



Análise da acessibilidade em interfaces criadas por modelos de inteligência artificial generativa

Rafael Lima Dalmagro | rafael.d29@ifsc.edu.br¹
Cristiano M. Garcia | cristiano.garcia@ifsc.edu.br²
Taynara Cerigueli Dutra | taynara.dutra@ifsc.edu.br³

EDITAL Nº 12/2025/PROPPI/DAE

Trabalho de Pesquisa

DT 4 - Inovação Tecnológica, Processos Produtivos e Futuro das Tecnologias

Resumo:

Este estudo investiga como modelos de Inteligência Artificial Generativa (IAG) incorporam critérios de acessibilidade digital na criação de interfaces web, identificando quais elementos das Diretrizes WCAG 2.2 são aplicados durante a geração automática de *layouts*. A pesquisa compara interfaces geradas via prompts *Zero-shot* e *Retrieval-Augmented Generation* (RAG) com modelos como ChatGPT e Claude avaliadas por ferramentas automáticas e análise manual baseada na metodologia WCAG-EM. Os resultados preliminares indicam que a qualidade acessível das interfaces depende diretamente da formulação do *prompt* e que os modelos ainda falham em atender completamente aos critérios mínimos das diretrizes, sendo a revisão humana indispensável.

Palavras-chave: inteligência artificial generativa; acessibilidade digital; engenharia de prompt; web content accessibility guidelines; large language models;

Introdução

A qualidade da experiência entre humanos e máquinas depende de transparência, sensibilidade (considera necessidades e limitações dos usuários) e percepção de confiança. Quando esses fatores coexistem, as pessoas tendem a aceitar e manter o uso da tecnologia (Sundar; Cao; Machleit, 2020). O avanço da inteligência artificial (IA) transformou o design de interfaces digitais, possibilitando que ferramentas automatizadas

¹ Rafael Lima Dalmagro, Sistemas de Informação do Instituto Federal de Santa Catarina Campus Caçador.

² Cristiano M. Garcia, docente de Informática do Instituto Federal de Santa Catarina, Câmpus Caçador.

³ Taynara Cerigueli Dutra, docente de Informática do Instituto Federal de Santa Catarina, Câmpus Caçador.

criem *layouts* e elementos visuais antes exclusivos de profissionais especializados. Esse fenômeno levanta questões sobre qualidade e acessibilidade das criações, pois a geração por IA deve atender a critérios como contraste adequado, navegação por teclado e compatibilidade com leitores de tela, conforme as WCAG 2.2 (W3C, 2023).

Destaca-se também o papel do *prompt*: instrução textual que orienta modelos de IAG na geração de saídas alinhadas às intenções humanas. A forma como é estruturado pode determinar a aderência do resultado às Web Content Accessibility Guidelines (WCAG) (Oppenlaender; Linder; Silvennoinen, 2024). Apesar do crescente uso dessas tecnologias, muitos modelos podem produzir páginas visualmente agradáveis, mas com falhas na acessibilidade.

Desse modo, o presente trabalho objetiva identificar quais critérios de acessibilidade são considerados por modelos de IAG ao criar páginas web acessíveis, usando técnicas *Zero-shot* e RAG. Os objetivos específicos são: (i) desenvolver *prompts* orientados à acessibilidade; (ii) solicitar interfaces web a diferentes modelos de IAG; (iii) classificar critérios de acessibilidade presentes nas interfaces; (iv) avaliar a consistência dos elementos acessíveis; (v) analisar a influência dos *prompts* na qualidade das interfaces; e (vi) identificar limitações e padrões dos modelos diante das demandas de acessibilidade.

Fundamentação teórica

Acessibilidade digital consiste no desenvolvimento de conteúdos e interfaces utilizáveis por todas as pessoas de forma autônoma, incluindo aquelas com deficiência (W3C, 2023). Segundo a OMS, 16% da população mundial possui algum tipo de deficiência (Chemnad; Othman, 2024). O relatório WebAIM Million 2024 revela que 95,9% das páginas analisadas contêm erros de acessibilidade, com média de 56,8 erros por página (Branham *et al.*, 2024). Tecnicamente, a acessibilidade exige HTML semântico, e atributos ARIA; contraste mínimo de 4,5:1 (WCAG 1.4.3), redimensionamento de texto (WCAG 1.4.4), e foco visível (WCAG 2.4.7). No Brasil, a LBI determina que serviços digitais assegurem acesso igualitário (Lei nº 13.146/2015).

