando essa troca de informações não ocorre de forma adequada, o processo perde eficiência e pode até deixar de funcionar corretamente, caracterizando um risco operacional relevante.

Conforme Kerievsky (2016), um aspecto importante na utilização do método ágil é a grande importância dada ao cliente. Essa dependência reforça que a eficácia das metodologias ágeis não está somente no processo, mas no comportamento das partes envolvidas. O avanço dos ciclos está atrelado à obtenção de feedbacks frequentes ao envolvimento contínuo do cliente. Se ele não estiver presente ou se os requisitos não estiverem claros, o trabalho de qualidade da equipe de engenharia pode ser comprometido.

Além disso, a receptividade às mudanças que é uma das maiores vantagens do método ágil pode resultar no *"scope creep"*, situação na qual o escopo do projeto se expande para além dos limites originais sobrecarregando os recursos disponíveis. Finalmente, a ênfase na rapidez da entrega pode às vezes obscurecer a importância de testes mais rigorosos colocando em risco o lançamento de um software com defeitos em vigor não resolvidos (PARGAONKAR, 2023). Esse cenário reforça que o método ágil não deve ser interpretado como uma solução universal, pois sua eficácia depende do contexto organizacional e da capacidade das equipes de manter disciplina e alinhamento contínuo ao longo do projeto.

A aplicação de metodologias ágeis em grande escala muitas vezes encontra obstáculos relacionados a questões humanas e organizacionais para além dos desafios procedimentais, conforme evidenciado por um estudo conduzido por Feitosa e Ferreira (2021). O caso analisado em uma instituição financeira no Brasil, cujo nome foi preservado pelos autores, ressaltou que a cultura empresarial pode representar uma barreira significativa ao êxito dos processos ágeis implementados em largas proporções.

Há desafios fundamentais mencionados na implementação e expansão das metodologias ágeis como resistência internamente às mudanças e falta de comprometimento efetivo por parte das lideranças em destaque. Além disso, o desalinhamento nas práticas de gestão como sistemas de incentivo que priorizam o desempenho individual em detrimento do coletivo pode prejudicar os princípios colaborativos das abordagens ágeis de

desenvolvimento de software, causando conflitos e desânimo na equipe.

A análise dos dados indica que não se deve adotar a metodologia de forma dogmatizada; pelo contrário, é preciso considerar o contexto em que será aplicada. O modelo Cascata continua sendo uma opção válida em projetos com requisitos consistentes e escopo bem delimitado que exigem documentação detalhada obrigatória em setores regulamentados específicos (PARGAONKAR, 2023).

Feitosa e Ferreira (2021) identificaram que a complexidade de adaptar a metodologia Ágil para projetos de grande escala está relacionada à estrutura original dessas práticas ágeis mencionadas anteriormente. Modelos como *Scrum* e *Extreme Programming* (XP), inicialmente desenvolvidos para equipes pequenas e bem integradas com comunicação direta e ágil, podem apresentar dificuldades quando aplicados a projetos maiores, devido à necessidade de coordenação entre múltiplas equipes, aumento da burocracia e desafios na manutenção de feedback contínuo.

Quando aplicado em ambientes com grande quantidade de recursos e equipes distribuídas em vários locais diferentes e com extenso volume de código-fonte implementado surge uma situação desafiadora para alcançar padrões adequados tanto em qualidade quanto em desempenho. Thesing, Feldmann e Burchardt (2021) destacam que coordenar múltiplas equipes dispersas geograficamente torna-se um obstáculo que requer novos modelos organizacionais para lidar com a interconexão entre diversos projetos complexos e assim facilitar sua gestão eficiente.

Em contrapartida, na maioria dos projetos de software atuais que se destacam pela incerteza e pela necessidade de se adaptarem rapidamente ao mercado, há uma clara tendência para a adoção de metodologias ágeis, conforme observado por Pargaonkar (2023).

3 CONCLUSÃO

O principal objetivo deste estudo foi descrever as discrepâncias, restrições e possibilidades das metodologias ágeis e tradicionais na engenharia de software. Combinando os dados coletados a partir da revisão bibliográfica realizada, foi possível reconhecer que os objetivos propostos foram alcançados, uma vez que foi possível descrever as principais diferenças entre as abordagens, reconhecendo suas vantagens e limitações e compreender em quais situações cada uma se sobressai. A pesquisa permitiu através da análise, responder ao problema central do estudo de forma clara, evidenciando que a adequação metodológica depende da estabilidade dos requisitos, o ambiente de desenvolvimento e grau de flexibilidade ao decorrer do projeto.

Os resultados discutidos revelam que a metodologia tradicional é a mais indicada em situações nos quais a previsibilidade é fundamental, estruturas bem definidas que não necessitam de mudanças bruscas, enquanto as metodologias ágeis se mostraram mais eficazes em ambientes de grande dinamismo, cooperativos e sujeitos a constantes adaptações. Essas conclusões evidenciam que a escolha da metodologia não é definitiva ou universal, mas uma decisão que deve ser planejada considerando os contextos, a experiência da equipe e o objetivo do projeto. A reflexão conclusiva desse estudo mostrou que a compreensão dessas características é fundamental para que a eficiência e os resultados obtidos no processo sejam maximizados.

Este estudo trouxe uma contribuição significativa tanto para a academia quanto para o ambiente profissional ao resumir informações que podem orientar as decisões estratégicas em projetos de engenharia de software. A seleção da abordagem metodológica não

surge apenas como uma questão de preferência, ela configura-se como um aspecto fundamental para o êxito do projeto e deve ser avaliada com atenção considerando a natureza do projeto em si, a cultura da organização e a experiência da equipe envolvida.

Apesar da análise proporcionar diversos avanços, o estudo apresenta restrições, principalmente pelo fato de se concentrar apenas em características teóricas e comparativas, não incluindo avaliações práticas e pesquisas de campo. Diante deste fato, é recomendado que pesquisas futuras explorem a aplicação das práticas tradicionais e ágeis em conjunto, apurem metodologias híbridas ou examinem estudos de caso que permitam de maneira prática, observar os impactos de cada abordagem nos prazos, qualidade do produto e na satisfação da equipe e do cliente. Tais desdobramentos poderão incrementar os resultados apresentados e ampliar a compreensão sobre a escolha da metodologia mais adequada para diferentes tipos de projetos.

REFERÊNCIAS

ALAMI, Adam; ZAHEDI, Mansooreh; KRANCHER, Oliver. **The role of psychological safety in promoting software quality in agile teams.** Empirical Software Engineering, v. 29, n. 119, p. 1–50, 2024. DOI: 10.1007/s10664-024-10512-1.

DIGITAL.AI. **17th Annual State of Agile Report.** 2023. Disponível em: https://stateofagile.com. Acesso em: 13 set. 2025.

FEITOSA, Leonardo Augusto; FERREIRA, Wagner Solivan. **Desafios da aplicação do ágil escalado em projetos de software: estudo de caso em uma organização financeira.** Revista de Gestão e Projetos (GeP), [S. l.], v. 12, n. 1, p. 195–221, 2021. Disponível em: https://doi.org/10.5585/gep.v12i1.17825. Acesso em: 12 set. 2025.

INDRIASARI, A. et al. **Adoption of Design Thinking, Agile Software Development and Co-creation:** A Qualitative Study towards Digital Banking Innovation Success**.** Procedia Computer Science, 2022. Disponível em: https://ieeexplore.ieee.org/document/9680763/figures#figures. Acesso em: 20 out. 2025.

KERIEVSKY, Joshua. **Modern Agile.** 2016. Disponível em: https://modernagile.org. Acesso em: 23 out. 2025.

MARCONI, Maria A.; LAKATOS, Eva M. **Fundamentos de metodologia científica**. 8. ed. São Paulo: Atlas, 2017.

MISHRA, A.; ALZOUBI, Y. I. **Structured software development versus agile software development: a comparative analysis.** 2023. Disponível em: https://link.springer.com/article/10.1007/s13198-023-01958-5. Acesso em: 12 set. 2025.

NIZAM, A**. Software Project Failure Process Definition.** *IEEE Access*, v. 10, p. 34428–34441, 2022. DOI: 10.1109/ACCESS.2022.3162878.

PARGAONKAR, S. **A Comprehensive Research Analysis of Software Development Life Cycle (SDLC) Agile & Waterfall Model Advantages, Disadvantages, and Application Suitability in Software Quality Engineering.** International Journal of Scientific and Research Publications, v. 13, n. 8, 2023. DOI: 10.29322/IJSRP.13.08.2023.p14015.

PRESSMAN, Roger S.; MAXIM, Bruce R**. Engenharia de software:** uma abordagem profissional. 8. ed. São Paulo: McGraw-Hill, 2016.

SCRUM.ORG. **Scrum Framework Poster.** 2023. Disponível em: https://www.scrum.org/resources/scrum-framework-poster. Acesso em: 23 out. 2025.

STANDISH GROUP. **CHAOS Report 2015: The law of diminishing returns.** 2015. Disponível em: https://www.standishgroup.com/sample_research_files/CHAOSReport2015_rev.pdf Acesso em: 13 set. 2025.

THESING, T.; FELDMANN, C.; BURCHARDT, M. **Agile versus Waterfall Project Management: Decision Model for Selecting the Appropriate Approach to a Project.** 2021. Disponível em: https://www.sciencedirect.com/science/article/pii/S1877050921002702. Acesso em: 12 out. 2025.

8

DESENVOLVIMENTO DE UM SISTEMA WEB PARA AGENDAMENTO DE CLIENTES: UMA REVISÃO DE LITERATURA SOBRE O USO DE BACKEND AS A SERVICE (BAAS)

DEVELOPMENT OF A WEB-BASED CLIENT SCHEDULING SYSTEM: A LITERATURE REVIEW ON THE USE OF BACKEND AS A SERVICE (BAAS)

Marcio Willian Chaves Cardoso
Miriam Nunes de Carvalho Nunes
Tayssara Elizavieta Martins Varão

Resumo

Este artigo apresenta uma revisão de literatura sobre o desenvolvimento de sistemas web para agendamento de clientes, com foco no modelo Backend as a Service (BaaS) e em tecnologias modernas de front-end. O estudo analisou publicações dos últimos dez anos, identificando como tais soluções têm sido utilizadas para otimizar processos de atendimento, reduzir aglomerações e aprimorar a experiência do usuário. A pesquisa evidenciou que sistemas de agendamento online evoluíram significativamente, tornando-se mais responsivos, escaláveis e integrados a serviços em nuvem. Além disso, observou-se que o modelo BaaS contribui para a redução da complexidade do backend, embora apresente limitações relacionadas à dependência de provedores externos e à segurança dos dados. Os resultados demonstram a relevância dessas tecnologias no contexto da transformação digital e destacam desafios futuros para ampliar acessibilidade, flexibilidade e proteção das informações.

Palavras-chave: Sistemas Web; Agendamento de Clientes; Backend as a Service; BaaS; Experiência do Usuário.

Abstract

This article presents a literature review on the development of web systems for client scheduling, with emphasis on Backend as a Service (BaaS) and modern front-end technologies. The study analyzed publications from the last ten years, identifying how such solutions have been used to optimize service processes, reduce crowding, and enhance user experience. The research indicated that online scheduling systems have evolved significantly, becoming increasingly responsive, scalable, and integrated with cloud services. In addition, it was observed that the BaaS model contributes to reducing backend complexity, although it presents limitations related to provider dependency and data security. The results demonstrate the relevance of these technologies in the context of digital transformation and highlight future challenges in expanding accessibility, flexibility, and information protection.

Keywords: Web Systems, Client Scheduling, Backend as a Service, BaaS, User Experience.

1 INTRODUÇÃO

Nas últimas décadas, os sistemas web passaram a desempenhar papel essencial na transformação digital de empresas e instituições públicas. A expansão da internet e a popularização de dispositivos conectados permitiram que diversas atividades presenciais fossem migradas para plataformas digitais, incluindo compras, pagamentos, serviços administrativos e agendamentos. A crescente necessidade por soluções rápidas e acessíveis impulsionou a adoção de sistemas capazes de organizar o fluxo de atendimento e reduzir gargalos operacionais, tornando-se ferramentas fundamentais em diferentes setores.

Esse cenário ganhou destaque particular durante a pandemia de COVID-19, quando a necessidade de distanciamento físico tornou indispensável o uso de tecnologias para evitar aglomerações. Os sistemas de agendamento foram amplamente utilizados por clínicas, comércios e prestadores de serviço, viabilizando a organização do atendimento e a proteção da população. Mesmo após o período crítico, essas tecnologias mantiveram relevância ao oferecer praticidade, otimização do tempo e melhoria na experiência do usuário.

A importância desse tema se justifica pela crescente demanda por soluções tecnológicas que apoiem a organização de serviços, melhorem o relacionamento com o cliente e ampliem a eficiência operacional. Em um ambiente cada vez mais competitivo, empresas buscam ferramentas capazes de oferecer acessibilidade, flexibilidade e integração com múltiplas plataformas. Tecnologias como o Backend as a Service (BaaS) surgem nesse contexto como alternativa moderna para acelerar o desenvolvimento e reduzir custos, representando um recurso relevante tanto para desenvolvedores quanto para gestores de negócios.

Diante disso, este artigo tem como objetivo geral analisar, a partir de uma revisão de literatura, como os sistemas web de agendamento de clientes tem sido desenvolvidos utilizando tecnologias modernas como BaaS, frameworks de front-end e bancos de dados em nuvem. O estudo busca compreender os impactos dessas soluções em termos de eficiência, escalabilidade, segurança e experiência do usuário.

Como objetivos específicos, pretende-se: (i) apresentar conceitos fundamentais sobre sistemas web e sua aplicação no contexto de agendamento; (ii) examinar as vantagens e limitações do modelo Backend as a Service; (iii) identificar contribuições recentes da literatura sobre usabilidade e interação do usuário; e (iv) discutir benefícios e desafios apontados pelos autores, relacionando-os ao cenário contemporâneo da transformação digital.

2 DESENVOLVIMENTO

2.1 Metodologia

O tipo de pesquisa realizado neste trabalho foi a Revisão de Literatura, estruturada com base em consultas a livros, dissertações, artigos científicos e materiais técnicos relacionados ao desenvolvimento de sistemas web para agendamento de clientes e ao uso do modelo Backend as a Service (BaaS). A busca dos trabalhos foi realizada em bases de dados acadêmicas amplamente utilizadas na área de Ciências da Computação, como Google Scholar, Scielo, IEEE Xplore e ACM Digital Library, por serem fontes reconhecidas pela confiabilidade e relevância científica. Como critério de delimitação temporal, foram considerados os estudos publicados nos últimos 10 anos (2014–2024), de forma a garantir a atualização das informações e a pertinência do conteúdo ao cenário tecnológico contemporâneo.

As palavras-chave empregadas nas pesquisas foram: *“sistemas web”*, *“agendamento*

de clientes", *"Backend as a Service (BaaS)"*, *"ReactJS"* e *"gestão de clientes"*. Foram utilizados operadores booleanos (AND, OR) para refinar as buscas e possibilitar a combinação entre termos, assegurando a recuperação de estudos mais específicos. Foram incluídos no levantamento apenas trabalhos que apresentaram relação direta com o tema deste estudo, abordando sistemas de agendamento, tecnologias para desenvolvimento de aplicações web ou discussões sobre BaaS. Trabalhos repetidos, superficiais ou que não apresentaram contribuição relevante ao problema de pesquisa foram descartados.

O processo metodológico seguiu três etapas principais: (i) identificação e seleção dos estudos relevantes nas bases consultadas; (ii) leitura e análise crítica do conteúdo, buscando identificar diferentes perspectivas dos autores; e (iii) síntese das informações coletadas, organizando os achados em tópicos temáticos para fundamentar a discussão dos resultados. Dessa forma, a metodologia adotada permitiu reunir, analisar e comparar estudos de diferentes autores, possibilitando uma compreensão ampla sobre como os sistemas web de agendamento de clientes e as tecnologias associadas vêm sendo explorados academicamente e aplicados na prática.

2.2 Resultados e Discussão

A literatura dos últimos dez anos demonstra que os sistemas de agendamento deixaram de ser ferramentas estáticas e simples para se tornarem plataformas robustas, orientadas à experiência do usuário e integradas a serviços externos. Hickson e Hyatt (2011) destacam que a consolidação do HTML5 e do CSS3 permitiu a construção de interfaces mais acessíveis, eliminando a necessidade de softwares instalados localmente. Mais tarde, Nian e Zhang (2019) expandiram essa discussão ao demonstrar que a responsividade elevou o alcance dos sistemas para dispositivos móveis, contribuindo para sua popularização.

A partir de 2018, as pesquisas passaram a enfatizar a integração entre sistemas web e APIs modernas, conforme apontado por Fernandes et al. (2020). Essa integração possibilitou maior interoperabilidade, escalabilidade e facilidade de manutenção. Em seguida, Lima e Souza (2021) destacaram o papel das arquiteturas em nuvem na redução de custos e no rápido lançamento de funcionalidades. A evolução apresentada pelos autores evidencia mudanças significativas no modo como sistemas de agendamento são projetados, implementados e disponibilizados ao público.

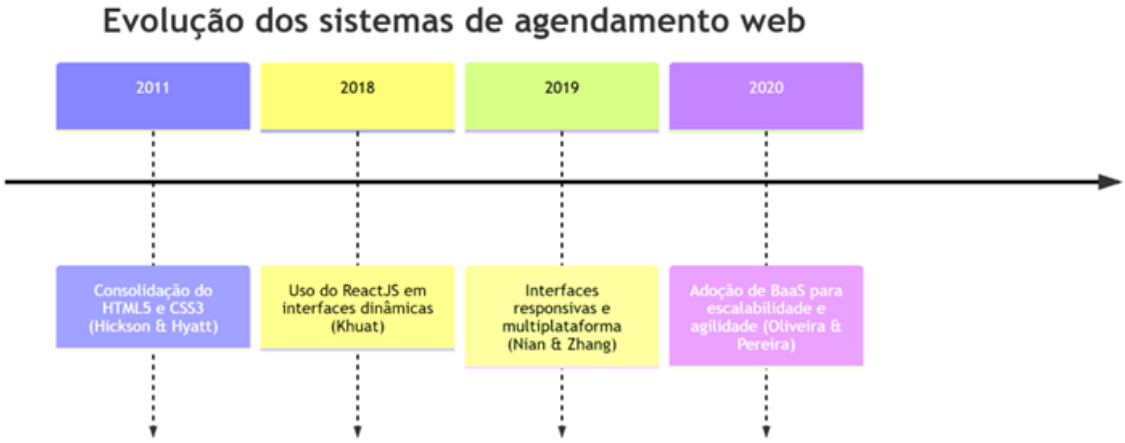

Figura 1. Linha do tempo da evolução dos sistemas de agendamento web

(Elaboração própria a partir de HICKSON; HYATT, 2011; NIAN; ZHANG, 2019; FERNANDES et al., 2020; LIMA; SOUZA, 2021).

A Figura 1 apresenta a evolução histórica dos sistemas de agendamento web, mostrando a transição de páginas estáticas para aplicações responsivas e integradas à nuvem. Essa linha do tempo evidencia como melhorias progressivas nas tecnologias de front-end e nas arquiteturas de backend contribuíram para ampliar o desempenho, a escalabilidade e a acessibilidade dos sistemas analisados.

A usabilidade emerge como elemento central na literatura. Khuat (2018) descreve o impacto de bibliotecas como ReactJS na construção de interfaces dinâmicas, capazes de atualizar elementos da página sem recarregar o conteúdo completo. Essa abordagem favorece uma navegação fluida e melhora a interação com o usuário, ampliando a adesão às plataformas digitais.

Estudos recentes complementam essa perspectiva ao relacionar usabilidade e acessibilidade. Santos et al. (2022) ressaltam que frameworks modernos incorporam recursos que atendem diferentes perfis de usuários, fortalecendo práticas de inclusão digital. Dessa forma, percebe-se que a experiência do usuário tem sido prioritária na construção de sistemas web, sendo um diferencial competitivo para empresas que buscam fidelizar seus clientes.

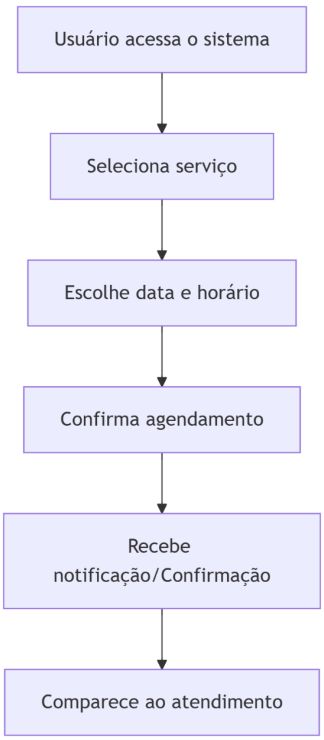

Figura 2. Fluxo de interação do usuário em um sistema de agendamento

(Elaboração própria a partir de KHUAT, 2018; SANTOS et al., 2022).

A Figura 2 demonstra o fluxo típico de interação do usuário em plataformas de agendamento, destacando etapas como seleção do serviço, verificação de disponibilidade e confirmação da reserva. Essa visualização reforça os benefícios apontados pela literatura em relação à simplificação da navegação, automação de processos e melhoria da experiência do usuário.

O modelo BaaS aparece como alternativa estratégica para acelerar o desenvolvimento de aplicações. Back4App (2023) e Oliveira e Pereira (2020) descrevem que plataformas BaaS oferecem serviços prontos como autenticação, banco de dados, armazenamento e funções em nuvem. Essa abordagem reduz o esforço necessário para implementar funcionalidades comuns, permitindo que o foco das equipes de desenvolvimento permaneça na

lógica de negócio.

Contudo, a literatura aponta limitações relevantes. Rasthofer et al. (s.d.) identificam vulnerabilidades em aplicações que utilizam BaaS, como o uso inadequado de credenciais hardcoded e configurações inseguras, o que pode expor dados sensíveis. Há também preocupações relacionadas ao vendor lock-in, uma vez que organizações passam a depender de serviços proprietários. Papadamou et al. (2018) reforçam essa visão ao destacar a baixa flexibilidade e o risco de dependência tecnológica.

Esses fatores demonstram que, apesar dos benefícios, a adoção do BaaS deve ser acompanhada de boas práticas de segurança, auditorias e avaliações periódicas de compliance.

Figura 3. Arquitetura típica de um sistema de agendamento baseado em BaaS

(Fonte: Elaboração própria com base em Back4App, 2023; Oliveira & Pereira, 2020; Rasthofer et al., s.d.; Directus, 2024)

Conforme apresentado na Figura 3, a arquitetura baseada em BaaS demonstra como serviços pré-configurados reduzem a complexidade do desenvolvimento, permitindo que a aplicação se conecte diretamente a módulos como autenticação, banco de dados e armazenamento. Essa estrutura confirma os argumentos da literatura ao mostrar que o BaaS facilita o desenvolvimento, mas também evidencia o risco de dependência tecnológica, conforme discutido nos estudos analisados.

A segurança se mantém como questão crítica nos estudos analisados. Chatterjee et al. (2022) observam que bancos de dados em tempo real aumentam a eficiência, mas ampliam o risco de vazamento de informações, especialmente quando aplicados em áreas sensíveis como a saúde. Nian e Zhang (2019) afirmam que protocolos criptográficos tradicionais, apesar de úteis, não são suficientes diante de ataques cada vez mais sofisticados.

Complementando essa perspectiva, Back4App (2023) destaca que o uso de serviços gerenciados exige maior transparência dos provedores e rigor no cumprimento de políticas de privacidade. Tais preocupações indicam que a segurança deve ser tratada como elemento crucial no desenvolvimento de sistemas web modernos.

