AVALIAÇÃO DE AUDIODESCRIÇÃO GERADA POR INTELIGÊNCIA ARTIFICIAL: USABILIDADE, CONTEÚDO E CONTEXTO.

Bruna Cardoso Giachetto (PIBIC/CNPq/FA/UEM), Bruno Montanari Razza (Orientador). E-mail: bmrazza@uem.br

Universidade Estadual de Maringá, Centro de Tecnologia, Cianorte, PR.

Área e subárea do conhecimento: Desenho industrial /Programação visual

Palavras-chave: Acessibilidade digital; Visão computacional; Design de interação.

RESUMO

Este trabalho avalia a qualidade e a usabilidade da audiodescrição (AD) gerada por Inteligência Artificial (IA) para pessoas com deficiência visual em ambientes digitais e produtos. A metodologia envolveu revisão sistemática da literatura na plataforma CAPES e avaliação de aproximadamente 30 legendas em redes sociais (Instagram e TikTok). Os resultados apontaram alucinações, omissão de detalhes essenciais e complexidade dos leitores de tela. Nos produtos, houve falhas graves e omissão de alertas sobre alérgenos e segurança nutricional. Na avaliação de cenas, os modelos de IA mostraram precisão espacial e técnica, enquanto a AD humana destacou-se pela sensibilidade subjetiva em obras de arte. Nas redes sociais, constatou-se falta de diversidade em posts não governamentais e baixa qualidade das legendas. Conclui-se que a acessibilidade efetiva requer a convergência entre rigor normativo, design centrado no humano e padronização de dados, indicando que a automação por IA deve complementar a mediação humana e as diretrizes das comunidades com deficiência visual.

INTRODUÇÃO

O progresso da Inteligência Artificial (IA), fomentado pela Visão Computacional e pelos Modelos de Linguagem de Grande Escala (LLMs), reconfigurou o panorama das Tecnologias Assistivas (TA). Previamente restritas a leitores de tela estruturais como o *VoiceOver* e o *TalkBack*, tais recursos transmutaram-se em sistemas multimodais capazes de decodificar e interpretar componentes visuais do cotidiano.
Entretanto, com a rápida expansão de conteúdos dinâmicos no meio digital (especialmente em redes sociais) acaba sendo exposta uma lacuna entre a capacidade de geração automatizada de dados e a entrega de uma audiodescrição adequada. Nesse contexto, apesar da popularização de hashtags inclusivas como #PraCegoVer, #DescricaodaImagem e #PraTodosVerem, criadas para acompanhar legendas precisas explicando a imagem e/ou vídeo aos público deficientes, a constante adesão informal de criadores de conteúdo sem conscientização sobre a acessibilidade ainda permanece sendo um problema, acarretando dependência de rotulagens automáticas genéricas presentes no Instagram ou até mesmo a falta delas, como o TikTok, ao usuário cego.
Diante desse cenário, esta pesquisa teve como objetivo geral avaliar a qualidade e a usabilidade da audiodescrição gerada por IA direcionada a pessoas com deficiência visual, com apoio de literaturas encontrada na revisão sistemática como a filosofia do Design Centrado no Humano (Tosi *et al.*, 2020) e estruturando a investigação a

partir da revisão da literatura e da análise das redes sociais, com base na tríade: usabilidade, conteúdo e contexto.