Large Language Models (LLMs) são modelos de aprendizado baseados na arquitetura Transformer, que sustenta os principais sistemas de IAG na geração de código HTML, CSS e JavaScript (Stryker, 2024). Apresentam capacidades emergentes como

geração de código funcional, mas negligenciam critérios de acessibilidade por padrão (Mowar *et al.*, 2025). A engenharia de *prompt* otimiza essa interação por instruções textuais cuidadosas. As técnicas usadas serão: *Zero-shot Prompting*, sem exemplos, verificando o comportamento padrão; RAG, que combina recuperação de documentos externos com geração textual, reduzindo alucinações e desatualizações (Lewis *et al.*, 2020).

Procedimentos metodológicos

A pesquisa adota abordagem exploratória de estratégia experimental (Gil, 2010), combinando análise quantitativa (avaliação técnica com base nas WCAG 2.2) e qualitativa (clareza visual, facilidade de uso e percepção de acessibilidade). As interfaces serão geradas sob duas condições: (i) *Zero-shot*, com descrições genéricas sem referências à acessibilidade e (ii) RAG, *prompts* enriquecidos com trechos das WCAG 2.2 e boas práticas de design inclusivo. Os *prompts* seguem diretrizes metodológicas de engenharia de *prompt* (Oppenlaender; Linder; Silvennoinen, 2024). Até o momento, foram aplicados três *prompts* baseados na técnica *Zero-shot* aos modelos ChatGPT 5.3, Claude Opus 4.6 e MiniMax 2.7 usando a ferramenta Ollama.

A avaliação de acessibilidade segue a metodologia WCAG-EM (WCAG-EM, 2014), que combina verificações automatizadas e validações manuais. Na etapa automatizada, utilizam-se as ferramentas axe DevTools, WAVE, Google Lighthouse e modelos de LLM, que identificam falhas como contraste insuficiente, ausência de rótulos, links, etc. Os relatórios são consolidados em planilha comparativa (Mowar *et al.*, 2025).

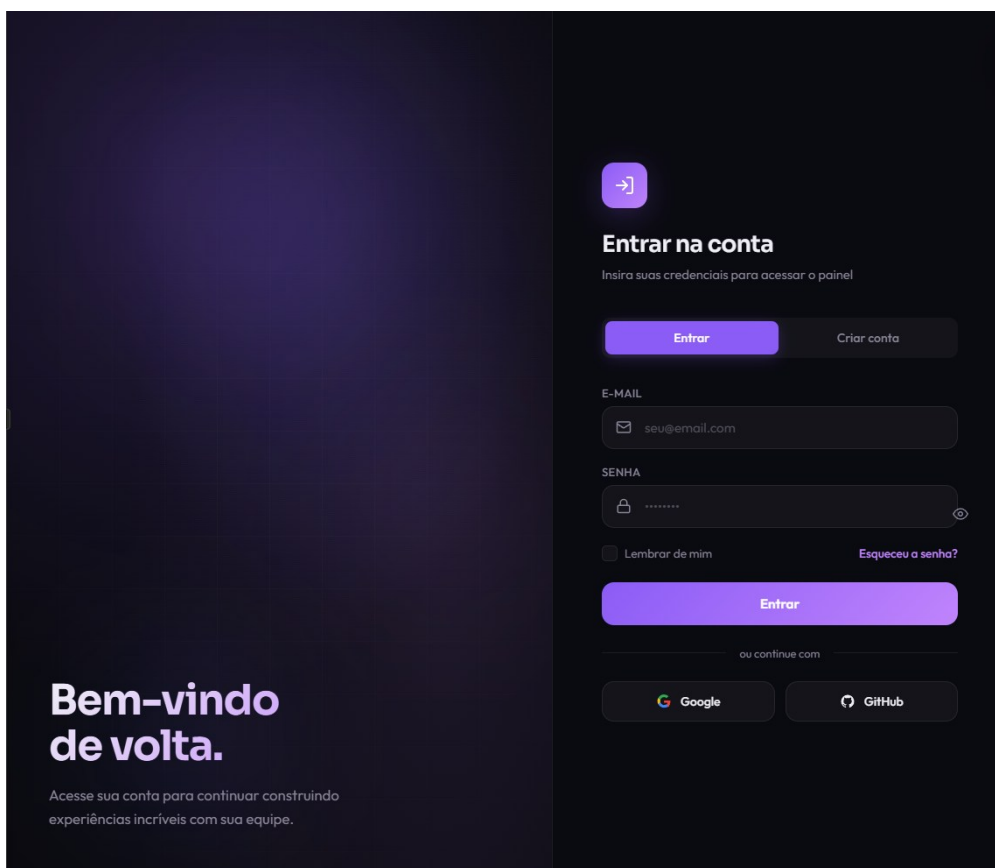
A avaliação manual verifica aspectos como: navegabilidade via teclado, organização lógica do código HTML e presença de foco visível em elementos interativos, garantindo conformidade com os princípios POUR (Perceptível, Operável, Compreensível e Robusto) (Kulkarni, 2019). Cada interface é analisada e classificada em Atende ou Não atende. Ciclos de revisão e ajuste dos *prompts* são adotados conforme Al Thinyan, Al Wohaibi e Al Shehri (2025), assegurando confiabilidade aos resultados.