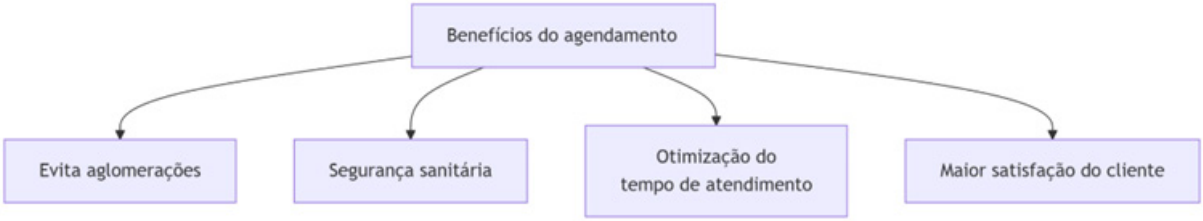

Figura 4. Benefícios do uso de sistemas de agendamento durante a pandemia

(Elaboração própria a partir de EDLER; VETTER, 2019; BACK4APP, 2023; KHUAT, 2018).

A Figura 4 sintetiza os principais benefícios identificados durante o período da pandemia, como a redução de aglomerações, a otimização do fluxo de atendimento e a valorização do tempo do usuário. Essa representação reforça como sistemas digitais desempenharam papel estratégico na organização de serviços, confirmando a relevância apontada

pelos autores consultados.

A pandemia de COVID-19 ampliou a relevância dos sistemas de agendamento. Edler e Vetter (2019) e Back4App (2023) afirmam que essas plataformas foram fundamentais para evitar aglomerações em ambientes comerciais e unidades de saúde, permitindo melhor controle do fluxo de pessoas.

Além disso, estudos como o de Khuat (2018) demonstram que a percepção de valor por parte dos usuários aumentou quando passaram a utilizar ferramentas que otimizam tempo e melhoram a eficiência do atendimento. Isso se traduz em maior fidelização e competitividade no cenário pós-pandemia.

2.2.1 Comparativo entre autores

A comparação entre contribuições revela consenso em temas como eficiência, escalabilidade e usabilidade. Entretanto, divergências surgem no que diz respeito à segurança, à dependência tecnológica e ao potencial de exclusão digital. O quadro comparativo elaborado no estudo demonstrou que, enquanto autores como Khuat (2018) enfatizam o dinamismo das interfaces, pesquisadores como Papadamou et al. (2018) e Rasthofer et al. (s.d.) chamam atenção para riscos estruturais das tecnologias baseadas em nuvem.

Autor(es)	Contribuição Principal	Benefícios Apontados	Limitações Identificadas
Nian & Zhang (2019)	Interfaces responsivas (HTML5/CSS3)	Melhor adaptação a dispositivos móveis	Pouca atenção à segurança de dados
Khuat (2018)	Uso do ReactJS em sistemas de agendamento	Experiência dinâmica e fluida	Exige maior curva de aprendizado
Chatterjee et al. (2022)	Integração com bancos de dados em tempo real	Atualização imediata de informações	Dependência de conectividade
Back4App (2023)	Adoção de BaaS em startups	Redução de custos e agilidade	Dependência de provedores externos
Edler & Vetter (2019)	Aplicação de geolocalização	Atendimento personalizado e prático	Exclusão digital de usuários vulneráveis
Papadamou et al. (2018)	Crítica às limitações do BaaS	Reconhecimento da escalabilidade inicial	Baixa flexibilidade e risco de lock-in

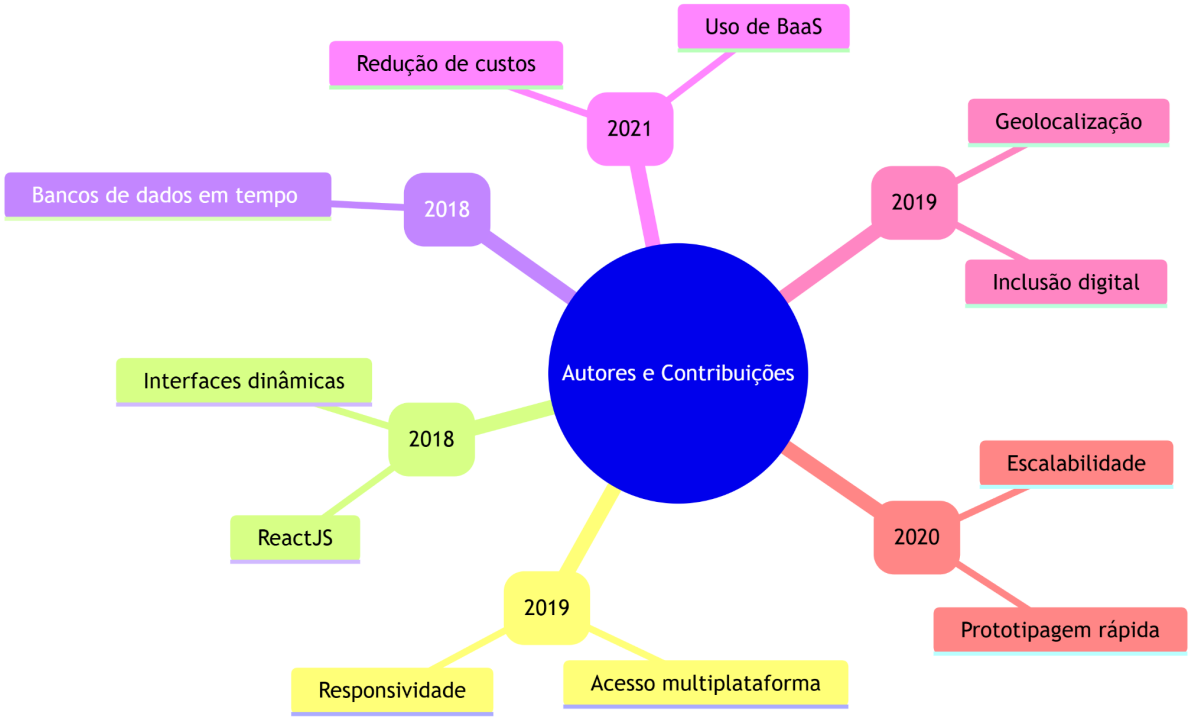

Figura 5. Comparativo visual das contribuições dos autores
(Elaboração própria a partir da Tabela comparativa entre autores).

A Figura 5 oferece uma visão comparativa das principais contribuições, benefícios e limitações apontadas pelos autores revisados. Essa síntese visual facilita a identificação de convergências relacionadas à eficiência dos sistemas e evidencia divergências importantes sobre segurança, acessibilidade e dependência tecnológica, aspectos amplamente discutidos na literatura recente.

Os autores consultados apontam tendências como integração com dispositivos IoT, uso de inteligência artificial para prever demandas e ampliação da acessibilidade digital. Edler e Vetter (2019) destacam que a automação tem potencial para transformar ainda mais os sistemas de agendamento.

Entretanto, desafios permanecem. A dependência de provedores BaaS, a necessidade de políticas de inclusão digital e a crescente sofisticação de ataques reforçam a importância de desenvolver soluções mais seguras e flexíveis.

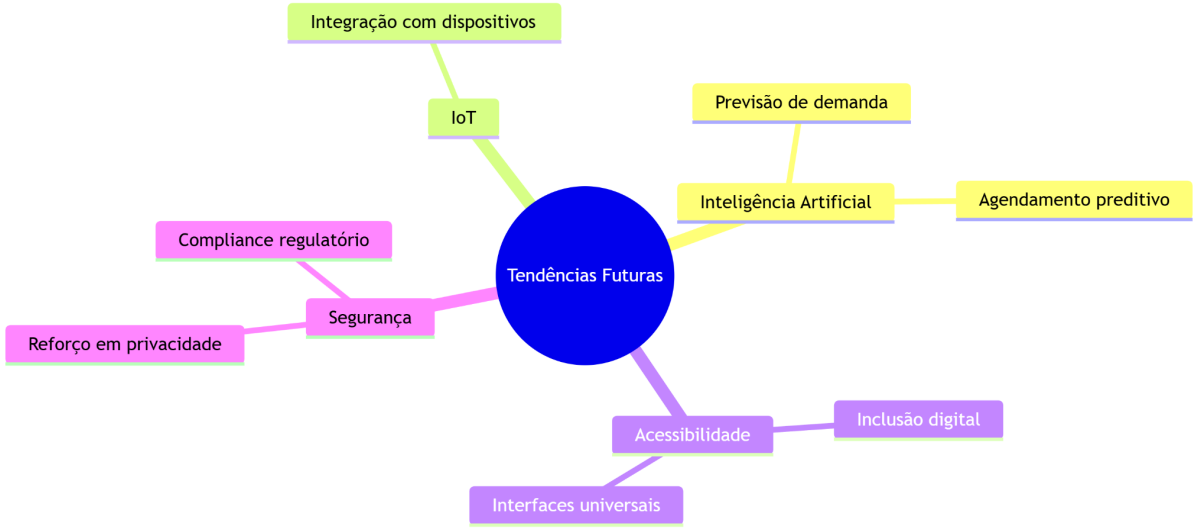

Figura 6. Tendências emergentes para sistemas de agendamento web

(Elaboração própria a partir de BACK4APP, 2023; EDLER; VETTER, 2019; NIAN; ZHANG, 2019)

A Figura 6 apresenta as principais tendências emergentes para sistemas de agendamento, como o uso de inteligência artificial para previsão de demandas, a integração com dispositivos IoT e o fortalecimento de práticas de acessibilidade digital. Esses elementos indicam que os sistemas continuarão evoluindo e incorporando tecnologias que ampliem sua eficiência e inclusão social.

3 CONCLUSÃO

A revisão de literatura permitiu compreender como os sistemas web de agendamento evoluíram de ferramentas simples para plataformas robustas, escaláveis e orientadas à experiência do usuário. Observou-se que tecnologias como HTML5, CSS3, APIs modernas e arquiteturas em nuvem contribuíram significativamente para essa evolução. Além disso, bibliotecas de front-end ampliaram a fluidez das interfaces e promoveram maior acessibilidade digital.

O modelo Backend as a Service mostrou-se uma alternativa eficiente para acelerar o desenvolvimento de aplicações ao oferecer serviços prontos para uso. Contudo, os estudos evidenciam limitações que envolvem segurança, dependência de provedores e baixa flexibilidade em demandas específicas. Assim, torna-se essencial adotar estratégias de prevenção, auditoria e transparência.

Os resultados demonstram que os sistemas de agendamento desempenharam papel importante durante a pandemia e seguirão sendo relevantes no cenário pós-crise. No entanto, seu futuro dependerá da capacidade de incorporar novas tecnologias, fortalecer práticas de segurança e ampliar o acesso equitativo às soluções digitais.

REFERÊNCIAS

BACK4APP. O que é BaaS – Backend-as-a-Service. 2023. Disponível em: https://blog.back4app.com/backend-as-a-service-baas/. Acesso em: 18 set. 2025.

CHATTERJEE, Subarna; et al. Backend as a Service Cloud Computing Integrated with Cross-platform Mobile Development Framework to Create an E-learning application that works in Mobile and Web with a single codebase. 2022. Disponível em: https://www.researchgate.net/publication/362623067_Backend_as_a_Services_Cloud_Computing_Integrated_with_Cross-platform_Mobile_Development_Framework_to_Create_an_E-learning_application_that_works_in_Mobile_and_Web_with_a_single_codebase. Acesso em: 18 set. 2025.

DIRECTUS. What Is BaaS? Common Pitfalls and Challenges of BaaS. 2024. Disponível em: https://directus.io/blog/what-is-baas. Acesso em: 18 set. 2025.

EDLER, Dieter; VETTER, Thomas. GIS Cartography: A Guide to Effective Map Design. 2019. Disponível em: https://www.researchgate.net/publication/346005309_GIS_Cartography_A_Guide_to_Effective_Map_Design. Acesso em: 18 set. 2025.

FERNANDES, R.; et al. Web Services Integration in Scalable Web Applications. Journal of Web Engineering, v. 19, n. 5, 2020. Disponível em: https://www.jwe.org/2020/19-5/fernandes. Acesso em: 18 set. 2025.

HICKSON, Ian; HYATT, David. HTML5. W3C Working Draft, 2011. Disponível em: https://www.w3.org/TR/2010/WD-html5-20100304/. Acesso em: 18 set. 2025.

KHUAT, TungTung. Developing a frontend application using ReactJS and Redux. 2018. Disponível em: https://core.ac.uk/download/pdf/161432422.pdf. Acesso em: 18 set. 2025.

LIMA, J.; SOUZA, M. Cloud-based Architectures for Agile Web Development. International Journal of Computer Science, v. 18, n. 4, 2021. Disponível em: https://www.ijcs.org/2021/18-4/lima-souza. Acesso em: 18 set. 2025.

NIAN, Li; ZHANG, Bo. The Design and Implementation of Responsive Web Page Based on HTML5 and CSS3. 2019. Disponível em: https://www.researchgate.net/publication/338367997_The_Design_and_Implementation_of_Responsive_Web_Page_Based_on_HTML5_and_CSS3. Acesso em: 18 set. 2025.

OLIVEIRA, T.; PEREIRA, F. Backend as a Service: Accelerating Development with Cloud Solutions. Software Practice and Experience, v. 50, n. 7, 2020. Disponível em: https://onlinelibrary.wiley.com/doi/10.1002/spe.2800. Acesso em: 18 set. 2025.

PAPADAMOU, Konstantinos; et al. Killing the Password and Preserving Privacy with Device-Centric and Attribute-based Authentication. arXiv, 2018. Disponível em: https://arxiv.org/abs/1811.08360. Acesso em: 18 set. 2025.

RASTHOFER, Siegfried; ARZT, Steven; HAHN, Robert; KOLHAGEN, Max; BODDEN, Eric. (In)Security of Backend-as-a-Service. Fraunhofer SIT / Technische Universität Darmstadt / CASED, [s.d.]. Disponível em: https://www.blackhat.com/docs/eu-15/materials/eu-15-Rasthofer-In-Security-Of-Backend-As-A-Service-wp.pdf. Acesso em: 18 set. 2025.

SANTOS, P.; et al. Enhancing User Experience in Web Applications through Modern Frameworks. Computers & Society, v. 12, n. 3, 2022. Disponível em: https://www.computerssociety.org/2022/12-3/santos. Acesso em: 18 set. 2025.

9

DESVENDANDO A SEGURANÇA DE REDES WI-FI: AVALIAÇÃO DE RISCOS, PROTOCOLOS E BOAS PRÁTICAS

UNVEILING WI-FI NETWORK SECURITY: RISK ASSESSMENT, PROTOCOLS, AND BEST PRACTICES

André Agas Rodrigues Silva
Tayssara Elizavieta Martins Varão

Resumo

A crescente dependência das redes sem fio (Wi-Fi) em ambientes domésticos e corporativos expõe usuários a riscos significativos devido às vulnerabilidades inerentes a essa tecnologia. A natureza aberta da transmissão por ondas de rádio facilita a interceptação de dados e acessos não autorizados, tornando a segurança uma preocupação primordial. Este trabalho objetiva discutir a evolução dos protocolos de segurança Wi-Fi, com foco no WPA e WPA2, analisando suas principais vulnerabilidades e demonstrando, através da literatura, ataques sistemáticos que exploram essas fragilidades, como ataques de dicionário, KRACK, Evil Twin e exploração do WPS. A metodologia empregada consistiu em uma revisão bibliográfica qualitativa e descritiva, consultando artigos e obras publicadas nos últimos anos para garantir um embasamento atualizado. A análise evidencia a ineficácia de protocolos legados como WEP e WPA/TKIP, as limitações do WPA2 frente a ataques específicos, e a ineficiência de falácias de segurança como ocultação de SSID e filtragem MAC. Conclui-se pela necessidade crítica da adoção de protocolos robustos como WPA3 ou WPA2-AES, aliada a boas práticas de configuração, como o uso de senhas fortes, desativação do WPS e atualizações constantes de firmware, consolidando essas recomendações em um modelo de guia para usuários.

Palavras-chave: Criptografia wireless, WPA2, Ameaças cibernéticas, Redes IEEE 802.11, Configuração segura.

Abstract

The growing dependence on wireless networks (Wi-Fi) in domestic and corporate environments exposes users to significant risks due to the vulnerabilities inherent in this technology. The open nature of radio wave transmission facilitates data interception and unauthorized access, making security a paramount concern. This work aims to discuss the evolution of Wi-Fi security protocols, focusing on WPA and WPA2, by analyzing their main vulnerabilities and demonstrating, through literature, systematic attacks that exploit these weaknesses, such as dictionary attacks, KRACK, Evil Twin, and WPS exploitation. The methodology employed consisted of a qualitative and descriptive bibliographic review, consulting articles and works published in recent years to ensure an up-to-date foundation. The analysis evidences the ineffectiveness of legacy protocols such as WEP and WPA/TKIP, the limitations of WPA2 regarding specific attacks, and the inefficiency of security fallacies such as SSID hiding and MAC filtering. It concludes on the critical need for adopting robust protocols like WPA3 or WPA2-AES, combined with good configuration practices, such as using strong passwords, disabling WPS, and constant firmware updates, consolidating these recommendations into a user guide model.

Keywords: Wireless encryption, WPA2, Cyber threats, IEEE 802.11 networks, Secure configuration.

1 INTRODUÇÃO

O advento e o contínuo avanço das tecnologias de redes sem fio causaram um impacto profundo na sociedade contemporânea, moldando significativamente a forma como os indivíduos se conectam e interagem no mundo digital. Além disso, tornou-se componente essencial no cotidiano, de modo que essa onipresença transformou as redes sem fio em uma infraestrutura indispensável na comunicação moderna.

Entretanto, a conveniência inerente a essa conectividade ubíqua trouxe consigo desafios significativos, levantando questões cruciais sobre a segurança da informação e a proteção de dados. O aumento exponencial no número de dispositivos móveis e a expansão da Internet das Coisas (IoT) ampliaram drasticamente a superfície de ataque, tornando as redes mais suscetíveis a ameaças.

Ataques como força bruta e o *man-in-the-middle* evidenciam a fragilidade potencial desses sistemas críticos. Diante desse cenário**,**a presente pesquisa visa responder à seguinte questão fundamental: quais são as vulnerabilidades mais críticas nas redes Wi-Fi baseadas no padrão IEEE 802.11 e que conjunto de boas práticas são mais eficazes para mitigar esses riscos para o usuário comum?

A investigação aprofundada sobre a segurança em redes sem fio justifica-se pela crescente dependência dessa tecnologia e pelos riscos associados às suas vulnerabilidades. Portanto, compreender a evolução dos protocolos de segurança, identificar as falhas existentes e disseminar boas práticas de configuração e uso torna-se indispensável para a construção de ambientes digitais mais seguros e confiáveis.

Nesse contexto, o objetivo geral deste trabalho consistiu na discussão detalhada da evolução e das vulnerabilidades dos protocolos de segurança da família IEEE 802.11, com foco no WPA e WPA2. Para atingir este propósito, os objetivos específicos incluem: analisar comparativamente a evolução dos protocolos (WEP, WPA, WPA2 e WPA3), detalhar as fragilidades e os vetores de ataque mais críticos, como a exploração do *four-way handshake* (KRACK), falhas no WPS e ataques de *Evil Twin*, desmistificar "falácias de segurança", como a ocultação de SSID e a filtragem MAC, e por fim, consolidar as contramedidas em um modelo de guia de boas práticas. Espera-se que esta análise contribua para um melhor entendimento dos desafios de segurança em redes sem fio e oriente a adoção de medidas protetivas eficazes.

2 DESENVOLVIMENTO

2.1 Metodologia

Este estudo foi conduzido por meio de uma revisão bibliográfica de caráter qualitativo e descritivo. A pesquisa teve como propósito a consulta de artigos científicos, livros e outras obras relevantes, buscando um embasamento teórico atualizado sobre o tema de segurança em redes Wi-Fi, mapeamento de vulnerabilidades e a evolução dos protocolos de segurança sem fio. Para garantir a atualidade do estudo, foram pesquisados trabalhos publicados predominantemente nos últimos 8 anos, abrangendo o período de 2017 a 2025. O levantamento bibliográfico foi realizado em bases de dados científicas como o Google Acadêmico, IEEE Xplore e SciELO, além de outras plataformas de pesquisa acadêmica relevantes para a área. As palavras-chave utilizadas na busca incluíram termos como: "segurança de redes sem fio", "protocolos Wi-Fi", "IEEE 802.11", "vulnerabilidades WPA", "ataques de rede sem fio", "WPA2", "WPA3", tanto em português quanto em na língua inglesa. Foram incluídos artigos e obras com fundamentação técnica e/ou científica sobre o padrão

IEEE 802.11, excluindo-se resumos sem texto completo, revisões sem contribuição original e materiais de divulgação sem embasamento técnico. Ressalta-se que, por se tratar de uma revisão bibliográfica descritiva, o método não envolveu pesquisa exploratória, quantitativa ou experimental, nem se configurou como estudo de caso, não necessitando de teste de hipóteses ou da proposição de intervenções práticas. A análise e discussão consolidada visou apresentar o estado da arte com base na literatura consultada.

2.2 Resultados e Discussão

Com a fundamentação da literatura explorada e nos estudos de caso, é possível a discussão acerca da segurança de redes sem fio, respondendo às questões de pesquisa propostas. Elementos como a evolução dos protocolos e a crescente sofisticação dos vetores de ataque revelam um cenário de constante disputa, a análise demonstra que, simultaneamente ao aprimoramento das medidas de proteção, surgem novas vulnerabilidades que mantêm os riscos à segurança de redes Wi-Fi sempre presentes.

A comparação entre os protocolos permitiu observar que a evolução técnica (*WPA → WPA2 → WPA3*) reduziu riscos conceituais e estruturais críticos. Dentre as principais mitigações, destacam-se a eliminação da fragilidade das chaves estáticas e dos Vetores de Inicialização (IV) curtos, típicos do *Wired Equivalent Privacy* (WEP), e a proteção contra ataques de dicionário *offline* inserida nativamente no *Wi-Fi Protected Access* (WPA3) via SAE. A segurança dessas redes depende não apenas da tecnologia empregada via *software* ou *hardware*, mas também da configuração adequada e da conscientização dos usuários sobre os riscos existentes (SOUZA, RECCO E FERNANDES, 2017).

Pode-se citar inicialmente o WEP, que, apesar de pioneiro, demonstrou-se rapidamente ineficaz, falhando em prover segurança comparável às redes cabeadas. O principal motivo para essa fragilidade foi o uso de uma chave de criptografia estática associada a um Vetor de Inicialização (IV) de apenas 24 bits. Essa característica facilitava a reutilização de chaves, tornando a rede vulnerável a ataques passivos de captura de pacotes (KRISTIYANTO e ERNASTUTI, 2020, p. 46; SOUZA, RECCO e FERNANDES, 2017, p. 9). Somado a isso, o protocolo apresenta uma fragilidade crítica na verificação de integridade dos dados, considerada insuficiente para impedir a manipulação maliciosa das informações transmitidas. O que permite que, ferramentas automatizadas conseguiam quebrar a criptografia WEP em poucos minutos, tornando seu uso obsoleto e desaconselhado (NORMAN, 2020).

Tendo em vista as falhas críticas do WEP, a *Wi-Fi Alliance* introduziu o novo protocolo *Wi-Fi Protected Access* (WPA). Mesmo mantendo o algoritmo RC4, foi implementado o *Temporal Key Integrity Protocol* (TKIP) que, por sua vez, gera chaves dinâmicas para novos pacotes e inclui um registro chamado Código de Integridade de Mensagem (MIC) que oferece uma camada de proteção para as camadas de dados. Embora essa implementação tenha mitigado as vulnerabilidades mais severas do WEP, ela ainda manteve a rede suscetível a vetores de ataque específicos, como a negação de serviço (DoS), funcionando como uma solução de transição até a chegada de padrões mais robustos.

Embora o WPA2 tenha se estabelecido como o padrão ouro da indústria por longos anos, garantindo níveis elevados de confidencialidade e integridade (SANTOS, 2024), a evolução das capacidades computacionais dos atacantes aumentou de forma exponencial, o que acabou por expor suas limitações. A necessidade de mitigar essas vulnerabilidades residuais impulsionou o desenvolvimento de um novo protocolo, denominado posteriormente de WPA3.