MATERIAIS E MÉTODOS

A metodologia integrou revisão sistemática da literatura por meio da utilização da plataforma Periódicos CAPES, realizando o levantamento bibliográfico por meio da delimitação temporal entre os anos de 2015 e 2026. Ademais, foi utilizada combinações booleanas relacionadas ao tema como "*visual impairment*" e "*digital accessibility*" além da triagem da pesquisa ter sido auxiliada por ferramentas de IA Gemini e a organização dos textos selecionados foram executadas pelo *NotebookLM*, resultando em um núcleo de estudos analisados em profundidade e bem agregados à temática de pesquisa.
Complementando o levantamento bibliográfico, realizou-se uma etapa exploratória e qualitativa de campo para avaliar o desempenho das tecnologias assistivas no fornecimento de audiodescrição, além de analisar se os criadores de conteúdos que utilizaram as hashtags inclusivas adotaram legendas acessíveis. As legendas e textos alternativos dessas publicações foram acessadas por meio da audiodescrição do *VoiceOver* (em um IPhone 11, no sistema IOS). Ademais, a investigação focou na navegação pelas redes sociais Instagram e TikTok, analisando publicações marcadas com as *hashtags* #PraCegoVer, #PraTodosVerem e #DescricaoDaImagem, o que permitiu averiguar a fluidez da leitura, o grau de detalhamento das descrições geradas automaticamente e a experiência de uso no ambiente digital.

RESULTADOS E DISCUSSÃO

A revisão sistemática mostrou um volume expressivo de estudos sobre acessibilidade digital e AD, em contrapartida, uma lacuna acentuada na interseção entre texto alternativo e IA, pois o cruzamento dos descritores "automatic alt text", "AI-generated alt text" e "automatic alternative text" retornou apenas um resultado nas bases consultadas. Do núcleo de artigos selecionados, é possível destacar a filosofia do Design Centrado no Humano (Tosi *et al.*, 2020), segundo a qual a avaliação de qualquer tecnologia assistiva deve considerar não apenas a precisão técnica, mas a satisfação e a facilidade de compreensão do usuário final.
Nesse panorama, na dimensão de usabilidade, Gonçalves et al. (2019) sistematizam requisitos de interface relevantes para a audiodescrição por IA, tendo como destaque: a interrupção imediata da fala anterior ao iniciar uma nova leitura, evitando o empilhamento de áudios, e o fornecimento de feedback de situação, confirmando ao usuário que a imagem foi corretamente capturada antes de detalhar a descrição. Tais requisitos reforçam que a usabilidade de uma audiodescrição automática depende tanto da qualidade do texto gerado quanto da forma como esse conteúdo é entregue ao usuário.
No sentido de conteúdo, Leotta, Mori e Ribaudo (2022) compararam descrições humanas e geradas por quatro ferramentas de IA em 60 imagens da Wikipedia, avaliadas às cegas por 76 estudantes e concluíram que as descrições humanas mantiveram a maior pontuação em todas as categorias, enquanto o desempenho IA variou conforme o tipo de imagem, com falhas recorrentes e legendas genéricas, como a catedral Nossa Senhora da Assunção em São Paulo ser reconhecida por um

"prédio de concreto". Em continuidade a isso, Pieriboni, Buzzi e Leporini (2026) confirmam limitação semelhante em conteúdos complexos, nos quais os modelos generativos ainda dependem de validação para superar a etapa de protótipo. Em conjunto, esses achados indicam que a IA já viabiliza a automação em larga escala da AD, mas ainda requer mediação humana em contextos de maior densidade informacional ou subjetividade.

Na dimensão de contexto, a ausência de norma brasileira específica para audiodescrição em redes sociais faz com que a acessibilidade via AD tenha maior incidência por tipos de perfis que a Lei Brasileira de inclusão da Pessoa com Deficiência (LBI) e da ABNT NBR 17225:2025 (ABNT, 2025) conseguem abranger. Como referencial complementar, a WCAG 2.1 (W3C, 2018) organiza a acessibilidade pelos princípios POUR: perceptível, operável, compreensível e robusto. Em contrapartida, pela aplicação da AD concentrar-se em contas governamentais e jornalísticas graças à pressões legislativas, o conteúdo gerado por usuários comuns (criadores de conteúdos) acaba ficando de fora do alcance efetivo da fiscalização.