Resultados e discussões

Os resultados apresentados são parciais, referentes à primeira fase da pesquisa. As interfaces produzidas consistem em páginas de login em HTML, CSS e JavaScript, geradas pelos modelos ChatGPT 5.3, Claude Opus 4.6 e MiniMax 2.7 a partir de três *prompts Zero-shot*.

No primeiro *prompt*, "Desenvolva uma interface web para um sistema de login", os três modelos produziram interfaces funcionais sem menção explícita à acessibilidade. O ChatGPT gerou código compacto com rótulos e contraste implícito, mas sem atributos ARIA ou estrutura semântica. O Claude entregou a interface visualmente mais elaborada, com tema escuro, painel lateral e login social, mas sem referência às WCAG, conforme Figura 1. O MiniMax 2.7 trouxe elementos de contexto de negócio, como métricas, igualmente sem preocupação com acessibilidade.

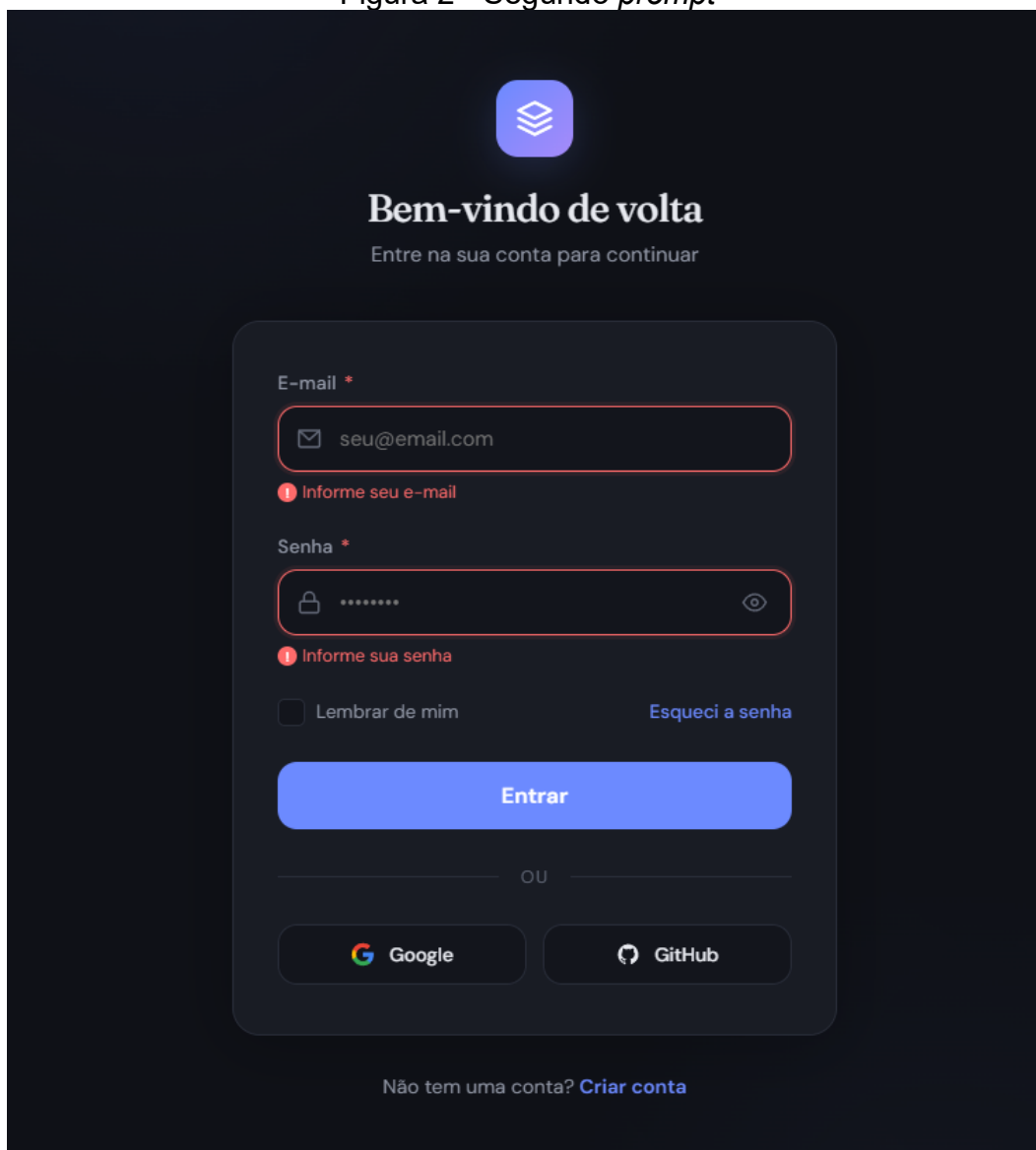
Figura 1 - Primeiro *Prompt*



Fonte: Gerada pelo Claude (2026)

No segundo *prompt*, com a inclusão do termo "acessível", foi observada mudança significativa nos três modelos. O ChatGPT passou a aplicar *aria-live="polite"*, foco visível e contraste adequado. O Claude citou diretamente a WCAG nível AA e implementou *skip link* e *aria-describedby*, conforme Figura 2. O MiniMax 2.7 foi o único a referenciar explicitamente tecnologias assistivas como NVDA e JAWS. Esses resultados evidenciam que a simples inclusão do termo "acessível" no *prompt* orienta os modelos a práticas mais inclusivas, corroborando Oppenlaender, Linder e Silvennoinen (2024).