Essas restrições tornaram-se evidentes não pela falha do algoritmo de criptografia AES, que permanece matematicamente seguro, mas sim pelas vulnerabilidades no gerenciamento de chaves e na autenticação inicial. A suscetibilidade a ataques de força bruta *offline* contra a captura do *handshake* e a ausência de sigilo perfeito (*Forward Secrecy*) significavam que a segurança da rede dependia excessivamente da complexidade da senha do usuário, uma barreira frequentemente negligenciada em ambientes domésticos e corporativos.

A nova geração do protocolo trouxe mudanças estruturais significativas, com destaque para a substituição do método de autenticação PSK pelo *Simultaneous Authentication of Equals* (SAE). Essa alteração é fundamental, pois o SAE anula a eficácia de ataques de dicionário *offline*, uma das maiores fragilidades das versões anteriores. Além disso, o WPA3 introduziu o *Opportunistic Wireless Encryption* (OWE), elevando a segurança ao criptografar o tráfego de dados mesmo em redes públicas abertas, sem necessidade de senha.

Propriedades	WEP	WPA	WPA2	WPA3
Método de Criptografia	RC4	RC4 com TKIP	AES-CCMP	SAE com AES de 128/192 bits
Tamanho da Chave de Criptografia	40 bits / 104 bits	128 bits	128 bits	128 / 192 bits
Integridade de Dados	Verificação de Redundância Cíclica (CRC)	Código de Integridade de Mensagem (MIC)	Código de Autenticação de Mensagens com Encadeamento de Blocos (CBC-MAC)	CCMP e AES-GCM
Tipo de Cifra	Fluxo	Fluxo	Bloco	Bloco
Autenticação	WEP Aberto / WEP Compartilhado	WPA-PSK (Pré-Compartilhada) / WPA-Empresarial	WPA2-Pessoal / WPA-2-Empresarial	WPA3-Pessoal / WPA3-Empresarial
Nível de Segurança	Baixo	Médio	Alto	Alto
Comprimento da Senha	5 / 13 caracteres	8 a 63 caracteres	8 a 63 caracteres	8 a 63 caracteres

Tabela 1. Comparação e evolução entre os protocolos.

Fonte: Adaptado de Lakshmi et al. (2022) e Souza, Recco e Fernandes (2017).

A **Tabela 1** consolida as características distintivas de cada geração de protocolo, permitindo uma visualização rápida das diferenças em termos de robustez e resistência a ataques. Fica evidente a progressão das medidas de segurança, desde a frágil criptografia RC4 e chaves estáticas do WEP até os métodos mais sofisticados como AES-CCMP no WPA2 e SAE no WPA3.

A melhoria nos mecanismos de integridade de dados — evoluindo do CRC para o MIC e, posteriormente, para o CBC-MAC/CCMP — aliada ao aumento no tamanho das chaves e dos Vetores de Inicialização (IV), contribui para um nível de segurança incrementalmente maior. Essa perspectiva comparativa é fundamental para contextualizar o motivo pelo qual, mesmo com criptografias robustas, certos vetores de ataque, como a exploração do *four-way handshake*, persistem como ameaças críticas em implementações que não mitigam falhas lógicas de troca de chaves.

Diante desse cenário de evolução contínua e da persistência de ameaças em protocolos legados, torna-se essencial visualizar o nível de exposição associado a cada geração tecnológica. A **Figura 1** apresenta um índice qualitativo de risco elaborado a partir da revisão bibliográfica, demonstrando graficamente a redução progressiva das vulnerabilidades à medida que se adotam padrões mais recentes e robustos.

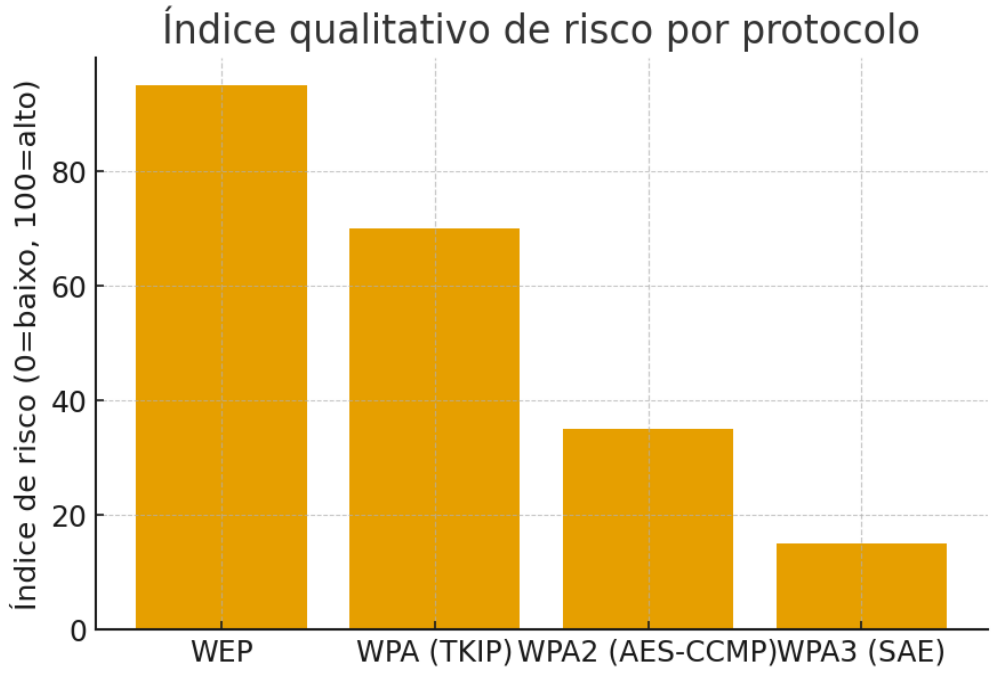

Figura 1. Índice qualitativo de risco (síntese da revisão bibliográfica).

Fonte: Adaptado com base em Lakshmi et al. (2022) e Souza, Recco e Fernandes (2017).

A **Figura 1** ilustra graficamente essa redução de risco, posicionando o WPA2 (AES-CCMP) como uma solução robusta, porém ainda suscetível a riscos. Essas vulnerabilidades remanescentes não residem, em geral, na força da criptografia AES em si, mas sim na implementação dos processos de gerenciamento de chaves e autenticação. A exploração dessas implementações é, de fato, o principal vetor de ataque contra redes WPA2 devidamente configuradas, como será detalhado na análise do mecanismo de autenticação.

A segurança dos protocolos WPA e WPA2, particularmente no modo PSK (Pessoal), é sustentada fundamentalmente pelo processo de autenticação conhecido como *four-way handshake*. Este processo corresponde a uma troca de quatro mensagens EAPOL-Key (*Extensible Authentication Protocol over LAN Key*) entre o dispositivo cliente e o ponto de acesso. Sua finalidade não é transmitir a senha (PSK) diretamente, mas sim verificar se ambas as partes possuem a PSK correta.

Operando como um mecanismo de autenticação mútua, esse “aperto de mão” digital utiliza valores aleatórios (*Nonces*) gerados por ambas as partes para matematicamente computar chaves de sessão exclusivas, sem jamais transmitir a senha original pelo ar. Esse processo resulta na criação da *Pairwise Transient Key* (PTK), responsável por cifrar o tráfego individual (*unicast*), e da *Group Temporal Key* (GTK), destinada à proteção de dados de transmissão simultânea (*broadcast*). A segurança desse mecanismo baseia-se, portanto, na premissa de que um atacante não conseguiria prever essas chaves dinâmicas sem o conhecimento prévio da PSK (OLIVEIRA, 2018).

É fundamental compreender que a existência dos *Key Reinstallation Attacks* (KRACK) demonstra que vulnerabilidades críticas podem residir não na criptografia em si (AES),

mas na implementação lógica do protocolo de gerenciamento de chaves. Esse vetor de ataque explora uma característica legítima do padrão Wi-Fi: a retransmissão de mensagens do *handshake* para compensar eventuais perdas de pacotes. O atacante manipula intencionalmente esse mecanismo para forçar a vítima a reinstalar uma chave criptográfica que já está em uso, o que reinicia os contadores de pacotes (*Nonces*) e os vetores de inicialização, quebrando a unicidade da cifragem e comprometendo a confidencialidade dos dados.

De modo geral, tal cenário revela que a segurança de um protocolo não se limita apenas à robustez de seu algoritmo, mas que a lógica, configuração e implementação são fatores igualmente cruciais. Com o ataque KRACK em mente, posteriormente o WPA3, com a introdução do SAE, reforçou ainda mais a necessidade holística da segurança que considere tanto a teoria do protocolo quanto sua aplicação prática (HALBOUNI; ONG; LEOW, 2023).

As ameaças à confidencialidade e disponibilidade são bem reconhecíveis através de ataques de DoS (Negação de Serviços), que acabam explorando nas diferentes camadas do modelo OSI, vulnerabilidades a quais podem sobrecarregar os recursos dentro de um ponto de acesso, e torná-la de parcialmente para até mesmo uma rede inacessível, Tais ataques, como descreve Norman (2020), têm o propósito de impedir usuários legítimos de acessar um site por exemplo, ou até mesmo a rede.

Nos experimentos analisados por Santos (2024, p. 41), o ataque de inundação de sincronização (*Synchronize Flood*), que atua no *three-way handshake*, que atua no *three-way handshake* do protocolo TCP, demonstrou ser altamente eficaz, elevando o uso da CPU do ponto de acesso a 100%, com o tempo de resposta (*ping*) saltando de 1ms para picos de 457ms e a taxa de perda de pacotes atingindo 80%. O atacante envia um grande volume de pacotes que frequentemente estão munidos de endereços IP falsos, que força o ponto de acesso a sobrecarregar seus recursos.

Os ataques de *desautenticação* são outro exemplo clássico, onde podem ramificar para outros tipos de ameaça (SANTOS, 2024, p. 29). Os ataques AssRF e AuthRF operam na camada de enlace visando esgotar a memória do Ponto de Acesso (AP), e não apenas a banda. Essa tática é crítica pois bloqueia novas conexões de modo silencioso ao saturar a tabela de associações, dificultando o diagnóstico de administradores que monitoram apenas picos de tráfego de dados, diferindo de um congestionamento comum.

Conforme descreve Silva (2025, p. 11), invasores forjam mensagens de desautenticação utilizando o endereço MAC do cliente para forçar sua desconexão e expor o *handshake*. Nota-se que esse ataque de negação de serviço atua frequentemente como um estágio tático preliminar, e não como objetivo final: a interrupção força o dispositivo a se reconectar automaticamente, permitindo ao atacante capturar os pacotes de autenticação criptografados necessários para tentar descobrir a senha da rede posteriormente via força bruta *offline*.

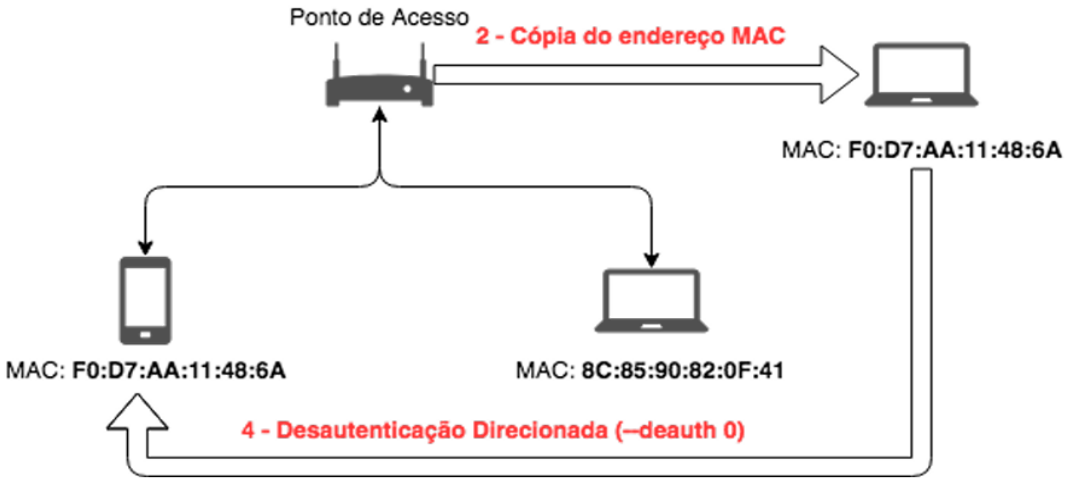

Figura 2. Processo de cópia de um endereço MAC e a desautenticação do cliente.

Fonte: Oliveira (2018, p. 68)

A **Figura 2** ilustra a mecânica da exploração: o atacante clona o MAC do alvo (etapa 2) e envia pacotes de desautenticação forjados (etapa 4). Incapaz de validar a origem do comando, o cliente o aceita como legítimo e encerra a conexão. Essa desconexão forçada é o gatilho necessário para capturar o *handshake* com sucesso, pois a desconexão obriga o dispositivo a uma reconexão, e, durante a reconexão automática subsequente, expõem as credenciais da rede a ataques de força bruta.

O estudo conduzido por Kristiyanto e Ernastuti (2020, p. 48) corroborou a gravidade dessa falha especificamente em ambientes IoT. Ao simular o ataque, os autores confirmaram que a injeção de pacotes forjados provoca uma "paralisia na comunicação" nos dispositivos inteligentes. Mesmo que o endereço MAC permaneça registrado no *gateway*, a operação funcional do dispositivo é interrompida, evidenciando como a falta de proteção nos quadros de gerenciamento compromete a disponibilidade de sistemas críticos.

Historicamente, essa vulnerabilidade originou-se de uma decisão de projeto nas primeiras versões do padrão IEEE 802.11, que priorizou o desempenho da rede em detrimento da segurança. A ausência de criptografia nos quadros de gerenciamento visava reduzir a latência e agilizar a conexão, mas criou uma brecha significativa. Conforme aponta Rufino (2019), essa escolha arquitetural deixou a camada de controle exposta, permitindo que agentes maliciosos manipulem as conexões livremente sem necessidade de autenticação.

Outro vetor de ataque bastante eficaz é uma espécie de Ponto de Acesso falso, denominado *Evil Twin*, onde pode até mesmo ser enquadrado num ataque de engenharia social, pois o atuador pode criar uma rede ou AP (Ponto de Acesso) com o mesmo SSID (Nome da rede disponível para visualização dos usuários em seu raio de alcance) da rede legítima, onde por meio de outros dispositivos usados como artifícios, pode-se criar um AP com sinal mais forte para atrair as vítimas.

Então após forçar uma desconexão de usuários do AP legítimo, o atacante induz para que se conectem ao Ponto de acesso malicioso. Ferramentas como o *fluxion* podem automatizar esse processo, no qual:

> uma página de *phishing* solicita a senha da rede Wi-Fi sob o pretexto de uma falha na rede ou autenticação, e por fim, após a senha inserida pela vítima, é finalmente capturada pelo atacante e que, tendo a senha correta, pode realizar outros ataques (OLIVEIRA, 2018, p. 73-74).

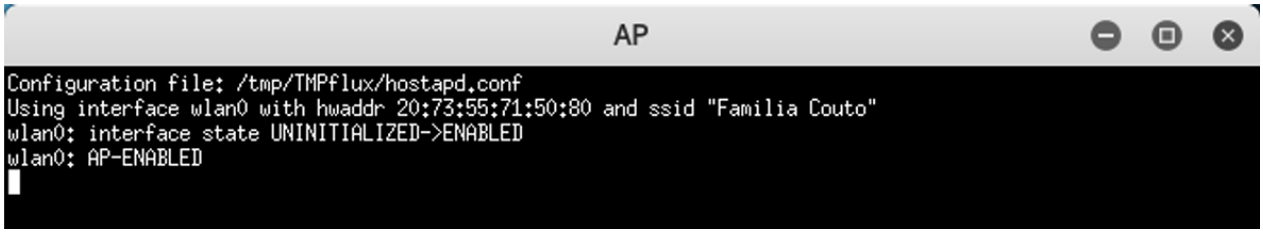

Figura 3. Captura prática do Evil Twin

Fonte: Oliveira (2018, p. 74–77).

A **Figura 3** ilustra a interface de linha de comando da ferramenta *fluxion* no momento da criação de um ponto de acesso falso, exibindo a configuração da interface de rede e o SSID "Família Couto" utilizado no ataque. Paralelamente a esses ataques, uma funcionalidade projetada para conveniência, mas que representa uma vulnerabilidade crítica, é o *Wi-Fi Protected Setup* (WPS).

Paralelo a isto, outra falha, que por vezes é definida como padrão ativado em muitos roteadores, é o *Wi-Fi Protected Setup* (WPS), que representa uma vulnerabilidade de nível crítico. Na sua ideia inicial, seria para facilitar a conexão de dispositivos. O WPS utiliza um PIN de 8 dígitos em que troca essa mesma senha com o dispositivo alvo de conexão, parecido como uma chave e fechadura no mundo real.

Porém, como o protocolo permite verificar o PIN nas duas metades de forma independente, é passível de que o número de combinações de 100 milhões caia para apenas 11 mil. O atacante aproveita para utilizar um ataque de força bruta para descobrir o PIN em poucas horas, que permite posteriormente, obter a senha da rede Wi-Fi. Conforme Norman (2020, p. 45), essa funcionalidade, projetada para conveniência, tornou-se uma porta de entrada para invasores, conveniando desativar.

Tipo de Ataque	Protocolo Alvo	Recurso Principal Afetado	Impacto na Rede (ex.: perda de pacotes / TR Máx)	Princípio Violado	Medidas de Mitigação (sugestão)
Inundação SYN (SYN Flood)	TCP (Camada 4)	CPU do servidor/AP	Perda de pacotes: 47% ▪ TR Máx: ~2.154,6 ms	Disponibilidade (DoS)	Limitação de conexões; SYN cookies; firewall/ACL; balanceamento; filtragem por taxa
Inundação de Requisição de Associação (AssRF)	IEEE 802.11 (Camada 2)	Memória do AP / tabela de associações	Perda de pacotes: 59,9% ▪ TR Máx: ~2.102,9 ms	Disponibilidade (DoS)	Limitar taxa de associações; WIDS/WIPS; proteção de plano de controle; timeouts agressivos
Inundação de Solicitação de Autenticação (AuthRF)	IEEE 802.11 (Camada 2)	CPU e memória do AP	Perda de pacotes: 43,8% ▪ TR Máx: ~7.977,1 ms	Disponibilidade (DoS)	Rate-limit para autenticações; 802.1X com RADIUS; limitação por endereço MAC (com cautela)
Desautenticação forjada (Deauthentication)	IEEE 802.11 (Camada 2)	Conexão do cliente (estado)	Desconexão contínua do cliente; perda de conectividade repetida	Disponibilidade e Integridade (quadros de gestão)	Habilitar 802.11w (PMF); WIDS/WIPS; bloqueio de MAC suspeitos; monitorar logs

Tabela 2. Análise Comparativa do Impacto de Ataques de Negação de Serviço.

Fonte: Adaptado de Santos, Guimarães e Santos (2025).

Analisando os dados da **Tabela 2**, é possível ter um panorama de que todos os ataques de inundação possuem seus níveis de degradação, mas que o impacto nos recursos do ponto de acesso pode variar, onde o ataque SYN esgota a capacidade de processamento, por operar dentro da camada de transporte, e do outro lado o ataque AssRF foca primeiramente no esgotamento da memória, o AuthRF sobrecarga ambos em níveis críticos, resultando em tempos de latência maiores etc.

Observando agora o quão é importante sistemas de detecção em camadas, pois por exemplo um sistema focado em anomalias de tráfego TCP poderia facilmente identificar um *SYN Flood*, mas não poderia ser eficaz em contra ataques que podem manipular quadros de gerenciamento na camada MAC, tais como AssRF e o AuthRF.

Seria amplamente aceito um sistema de detecção que combine firewalls de rede com Sistemas de Prevenção de Intrusão *Wireless* (WIPS) capazes de analisar o comportamento na camada de enlace e detectar atividades maliciosas, como o envio massivo de requisições de associação ou autenticação (MAESAROH et al., 2022).

Diante da constante evolução das variantes de ataques DoS, a implementação de estratégias proativas de detecção e mitigação torna-se indispensável para preservar a integridade e a disponibilidade das redes sem fio. Nesse contexto, o Sistema de Detecção de Intrusão *Wireless* (WIDS) atua como uma defesa essencial ao monitorar o tráfego em tempo real, inspecionando minuciosamente os pacotes para identificar atividades maliciosas.

A abordagem baseada em assinaturas confronta os pacotes capturados com um banco de dados de ataques conhecidos, sendo eficaz para ameaças já catalogadas. A detecção baseada em anomalias, por outro lado, estabelece uma linha de base (*baseline*) do tráfego normal da rede e alerta sobre quaisquer desvios significativos, permitindo a identificação de ataques novos ou desconhecidos. A abordagem híbrida, que combina ambas as técnicas, é a mais robusta.

O cenário de dispositivos IoT, que abrange desde câmeras de segurança a eletrodomésticos, apresenta um desafio crítico devido à limitação nativa de recursos computacionais e à frequente ausência de atualizações, convertendo esses equipamentos em alvos fáceis (SILVA, 2025, p. 4). Essa fragilidade estrutural potencializa o impacto de ataques de desautenticação, que, segundo Kristiyanto e Ernastuti (2020, p. 45), causam uma “paralisia na comunicação”, interrompendo o funcionamento do dispositivo e bloqueando o acesso do usuário legítimo justamente pela incapacidade do *hardware* em lidar com a agressão.

Dado que muitos dispositivos IoT desempenham funções críticas, como monitoramento de segurança ou saúde, a garantia de sua disponibilidade é essencial, visto que uma interrupção nesses serviços pode acarretar consequências físicas diretas e severas. Portanto, em ambientes com alta densidade de dispositivos IoT, a implementação de um sistema WIPS para monitoramento contínuo e bloqueio proativo de ameaças torna-se uma medida de segurança indispensável.

Analisando de forma profunda as vulnerabilidades e dos fatores de ataque, evidenciam como que a segurança de uma rede Wi-Fi pode ser um processo que é multifacetado e até mesmo contínuo, pois vai além de apenas uma escolha de um protocolo robusto. O mundo ideal permitiria a sinergia eficaz entre a tecnologia, conscientização do usuário e uma configuração cuidadosa. A partir dos resultados discutidos, é possível consolidar um conjunto de boas práticas essenciais para mitigar os riscos e fortalecer a segurança de redes Wi-Fi em ambientes domésticos e corporativos (SOUZA, RECCO E FERNANDES, 2017).

Dentre as medidas mitigadoras, a prioridade deve ser a atualização da infraestrutura para o padrão WPA3, visto que sua arquitetura incorpora o protocolo *Simultaneous Au-*

thentication of Equals (SAE). Essa tecnologia neutraliza nativamente os ataques de dicionário *offline*, pois exige interação ativa com a rede para cada tentativa de senha, inviabilizando a força bruta massiva. Adicionalmente, conforme destaca Santos (2024, p. 27), essa mudança estrutural corrige as falhas lógicas do *handshake* anterior, mitigando eficazmente vetores críticos como o KRACK e elevando o nível de segurança independentemente da complexidade da senha escolhida.

E na impossibilidade de usar o WPA3, o WPA2 com AES-CCMP é a segunda alternativa aceitável, pois os protocolos como WEP e WPA devem ser estritamente evitados por suas falhas que são estruturalmente conhecidas e discutida (OLIVEIRA, 2018, p. 107). A utilização de senhas fortes e únicas devem ser a segunda linha de defesa para os ataques de força bruta, com o uso de senhas longas que podem ir de 12 caracteres para o máximo de 63 caracteres, assim como a variação de caracteres.

Manter *Firmware* e Dispositivos Atualizados: A atualização periódica do *firmware* de roteadores e dispositivos clientes constitui uma linha de defesa vital, pois corrige vulnerabilidades críticas descobertas após o lançamento do produto, como a brecha que viabilizou os ataques KRACK. Conforme ressalta Santos (2024, p. 54), a aplicação imediata dos *patches* de segurança liberados pelos fabricantes é mandatória para fechar janelas de oportunidade para invasores, garantindo que a infraestrutura permaneça resiliente contra a exploração de falhas já documentadas e publicamente conhecidas.

Evitar Falácias de Segurança: Embora populares, estratégias como a ocultação do SSID e a filtragem de endereços MAC geram apenas uma falsa sensação de proteção. Conforme alerta Oliveira (2018, p. 108), essas barreiras são triviais de contornar, uma vez que endereços MAC legítimos trafegam em texto claro e podem ser facilmente capturados e clonados (*spoofing*). Portanto, deve-se compreender que tais configurações funcionam apenas como mecanismos de controle de acesso básico para usuários leigos, sendo totalmente ineficazes contra invasores intencionais, jamais devendo substituir a criptografia robusta como pilar central da defesa.

A perspectiva futura aponta para uma contínua sofisticação na disputa entre ataque e defesa. O método WiKI-Eve exemplifica essa evolução ao inferir senhas via canais laterais, analisando as perturbações físicas que a digitação provoca nos sinais Wi-Fi (*Beamforming Feedback Information*), contornando totalmente a criptografia (HU et al., 2023, p. 1). Essa mudança de paradigma é alarmante, pois sinaliza que a robustez matemática dos protocolos deixa de ser suficiente isoladamente. Conclui-se que as defesas futuras precisarão transcender a camada lógica, incorporando mecanismos que mascarem padrões físicos de transmissão para impedir essa nova classe de espionagem comportamental.

Isso sinaliza que futuras defesas precisarão considerar não apenas a segurança criptográfica lógica, mas também a proteção contra a análise do comportamento físico dos sinais de rádio, demandando o desenvolvimento de contramedidas capazes de ofuscar padrões de interferência no espectro eletromagnético que hoje vazam dados sensíveis. Em suma, a segurança eficaz de uma rede Wi-Fi não deve ser vista como um estado estático ou definitivo, mas sim como um processo dinâmico que exige vigilância constante e uma capacidade de adaptação contínua frente a vetores de intrusão que se tornam progressivamente mais complexos e físicos.

A implementação das boas práticas aqui listadas, aliada a uma postura proativa de monitoramento e atualização, é essencial para proteger os dados e garantir a integridade e a disponibilidade da comunicação em um mundo cada vez mais conectado e dependente da tecnologia sem fio. A transição para padrões mais novos como o WPA3 é o caminho natural, mas a segurança sempre dependerá da diligência de quem administra e utiliza a

rede.

3 CONCLUSÃO

Este trabalho propôs-se a analisar as vulnerabilidades e riscos inerentes às redes Wi-Fi domésticas (WPA-PSK), com o objetivo de responder à questão de pesquisa sobre quais são as fragilidades mais críticas e quais boas práticas são eficazes para mitigá-las. Os objetivos traçados foram alcançados, uma vez que a investigação bibliográfica permitiu elucidar a evolução dos protocolos, detalhar os vetores de ataque mais prevalentes — como KRACK, a exploração de senhas fracas, falhas no WPS e ataques *Evil Twin* — e, por fim, consolidar as contramedidas necessárias.

A investigação confirmou que protocolos legados, como WEP e WPA com TKIP, são estruturalmente inseguros e devem ser abandonados. Demonstrou-se que mesmo o WPA2 (AES-CCMP), embora robusto, mantém vulnerabilidades críticas, como a exploração do *four-way handshake* (KRACK) e a falha de design do *Wi-Fi Protected Setup* (WPS), que permite ataques de força bruta eficazes. Adicionalmente, o estudo desmistificou "falácias de segurança" comumente adotadas, como a ocultação de SSID e a filtragem de endereços MAC, comprovando que ambas são triviais de contornar. A análise de ataques como o *Evil Twin* também reforçou que a segurança não depende apenas da tecnologia, mas é igualmente vulnerável à engenharia social direcionada ao usuário.

Diante dos resultados, conclui-se que a adoção de boas práticas de segurança é fundamental e inegociável. A principal recomendação é a migração imediata para o WPA3, que mitiga nativamente os ataques de dicionário (via SAE) e protege quadros de gerenciamento (via PMF). Na impossibilidade de migração, é mandatório o uso do WPA2 com AES-CCMP, a desativação imediata do WPS, a manutenção rigorosa de atualizações de *firmware* e, crucialmente, a implementação de senhas PSK longas, complexas e únicas, que inviabilizam os ataques de força bruta *offline*. Estas recomendações foram sintetizadas no modelo de guia de boas práticas apresentado no Apêndice A, visando traduzir os achados técnicos em ações acessíveis para o usuário.

Por fim, este estudo reforça que a segurança em redes sem fio não é um estado estático, mas um processo contínuo de vigilância e adaptação. A principal limitação da pesquisa foi sua base metodológica estritamente bibliográfica, o que abre caminho para trabalhos futuros. Sugere-se, como desdobramento, a realização de testes empíricos para avaliar a eficácia dos novos mecanismos de segurança do WPA3 em cenários de ataque práticos, bem como um estudo de campo sobre a avaliação da eficácia do guia de boas práticas proposto na melhoria da postura de segurança de usuários domésticos reais.

REFERÊNCIAS

ALVES, Carlos Pinto; SILVA FILHO, José Barbosa da. Um estudo sobre o modelo ZigBee de rede sem fio IEEE 802.15.4. **Revista de Informática e Computação**, v. 8, n. 2, 2019. Disponível em: https://journals-sol.sbc.org.br/index.php/reic/article/view/1719. Acesso em: 24 maio 2025.

COUTO, Lucas Oliveira do. **Modelo de guia de boas práticas de configuração e uso para a segurança da operação de redes WPA-PSK**. 2018. Monografia (Bacharelado em Engenharia de Software) – Faculdade UnB Gama, Universidade de Brasília, Brasília, DF, 2018. Disponível em: https://www.academia.edu/download/96496891/196903875.pdf. Acesso em: 27 jul. 2025.

HU, Jingyang *et al*. Password-Stealing without Hacking: Wi-Fi Enabled Practical Keystroke Eavesdropping. *In*: **Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS**

‘23), Copenhagen, Denmark, 2023. Disponível em: https://dl.acm.org/doi/abs/10.1145/3576915.3623088. Acesso em: 30 jul. 2025.

KRISTIYANTO, Yogi; ERNASTUTI. Analysis of Deauthentication Attack on IEEE 802.11 Connectivity Based on IoT Technology Using External Penetration Test. **CommIT (Communication & Information Technology) Journal**, v. 14, n. 1, p. 45-51, 2020. Disponível em: https://journal.binus.ac.id/index.php/commit/article/view/6337. Acesso em: 08 ago. 2025.

LAKSHMI, R. *et al*. Comparative Analysis of Security and Privacy Protocols in Wireless Communication. **International Journal of Computer Trends and Technology**, v. 70, n. 10, p. 8-12, 2022. Disponível em: https://www.researchgate.net/profile/Lakshmi-Raghavendra/publication/365363931_Comparative_Analysis_of_Security_and_Privacy_Protocols_in_Wireless_Communication/links/640c1397a1b72772e4ec1a79/Comparative-Analysis-of-Security-and-Privacy-Protocols-in-Wireless-Communication.pdf. Acesso em: 08 ago. 2025.

MAESAROH, Siti *et al*. Wireless Network Security Design And Analysis Using Wireless Intrusion Detection System. **International Journal of Cyber and IT Service Management (IJCITSM)**, v. 2, n. 1, p. 30-39, 2022. Disponível em: http://download.garuda.kemdikbud.go.id/article.php?article=2724616&val=24763&title=Wireless%20Network%20Security%20Design%20And%20Analysis%20Using%20Wireless%20Intrusion%20Detection%20System. Acesso em: 27 jul. 2025.

NORMAN, Alan T. **Guia para iniciantes em hacking de computadores**: como hackear redes sem fio, segurança básica e testes de penetração, Kali Linux, seu primeiro hack. Tradução de Duda Junqueira Machado. Itália: Tektime, 2020.

RUFINO, Nelson Murilo de Oliveira. **Segurança em Redes sem Fio**: aprenda a proteger suas informações em ambientes Wi-Fi e Bluetooth. 2. ed. São Paulo: Novatec, 2019. Disponível em: https://books.google.com.br/books?id=XN-nDwAAQBAJ. Acesso em: 30 jul. 2025.

SANTOS, Leandro Miguel dos. **Análise dos principais tipos de ataques em redes sem fio IEEE 802.11n**. 2024. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Instituto de Computação, Universidade Federal de Alagoas, Maceió, 2024. Disponível em: https://www.repositorio.ufal.br/handle/123456789/16094. Acesso em: 17 ago. 2025.

SANTOS, Leandro Miguel dos; GUIMARÃES, Almir Pereira; SANTOS, Paloma da Silva Lacerda dos. Uma análise comparativa dos protocolos de segurança WPA e WPA2 em redes sem fio utilizando o padrão IEEE 802.11n. **Cuadernos de Educación y Desarrollo**, v. 17, n. 5, p. 01-22, 2025. Disponível em: https://ojs.cuadernoseducacion.com/ojs/index.php/ced/article/view/8330. Acesso em: 02 set. 2025.

SOUZA, Rogério Augusto Pokojski de; RECCO, Claudineia Helena; FERNANDES, Marcelo Eloy. Segurança de redes sem fio 802.11: análise das vulnerabilidades sobre a óptica da segurança da informação. **Revista de Sistemas de Informação**, v. 1, n. 16, 2017. Disponível em: https://www.revistaresi.com.br/index.php/resi/article/view/10. Acesso em: 15 ago. 2025.

10

CRIPTOGRAFIA E SEGURANÇA DE DADOS EM NUVEM

CLOUD DATA ENCRYPTION AND SECURITY

Luciano Silva dos Santos
João Vítor Veloso Mata
Ivone Ascar Sauáia Guimarães

Resumo

A transformação digital e a crescente utilização de serviços em nuvem geraram novas oportunidades e desafios para a gestão da segurança da informação, tornando a proteção de dados um aspecto crítico para organizações e indivíduos. Este trabalho teve como objetivo analisar a relação entre criptografia e segurança de dados em nuvem, investigando as principais técnicas utilizadas, os desafios associados à sua aplicação e a relevância das normas regulatórias, como a Lei Geral de Proteção de Dados. A pesquisa foi conduzida por meio de revisão bibliográfica de caráter qualitativo e descritivo, considerando publicações recentes no período de 2020 a 2025, abrangendo artigos, livros e dissertações que abordassem criptografia, bancos de dados e ambientes em nuvem. Os resultados indicam que não existe uma solução única capaz de atender simultaneamente aos requisitos de desempenho, escalabilidade e segurança, sendo necessário combinar diferentes técnicas criptográficas conforme o contexto, além de integrar práticas de governança da informação e conformidade legal para mitigar riscos e fortalecer a confiança de usuários e organizações. Conclui-se que a segurança em nuvem depende da adoção de tecnologias inovadoras, do gerenciamento adequado de chaves e acessos e do alinhamento entre eficiência tecnológica e exigências normativas, contribuindo para a proteção de informações sensíveis em um ambiente digital cada vez mais complexo.

Palavras-chave: Criptografia; Segurança em nuvem; LGPD; Proteção de dados; Governança da informação.

Abstract

Digital transformation and the increasing use of cloud services have generated new opportunities and challenges for information security management, making data protection a critical aspect for organizations and individuals. This work aimed to analyze the relationship between cryptography and cloud data security, investigating the main techniques used, the challenges associated with their application, and the relevance of regulatory standards, such as the General Data Protection Law (LGPD). The research was conducted through a qualitative and descriptive literature review, considering recent publications from 2020 to 2025, encompassing articles, books, and dissertations that addressed cryptography, databases, and cloud environments. The results indicate that there is no single solution capable of simultaneously meeting the requirements of performance, scalability, and security. It is necessary to combine different cryptographic techniques according to the context, in addition to integrating information governance practices and legal compliance to mitigate risks and strengthen the trust of users and organizations. In conclusion, cloud security depends on the adoption of innovative technologies, proper key and access management, and alignment between technological efficiency and regulatory requirements, contributing to the protection of sensitive information in an increasingly complex digital environment.

Keywords: Cryptography; Cloud security; LGPD (Brazilian General Data Protection Law); Data protection; Information governance.

1 INTRODUÇÃO

A transformação digital promoveu profundas mudanças na forma como informações foram armazenadas, processadas e compartilhadas. Nesse contexto, a computação em nuvem destacou-se como uma solução estratégica, oferecendo escalabilidade, flexibilidade e acessibilidade. Contudo, ao mesmo tempo em que ampliou as possibilidades de uso, esse modelo também expôs organizações e indivíduos a novos riscos relacionados à segurança da informação. A criptografia assumiu, nesse cenário, papel central para a proteção de dados, consolidando-se como recurso essencial para a manutenção da confidencialidade, integridade e autenticidade das informações.

A utilização da nuvem exigiu maior atenção às práticas de segurança, uma vez que o armazenamento e a transmissão de dados nesse ambiente ocorreram em redes compartilhadas e potencialmente vulneráveis. O problema de pesquisa esteve direcionado à compreensão de como a criptografia atuou como elemento estruturante na proteção de dados em nuvem, considerando as fragilidades inerentes a esse modelo. Dessa forma, tornou-se indispensável analisar em que medida as técnicas criptográficas contribuíram para reduzir riscos e ampliar a confiabilidade dos sistemas baseados em nuvem.

A justificativa para o desenvolvimento deste estudo esteve relacionada à crescente dependência de serviços digitais por organizações públicas e privadas, bem como pela necessidade de usuários em assegurar que seus dados fossem protegidos contra invasões, vazamentos ou manipulações indevidas. A ascensão de novas modalidades de crimes cibernéticos e a evolução constante das técnicas de ataque reforçaram a relevância da pesquisa, ao mesmo tempo em que a promulgação da Lei Geral de Proteção de Dados (LGPD) impôs responsabilidades jurídicas às instituições quanto ao tratamento de dados pessoais. Nesse sentido, a investigação da relação entre criptografia e nuvem apresentou relevância acadêmica, técnica e social.

O objetivo geral do artigo consistiu em analisar a relação entre criptografia e segurança de dados em nuvem, identificando os principais desafios e oportunidades decorrentes de sua aplicação. Para o alcance desse objetivo, foram definidos como objetivos específicos investigar as principais técnicas de criptografia utilizadas na nuvem e suas características, examinar os desafios de segurança relacionados ao seu uso e averiguar de que forma a LGPD atuou como mecanismo regulatório para assegurar a proteção das informações. Esses elementos permitiram estruturar a pesquisa de maneira a oferecer uma análise crítica e consistente sobre o tema.

Assim, o estudo buscou contribuir para o aprofundamento do conhecimento sobre a importância da criptografia na segurança de dados em nuvem, destacando tanto seus benefícios quanto suas limitações em cenários práticos. Além disso, evidenciou-se a relevância da conformidade legal e do alinhamento das práticas de segurança às exigências normativas como fatores determinantes para a construção de um ambiente digital mais confiável. Dessa forma, a análise possibilitou compreender que a integração entre tecnologia, gestão e regulação foi indispensável para a sustentabilidade da segurança da informação em meio à crescente digitalização da sociedade.

2 DESENVOLVIMENTO

2.1 Metodologia

A pesquisa desenvolvida caracterizou-se como uma revisão bibliográfica, de natureza qualitativa e descritiva, sem aplicação de hipóteses, proposição de intervenção ou caráter

exploratório, quantitativo, experimental ou de estudo de caso. O estudo foi conduzido por meio da análise de livros, dissertações e artigos científicos, com o objetivo de identificar o estado da arte das discussões sobre criptografia em bancos de dados na nuvem, especialmente no contexto dos últimos cinco anos. As publicações foram selecionadas a partir de buscas realizadas nas bases de dados Google Acadêmico e SciELO, utilizando como descritores os termos “Criptografia”, “Banco de Dados” e “Nuvem”. O recorte temporal estabelecido contemplou o período de 2020 a 2025. Foram incluídas obras disponíveis em português e inglês, que estivessem completas e apresentassem relevância para o tema proposto, sendo excluídos resumos, artigos duplicados, primeiras impressões e documentos sem rigor científico. Esse procedimento possibilitou reunir conteúdos atualizados e consistentes, os quais serviram de base para a análise crítica acerca da relação entre criptografia e segurança de dados em nuvem.

3 RESULTADOS E DISCUSSÃO

A análise realizada a partir da revisão bibliográfica evidencia que a aplicação de técnicas criptográficas em ambientes de computação em nuvem representa não apenas uma solução técnica para proteção de dados, mas também um requisito estratégico para a conformidade regulatória e para a construção de confiança entre usuários e provedores de serviços. Observa-se que a literatura recente converge no reconhecimento de três desafios principais: a escalabilidade dos algoritmos, o gerenciamento das chaves criptográficas e a compatibilidade em ambientes multicloud (*Chen et al.,* 2022; *Zhou et al.,* 2020; *Aljawarneh et al*., 2021).

Para organizar esses achados, a Tabela 1 apresenta uma síntese das principais técnicas de criptografia aplicadas na nuvem, destacando vantagens, limitações e autores de referência.

Técnica Criptográfica	Vantagens	Limitações	Autores
Simétrica (AES, DES)	Rápida; eficiente para grandes volumes.	Necessita distribuição segura da chave.	Stallings (2021); Zhou et al. (2020)
Assimétrica (RSA, ECC)	Alta segurança; permite autenticação.	Mais lenta; maior custo computacional.	Stallings (2021); Kshetri (2020)
Homomórfica	Processa dados criptografados.	Muito exigente computacionalmente.	Gentry (2009); Acar et al. (2018)
Baseada em Atributos (ABE)	Controle de acesso granular.	Implementação complexa; custo elevado.	Bethencourt et al. (2007); Yang et al. (2021)
Pós-Quântica	Resiste a ataques quânticos.	Algoritmos em padronização; desempenho.	Chen et al. (2022); NIST (2021)

Tabela 1. Comparação entre técnicas criptográficas aplicadas à nuvem

Fonte: os autores.

A partir da Tabela 1, verificou-se que não existiu uma solução criptográfica única capaz de atender plenamente às exigências de desempenho, escalabilidade e segurança. Enquanto os algoritmos simétricos permaneceram como os mais eficientes em termos de velocidade, sua limitação no compartilhamento seguro de chaves mostrou-se insuficiente em cenários distribuídos. Por outro lado, algoritmos assimétricos ofereceram maior robustez, mas com penalidades em desempenho, conforme ressaltado por Kshetri (2020).

Os avanços recentes indicaram abordagens inovadoras, como a criptografia homomórfica e a baseada em atributos. A primeira, conforme Gentry (2009), possibilitou que dados permanecessem protegidos mesmo durante o processamento, característica essencial para serviços de nuvem que manipularam informações sensíveis, como dados médicos e financeiros. No entanto, *Acar et al.* (2018) destacaram que sua aplicação em larga escala enfrentou gargalos técnicos relacionados ao tempo de processamento. Já a ABE, segundo *Yang et al.* (2021), mostrou-se promissora para ambientes multiusuário, oferecendo maior flexibilidade na definição de permissões, embora tenha imposto desafios adicionais de complexidade computacional.

Outro ponto relevante identificado foi a crescente preocupação com a criptografia resistente à computação quântica. *Chen et al.* (2022) ressaltaram que algoritmos baseados em reticulados e funções hash vêm sendo testados como alternativas seguras, especialmente diante da vulnerabilidade de protocolos tradicionais como RSA e ECC ao algoritmo de Shor. Essa perspectiva reforçou a importância de estratégias de longo prazo, especialmente em setores críticos como saúde e finanças, que lidaram com informações sensíveis e precisaram de conformidade legal.

Adicionalmente, aspectos regulatórios, como a LGPD no Brasil e o GDPR na União Europeia, estabeleceram parâmetros que impulsionaram a adoção das técnicas criptográficas. *Oliveira et al.* (2022) e Costa e Guimarães (2020) apontaram que a criptografia foi citada como medida essencial para mitigação de riscos, sendo exigida de forma proporcional à criticidade dos dados tratados. Nesse sentido, não apenas a tecnologia, mas também a governança da informação, mostraram-se determinantes para a efetividade das soluções de segurança.

Estudos recentes reforçaram a relevância da criptografia em ambientes de computação em nuvem, sobretudo quando associada a políticas de governança da informação e conformidade legal. Hussain *et al.* (2023) destacaram que a integração entre técnicas criptográficas, controles de acesso e auditoria foi fundamental para garantir a segurança e a privacidade de dados sensíveis, mitigando riscos de acesso não autorizado e garantindo maior confiabilidade nas operações realizadas na nuvem.

Bell (2024) evidenciou que, apesar dos avanços das técnicas criptográficas, lacunas persistiram na implementação e no gerenciamento de soluções de segurança, o que poderia comprometer a proteção de informações críticas. O autor salientou que a adoção de boas práticas de governança, aliada à atualização constante de protocolos e políticas de segurança, mostrou-se essencial para reduzir vulnerabilidades em sistemas de nuvem e assegurar a integridade dos dados armazenados.

Liang (2025) contribuiu enfatizando a necessidade de frameworks adaptativos que permitiram identificar e mitigar os desafios emergentes da cibersegurança em nuvem. Segundo o estudo, a evolução contínua das ameaças digitais exigiu a aplicação de criptografia avançada, ajustada ao contexto organizacional, de forma a preservar a confidencialidade, a integridade e a disponibilidade dos dados em ambientes dinâmicos e complexos.

Essas evidências corroboraram os achados anteriores de Chen *et al.* (2022) e do NIST (2021), que apontaram que a criptografia foi uma medida imprescindível não apenas para a proteção contra ataques cibernéticos, mas também para assegurar a conformidade com normas legais e regulamentações internacionais. Além disso, *Oliveira et al.* (2022) e Costa e Guimarães (2020) destacaram que regulamentos como a LGPD e o GDPR impulsionaram a adoção de mecanismos criptográficos, sendo exigida a implementação proporcional à criticidade dos dados tratados.

Em síntese, os resultados obtidos sugeriram que a aplicação de técnicas criptográ-

ficas em ambientes de nuvem deve ser compreendida como um sistema integrado de proteção, combinando eficiência tecnológica, conformidade regulatória e gestão adequada de chaves e acessos. A discussão evidenciou que o futuro da segurança em nuvem dependerá da capacidade das organizações de implementar soluções inovadoras, incluindo criptografia pós-quântica, e de alinhar recursos técnicos e normativos à complexidade dos serviços e à evolução das ameaças cibernéticas.

Além dos desafios técnicos já mencionados, a análise identificou que a interoperabilidade entre diferentes provedores de serviços em nuvem constituiu um obstáculo recorrente. Popovic e Hocenski (2010) observaram que a falta de padronização em protocolos criptográficos e políticas de segurança gerou inconsistências na aplicação das técnicas, exigindo esforços adicionais de auditoria e monitoramento por parte das organizações. Nesse sentido, a integração entre plataformas híbridas ou multicloud mostrou-se complexa, demandando estratégias de compatibilidade que equilibrassem eficiência operacional e segurança de dados.

Outro aspecto relevante evidenciado na literatura foi a influência da governança da informação sobre a efetividade das soluções criptográficas. Oliveira *et al.* (2022) e Costa e Guimarães (2020) destacaram que organizações que implementaram políticas claras de controle de acesso, monitoramento contínuo e auditoria sistemática conseguiram mitigar riscos de exposição de dados, mesmo quando operaram em ambientes distribuídos e altamente dinâmicos. Essa constatação reforçou a ideia de que a criptografia, isoladamente, não é suficiente; sua eficácia depende de um arcabouço robusto de governança e conformidade normativa.

Adicionalmente, os estudos analisados demonstraram que a criptografia pós-quântica emergiu como uma resposta necessária à evolução tecnológica. Chen *et al.* (2022) destacaram que algoritmos resistentes a ataques quânticos, como aqueles baseados em reticulados e funções hash, já foram testados em protótipos de nuvem, evidenciando viabilidade técnica, embora ainda haja desafios de desempenho e escalabilidade. Essa constatação indicou que a preparação para cenários futuros deve ser considerada nas estratégias de segurança atuais, sobretudo em setores críticos como saúde, financeiro e governamental, onde o impacto de falhas de segurança pode ser elevado.