Figura 2 - Segundo *prompt*

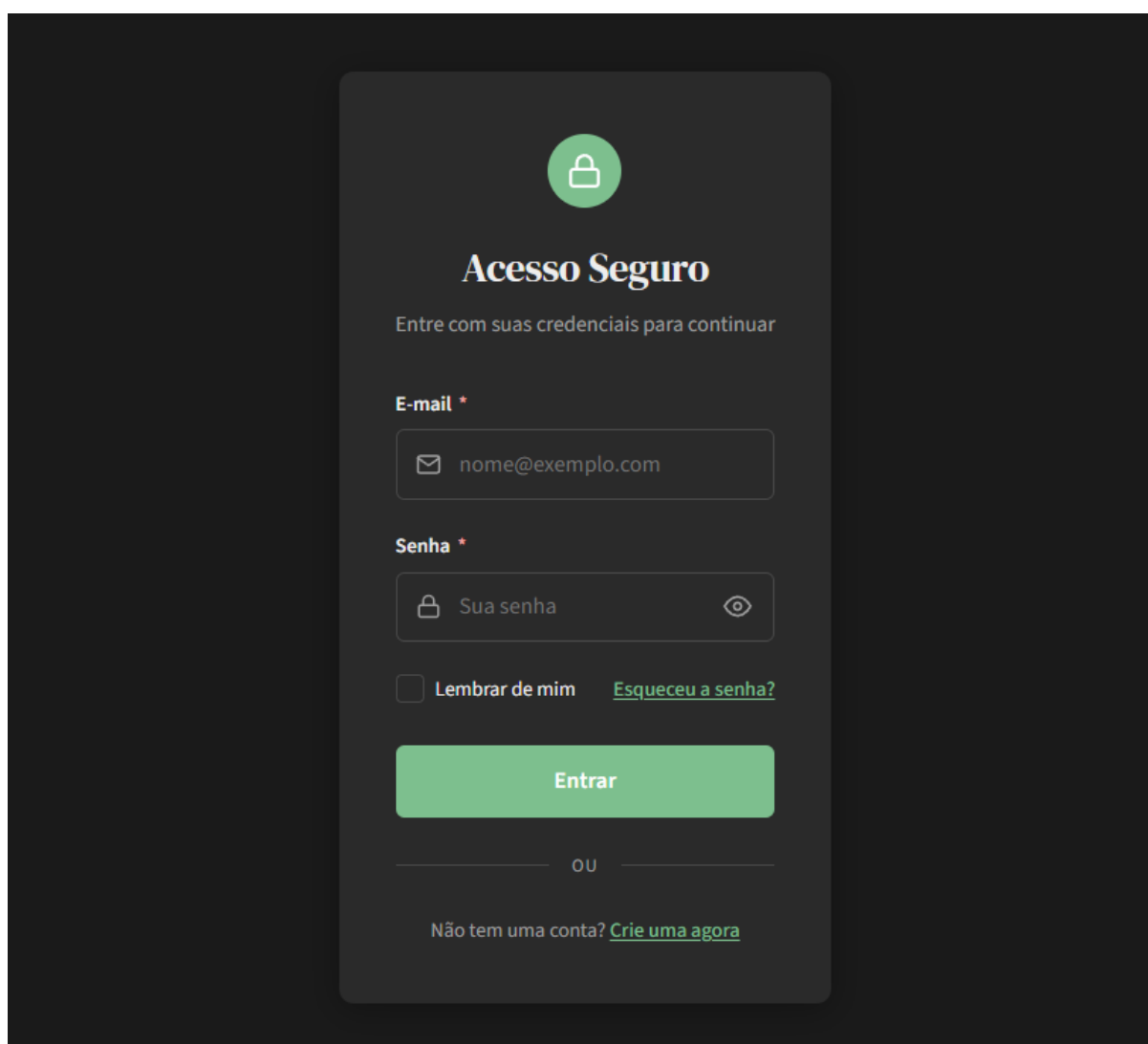


The image shows a login interface with a dark background. At the top, there is a purple icon with three stacked squares. Below it, the text "Bem-vindo de volta" (Welcome back) is displayed in white, followed by "Entre na sua conta para continuar" (Log in to your account to continue). The main form has two input fields: "E-mail *" with a placeholder "seu@email.com" and a red error message "Informe seu e-mail" (Provide your email); and "Senha *" (Password) with a masked password "....." and a red error message "Informe sua senha" (Provide your password). There is a checkbox for "Lembrar de mim" (Remember me) and a link "Esqueci a senha" (I forgot my password). A large blue button labeled "Entrar" (Log in) is below the fields. Below the button is a horizontal line with "OU" (OR) in the center. At the bottom of the form are two buttons for social login: "Google" and "GitHub". At the very bottom, there is a link "Não tem uma conta? Criar conta" (Don't have an account? Create account).

Fonte: Gerada pelo Claude (2026)

No terceiro *prompt*, que impunha estrutura de resposta com seções definidas, o ChatGPT entregou resposta mais organizada, porém com código reduzido e sem JavaScript. O Claude apresentou a documentação mais precisa, com referências aos WCAG e justificativas técnicas detalhadas resultando na interface demonstrada na Figura 3. O MiniMax 2.7 obteve maior nível de detalhamento, citando critérios WCAG e documentando suporte a modo escuro.