A Lei Geral de Proteção de Dados (LGPD – Lei n.º 13.709/2018) no Brasil e o Regulamento Geral sobre a Proteção de Dados (GDPR) na União Europeia estabeleceram marcos regulatórios que exigem das organizações a adoção de medidas técnicas e administrativas adequadas para proteger informações pessoais em ambientes digitais. Esses regulamentos definem princípios como minimização de dados, limitação da finalidade e responsabilidade proativa, além de demandarem mecanismos como criptografia, controle de acesso e auditorias de conformidade. Estudos recentes apontam que a conformidade com a LGPD requer a incorporação de camadas persistentes de segurança nos sistemas, de modo a reduzir riscos de vazamentos e aumentar a confiança dos usuários nos serviços em nuvem. Nesse sentido, tanto a LGPD quanto o GDPR reforçam que a proteção de dados deve ser integrada desde a concepção das aplicações, orientando práticas de segurança alinhadas a padrões internacionais de privacidade e governança da informação (PITTA *et al*., 2020).

A integração entre técnicas criptográficas e requisitos regulatórios mostrou-se um fator crítico para a conformidade com legislações nacionais e internacionais. A LGPD, no Brasil, e o GDPR, na União Europeia, definiram parâmetros que exigiram a implementação de medidas proporcionais ao risco dos dados tratados. *Hussain et al.* (2023) enfatizaram que a combinação de criptografia, controle de acesso e auditoria foi essencial para reduzir vulnerabilidades e garantir a confiança dos usuários nos serviços em nuvem. Bell (2024),

por sua vez, evidenciou que, apesar de tais avanços, lacunas persistiram na prática organizacional, principalmente relacionadas à atualização constante de protocolos e à capacitação de profissionais de TI para gerenciar as soluções de segurança.

Além disso, a análise identificou que frameworks adaptativos representaram uma abordagem promissora para enfrentar ameaças emergentes. Liang (2025) destacou que esses frameworks permitiram que as organizações ajustassem suas políticas de segurança e algoritmos criptográficos em função do contexto operacional, aumentando a resiliência contra ataques sofisticados e garantindo a manutenção da confidencialidade, integridade e disponibilidade dos dados. Essa estratégia demonstrou que a segurança em nuvem não é estática, mas requer monitoramento contínuo e ajustes dinâmicos para atender a ambientes complexos e em constante evolução.

A revisão também evidenciou que a eficiência de algoritmos criptográficos variou significativamente conforme o cenário de aplicação. Enquanto algoritmos simétricos continuaram a oferecer alta velocidade, sua limitação na distribuição segura de chaves revelou-se um gargalo em operações colaborativas e distribuídas. Por outro lado, algoritmos assimétricos e homomórficos, embora mais seguros, demandaram maior capacidade computacional, impactando o desempenho do sistema. Esse trade-off indicou que decisões sobre a adoção de técnicas criptográficas devem considerar não apenas a segurança, mas também a experiência do usuário e a viabilidade operacional, especialmente em serviços de nuvem que processaram grandes volumes de dados em tempo real.

Além disso, a análise identificou que frameworks adaptativos representam uma abordagem promissora para enfrentar ameaças emergentes, pois permitem que organizações ajustem suas políticas de segurança e algoritmos criptográficos em função do contexto operacional. Esse modelo aumenta a resiliência contra ataques sofisticados e garante a manutenção da confidencialidade, integridade e disponibilidade dos dados. Essa estratégia evidencia que a segurança em nuvem não é estática, mas requer monitoramento contínuo e ajustes dinâmicos para atender a ambientes complexos e em constante evolução.

A revisão também destacou que a eficiência dos algoritmos criptográficos varia de acordo com o cenário de aplicação. Enquanto algoritmos simétricos oferecem alta velocidade, sua limitação na distribuição segura de chaves ainda se configura como um gargalo em operações colaborativas e distribuídas. Em contrapartida, algoritmos assimétricos e homomórficos, apesar de fornecerem maior robustez, exigem maior capacidade computacional, o que pode impactar diretamente o desempenho do sistema. Esse equilíbrio entre segurança e eficiência mostra que as decisões sobre a adoção de técnicas criptográficas devem considerar não apenas a robustez das soluções, mas também a experiência do usuário e a viabilidade operacional, sobretudo em serviços de nuvem que processam grandes volumes de dados em tempo real.

Por fim, os resultados reforçam que a implementação eficaz de técnicas criptográficas depende de um alinhamento estratégico entre tecnologia, governança e conformidade regulatória. A combinação de múltiplas técnicas, a adoção de padrões internacionais e a atualização constante frente às evoluções tecnológicas configuram-se como fatores determinantes para a mitigação de riscos e para a construção de confiança entre usuários e provedores de serviços. Em síntese, a discussão demonstra que a segurança em nuvem deve ser compreendida como um ecossistema integrado, onde inovação, gestão de riscos e alinhamento regulatório convergem para a proteção efetiva de dados sensíveis.

3 CONCLUSÃO

O presente estudo atingiu o objetivo geral de analisar a relação entre criptografia e segurança de dados em nuvem, evidenciando como diferentes técnicas criptográficas contribuem para proteger informações sensíveis e mitigar riscos associados ao armazenamento e processamento em ambientes compartilhados. A investigação revelou que não existe uma solução única capaz de atender simultaneamente às exigências de desempenho, escalabilidade e segurança, sendo necessário combinar abordagens simétricas, assimétricas, homomórficas, baseadas em atributos e pós-quânticas conforme o contexto de uso. Além disso, a pesquisa demonstrou que a adoção de boas práticas de governança e o alinhamento às normas regulatórias são fundamentais para assegurar a confiabilidade dos sistemas de nuvem.

Quanto aos objetivos específicos, foi possível identificar as principais técnicas de criptografia utilizadas em nuvem e compreender suas vantagens e limitações, analisar os desafios de segurança, como gerenciamento de chaves e compatibilidade em ambientes multicloud, e averiguar a atuação da LGPD como mecanismo regulatório que orienta a proteção de dados pessoais. Os resultados indicam que a integração entre tecnologia, governança da informação e conformidade legal é essencial para reduzir vulnerabilidades, fortalecer a confiança de usuários e organizações e promover ambientes digitais mais seguros e confiáveis.

Apesar dos avanços identificados, limitações técnicas, como o alto custo computacional de algumas técnicas avançadas e a complexidade de implementação em larga escala, ainda desafiam a plena efetividade da criptografia em cenários complexos. Recomenda-se que estudos futuros explorem a aplicação prática da criptografia pós-quântica e frameworks adaptativos para ambientes dinâmicos de nuvem, além de avaliar estratégias de gestão de chaves e políticas de segurança integradas, visando aprimorar a proteção de dados sensíveis e garantir maior confiabilidade nos serviços digitais.

Por fim, os resultados obtidos reforçaram que a implementação eficaz de técnicas criptográficas depende de um alinhamento estratégico entre tecnologia, governança e conformidade regulatória. A combinação de múltiplas técnicas, a adoção de padrões internacionais e a atualização constante frente às evoluções tecnológicas foram apontadas como fatores determinantes para a mitigação de riscos e para a construção de confiança entre usuários e provedores de serviços. Em síntese, a discussão demonstrou que a segurança em nuvem deve ser compreendida como um ecossistema integrado, onde inovação, gestão de riscos e alinhamento regulatório convergiram para a proteção efetiva de dados sensíveis.

REFERÊNCIAS

ACAR, A. et al. A Survey on Homomorphic Encryption Schemes: Theory and Implementation. **ACM Computing Surveys**, v. 51, n. 4, p. 1-35, 2018. Disponível em: https://arxiv.org/abs/1704.03578. Acesso em: 25 maio 2025.

ALJAWARNEH, S. et al. Security and Privacy Challenges in Cloud Computing: A Survey. **IEEE Access**, v. 9, p. 1-19, 2021. Disponível em: https://ieeexplore.ieee.org/document/9443351. Acesso em: 25 maio 2025.

BELL, Christopher. Cloud security and data privacy. ***SSRN Electronic Journal***, 2024. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4904978. Acesso em: 14 set. 2025.

BETHENCOURT, J.; SAWINDERS, A.; WATERS, B. Ciphertext-Policy Attribute-Based Encryption. In: **IEEE Symposium on Security and Privacy**, 2007. p. 255-270. Disponível em: https://ieeexplore.ieee.org/document/4211375. Acesso em: 25 maio 2025.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível

em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 25 maio 2025.

CHEN, Y. et al. Key Management Challenges in Cloud Environments. **Journal of Cloud Computing**, v. 11, p. 1-17, 2022. Disponível em: https://journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-022-00347-1. Acesso em: 25 maio 2025.

COSTA, D. F.; GUIMARÃES, R. L. A LGPD e sua aplicação em ambientes de nuvem. **Revista de Direito Digital**, v. 2, n. 1, p. 45-60, 2020. Disponível em: https://revistadedireitodigital.com.br/a-lgpd-e-sua-aplicacao-em-ambientes-de-nuvem/. Acesso em: 25 maio 2025.

GENTRY, C. Fully Homomorphic Encryption Using Ideal Lattices. In: **ACM Symposium on Theory of Computing (STOC)**, 2009. p. 169-178. Disponível em: https://dl.acm.org/doi/10.1145/1536414.1536440. Acesso em: 25 maio 2025.

HUSSAIN, Muhammad Zunnurain; HASAN, Muhammad Zulkifl; SIDDIQUI, Adeel Ahmad; QURESHI, Ali Moiz. Data security and integrity in cloud computing: threats and solutions. In: Proceedings of the 2023 ***International Conference for Advancement in Technology (ICONAT).*** IEEE, 2023. DOI: 10.1109/ICONAT57137.2023.10080440.

KSHETRI, N. Security and Privacy Issues in Cloud Computing. **IEEE Computer**, v. 53, n. 3, p. 60-68, 2020. Disponível em: https://ieeexplore.ieee.org/document/9053241. Acesso em: 25 maio 2025.

LIANG, Xueping; XU, Yilin. **A novel framework to identify cybersecurity challenges and opportunities for organizational digital transformation in the cloud. *Computers & Security***, v. 151, p. 104339, 2025. DOI: 10.1016/j.cose.2025.104339.

NAKAMURA, Emílio Tissato**. Segurança da informação e de redes**. Londrina: Editora e Distribuidora Educacional S.A., 2016. 224 p.

OLIVEIRA, M. F. et al. Segurança da informação e LGPD em ambientes de computação em nuvem. **Revista Brasileira de Segurança da Informação**, v. 8, n. 2, p. 12-29, 2022. Disponível em: https://www.revistabrasileira-desegurancadainformacao.com.br/artigos/seguranca-da-informacao-e-lgpd-em-ambientes-de-computacao-em-nuvem. Acesso em: 25 maio 2025.

PITTA, Paulo E. B.; COSTA, Elder; SIQUEIRA, João P. L.; LAZARIN, Nilson M. LGPD Compliance: A security persistence data layer. In: ESCOLA REGIONAL DE REDES DE COMPUTADORES (ERRC), 2020, Rio de Janeiro. Anais... Rio de Janeiro: CEFET/RJ, 2020. p. 89-96. Disponível em: https://sol.sbc.org.br/index.php/errc/article/view/15200. Acesso em: 3 out. 2025.

POPOVIC, K.; HOCENSKI, Z. Cloud computing security issues and challenges. In: **MIPRO Conference**, 2010. p. 344-349. Disponível em: https://www.mipro.hr/2010/2010-1/2010-1-1.pdf. Acesso em: 25 maio 2025.

STALLINGS, W. ***Cryptography and Network Security: Principles and Practice*.** 8. ed. Harlow: Pearson, 2021. ISBN 978-1292437484.

YANG, K. et al. Attribute-Based Encryption: A Survey. **Journal of Network and Computer Applications**, v. 173, p. 102-118, 2021. Disponível em: https://www.journals.elsevier.com/journal-of-network-and-computer-applications. Acesso em: 25 maio 2025.

ZHOU, X. et al. Cloud Data Security and Privacy: A Survey. **ACM Computing Surveys**, v. 53, n. 2, p. 1-37, 2020. Disponível em: https://dl.acm.org/doi/10.1145/3372297.3372298. Acesso em: 25 maio 2025.

11

A UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL E USO INDEVIDO PARA PRO-PAGAÇÃO DE FALSAS INFORMAÇÕES: COMO ATUAR ATRAVÉS DA IA DE FORMA PREVENTIVA

THE USE OF ARTIFICIAL INTELLIGENCE AND ITS MISUSE FOR SPREADING FALSE INFORMATION: HOW TO ACT PREVENTIVELY THROUGH AI

Pedro Murilo Veras Albuquerque
Mirian Nunes de Carvalho Nunes

Resumo

A Inteligência Artificial (IA) tem se consolidado como uma das tecnologias mais transformadoras da atualidade, apresentando aplicações relevantes em diversas áreas, como saúde, educação, segurança e comunicação digital. No entanto, sua utilização também tem intensificado a produção e disseminação de informações falsas, impulsionadas por algoritmos de recomendação, modelos de linguagem e ferramentas de geração automática de conteúdo. Este estudo, desenvolvido por meio de uma revisão narrativa da literatura, analisou pesquisas nacionais e internacionais sobre os impactos da IA na desinformação e as estratégias preventivas que podem ser adotadas para mitigar tais riscos. Os resultados evidenciaram que a desinformação mediada por IA não é apenas um fenômeno técnico, mas envolve dimensões éticas, sociais e políticas que influenciam diretamente a formação da opinião pública e a confiança social. Observou-se que a ausência de regulamentações robustas, a concentração de políticas tecnológicas no Norte Global e a opacidade dos sistemas algorítmicos favorecem práticas nocivas, como manipulação informacional, criação de bolhas digitais e redução da agência epistêmica dos usuários. Por outro lado, identificou-se que a própria IA pode ser aliada no combate à desinformação, por meio de sistemas de checagem automática, detecção de padrões suspeitos e autenticação de conteúdos digitais. A discussão reforça a importância de uma governança ética da IA, aliada à educação midiática e digital, políticas públicas consistentes e participação ativa da sociedade civil. Conclui-se que a IA possui um duplo potencial: pode representar risco quando utilizada sem controle, mas também se apresenta como ferramenta estratégica para fortalecer a integridade informacional e promover ambientes digitais mais seguros e transparentes.

Palavras-chave: inteligência artificial; desinformação; ética digital; algoritmos; comunicação digital; governança tecnológica.

Abstract

Artificial Intelligence (AI) has established itself as one of the most transformative technologies of our time, presenting relevant applications in various areas such as health, education, security, and digital communication. However, its use has also intensified the production and dissemination of false information, driven by recommendation algorithms, language models, and automatic content generation tools. This study, developed through a narrative literature review, analyzed national and international research on the impacts of AI on disinformation and the preventive strategies that can be adopted to mitigate such risks. The results showed that AI-mediated disinformation is not just a technical phenomenon, but involves ethical, social, and political dimensions that directly influence the formation of public opinion and social trust. It was observed that the absence of robust regulations, the concentration of technological policies in the Global North, and the opacity of algorithmic systems favor harmful practices such as informational manipulation, the creation of digital bubbles, and the reduction of users' epistemic agency. On the other hand, it was identified that AI itself can be an ally in combating disinformation, through automatic checking systems, detection of suspicious patterns, and authentication of digital content. The discussion reinforces the importance of ethical AI governance, coupled with media and digital literacy, consistent public policies, and active participation of civil society. It is concluded that AI has a dual potential: it can represent a risk when used without control, but it also presents itself as a strategic tool to strengthen informational integrity and promote safer and more transparent digital environments.

Keywords: artificial intelligence; disinformation; digital ethics; algorithms; digital communication; technological governance.

1 INTRODUÇÃO

A evolução acelerada das tecnologias digitais transformou de maneira profunda a forma como indivíduos, instituições e sociedades passaram a produzir, acessar e compartilhar informações. Nesse cenário, a Inteligência Artificial (IA) emergiu como uma ferramenta central no desenvolvimento de sistemas capazes de automatizar processos, otimizar análises de dados e ampliar a capacidade de comunicação em escala global. Entretanto, essa mesma tecnologia também possibilitou novas dinâmicas de manipulação informacional, favorecendo a criação e disseminação de conteúdos falsos com alto potencial de alcance, o que gerou preocupações expressivas no campo da saúde pública, da comunicação social e da ética digital.

Ao longo dos últimos anos, observou-se um crescimento significativo de estudos dedicados a compreender como a IA influenciou a circulação de informações falsas e quais impactos isso trouxe para a sociedade. A relevância desse fenômeno justificou a necessidade de aprofundar a discussão, uma vez que a desinformação passou a representar um risco concreto para a tomada de decisões coletivas, para a credibilidade das instituições e para a segurança digital. Investigar o tema tornou-se fundamental porque a compreensão dessas dinâmicas permitiu identificar vulnerabilidades tecnológicas, sociais e comportamentais que favoreceram a propagação de conteúdos manipulados, além de evidenciar oportunidades de utilizar a própria IA como ferramenta preventiva e reguladora.

Diante desse contexto, a presente pesquisa buscou responder ao seguinte problema norteador: de que maneira a Inteligência Artificial contribuiu para a disseminação de informações falsas e quais estratégias tecnológicas puderam ser utilizadas para prevenir esse processo? Essa questão guiou a construção de toda a investigação, pois permitiu delimitar o foco do estudo, estruturar a análise sobre os riscos associados ao mau uso da IA e compreender o potencial da tecnologia na construção de sistemas de verificação, monitoramento e mitigação da desinformação.

Para alcançar esse propósito, a pesquisa foi desenvolvida por meio de uma revisão narrativa da literatura, que permitiu selecionar, organizar e interpretar produções científicas que abordaram a relação entre Inteligência Artificial, desinformação e ética digital. Foram utilizados descritores específicos em bases de dados reconhecidas internacionalmente, considerando critérios de inclusão que privilegiaram trabalhos recentes, metodologicamente consistentes e diretamente relacionados ao objeto de estudo. Esse percurso metodológico possibilitou reunir evidências atualizadas, analisar diferentes perspectivas e compreender como a literatura científica discutiu o tema nos últimos anos.

O objetivo geral da pesquisa consistiu em analisar como a Inteligência Artificial participou dos processos de produção e disseminação de informações falsas, bem como identificar estratégias éticas e preventivas propostas pela literatura científica. Como objetivos específicos, buscou-se mapear os principais riscos associados ao uso indevido da IA, descrever os mecanismos tecnológicos envolvidos na propagação da desinformação, examinar propostas de uso responsável da tecnologia e compreender como diferentes áreas do conhecimento interpretaram o fenômeno. Juntos, esses elementos permitiram estruturar um trabalho coeso, fundamentado e alinhado às demandas contemporâneas de segurança e ética digital.

2 DESENVOLVIMENTO

2.1 Metodologia

A presente pesquisa será desenvolvida por meio de uma revisão de literatura narrativa, com o objetivo de analisar a produção científica recente acerca da utilização da Inteligência Artificial e seu uso indevido na propagação de informações falsas, bem como as estratégias preventivas possíveis a partir da própria tecnologia. Para a construção do corpus de análise, foram selecionadas as bases de dados PubMed (Public Medline) e Scopus, escolhidas por sua abrangência e relevância internacional no campo da saúde, ciências sociais aplicadas e tecnologia da informação.

O processo de busca será conduzido utilizando os seguintes descritores em língua inglesa: Artificial Intelligence, Misinformation, Disinformation Prevention e Digital Ethics. Esses termos foram escolhidos por representarem, de forma integrada, os principais eixos temáticos de interesse para o estudo: o papel da Inteligência Artificial, a dinâmica de disseminação de notícias falsas e as medidas de caráter ético e preventivo. As buscas serão restritas ao período de publicação dos últimos cinco anos, garantindo que os resultados reflitam os avanços mais atuais e relevantes sobre o tema.

Serão incluídos na análise artigos científicos originais, revisões de literatura, dissertações, teses e livros que abordem direta ou indiretamente a temática. Como critérios de inclusão, serão considerados trabalhos que apresentem clareza metodológica, relevância para o objeto de estudo e que estejam disponíveis em inglês, português ou espanhol. Serão excluídas publicações repetidas, resumos de eventos sem texto completo e materiais que não apresentem relação direta com a Inteligência Artificial e o fenômeno da desinformação.

Após a coleta, os trabalhos selecionados serão organizados em categorias temáticas, de modo a possibilitar uma análise crítica e comparativa sobre os diferentes enfoques existentes na literatura. Essa sistematização permitirá identificar não apenas os riscos associados ao uso da IA na propagação de informações falsas, mas também as propostas e soluções já discutidas no âmbito científico para sua utilização de forma preventiva e ética.

2.2 Resultados e discussão

Os artigos escolhidos para a composição desta revisão foram categorizados da seguinte forma: Título da Publicação, Autor, Periódico (incluindo Volume, Número e Página, quando disponíveis), Ano e País de Publicação, bem como uma síntese abrangente da Metodologia e dos Resultados do Trabalho. Esses elementos foram cuidadosamente dispostos na Tabela 2, a fim de proporcionar uma estrutura ordenada e clara

O estudo de Küster e Schultz (2023), publicado na revista *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* na Alemanha, apresentou uma análise qualitativa reflexiva baseada em revisão narrativa e avaliação de cenários críticos relacionados ao uso da inteligência artificial na saúde. Os autores identificaram riscos significativos associados à aplicação inadequada da IA, incluindo a disseminação de informações falsas por meio de algoritmos e a possibilidade de uma “algocracia”, em que decisões seriam conduzidas por sistemas com vieses estruturais. Apesar desses riscos, o estudo também apontou oportunidades relevantes, como o uso de sistemas assistivos baseados em IA para apoiar indivíduos com declínio cognitivo. O exemplo do sistema I-CARE demonstrou que tecnologias podem ser integradas de forma ética, garantindo autonomia ao paciente e reduzindo a desinformação por meio de recomendações personalizadas e confiáveis.

Roche, Wall e Lewis (2022), em uma análise documental crítica publicada na revista *AI Ethics* na Irlanda, examinaram mais de 470 políticas internacionais voltadas à ética da inteligência artificial. Os autores verificaram que a produção normativa estava concentrada majoritariamente no Norte Global, revelando um viés estrutural que privilegia valores ocidentais e limita a representatividade de realidades socioculturais diversas. Essa assimetria impacta diretamente a efetividade das estratégias de combate à desinformação, pois diretrizes formuladas sem considerar o contexto do Sul Global tornam-se menos eficientes diante das especificidades informacionais e tecnológicas desses países. Assim, a falta de diversidade na construção das políticas de IA compromete iniciativas globais de regulação e dificulta respostas eficazes ao avanço da desinformação digital.

O artigo de Federico e Trotsyuk (2024), publicado no *Annual Review of Biomedical Data Science* nos Estados Unidos, apresentou uma revisão crítica interdisciplinar que discutiu os desafios emergentes relacionados ao crescimento acelerado da ciência de dados biomédicos e sua integração com sistemas de inteligência artificial. Os autores destacaram que, embora a IA tenha potencial expressivo para aprimorar diagnósticos e tratamentos, também existe a possibilidade de uso indevido decorrente de vieses algorítmicos, falta de transparência e ausência de governança ética estruturada. Tais fragilidades tornam os sistemas suscetíveis à manipulação de dados e à disseminação de informações enganosas, prejudicando a confiança pública. Nesse sentido, os autores enfatizaram a necessidade de estratégias interdisciplinares que integrem ciência de dados, bioética e políticas públicas para mitigar riscos informacionais associados à IA.