Figura 3 - Terceiro *prompt*



Fonte: Gerada pelo Claude (2026)

De maneira geral, a progressão na especificidade dos *prompts* resultou em maior presença de elementos de acessibilidade. No entanto, também houve redução na complexidade visual das interfaces, sugerindo que os modelos tendem a priorizar a conformidade técnica em detrimento da elaboração visual quando instruídos de forma mais específica (Sarkar, 2023).

Considerações finais

Os resultados parciais confirmam que nenhum dos modelos atende de forma plena aos critérios mínimos de acessibilidade definidos pelas WCAG 2.2, especialmente quando operados com *prompts* genéricos. A ausência de atributos como *lang* no elemento `<html>`, de mecanismos de navegação por *skip link* e de validação semântica completa nos *prompts* mais neutros reforça que a engenharia de *prompt* atua como fator determinante para a qualidade acessível das interfaces geradas. Com a continuidade da pesquisa e a aplicação da técnica RAG, espera-se identificar como o enriquecimento dos *prompts* com trechos documentais das WCAG 2.2 pode ampliar a cobertura dos critérios de acessibilidade, permitindo uma comparação mais completa entre as abordagens. Ainda assim, mesmo com *prompts* mais elaborados, a revisão humana permanece indispensável para assegurar conformidade e usabilidade (Mowar *et al.*, 2025).

Referência ao fomento recebido

EDITAL Nº 12/2025/PROPPI/DAE

Referências

AL THINYAN, Khalid; AL WOHAIBI, Mohammad; AL SHEHRI, Abdullah. Prompt Engineering Methodology. *Int. J. Innov. Sci. Res. Technol.*, v. 10, n. 4, p. 1278-1282, abr. 2025.

BRANHAM, Stacy M. *et al.* The state of digital accessibility. *ACM Queue*, v. 22, n. 4, set./out. 2024.



BRASIL. Lei nº 13.146, de 6 de julho de 2015. Lei Brasileira de Inclusão da Pessoa com Deficiência. Disponível em: https://www.planalto.gov.br/ccivil_03/ato2015-2018/2015/lei/l13146.htm.

CHEMNAD, Khansa; OTHMAN, Achraf. Digital accessibility in the era of artificial intelligence—Bibliometric analysis and systematic review. **Front. Artif. Intell.**, v. 7, art. 1349668, fev. 2024.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

KULKARNI, Mukta. Digital accessibility: Challenges and opportunities. **IIMB Management Review**, v. 31, p. 91-98, 2019.

LEWIS, Patrick *et al.* Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *In: Proceedings of 2020 NeurIPS*, 34., 2020, Vancouver. 2020. p. 1-16.

MOWAR, Peya *et al.* CodeA11y: making AI coding assistants useful for accessible web development. *In: Proceedings of CHI Conf. on Hum. Factors in Comput. Syst.*, 2025, Yokohama.

OPPENLAENDER, Jonas; LINDER, Rhema; SILVENNOINEN, Johanna. Prompting AI art: an investigation into the creative skill of prompt engineering. **arXiv**, [S.l.], 2024. Disponível em: <https://arxiv.org/abs/2303.13534>. Acesso em: 29 jun. 2025.

STRYKER, Cole. **Large language models (LLMs)**. 2024. Disponível em: <https://www.ibm.com/think/topics/large-language-models>. Acesso em: 25 nov. 2025.

SUNDAR, Aparna; CAO, Edita S.; MACHLEIT, Karen A. How product aesthetics cues efficacy beliefs of product performance. **Psychology & Marketing**, v. 37, n. 9, p. 1246-1262, 2020.

WORLD WIDE WEB CONSORTIUM (W3C). **Website Accessibility Conformance Evaluation Methodology (WCAG-EM) 1.0**. 2014. Disponível em: <https://www.w3.org/TR/2014/NOTE-WCAG-EM-20140710>. Acesso em: 12 nov. 2025.



WORLD WIDE WEB CONSORTIUM (W3C). **Web Content Accessibility Guidelines (WCAG) 2.2**. 2023. Disponível em: <https://www.w3.org/TR/WCAG22/>. Acesso em: 19 jun. 2025.