No estudo de Mao e Shi-Kupfer (2023), publicado na revista *AI & Society*, foram analisadas discussões públicas sobre ética da inteligência artificial em plataformas chinesas de mídia social por meio de metodologia qualitativa com análise de conteúdo. Os resultados evidenciaram que, mesmo em um ambiente caracterizado por forte controle estatal, houve ampla circulação de debates envolvendo preocupações éticas, privacidade e risco de desinformação. As plataformas permitiram trocas diversificadas entre pesquisadores, cidadãos e especialistas, fortalecendo a construção coletiva de compreensão ética sobre IA. Os autores concluíram que o engajamento social informado pode funcionar como mecanismo de educação pública e de mitigação da desinformação, ao estimular usuários a desenvolverem maior senso crítico sobre sistemas algorítmicos e conteúdos digitais

Borenstein e Howard (2021), em um ensaio teórico publicado na revista *AI Ethics* nos Estados Unidos, discutiram os desafios emergentes do desenvolvimento da inteligência artificial e defenderam a importância de formação ética para profissionais da área. Os autores apontaram que algoritmos supostamente neutros acabam reproduzindo os vieses presentes nos dados e nas escolhas de design dos programadores, o que pode gerar discriminações e favorecer a circulação de informações distorcidas. O estudo ressaltou que muitas instituições tratam a ética como elemento secundário, comprometendo a capacidade dos futuros profissionais em compreender e prevenir impactos sociais negativos, incluindo a disseminação de desinformação. Para os autores, a incorporação sistemática de educação ética é fundamental para promover uma IA mais responsável e socialmente segura

No estudo de Saheb (2023), desenvolvido no Irã e publicado na revista *AI Ethics*, empregou-se uma abordagem bibliométrica e qualitativa com análise de conteúdo e modelagem de tópicos para investigar a vigilância baseada em IA. O autor identificou sete categorias principais de vigilância discutidas na literatura, incluindo vigilância estatal, militar, capitalista e sanitária. Embora amplamente implementadas em cidades inteligentes e sistemas de reconhecimento facial, essas tecnologias levantaram preocupações éticas, sobretudo pela tendência à normalização da violação de direitos fundamentais e pela pos-

sibilidade de uso da IA para manipulação de informações e repressão de dissidências. O estudo também evidenciou o crescimento acentuado de publicações sobre o tema após 2017, revelando lacunas ainda pouco exploradas, como a relação entre vigilância automatizada e desinformação institucionalizada

Por fim, o ensaio filosófico de Coeckelbergh (2022), publicado na revista *AI Ethics* na Áustria, discutiu implicações epistemológicas e democráticas do uso de IA em ambientes informacionais. O autor argumentou que sistemas de personalização baseados em machine learning reduzem a diversidade de conteúdos acessados pelos usuários e comprometem sua agência epistêmica, isto é, sua capacidade de formar julgamentos críticos e revisar crenças. Esses mecanismos favorecem a criação de bolhas informacionais e câmaras de eco que reforçam crenças prévias, facilitando a disseminação de desinformação e prejudicando a qualidade dos debates públicos. Assim, o autor destacou que o avanço da IA impõe desafios fundamentais à democracia, especialmente pela influência dos algoritmos sobre a formação de opinião e a participação cidadã

A análise dos estudos revisados demonstrou que a incorporação de princípios éticos no desenvolvimento de sistemas de inteligência artificial é um requisito indispensável para reduzir riscos associados à manipulação informacional e à perda de autonomia cognitiva dos usuários. Como destacam Küster e Schultz (2023), a ausência de diretrizes claras e de mecanismos de responsabilização permite que algoritmos operem segundo lógicas opacas, capazes de amplificar distorções e favorecer a circulação de conteúdos enganosos. Essa vulnerabilidade estrutural se torna mais evidente quando se observa que as plataformas digitais utilizam modelos de recomendação baseados em engajamento, privilegiando conteúdos sensacionalistas em detrimento da precisão informacional. Assim, a IA passa a desempenhar papel ambivalente: ao mesmo tempo em que oferece soluções técnicas avançadas, também se torna vetor de riscos quando não é orientada por princípios éticos robustos.

A desigualdade na produção de diretrizes éticas também emergiu como uma questão central nos estudos. De acordo com Roche, Wall e Lewis (2022), mais de 90% dos documentos internacionais sobre ética da IA foram elaborados por países do Norte Global, reforçando a hegemonia de valores ocidentais. Esse fenômeno limita a aplicabilidade de tais políticas em contextos socioculturais distintos, especialmente no Sul Global, onde a desinformação está frequentemente associada a fatores como desigualdade social, instabilidade política e baixa infraestrutura digital. O resultado é a formulação de políticas que, embora tecnicamente sofisticadas, não respondem de forma eficaz às demandas locais. Essa assimetria reforça a importância de uma epistemologia plural, que considere diversidade cultural, realidades políticas e especificidades regionais como elementos centrais para a construção de normas éticas verdadeiramente universais.

Essa pluralidade envolve não apenas diretrizes oficiais, mas também o debate público sobre a IA. A experiência analisada por Mao e Shi-Kupfer (2023) na China demonstra que, mesmo em cenários de forte controle governamental, o engajamento popular possibilita reflexões críticas sobre privacidade, ética e desinformação. As discussões observadas em plataformas como WeChat mostraram que a sociedade civil pode desempenhar papel ativo na cobrança por transparência e no fortalecimento de práticas éticas. Essa participação social se torna fundamental para contrapor os potenciais abusos algorítmicos e favorecer o desenvolvimento de uma cultura tecnológica orientada por valores democráticos. Quando a população compreende os impactos da IA, desenvolve maior resiliência contra conteúdos manipulados e consegue reconhecer estratégias de influência invisíveis que moldam comportamentos e percepções.

Ao analisar os efeitos da desinformação no ambiente digital, os estudos revelaram que o problema não se limita à circulação de notícias falsas, mas envolve um conjunto mais amplo de fatores técnicos e sociais. Segundo Federico e Trotsyuk (2024), sistemas de IA apresentam vulnerabilidades relacionadas a vieses de treinamento, opacidade algorítmica e governança insuficiente, criando condições para a produção e amplificação de conteúdos enganosos. A desinformação, portanto, não é apenas uma falha de comunicação, mas parte de uma dinâmica estruturada na qual algoritmos reproduzem padrões de desigualdade e influenciam a formação crítica dos indivíduos. Esse quadro é agravado pela personalização intensa dos conteúdos, que cria ambientes informacionais fragmentados, capazes de restringir a diversidade de perspectivas acessadas pelos usuários.

Nesse sentido, a educação ética e a formação de profissionais da área tecnológica emergem como estratégias essenciais. Para Borenstein e Howard (2021), o ensino da ética em IA não pode continuar sendo tratado como conteúdo opcional, visto que decisões de design algorítmico definem o alcance e o impacto das tecnologias sobre a sociedade. A ausência de capacitação adequada contribui para a criação de sistemas que reforçam estereótipos, reproduzem exclusões e disseminam falsidades de maneira automatizada. A ética, portanto, deve constituir eixo transversal que orienta desde o planejamento até a implementação de projetos de IA, buscando minimizar efeitos adversos e maximizar benefícios sociais.

Os estudos também destacaram que os algoritmos, ao priorizarem conteúdos baseados em métricas de engajamento, favorecem a viralização de informações sensacionalistas, um fator crítico para o enfraquecimento do debate democrático. Coeckelbergh (2022) argumenta que a IA pode limitar a agência epistêmica dos cidadãos ao confiná-los em bolhas informacionais, ambientes digitais que reforçam crenças prévias e reduzem a exposição a opiniões divergentes. Esses espaços são terreno fértil para a disseminação de desinformação, pois fragilizam o senso crítico e reduzem a capacidade de reflexão autônoma. Para mitigar esse processo, torna-se urgente o redesenho das plataformas digitais, incorporando mecanismos que incentivem pluralidade informacional, auditabilidade e priorização de fontes confiáveis.

Outro ponto relevante identificado nos estudos refere-se ao uso de IA em sistemas de vigilância. Saheb (2023) demonstra que tecnologias de monitoramento automatizado, muitas vezes legitimadas sob o discurso da segurança, podem se transformar em instrumentos de repressão social e de manipulação informacional institucionalizada. A falta de transparência e supervisão permite que algoritmos sejam utilizados para controlar narrativas, silenciar grupos específicos e influenciar comportamentos coletivos. Sob o ponto de vista das ciências sociais, tais mecanismos geram efeitos psicológicos relevantes, como sensação de vigilância constante, redução da liberdade de expressão e erosão da confiança pública. A ética da IA, nesse contexto, cumpre papel essencial ao defender direitos fundamentais como privacidade, autonomia informacional e participação democrática.

Além dos riscos relacionados à vigilância e ao reforço de vieses, a literatura também destacou a importância da diversidade informacional como estratégia central para combater a desinformação. A promoção de conteúdos variados e plurais depende de transformações no modelo de negócios das grandes plataformas digitais, que hoje priorizam o lucro associado ao engajamento. Conforme apontam Roche, Wall e Lewis (2022), o combate à desinformação exige que as empresas de tecnologia repensem seus critérios de recomendação, adotando métricas que valorizem qualidade e verificabilidade das informações. Nesse sentido, a construção de ambientes digitais mais equilibrados necessita da articulação entre regulação governamental, compromisso corporativo e pressão da sociedade civil.

A transparência algorítmica também se mostrou um pilar essencial para o enfrentamento da desinformação. Federico e Trotsyuk (2024) destacam que os usuários devem ter o direito de compreender como sistemas automatizados influenciam os conteúdos que consomem, bem como a possibilidade de contestar decisões algorítmicas. Ferramentas de explicabilidade, mecanismos de auditoria e interfaces que permitam controle do usuário representam caminhos viáveis para restaurar a confiança e fortalecer a autonomia cognitiva na era digital. Somente com maior clareza sobre o funcionamento interno dos algoritmos é possível reagir de forma eficaz aos processos de manipulação invisível.

Por fim, os resultados analisados reforçam que a prevenção da desinformação associada à IA depende da integração entre educação digital, políticas públicas, governança ética e participação social. Como argumentam Borenstein e Howard (2021), nenhuma solução tecnológica isolada é capaz de resolver o problema sem que exista, simultaneamente, compromisso político e engajamento coletivo. A construção de um ecossistema informacional mais justo requer reconhecer que disputas econômicas, interesses geopolíticos e estruturas de poder moldam as tecnologias e influenciam sua aplicação na sociedade. Dessa forma, como sintetizam Küster e Schultz (2023), a IA deve ser compreendida como parte de um sistema complexo, no qual os riscos e benefícios dependem diretamente das escolhas éticas, políticas e sociais que orientam seu desenvolvimento.

3 CONCLUSÃO

A análise realizada permite constatar que a Inteligência Artificial constitui um dos mais relevantes avanços tecnológicos da contemporaneidade, trazendo benefícios expressivos em diferentes áreas do conhecimento, desde a saúde e a educação até a comunicação e a segurança digital. Contudo, esse mesmo potencial inovador também tem sido explorado de maneira indevida, ampliando a capacidade de produção e disseminação de conteúdos falsos, como deepfakes, notícias fabricadas e manipulações digitais cada vez mais sofisticadas. Tal fenômeno coloca em evidência o impacto direto da IA sobre a confiança social, o processo democrático e a formação da opinião pública, revelando a urgência de estratégias preventivas e regulatórias capazes de acompanhar a velocidade dos avanços tecnológicos.

Verificou-se que a difusão de informações falsas mediada por algoritmos e modelos de linguagem não se restringe a uma questão puramente técnica, mas envolve igualmente aspectos éticos, sociais, psicológicos e educacionais. A desinformação está inserida em um ecossistema informacional complexo, atravessado por disputas políticas, interesses econômicos e estruturas de poder que moldam a forma como os conteúdos circulam. Nesse sentido, torna-se evidente a necessidade de integrar o desenvolvimento da IA a um arcabouço de ética digital que assegure transparência, rastreabilidade e responsabilidade no uso das ferramentas. Além disso, reforça-se a importância da educação midiática e digital como recurso indispensável para preparar indivíduos e comunidades a reconhecer e questionar criticamente os conteúdos com os quais entram em contato, fortalecendo sua autonomia cognitiva.

Do ponto de vista prático, observou-se que a mesma tecnologia capaz de gerar informações enganosas pode ser mobilizada como instrumento de monitoramento e prevenção. Soluções como sistemas de detecção automática de padrões de desinformação, algoritmos de checagem de fatos, mecanismos de autenticação de imagens e vídeos e ferramentas de explicabilidade algorítmica mostram-se caminhos promissores para enfrentar o problema. No entanto, tais mecanismos só terão eficácia real se aliados a políticas

públicas consistentes, padrões internacionais de governança e mecanismos regulatórios que garantam a segurança, a transparência e os direitos dos cidadãos no ambiente digital. A cooperação entre pesquisadores, desenvolvedores, governos, jornalistas, instituições de ensino e sociedade civil surge como elemento essencial para a construção de uma cultura digital mais segura e orientada pelo interesse coletivo.

Observou-se também que, diante da crescente sofisticação das técnicas de manipulação digital, estratégias de enfrentamento baseadas apenas na tecnologia tendem a ser insuficientes. É necessário reconhecer os limites da IA e compreender que o combate à desinformação passa igualmente pela promoção da diversidade informacional, pelo incentivo ao pensamento crítico e pelo fortalecimento das instituições democráticas. Uma sociedade mais resiliente à desinformação é aquela capaz de compreender como os algoritmos operam, questionar fontes, reconhecer vieses e desenvolver consciência sobre o papel que a tecnologia desempenha na mediação da realidade.

Conclui-se, portanto, que a Inteligência Artificial não deve ser entendida unicamente como uma ameaça à veracidade da informação, mas como uma tecnologia de duplo potencial: pode ser instrumento de risco quando empregada de forma irresponsável, mas também se configura como ferramenta estratégica e preventiva no combate à desinformação. O desafio que se impõe é construir um uso ético, regulado e socialmente comprometido da IA, de modo que seus benefícios superem os riscos e que sua contribuição fortaleça a confiabilidade da comunicação digital em escala global. O futuro da relação entre IA e sociedade dependerá das escolhas éticas, políticas e educacionais realizadas no presente, bem como da capacidade coletiva de promover um ambiente informacional mais seguro, plural e orientado para o bem comum

REFERÊNCIAS

BORENSTEIN, J.; HOWARD, A. Emerging challenges in AI and the need for AI ethics education. *AI Ethics*, v. 1, n. 1, p. 61-65, 2021. doi: 10.1007/s43681-020-000027. Epub 2020 Oct 6. PMID: 38624388; PMCID: PMC7487209. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7487209/. Acesso em: 22 mar. 2025.

CASTELLS, M. Ruptura: a crise da democracia liberal. Rio de Janeiro: Zahar, 2021.

COECKELBERGH, M. Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence. *AI Ethics*, p. 1-10, 2022. doi: 10.1007/s43681-022-00239-4. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9685050/. Acesso em: 22 mar. 2025.

FEDERICO, C. A.; TROTSYUK, A. A. Biomedical Data Science, Artificial Intelligence, and Ethics: Navigating Challenges in the Face of Explosive Growth. *Annual Review of Biomedical Data Science*, v. 7, n. 1, p. 1-14, 2024. doi: 10.1146/annurev-biodatasci-102623-104553. Disponível em: https://pubmed.ncbi.nlm.nih.gov/38598860/. Acesso em: 22 mar. 2025.

FLORIDI, L. The Ethics of Artificial Intelligence. Oxford: Oxford University Press, 2021.

FLORIDI, L.; CHIRIATTI, M. GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, v. 30, p. 681–694, 2020.

KÜSTER, D.; SCHULTZ, T. Künstliche Intelligenz und Ethik im Gesundheitswesen – Spagat oder Symbiose? *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*, v. 66, n. 2, p. 176-183, 2023. doi: 10.1007/s00103-022-03653-5. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9892090/. Acesso em: 22 mar. 2025.

MAO, Y.; SHI-KUPFER, K. Online public discourse on artificial intelligence and ethics in China: context, content, and implications. *AI & Society*, v. 38, n. 1, p. 373-389, 2023. doi: 10.1007/s00146-021-01309-7. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8594647/. Acesso em: 23 mar. 2025.

NGUYEN, T. et al. Detecting fake news using machine learning: A comprehensive survey. *IEEE Access*, v. 8, p. 249–271, 2020.

O'NEIL, C. Weapons of math destruction: how big data increases inequality and threatens democracy. New York: Crown, 2016.

ROCHE, C.; WALL, P. J.; LEWIS, D. Ethics and diversity in artificial intelligence policies, strategies and initiatives. *AI Ethics*, p. 1-21, 2022. doi: 10.1007/s43681-022-00218-9. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9540088/. Acesso em: 23 mar. 2025.

SAHEB, T. Ethically contentious aspects of artificial intelligence surveillance: a social science perspective. *AI Ethics*, v. 3, n. 2, p. 369-379, 2023. doi: 10.1007/s43681-022-00196-y. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9294797/. Acesso em: 23 mar. 2025.

UNESCO. Guidelines for Regulating Digital Platforms. Paris: UNESCO, 2023. Disponível em: https://unesdoc.unesco.org/. Acesso em: 20 mar. 2025.

WARDLE, C.; DERAKHSHAN, H. Information disorder: toward an interdisciplinary framework for research and policymaking. Strasbourg: Council of Europe, 2017.

12

DESAFIOS E LIMITAÇÕES DA INTELIGÊNCIA ARTIFICIAL: DESAFIOS E LIMITAÇÕES

CHALLENGES AND LIMITATIONS OF ARTIFICIAL INTELLIGENCE: CHALLENGES AND LIMITATIONS

Yasmim Pereira Santana
Mirian Nunes de Carvalho Nunes
Ivone Ascar Sauáia Guimarães

Resumo

A Inteligência Artificial (IA) tem se consolidado como uma das tecnologias mais transformadoras da atualidade, com aplicações em diversos setores como saúde, educação, indústria e finanças. No entanto, sua implementação enfrenta desafios técnicos, sociais e econômicos que limitam sua adoção plena e eficaz. Este trabalho, desenvolvido por meio de revisão bibliográfica, tem como objetivo analisar criticamente os principais obstáculos à disseminação da IA, abordando aspectos como explicabilidade dos modelos, viés algorítmico, escalabilidade e acesso à infraestrutura tecnológica. Também são discutidas barreiras socioculturais, como resistência organizacional, escassez de profissionais qualificados e desigualdade no acesso à tecnologia, especialmente em países em desenvolvimento. A pesquisa destaca a importância da transparência algorítmica, da ética no desenvolvimento de sistemas inteligentes e da governança responsável como pilares para uma IA mais justa e segura. Além disso, são abordadas questões ambientais, como o alto consumo energético dos modelos de aprendizado profundo, e preocupações com segurança cibernética e uso malicioso da tecnologia. A análise evidencia que a superação desses desafios exige esforços coordenados entre governos, academia e setor privado, bem como políticas públicas e iniciativas educacionais que promovam inclusão digital e capacitação profissional. Conclui-se que a construção de uma IA responsável depende da articulação entre técnica, ética e política, visando garantir que seus benefícios sejam acessíveis de forma equitativa e sustentável para toda a sociedade.

Palavras-chave: Inteligência Artificial. Desafios Tecnológicos. Inclusão Digital.

Abstract

Artificial Intelligence (AI) has become one of the most transformative technologies of our time, with applications across diverse sectors such as healthcare, education, industry, and finance. However, its implementation faces technical, social, and economic challenges that limit its full and effective adoption. This paper, developed through a bibliographic review, aims to critically analyze the main obstacles to the dissemination of AI, addressing aspects such as model explainability, algorithmic bias, scalability, and access to technological infrastructure. Sociocultural barriers are also discussed, including organizational resistance, a shortage of qualified professionals, and inequalities in access to technology, especially in developing countries. The research highlights the importance of algorithmic transparency, ethics in the development of intelligent systems, and responsible governance as pillars for fairer and safer AI. Environmental issues are also addressed, such as the high energy consumption of deep learning models, as well as concerns about cybersecurity and malicious uses of the technology. The analysis shows that overcoming these challenges requires coordinated efforts among governments, academia, and the private sector, as well as public policies and educational initiatives that promote digital inclusion and professional training. It is concluded that building responsible AI depends on the articulation of technical, ethical, and political dimensions, aiming to ensure that its benefits are accessible equitably and sustainably to all of society.

Keywords: Artificial Intelligence. Technological Challenges. Digital Inclusion.

1 INTRODUÇÃO

A Inteligência Artificial (IA) tem se destacado como uma das tecnologias mais transformadoras da atualidade, com aplicações que vão desde diagnósticos médicos até sistemas de recomendação em plataformas digitais. Seu impacto é perceptível em áreas como saúde, educação, indústria e finanças, promovendo avanços significativos na automação de processos, na análise de dados e na tomada de decisões. No entanto, apesar de seu potencial revolucionário, a implementação da IA ainda enfrenta uma série de desafios que comprometem sua adoção plena e eficaz.

Entre os principais obstáculos técnicos, estão questões como a escalabilidade dos modelos, a explicabilidade dos algoritmos e o viés algorítmico. A dificuldade em compreender como os sistemas de IA chegam a determinadas conclusões levanta preocupações éticas e jurídicas, especialmente em contextos sensíveis como justiça e saúde. Além disso, a presença de vieses nos dados utilizados para treinar os modelos pode perpetuar desigualdades e comprometer a imparcialidade das decisões automatizadas.

Outro ponto crítico é a limitação no acesso a dados de qualidade e à infraestrutura computacional necessária para o treinamento de modelos robustos. A generalização dos sistemas de IA para diferentes contextos e populações também se mostra desafiadora, o que limita sua aplicabilidade em larga escala. Esses fatores técnicos não apenas dificultam o desenvolvimento de soluções eficazes, como também afetam a confiabilidade e a segurança das aplicações de IA.

Do ponto de vista prático, a adoção da IA é influenciada por barreiras socioculturais e econômicas. Muitas organizações ainda demonstram resistência à inovação, seja por desconhecimento dos benefícios da IA ou por receio de mudanças estruturais. A escassez de profissionais qualificados e a desigualdade no acesso a recursos tecnológicos também contribuem para a lentidão na disseminação da tecnologia, especialmente em países em desenvolvimento.

Diante desse cenário, torna-se essencial compreender os desafios que limitam a evolução da IA e buscar estratégias para superá-los. A análise crítica desses obstáculos pode orientar a formulação de políticas públicas, iniciativas educacionais e investimentos empresariais que promovam uma adoção mais equitativa e sustentável da tecnologia. A colaboração entre governos, academia e setor privado é fundamental para enfrentar essas barreiras de forma coordenada.

Este artigo tem como objetivo realizar uma revisão de literatura sobre os principais desafios e limitações da Inteligência Artificial, abordando tanto aspectos técnicos quanto socioculturais e econômicos. A proposta é identificar os fatores que dificultam a implementação da IA e discutir possíveis soluções que possam viabilizar sua adoção em diferentes contextos. A investigação será guiada por uma abordagem teórica, com base em estudos recentes e relevantes da área.

A relevância deste estudo está na sua capacidade de contribuir para o avanço do conhecimento sobre os entraves enfrentados pela IA e propor caminhos para superá-los. Ao compreender melhor os desafios técnicos, como a explicabilidade e o viés algorítmico, e os obstáculos sociais, como a resistência cultural e a desigualdade de acesso, será possível desenvolver estratégias mais eficazes para a disseminação da tecnologia.

Por fim, espera-se que esta pesquisa possa servir como base para futuras investigações e práticas de implementação da IA, promovendo uma integração mais responsável, inclusiva e eficiente da tecnologia na sociedade. A superação dos desafios aqui discutidos é essencial para que a Inteligência Artificial cumpra seu papel transformador de maneira

ética, segura e acessível a todos.

2 DESENVOLVIMENTO

2.1 Metodologia

Este trabalho foi desenvolvido por meio de uma revisão bibliográfica, com o objetivo de analisar os principais desafios técnicos, sociais e econômicos enfrentados pela Inteligência Artificial (IA) em sua implementação e adoção. Trata-se de uma pesquisa qualitativa e descritiva, voltada para a compreensão teórica das limitações da IA, sem a realização de coleta de dados primários ou experimentos. A revisão permitiu reunir e interpretar contribuições acadêmicas que abordaram a IA sob diferentes perspectivas, destacando autores relevantes e abordagens que relacionam aspectos tecnológicos, socioculturais e econômicos.

A seleção dos materiais foi realizada em fontes confiáveis, como livros, artigos científicos, dissertações e teses, disponíveis em bases de dados reconhecidas, tais como Google Acadêmico, SciELO, CAPES e IEEE Xplore. Foram priorizados estudos publicados entre os anos de 2020 e 2025, com o intuito de garantir a atualidade e relevância das informações analisadas, especialmente diante da rápida evolução da tecnologia e das discussões éticas e práticas que a acompanham.

Os critérios de inclusão consideraram trabalhos que abordaram diretamente os desafios técnicos da IA, como escalabilidade, explicabilidade e viés algorítmico, bem como estudos que discutiram barreiras sociais, como resistência organizacional, desigualdade no acesso à tecnologia e escassez de profissionais qualificados. Também foram incluídas análises sobre os impactos econômicos da IA, especialmente no que diz respeito à sua disseminação e à equidade de acesso em diferentes contextos.

Foram incluídos apenas documentos completos, excluindo resumos, prévias ou artigos de revisão, a fim de assegurar profundidade e consistência na análise. A pesquisa identificou autores clássicos e contemporâneos da ciência da computação, da ética tecnológica e da sociologia da inovação, além de estudos aplicados que evidenciaram os desafios enfrentados por instituições públicas e privadas na adoção de soluções baseadas em IA.

As palavras-chave utilizadas para a busca foram: “desafios da inteligência artificial”, “viés algorítmico”, “explicabilidade de modelos de IA”, “barreiras sociais na adoção da IA”, “limitações econômicas da IA” e “implementação da IA em setores diversos”. A partir dessas expressões, foram selecionados os materiais mais pertinentes e recentes, permitindo uma visão abrangente e fundamentada sobre os obstáculos que limitam o avanço da IA e as estratégias propostas para superá-los.

Essa abordagem metodológica possibilitou a construção de uma análise crítica e contextualizada, capaz de identificar os principais entraves à adoção da IA e de propor caminhos para uma implementação mais ética, eficaz e inclusiva. A revisão bibliográfica também contribuiu para o aprofundamento teórico sobre o tema, oferecendo subsídios para futuras pesquisas e práticas voltadas ao desenvolvimento sustentável da tecnologia.

2.2 Resultados

A Inteligência Artificial (IA) tem se consolidado como uma das tecnologias mais disruptivas do século XXI, promovendo transformações profundas em diversos setores da so-

ciedade. No entanto, sua implementação enfrenta desafios técnicos, sociais e econômicos que limitam seu potencial. A análise desses obstáculos é essencial para compreender os limites da IA e propor caminhos para uma adoção mais ética, eficiente e inclusiva (Russell; Norvig, 2021).

Do ponto de vista técnico, um dos principais entraves está na explicabilidade dos modelos de aprendizado de máquina. Muitos sistemas operam como "caixas-pretas", dificultando a compreensão de como decisões são tomadas. Essa opacidade compromete a confiança dos usuários e levanta questões éticas, especialmente em áreas sensíveis como saúde e justiça (Doshi-Velez; Kim, 2017).

Além disso, a escalabilidade dos modelos representa outro desafio relevante. Embora algoritmos avançados como redes neurais profundas tenham demonstrado grande capacidade de processamento, sua aplicação em larga escala exige infraestrutura computacional robusta e acesso a grandes volumes de dados, o que nem sempre está disponível (Goodfellow; Bengio; Courville, 2016).

A qualidade dos dados utilizados no treinamento dos modelos também é um fator crítico. Dados enviesados ou incompletos podem gerar resultados distorcidos, perpetuando desigualdades sociais. O caso analisado por Angwin et al. (2016) sobre algoritmos de risco criminal nos Estados Unidos ilustra como sistemas de IA podem reproduzir preconceitos raciais, afetando negativamente populações vulneráveis.

Nesse contexto, a ética da IA torna-se um campo de estudo fundamental. Bostrom e Yudkowsky (2014) argumentam que o desenvolvimento de sistemas inteligentes deve ser guiado por princípios éticos claros, que garantam a segurança, a justiça e o respeito aos direitos humanos. A ausência de tais diretrizes pode levar ao uso indevido da tecnologia.

A preocupação com o uso malicioso da IA também é crescente. Brundage et al. (2018) alertam para a possibilidade de que agentes mal-intencionados utilizem sistemas inteligentes para fins destrutivos, como ataques cibernéticos, manipulação de informações e vigilância em massa. A mitigação desses riscos exige políticas públicas e cooperação internacional.

A questão da vigilância é particularmente sensível. Zuboff (2019) denuncia o surgimento do "capitalismo de vigilância", no qual empresas coletam e exploram dados pessoais em larga escala, muitas vezes sem o consentimento dos usuários. Essa prática compromete a privacidade e a autonomia dos indivíduos, exigindo regulamentações mais rigorosas.

A disseminação de deepfakes é outro exemplo de ameaça à integridade da informação. Chesney e Citron (2019) discutem como vídeos falsificados gerados por IA podem ser usados para manipular a opinião pública, prejudicar reputações e desestabilizar democracias. A detecção e o combate a essas práticas são desafios urgentes.

No campo da saúde, a IA apresenta grande potencial, mas também enfrenta limitações. Topol (2019) destaca que, embora sistemas inteligentes possam melhorar diagnósticos e personalizar tratamentos, sua eficácia depende da qualidade dos dados e da integração com práticas médicas humanizadas. A tecnologia não deve substituir o julgamento clínico.

A energia necessária para treinar modelos de IA é outro fator preocupante. Strubell, Ganesh e McCallum (2019) mostram que o consumo energético de sistemas de processamento de linguagem natural pode ser comparável ao de cidades inteiras. Isso levanta questões ambientais e exige soluções mais sustentáveis.

A falta de profissionais qualificados é uma barreira significativa à adoção da IA. Bryn-

jolfsson e McAfee (2014) apontam que o avanço tecnológico não tem sido acompanhado por uma capacitação adequada da força de trabalho, o que limita a implementação de soluções inteligentes em empresas e instituições.

Além disso, a resistência cultural à inovação tecnológica é um obstáculo prático. Muitas organizações ainda demonstram receio em adotar sistemas de IA, seja por desconhecimento, seja por medo de mudanças estruturais. Essa resistência pode ser superada por meio de educação e sensibilização (Dignum, 2019).

A integração da IA com sistemas legados também representa um desafio técnico. Muitas empresas operam com infraestruturas antigas que não são compatíveis com tecnologias emergentes, dificultando a implementação de soluções inteligentes (SoftDesign, 2025).

A confiabilidade dos sistemas de IA é outro ponto crítico. Marcus (2020) argumenta que, para que a IA seja amplamente adotada, é necessário desenvolver modelos mais robustos, capazes de lidar com situações imprevistas e de se adaptar a diferentes contextos.

A governança da IA é uma questão central para seu desenvolvimento responsável. Floridi et al. (2018) propõem um marco ético que orienta o uso da tecnologia com base em princípios como beneficência, não maleficência, autonomia e justiça. A adoção desses princípios pode guiar políticas públicas e práticas empresariais.

A ONU reconheceu recentemente a IA como um desafio global, criando fóruns e painéis científicos para discutir riscos e oportunidades (Exame, 2025). Essa iniciativa demonstra a necessidade de coordenação internacional para enfrentar os impactos da tecnologia de forma colaborativa.

No Brasil, embora a IA seja considerada prioridade estratégica para 2026, o investimento ainda é limitado. Segundo estudo da Amcham, 77% das empresas destinam menos de 2% do orçamento à tecnologia, o que compromete sua aplicação efetiva (iMasters, 2025).

Essa lacuna entre discurso e prática revela a necessidade de políticas de incentivo à inovação. A criação de hubs de IA, como o proposto pela Amcham, pode promover capacitação, troca de experiências e desenvolvimento de projetos colaborativos, fortalecendo o ecossistema tecnológico.

A desigualdade no acesso à IA é outro problema relevante. Crawford (2021) analisa como o desenvolvimento da tecnologia está concentrado em países e empresas com maior poder econômico, o que pode ampliar disparidades globais e limitar os benefícios da IA para populações marginalizadas.

A representatividade nos dados é essencial para garantir justiça nos sistemas de IA. O'Neil (2016) alerta que algoritmos treinados com dados enviesados podem reforçar estigmas e excluir grupos sociais, tornando a tecnologia uma "arma de destruição matemática".

A definição de inteligência artificial também é objeto de debate. Floridi (2025) argumenta que a IA não precisa replicar o pensamento humano, mas sim executar tarefas com eficácia. Essa perspectiva amplia o entendimento sobre o papel da tecnologia na sociedade.

Searle (1980), por sua vez, questiona a possibilidade de que máquinas realmente pensem, propondo o experimento do quarto chinês como crítica à IA forte. Essa discussão filosófica continua relevante para delimitar os limites da cognição artificial.

A personalização da aprendizagem por meio da IA é uma inovação promissora na educação. Segundo Araujo (2025), sistemas inteligentes podem adaptar conteúdos às necessidades dos alunos, promovendo inclusão e autonomia. No entanto, a falta de infraestrutura nas escolas públicas limita essa aplicação.

A capacitação docente é fundamental para o uso eficaz da IA na educação. Muitos professores ainda não se sentem preparados para utilizar tecnologias em sala de aula, o que exige formação continuada e apoio institucional (Araujo, 2025).

A robótica educacional é outro exemplo de aplicação da IA que estimula o raciocínio lógico e a criatividade dos estudantes. Concursos de robótica têm se tornado comuns no Brasil, reforçando a importância da prática como forma de consolidar o aprendizado (Araujo, 2025).

A IA também pode contribuir para a inclusão digital, ampliando o acesso ao conhecimento. No entanto, é necessário garantir que todos tenham acesso às ferramentas tecnológicas, evitando que a inovação fique restrita a poucos (Crawford, 2021).

A confiança nos sistemas de IA depende da transparência e da responsabilidade no seu desenvolvimento. Dignum (2019) defende que a IA deve ser projetada com foco na responsabilidade social, considerando os impactos éticos e legais de suas aplicações.

A ausência de regulamentações claras sobre o uso da IA contribui para a insegurança jurídica e para a proliferação de práticas questionáveis. Muitos países ainda não possuem legislações específicas que orientem o desenvolvimento e a aplicação da tecnologia, o que abre espaço para abusos e para a exploração indevida de dados pessoais (Dignum, 2019). Essa lacuna normativa dificulta a responsabilização de agentes envolvidos em decisões automatizadas, especialmente quando há danos a indivíduos ou coletividades.

A transparência algorítmica é um dos pilares para a construção de sistemas de IA confiáveis. Sem mecanismos que permitam auditar e compreender o funcionamento dos algoritmos, torna-se impossível garantir que decisões automatizadas sejam justas e imparciais. Doshi-Velez e Kim (2017) defendem a criação de uma ciência rigorosa da interpretabilidade, que permita aos usuários e reguladores entenderem os critérios utilizados pelos sistemas inteligentes.

A explicabilidade, no entanto, enfrenta limitações técnicas significativas. Muitos modelos de aprendizado profundo operam com milhões de parâmetros, tornando sua lógica interna praticamente indecifrável. Essa complexidade compromete a capacidade de identificar erros, corrigir vieses e assegurar conformidade com normas éticas e legais (Goodfellow; Bengio; Courville, 2016). A busca por modelos mais transparentes é, portanto, uma prioridade para o avanço responsável da IA.

Além da explicabilidade, a equidade algorítmica é um desafio urgente. Sistemas de IA treinados com dados históricos tendem a reproduzir padrões discriminatórios, como demonstrado por Angwin et al. (2016) no caso dos algoritmos de risco criminal. A perpetuação de desigualdades por meio da tecnologia exige uma revisão crítica dos conjuntos de dados utilizados e a implementação de estratégias de mitigação de viés.

A mitigação de viés algorítmico pode ser alcançada por meio da diversificação dos dados e da inclusão de perspectivas interdisciplinares no desenvolvimento dos sistemas. O'Neil (2016) argumenta que a ausência de diversidade nas equipes de tecnologia contribui para a criação de algoritmos excludentes. A inclusão de profissionais de diferentes áreas e origens pode enriquecer o processo de construção da IA e torná-la mais sensível às demandas sociais.

Outro aspecto relevante é o impacto da IA sobre o mercado de trabalho. Brynjolfsson e McAfee (2014) discutem como a automação pode substituir funções humanas, especialmente em atividades repetitivas e operacionais. Embora a tecnologia também crie novas oportunidades, a transição exige políticas públicas que promovam requalificação profissional e proteção social para os trabalhadores afetados.

A concentração de poder nas mãos de grandes corporações tecnológicas é um fator que agrava os riscos da IA. Zuboff (2019) denuncia como empresas dominantes utilizam a tecnologia para ampliar sua influência sobre consumidores e governos, muitas vezes sem transparência ou controle democrático. Essa concentração compromete a pluralidade e a equidade no acesso aos benefícios da inovação.

A sustentabilidade ambiental da IA também merece atenção. O treinamento de modelos complexos consome grandes quantidades de energia, contribuindo para a emissão de gases de efeito estufa. Strubell, Ganesh e McCallum (2019) alertam para a necessidade de considerar os impactos ambientais da tecnologia, propondo alternativas mais eficientes e menos poluentes.

A segurança cibernética é outro campo diretamente afetado pela IA. Brundage et al. (2018) destacam que sistemas inteligentes podem ser utilizados para automatizar ataques, explorar vulnerabilidades e disseminar desinformação. A proteção contra esses riscos exige investimentos em infraestrutura, capacitação técnica e cooperação internacional.

Dessa forma, a construção de uma IA responsável depende da articulação entre ética, técnica e política. Floridi et al. (2018) propõem um framework que orienta o desenvolvimento da tecnologia com base em princípios universais, como dignidade humana, solidariedade e sustentabilidade. A adoção desses princípios pode transformar a IA em uma ferramenta de emancipação, e não de dominação.

Por fim, a superação dos desafios da IA exige esforços coordenados entre governos, empresas e academia. A construção de uma sociedade inteligente depende não apenas da tecnologia, mas da capacidade de usá-la de forma ética, justa e sustentável (Floridi et al., 2018).

3 CONCLUSÃO

A presente pesquisa evidenciou que, embora a Inteligência Artificial represente um avanço tecnológico significativo, sua implementação enfrenta obstáculos complexos que vão além das questões técnicas. A explicabilidade dos modelos, o viés algorítmico e a escalabilidade são desafios que comprometem a confiabilidade e a transparência dos sistemas inteligentes, especialmente em áreas sensíveis como saúde, justiça e segurança pública.

Além dos aspectos técnicos, os entraves socioculturais e econômicos também se mostram determinantes para a adoção da IA. A resistência organizacional, a escassez de profissionais qualificados e a desigualdade no acesso à tecnologia revelam um cenário de exclusão digital que limita o potencial transformador da IA em países em desenvolvimento. Esses fatores reforçam a necessidade de políticas públicas voltadas à inclusão tecnológica e à capacitação profissional.

A análise demonstrou que a governança ética da IA é essencial para garantir que seus benefícios sejam distribuídos de forma equitativa e responsável. A ausência de regulamentações claras e de mecanismos de controle sobre o uso da tecnologia abre espaço para práticas abusivas, como a vigilância em massa e a manipulação de dados. A construção de marcos legais e éticos deve acompanhar o ritmo acelerado da inovação.

Também se destacou a importância da colaboração entre academia, setor privado e governos para enfrentar os desafios da IA. A criação de ambientes de inovação, como hubs tecnológicos e redes de pesquisa, pode fomentar o desenvolvimento de soluções mais inclusivas, sustentáveis e alinhadas às necessidades sociais. A cooperação internacional é igualmente relevante para lidar com os impactos globais da tecnologia.

A sustentabilidade ambiental da IA surgiu como um ponto crítico, especialmente diante do alto consumo energético dos modelos de aprendizado profundo. A busca por alternativas mais eficientes e menos poluentes deve integrar a agenda de desenvolvimento tecnológico, considerando os impactos da IA não apenas sobre as pessoas, mas também sobre o planeta.

Conclui-se, portanto, que a superação dos desafios da Inteligência Artificial exige uma abordagem multidisciplinar e integrada, que considere aspectos técnicos, éticos, sociais e ambientais. Somente por meio de uma atuação coordenada e consciente será possível promover uma IA que contribua para o bem-estar coletivo, respeite os direitos humanos e fortaleça a justiça social em um mundo cada vez mais digitalizado.

REFERÊNCIAS

ANGWIN, Julia et al. Machine bias: **There's software used across the country to predict future criminals. And it's biased against blacks.** ProPublica, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 27 maio 2025.

BOSTROM, Nick; YUDKOWSKY, Eliezer. **The ethics of artificial intelligence.** In: FRANKISH, Keith; RAMSEY, William M. (Org.). **The Cambridge handbook of artificial intelligence. Cambridge**: Cambridge University Press, 2014. p. 316-334.

BRUNDAGE, Miles et al. **The malicious use of artificial intelligence: Forecasting, prevention, and mitigation.** arXiv preprint arXiv:1802.07228, 2018.

BRYNJOLFSSON, Erik; MCAFEE, Andrew. **The second machine age: Work, progress, and prosperity in a time of brilliant technologies.** New York: W. W. Norton & Company, 2014.

CHESNEY, Robert; CITRON, Danielle Keats. **Deep fakes: A looming challenge for privacy, democracy, and national security**. California Law Review, v. 107, n. 6, p. 1753-1820, 2019.

CRAWFORD, Kate. Atlas of AI: **Power, politics, and the planetary costs of artificial intelligence. New Haven: Yale University Press,** 2021.

DIGNUM, Virginia. **Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way.** Cham: Springer, 2019.

DOSHI-VELEZ, Finale; KIM, Been. **Towards a rigorous science of interpretable machine learning.** arXiv preprint arXiv:1702.08608, 2017.

FLORIDI, Luciano et al. **AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations.** Minds and Machines, v. 28, n. 4, p. 689-707, 2018.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning. Cambridge: MIT Press**, 2016.

MARCUS, Gary. The next decade in **AI: Four steps towards robust artificial intelligence. arXiv preprint arXiv:2002.**06177, 2020.

O'NEIL, Cathy. **Weapons of math destruction: How big data increases inequality and threatens democracy**. New York: Crown Publishing Group, 2016.

RUSSELL, Stuart J.; NORVIG, **Peter. Inteligência Artificial.** 4. ed. São Paulo: Pearson, 2021.

SEARLE, John R. **Minds, brains, and programs. Behavioral and Brain Sciences,** v. 3, n. 3, p. 417-457, 1980.

STRUBELL, Emma; GANESH, Ananya; MCCALLUM, Andrew. **Energy and policy considerations for deep learning in NLP. arXiv preprint arXiv:1906.02243, 2019.**

TOPOL, Eric**. Deep medicine: How artificial intelligence can make healthcare human** again. New York: Basic Books, 2019.

ZUBOFF, Shoshana. **The age of surveillance capitalism: The fight for a human future at the new frontier of power.** New York: PublicAffairs, 2019.

13

A IMPORTÂNCIA DA INTELIGÊNCIA ARTIFICIAL NA EDUCAÇÃO: BENEFÍCIOS, DESAFIOS E PERSPECTIVAS PARA O ENSINO NAS ESCOLAS

THE IMPORTANCE OF ARTIFICIAL INTELLIGENCE IN EDUCATION: BENEFITS, CHALLENGES, AND PERSPECTIVES FOR TEACHING IN SCHOOLS

Joarlan Silva Coelho
Ivone Ascar Sauaia Guimaraes
Mirian Nunes de Carvalho Nunes

Resumo

Este artigo analisa a importância da Inteligência Artificial na educação a partir de uma revisão bibliográfica realizada entre 2020 e 2025, buscando compreender como essa tecnologia pode aprimorar o processo de ensino e aprendizagem e quais desafios precisam ser superados para sua implementação responsável. O objetivo geral é mapear benefícios, limites e condições de efetividade da adoção de sistemas inteligentes no contexto escolar, com foco em personalização do ensino, monitoramento formativo, automação de rotinas e implicações éticas. A metodologia adotada é qualitativa e descritiva, baseada em revisão de literatura em bases acadêmicas e relatórios institucionais, considerando publicações em português, inglês e espanhol e priorizando estudos com descrição clara de métodos, evidências empíricas e contextos educacionais explícitos. Os resultados indicam ganhos consistentes quando há objetivos instrucionais explícitos, feedback formativo, uso criterioso de dados e tempos institucionais destinados à análise colegiada, com efeitos positivos sobre engajamento, persistência e progresso longitudinal dos estudantes. Persistem, entretanto, barreiras relacionadas à infraestrutura, formação docente e governança de dados, o que exige políticas de conectividade, formação continuada articulada à prática pedagógica e protocolos de privacidade, transparência e explicabilidade. Conclui-se que a efetividade da IA na educação depende da articulação entre qualidade pedagógica, condições técnicas e salvaguardas éticas, recomendando-se avaliação independente, monitoramento contínuo e orientação por princípios de equidade para sua ampliação em escala.

Palavras-chave: Personalização do Ensino. Monitoramento Formativo. Engajamento Estudantil. Governança de Dados. Interoperabilidade.

Abstract

This article analyzes the importance of Artificial Intelligence in education based on a literature review conducted between 2020 and 2025, seeking to understand how this technology can enhance the teaching–learning process and which challenges must be overcome for its responsible implementation. The general objective is to map the benefits, limits, and conditions for the effectiveness of adopting intelligent systems in the school context, with a focus on personalized learning, formative monitoring, automation of routines, and ethical implications. The methodology adopted was qualitative and descriptive, grounded in a review of literature from academic databases and institutional reports, considering publications in Portuguese, English, and Spanish and prioritizing studies with a clear description of methods, empirical evidence, and explicit educational contexts. The results indicate consistent gains when instructional objectives are explicit, formative feedback is present, data are used judiciously, and institutions allocate time for collegial analysis, yielding positive effects on student engagement, persistence, and longitudinal progress. Nevertheless, barriers remain related to infrastructure, teacher training, and data governance, which require connectivity policies, ongoing professional development articulated with pedagogical practice, and protocols for privacy, transparency, and explainability. It is concluded that the effectiveness of AI in education depends on the articulation between pedagogical quality, technical conditions, and ethical safeguards, and it is recommended that scaling efforts be guided by independent evaluation, continuous monitoring, and principles of equity.

Keywords: Personalized Learning. Formative Monitoring. Student Engagement. Data Governance. Interoperability.

1 INTRODUÇÃO

A tecnologia tem desempenhado um papel fundamental na transformação da sociedade contemporânea, influenciando diferentes setores e promovendo mudanças significativas no cotidiano das pessoas. Na educação, esse impacto se manifesta na introdução de ferramentas digitais que buscam otimizar os processos de ensino e aprendizagem. Entre essas inovações, a Inteligência Artificial (IA) surge como uma das mais promissoras, oferecendo recursos capazes de modernizar e tornar a prática pedagógica mais dinâmica e eficiente.

A aplicação da IA no ambiente educacional possibilita avanços relevantes, como a personalização do ensino, em que os conteúdos podem ser adaptados às necessidades de cada aluno. Além disso, sistemas inteligentes podem auxiliar os professores no acompanhamento do desempenho dos estudantes e na automação de tarefas administrativas, permitindo que dediquem mais tempo ao planejamento de estratégias pedagógicas. Dessa forma, a IA não apenas transforma a experiência do aluno, mas também oferece suporte direto ao trabalho docente.

Apesar de seu potencial, a implementação da Inteligência Artificial nas escolas ainda enfrenta barreiras expressivas. Entre os principais desafios estão a falta de infraestrutura tecnológica adequada, a necessidade de capacitação dos profissionais da educação e as preocupações com privacidade e segurança dos dados coletados pelos sistemas. Esses fatores dificultam a adoção plena da IA, especialmente em instituições de ensino público que enfrentam restrições orçamentárias e estruturais.

Esse cenário justifica a escolha do tema, pois compreender o impacto da Inteligência Artificial na educação é essencial para identificar caminhos que promovam o equilíbrio entre inovação tecnológica e práticas pedagógicas inclusivas. Além disso, o estudo é relevante por contribuir com reflexões acadêmicas e sociais, auxiliando gestores, professores e formuladores de políticas públicas a tomarem decisões mais fundamentadas sobre a integração da IA no ensino.

Diante desse contexto, o problema de pesquisa que orienta este trabalho pode ser formulado da seguinte maneira: de que forma a Inteligência Artificial pode ser utilizada para aprimorar o processo de ensino-aprendizagem e quais desafios precisam ser superados para sua implementação eficaz nas escolas? Essa questão busca analisar não apenas os benefícios dessa tecnologia, mas também os limites e riscos que acompanham sua adoção no ambiente educacional.

O objetivo geral deste estudo é analisar a importância da Inteligência Artificial na educação, destacando seus benefícios, desafios e perspectivas para o ensino. Como objetivos específicos, pretende-se investigar as principais ferramentas e sistemas de IA aplicados em ambientes escolares, identificar os obstáculos relacionados à infraestrutura, formação docente, custos e ética, e avaliar os impactos do uso da IA no desempenho e na motivação dos alunos. Assim, busca-se oferecer subsídios para uma reflexão crítica e fundamentada sobre a adoção responsável dessa tecnologia no campo educacional.

2 DESENVOLVIMENTO

2.1 Metodologia

Este trabalho foi conduzido por meio de revisão bibliográfica, de natureza qualitativa e descritiva, com o objetivo de mapear e interpretar criticamente a produção acadêmica

sobre Inteligência Artificial na educação. O recorte temporal adotado concentrou-se nos últimos cinco anos, abrangendo publicações de 2020 a 2025, de modo a garantir aderência ao estado da arte recente sobre o tema investigado. As buscas foram realizadas nas bases de dados Google Acadêmico, SciELO e Periódicos CAPES, complementadas por repositórios institucionais de teses e dissertações. Foram utilizados descritores combinados por operadores booleanos, tais como inteligência artificial AND educação, ensino personalizado, aprendizagem adaptativa, analytics educacional, ética AND dados educacionais e formação docente AND tecnologia. A pesquisa restringiu-se a textos em português e em inglês que abordassem diretamente aplicações, impactos, desafios ou diretrizes de IA no ensino, contemplando artigos científicos, livros, teses, dissertações e relatórios técnicos oficiais produzidos por organismos nacionais e internacionais. Inicialmente, os resultados das buscas foram triados pela leitura de títulos e resumos, com aplicação de critérios de inclusão e exclusão. Foram excluídos materiais de caráter predominantemente opinativo, notícias, entradas enciclopédicas, resumos sem explicitação de método, pré-prints sem revisão por pares e revisões secundárias que não apresentassem evidências empíricas originais. Em seguida, os textos elegíveis foram lidos na íntegra e codificados tematicamente em três eixos analíticos: benefícios pedagógicos da IA na educação, barreiras e riscos associados ao seu uso e condições de implementação em contextos educacionais. Ao final do processo de busca, triagem e leitura na íntegra, foram identificados 23 materiais que atenderam plenamente aos critérios de inclusão e compuseram o corpus final da revisão, sendo 14 artigos científicos, 4 dissertações, 2 teses, 2 relatórios técnicos oficiais e 1 livro, todos publicados entre 2020 e 2025. Considerando o escopo definido para este trabalho, o método de revisão bibliográfica caracteriza-se como pesquisa qualitativa e descritiva, não exploratória, não sistemática, não quantitativa, não experimental e não configurada como estudo de caso; por essa razão, não foram formuladas hipóteses nem proposta qualquer intervenção prática, concentrando-se a investigação na síntese crítica das evidências disponíveis para responder ao problema de pesquisa e aos objetivos delineados no estudo.

2.2 Resultados e Discussão

A revisão dos últimos cinco anos, de 2020 a 2025, evidencia a passagem do discurso prospectivo para usos concretos de IA na educação. Resultados mostram maior consistência quando os objetivos instrucionais são claros, públicos e compartilhados. A qualidade e a disponibilidade dos dados sustentam ciclos regulares de análise, feedback e ajustes pedagógicos. Janelas institucionais para leitura colegiada favorecem interpretações robustas e decisões mais estáveis. Nesse ambiente, planejamento explícito e rotinas de acompanhamento reduzem ruído e variabilidade dos efeitos. Esse conjunto de condições cria base para ganhos duradouros de aprendizagem e equidade (OECD, 2021; UNESCO, 2021; Wang et al., 2024).

A literatura recente amplia métricas além de acerto e erro, incorporando engajamento, persistência e progresso longitudinal. Essa expansão melhora as inferências sobre aprendizagem e equidade e respalda decisões pedagógicas justificáveis. Persistem diferenças associadas à conectividade, ao suporte técnico e à governança de dados entre redes e escolas. Em diálogo com a introdução, os achados mostram que metas e critérios explícitos potencializam a contribuição da IA. Indicam também dependência de infraestrutura, formação e governança e a necessidade de explicabilidade, privacidade e avaliação independente. Esses pontos organizam a leitura dos eixos subsequentes e orientam prioridades de implementação (OECD, 2021; UNESCO, 2023; CETIC.br, 2023; Reina-Parrado et al., 2025).

Na personalização, sistemas adaptativos ajustam ritmo, sequência e complexidade com base em evidências contínuas de desempenho. Quando metas e critérios são explíci-

tos, observam-se aumentos de persistência e autorregulação dos estudantes. Em turmas heterogêneas, reduzem-se assimetrias sem romper o fluxo curricular coletivo. O feedback frequente organiza estudo deliberado e facilita transferências entre tarefas e contextos. O docente valida recomendações algorítmicas como hipóteses de ação e preserva o sentido pedagógico do percurso (Merino-Campos, 2025; Vorobyeva et al., 2025; Wang et al., 2024).

A automação pedagógico-administrativa reúne resultados estáveis: correção automática, consolidação de frequência, relatórios analíticos e alertas formativos. Ao transferir tarefas repetitivas para a IA, amplia-se o tempo docente para planejamento e devolutivas qualitativas. Os ganhos são mais consistentes quando a avaliação utiliza rubricas e há alinhamento entre ensinar, praticar e avaliar. As redes reportam melhoria de qualidade quando critérios e processos são transparentes e auditáveis (OECD, 2021; Ifenthaler, 2024; United States, 2023).

Dashboards e modelos preditivos permitem identificar sinais precoces de risco, orientar intervenções oportunas, promover reagrupamentos temporários e recuperação paralela. A leitura colegiada de dados torna decisões mais justificáveis, previsíveis e comunicáveis às famílias. Relatórios deixam de ser descritivos e passam a orientar escolhas didáticas sustentadas por evidências. O ciclo evidência–decisão–acompanhamento encurta e melhora a responsividade instrucional (Susnjak, 2022; Ramaswami et al., 2023; Rodríguez-Ortiz et al., 2025; Noroozi, 2025).

Para organizar os achados e suas consequências práticas, apresenta-se o Quadro 1, que sintetiza benefícios, evidências, implicações e exemplos. A estrutura em colunas facilita conectar cada efeito a decisões de planejamento e mediação docente. O recorte cobre o quinquênio analisado, abrange diferentes maturidades de adoção e apoia a leitura e a avaliação interna do capítulo. Esses achados convergem com sínteses internacionais sobre feedback frequente e personalização criterial e avançam ao incorporar métricas de engajamento e persistência (Wang et al., 2024; Garzón, 2025).

Quadro 1. Benefícios pedagógicos da IA na educação

BENEFÍCIO (RESULTADO)	EVIDÊNCIAS OBSERVADAS (2020–2025)	IMPLICAÇÕES PEDAGÓGICAS	EXEMPLOS DE USO
Personalização do ensino	Ajuste de ritmo, sequência e desafio; feedback imediato; recomendações baseadas em desempenho	Diferenciação didática; redução de assimetrias; progressão criterial	Trilhas adaptativas; tutores inteligentes; exercícios graduados
Automação de rotinas	Correção automática; consolidação de frequência/notas; geração de relatórios analíticos	Libera tempo do professor para mediação qualitativa; padronização de critérios	Rubricas automatizadas; relatórios por IA; lançamentos assistidos
Monitoramento formativo	Painéis de domínio e engajamento; alertas precoces; modelos preditivos	Intervenções oportunas; reagrupamentos temporários; recuperação paralela	Dashboards docentes; relatórios por aluno; sinais de risco
Engajamento e autorregulação	Metas visíveis; monitoramento de tempo produtivo; progressão gamificada	Aumento de tempo on-task e persistência; hábitos de estudo	Badges; metas semanais; feedback contínuo

Apoio ao planejamento	Análises de carga cognitiva; mapas de objetivos; sequência espiral de competências	Alinhamento ensino–prática–avaliação; design instrucional iterativo	Roteiros semanais; matrizes de referência; planejamento reverso

Fonte: Adaptado de OECD (2021); UNESCO (2021, 2023); Wang et al. (2024); Garzón (2025); Merino-Campos (2025); Ifenthaler (2024).

O **Quadro 1** mostra que personalização, monitoramento e automação atuam de modo interdependente para elevar a qualidade instrucional. Em conjunto, evidenciam como dados se convertem em decisões pedagógicas justificáveis e replicáveis no cotidiano. A robustez aumenta quando a escola define finalidades, critérios e limites para o uso de informações educacionais. Tempos protegidos de análise colegiada sustentam consistência e reduzem vieses interpretativos. Registrar escolhas fortalece transparência e aprendizagem organizacional (OECD, 2021; UNESCO, 2023; Ifenthaler, 2024).

A IA agrega valor quando conectada a problemas autênticos, investigações e projetos interdisciplinares. Nesses cenários, dados sustentam devolutivas formativas e planejamento incremental centrado em competências. O professor contextualiza recomendações, articula conceitos e dá sentido às escolhas didáticas. O equilíbrio entre desafio e apoio previne frustrações e desistências em trajetórias complexas. A produção autoral dos estudantes informa feedbacks e replanejamento responsivo (Merino-Campos, 2025; Vorobyeva et al., 2025; Wang et al., 2024).

Registros sucintos das decisões tornam explícito por que sugestões foram aceitas ou rejeitadas. A prática favorece aprendizagem organizacional e reduz improvisos em ciclos subsequentes. A continuidade se preserva diante de rotatividade de profissionais e turmas. As evidências geradas embasam avaliações internas e externas com maior rigor. A transparência amplia previsibilidade de ações de ensino e acompanhamento (UNICEF, 2021; OECD, 2021; Topali et al., 2025).

A efetividade dos resultados permanece condicionada por infraestrutura, formação e governança. Conectividade estável, parque atualizado e suporte técnico próximo reduzem interrupções e fricções de uso. Formação continuada integra didática, letramento de dados e ética para apropriação crítica e sustentável. Governança garante privacidade, minimização de coleta e rastreabilidade de acessos conforme diretrizes legais. Tais pilares explicam variações de efeito entre redes com perfis socioeconômicos distintos (CETIC.br, 2023; OECD, 2021; UNESCO, 2021, 2023; Reina-Parrado et al., 2025).

Para condensar desafios e respostas correspondentes, apresenta-se o Quadro 2, com barreiras, causas, estratégias e riscos. A visão panorâmica facilita priorização de ações e planejamento de escala responsável em redes. O foco recai em medidas exequíveis e verificáveis que sustentem qualidade e equidade. O recorte ancora-se em diagnósticos nacionais e internacionais recorrentes no período recente e dialoga com as evidências previamente discutidas.

Quadro 2. Barreiras à implementação de IA, causas prováveis, estratégias de mitigação e riscos se ignoradas

BARREIRA	CAUSAS PROVÁVEIS	ESTRATÉGIAS DE MITIGAÇÃO	RISCOS SE IGNORADA
Infraestrutura insuficiente	Conectividade instável; parque obsoleto; falta de suporte	Plano de conectividade; renovação gradual; help desk regional	Ampliação de desigualdades; baixa adoção

Formação docente frágil	Capacitações pontuais; foco instrumental; pouco letramento em dados	Formação continuada com didática, dados e ética; comunidades de prática	Uso superficial; resistência; impacto limitado
Governança de dados/privacidade	Coleta excessiva; opacidade algorítmica; segurança fraca	Minimização de dados; transparência; controle de acesso; aderência à LGPD	Perda de confiança; riscos legais; vieses
Vieses algorítmicos	Bases não representativas; ausência de auditorias	Auditorias periódicas; métricas de justiça; validação docente	Reforço de desigualdades; decisões injustas
Sustentabilidade e escala	Pilotos sem custeio; falta de interoperabilidade	Planejamento de TCO; critérios técnicos de compra; integração	Descontinuidade; "ilhas de inovação"

Fonte: Adaptado de Brasil (2018); CETIC.br (2022, 2023); European Commission (2019); OECD (2021); UNICEF (2021).

O **Quadro 2** reforça que infraestrutura, formação e governança são condições habilitadoras para estabilidade de efeitos. As estratégias priorizam conectividade, manutenção e suporte contínuo no cotidiano escolar. Formação crítica e continuada evita usos superficiais e amplia a qualidade da mediação docente. Protocolos de privacidade e explicabilidade elevam legitimidade e confiança no uso de dados educacionais. Auditorias e métricas de justiça mitigam vieses e protegem públicos vulneráveis (CETIC.br, 2023; OECD, 2021; UNESCO, 2023; UNICEF, 2021).

O debate ético-regulatório ganhou densidade ao longo dos últimos cinco anos em documentos internacionais e nacionais. Diretrizes recomendam finalidades claras, proporcionalidade, segurança e explicabilidade a não especialistas. Salvaguardas específicas são requeridas para crianças e adolescentes. Trilhas de auditoria e papéis definidos elevam a confiança pública no uso de dados educacionais. Essas orientações mitigam riscos sistêmicos e jurídicos nas instituições de ensino (UNESCO, 2021, 2023; European Commission, 2020; UNICEF, 2021).

A literatura de revisão mapeia benefícios, limites e lacunas metodológicas que exigem prudência interpretativa. Os ganhos são maiores onde o feedback é frequente, as rubricas são claras e os objetivos instrucionais são explícitos. Intervenções curtas e amostras restritas limitam generalização e estabilidade temporal dos efeitos. Tais limitações reduzem a validade externa e a estabilidade temporal dos achados; para mitigar, recomenda-se desenho longitudinal, avaliadores independentes e triangulação com observação de sala e produções estudantis (Garzón, 2025; Wang et al., 2024; Merino-Campos, 2025).

Resultados não esperados incluem fadiga digital e sobrecarga atencional quando notificações são excessivas ou pouco relevantes. Além da frequência, explicações alternativas envolvem desenho dos alertas (saliente demais, baixo valor informacional) e carga cognitiva contextual (atividades simultâneas, conectividade instável). Mitigar requer calibrar periodicidade e saliência das mensagens e simplificar fluxos de tarefa, vinculando comunicados a metas explícitas e ao momento da atividade (UNESCO, 2023; OECD, 2021; Rodríguez-Ortiz et al., 2025).

Entre 2020 e 2021, prevaleceram estudos focalizados em acurácia e tempo de resposta em tarefas específicas, bem como em análises de impacto inicial de sistemas inteligentes em contextos reais. Esses desenhos informaram ajustes finos em itens, pistas e sequências no cotidiano de sala. A escala limitada e os contextos controlados dificultaram extrapolações para políticas de rede, mas consolidaram princípios de feedback rápido, visibilidade de metas e progressão criterial, retomados nas etapas seguintes (OECD, 2021; Foster, 2020; Russell; Norvig, 2021).

Entre 2022 e 2023, pilotos ampliados testaram relatórios, alertas e orquestrações sob currículos reais. O acompanhamento incluiu tempo produtivo, autorregulação e proficiência com marcos intermediários. Docentes relataram maior previsibilidade na rotina e clareza de critérios decisórios. A integração com rubricas favoreceu o alinhamento ensino–prática–avaliação no ciclo. Sínteses setoriais registraram ganhos principalmente quando houve análise colegiada de dados (Susnjak, 2022; Ramaswami et al., 2023; OECD, 2021).

Em 2024 e 2025, cresceram iniciativas de governança, explicabilidade e mitigação de vieses nas redes. A interoperabilidade tornou-se meta para continuidade de trajetórias e portabilidade de evidências. Avaliação independente passou a orientar decisões sobre escala e investimentos públicos. A comunicação com famílias adotou linguagem clara para processos e finalidades do uso de dados. Documentos de referência reforçaram proporcionalidade, segurança e justiça na tomada de decisão (UNESCO, 2023; OECD, 2021; Ifenthaler, 2024; Reina-Parrado et al., 2025).

Observa-se convergência para conectividade universal, interoperabilidade e avaliação de impacto transparente ao longo do período analisado. Princípios de explicabilidade, segurança e proporcionalidade articulam-se à Lei Geral de Proteção de Dados. A formação continuada sustenta apropriações pedagógicas duradouras alinhadas a competências curriculares. Critérios técnicos de compra e integração reduzem redundâncias e desperdícios. Esses fundamentos viabilizam expansão responsável com qualidade, segurança e foco em equidade territorial (OECD, 2021; UNESCO, 2021, 2023; CETIC.br, 2023; Reina-Parrado et al., 2025).

A IA aprimora ensino e aprendizagem quando orientada por objetivos claros e avaliáveis. Práticas responsáveis de dados e validação docente preservam o sentido pedagógico e a confiança institucional. Infraestrutura, formação e governança sustentam consistência e escala dos efeitos observados. Triangulação de evidências e documentação de decisões qualificam justiça e eficácia das escolhas. Esses elementos respondem ao problema investigado sem antecipar conteúdos próprios da conclusão (UNESCO, 2023; OECD, 2021; Wang et al., 2024; Garzón, 2025).

3 CONCLUSÃO

Os objetivos do estudo foram atendidos ao analisar a importância da Inteligência Artificial na educação, identificar benefícios e desafios e discutir condições de efetividade. O problema de pesquisa foi respondido ao demonstrar que a tecnologia aprimora o ensino e a aprendizagem quando há metas instrucionais claras, feedback formativo e uso responsável de dados. Observou-se ganho consistente em personalização, monitoramento contínuo e automação de rotinas com apoio ao trabalho docente. A efetividade depende do alinhamento entre currículo, avaliação e prática pedagógica em ciclos de planejamento, ação e revisão. Em síntese, a IA agrega valor quando serve a finalidades educativas explicitadas, mensuráveis e comunicadas à comunidade escolar.

Persistem limites estruturais e organizacionais que condicionam a qualidade e a escala dos resultados. A conectividade inadequada, a obsolescência do parque tecnológico e o suporte técnico insuficiente geram fricções e descontinuidades. A formação docente, quando episódica e instrumental, tende a produzir usos superficiais e pouco sustentáveis no tempo. A proteção de dados, a explicabilidade dos sistemas e a justiça na decisão exigem governança clara, comunicável e auditável. Esses fatores explicam a variação de efeitos entre redes e contextos com diferentes condições de implementação.

Recomenda-se um conjunto integrado de ações que priorize conectividade, atualização de infraestrutura e suporte próximo ao cotidiano escolar. A formação continuada deve arti-

cular didática, letramento de dados e ética para orientar escolhas pedagógicas consistentes. Protocolos de uso e proteção de dados precisam definir finalidades, limites e responsabilidades com mecanismos de avaliação. A documentação das decisões instrucionais favorece a aprendizagem organizacional e a previsibilidade das práticas. A avaliação independente pode reduzir vieses e orientar investimentos com transparência e foco em equidade.

Para trabalhos futuros, sugerem-se estudos longitudinais que acompanhem impactos em aprendizagem, engajamento e bem-estar em diferentes etapas e redes. É pertinente investigar estratégias para mitigar fadiga digital e sobrecarga informacional em ambientes mediados por dados. Também merece atenção a análise de modelos de interoperabilidade que reduzam custos de transição entre plataformas e preservem portabilidade de evidências. Investigações multicêntricas com amostras diversas podem fortalecer a validade externa dos resultados. Essas agendas podem sustentar uma adoção responsável que amplie benefícios com qualidade, segurança e inclusão.

REFERÊNCIAS

CETIC.BR. Tic Educação 2022: pesquisa sobre o uso das tecnologias de informação e comunicação nas escolas brasileiras. São Paulo: Comitê Gestor da Internet no Brasil, 2023.

EUROPEAN COMMISSION. White paper on artificial intelligence: a European approach to excellence and trust. Brussels, 2020.

FOSTER, E.; SIDDLE, R.; CROWSON, P.; BONNE, P. It's all about the intervention: reflections on building staff capacity for using learning analytics to support student success. In: IFENTHALER, D.; GIBSON, D. (org.). Adoption of data analytics in higher education learning and teaching. Cham: Springer, 2020. p. 241–256.

GARZÓN, J.; PATIÑO, E.; MARULANDA, C. Systematic review of artificial intelligence in education: trends, benefits, and challenges. Multimodal Technologies and Interaction, v. 9, n. 8, p. 84, 2025.

IFENTHALER, D. Artificial intelligence in education: definitions, applications, and implications. Journal of Computer Assisted Learning, v. 40, n. 2, p. 1–15, 2024.

MERINO-CAMPOS, C. The impact of artificial intelligence on personalized learning in higher education. Education and Information Technologies, v. 30, n. 1, p. 1–20, 2025.

NOROOZI, O. Advancing peer learning with learning analytics and artificial intelligence. International Journal of Educational Technology in Higher Education, v. 22, n. 1, p. 1–18, 2025.

OECD. OECD Digital Education Outlook 2021: pushing the frontiers with artificial intelligence, blockchain and robots. Paris: OECD Publishing, 2021.

RAMASWAMI, G.; SUSNJAK, T.; MATHRANI, A.; UMER, R. Use of predictive analytics within learning analytics dashboards: a review of case studies. Tech Know Learn, v. 28, n. 4, p. 1045–1070, 2023.

REINA-PARRADO, M.; GARCÍA, F.; MARTÍNEZ, A. Integration of artificial intelligence and machine learning in education: a systematic review. Education and Information Technologies, v. 30, n. 2, p. 1–28, 2025.

RODRÍGUEZ-ORTIZ, M. Á. et al. Machine learning and generative AI in learning analytics for higher education: a systematic review. Applied Sciences, v. 15, n. 3, p. 1–24, 2025.

RUSSELL, S. J.; NORVIG, P. Artificial intelligence: a modern approach. 4. ed. Hoboken: Pearson, 2021.

SUSNJAK, T. Beyond predictive learning analytics modelling and onto explainable AI with prescriptive analytics and ChatGPT. International Journal of Artificial Intelligence in Education, v. 34, n. 2, p. 1–24, 2022.

TOPALI, P. et al. Designing human-centered learning analytics and artificial intelligence in education solutions: a systematic literature review. Behaviour & Information Technology, v. 43, n. 5, p. 1–21, 2024.

UNESCO. AI and education: guidance for policy-makers. Paris: UNESCO, 2021.

UNESCO. Guidance for generative AI in education and research. Paris: UNESCO, 2023.

UNICEF. Policy guidance on AI for children. New York: UNICEF, 2021.

UNITED STATES. DEPARTMENT OF EDUCATION. Artificial intelligence and the future of teaching and lear-

ning: insights and recommendations. Washington, DC: Office of Educational Technology, 2023.

VOROBYEVA, K. I. et al. Personalized learning through AI: pedagogical approaches and critical insights. Contemporary Educational Technology, v. 17, n. 1, p. 1–19, 2025.

WANG, S. et al. Artificial intelligence in education: a systematic literature review. Expert Systems with Applications, v. 244, p. 123–456, 2024.

O livro *Uma visão abrangente da computação – Volume 5* reúne estudos acadêmicos que discutem avanços, aplicações e desafios contemporâneos da área de computação. A obra aborda temas como inteligência artificial, segurança de redes, ética digital, desinformação mediada por tecnologias e aplicações computacionais em diferentes setores da sociedade. Os capítulos apresentam análises teóricas e aplicadas que evidenciam o impacto das tecnologias digitais na educação, na comunicação, na segurança da informação e na tomada de decisões. A coletânea contribui para ampliar o debate científico sobre inovação tecnológica e seus desdobramentos sociais, éticos e profissionais no campo da computação.

ISBN nº 978-65-6068-226-9

Pascal
Editora