Nesse viés, com a análise das publicações sob as hashtags inclusivas foi possível constatar que a maior parte das descrições consistentes partiu de páginas governamentais, com conteúdos formais como campanhas de vacinação, título de eleitor e eventos públicos, enquanto publicações informais (memes, cotidiano e viagens) permaneceram majoritariamente sem descrição ou dependentes de legendas automáticas da plataforma do tipo "pode ser uma imagem de...". Em continuidade a isso, embora o recurso de texto alternativo personalizado do Instagram esteja disponível, mostrou-se subutilizado pelos próprios criadores, acarretada pela ausência de campanhas de divulgação e fiscalização sobre seu uso; no TikTok, essa alternativa é praticamente inexistente. Essa lacuna dialoga com o conceito de *Auteur Description* de Thompson (2018): assim como o diretor de uma obra audiovisual é quem melhor pode descrevê-la fielmente, o próprio criador da publicação será o melhor sujeito a transmitir a intenção da imagem ao público, apesar de ser pouco explorado nas plataformas.

CONCLUSÕES

A pesquisa evidencia que a IA generativa e a visão computacional representam avanços promissores para a autonomia de pessoas com deficiência visual. No entanto, a acessibilidade digital efetiva não é alcançada apenas pela automação técnica. É indispensável aplicar os princípios do Design Centrado no Humano no desenvolvimento de interfaces móveis e instituir a obrigatoriedade da padronização de uma efetiva audiodescrição, especialmente no contexto dinâmico múltiplos tipos de conteúdos presentes nas redes sociais. A atuação conjunta entre a precisão descritiva das IAs multimodais e a ampliação de uma regularização efetiva em diversos canais de comunicação constituem um caminho indispensável para a inclusão digital de qualidade a todos os indivíduos que acessam aos meios digitais no dia a dia.

AGRADECIMENTOS

Gostaria de dedicar meus agradecimentos à Fundação Araucária (FA), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Universidade

Estadual de Maringá (UEM) pela concessão de bolsa de Iniciação Científica e pelo apoio institucional indispensável ao desenvolvimento desta pesquisa.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 17225**: Acessibilidade em conteúdo e aplicações web: Requisitos. Rio de Janeiro: ABNT, 2025.

BRASIL. **Lei nº 13.146, de 6 de julho de 2015**. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Brasília, DF: Presidência da República, 2015. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 10 out. 2026.

CORREIA, W. F. M. *et al.* A roadmap towards mobile applications with accessibility for visually impaired users: guideline and its evaluation. **Design e Tecnologia**, Porto Alegre, v. 9, n. 19, p. 87-111, 2019. Disponível em: https://doi.org/10.23972/det2019iss19pp87-111. Acesso em: 1 out. 2026.

LEOTTA, F.; MORI, G.; RIBAUDO, M. Evaluating the effectiveness of automatic image captioning for web accessibility. **Universal Access in the Information Society**, Berlin, v. 21, n. 1, p. 1-18, 2022. Disponível em: https://doi.org/10.1007/s10209-022-00906-7. Acesso em: 1 out. 2026.

PIERIBONI, G.; BUZZI, M.; LEPORINI, B. STEM education and ICT-enhanced tools for students with disabilities: a five-year review. **Universal Access in the Information Society**, v. 25, n. 3, 2026. Disponível em: https://doi.org/10.1007/s10209-025-01282-8

THOMPSON, H. Audio Description: Turning Access to Film into Cinema Art. **Disability Studies Quarterly**, Ohio, v. 38, n. 3, p. 1-15, 2018. Disponível em: https://doi.org/10.18061/dsq.v38i3.6487. Acesso em: 1 out. 2026.

TOSI, F.; SERRA, A.; BRISCHETTO, A.; IACONO, E. (org.). **Design for Inclusion, Gamification and Learning Experience**. Milano: FrancoAngeli, 2020. *E-book*. Disponível em: https://series.francoangeli.it/index.php/oa/catalog/book/512. Acesso em: 1 out. 2026.

WORLD WIDE WEB CONSORTIUM. **Web Content Accessibility Guidelines (WCAG) 2.1**. [S. l.]: W3C, 5 jun. 2018. W3C Recommendation. Disponível em: https://www.w3.org/TR/2018/REC-WCAG21-20180605/. Acesso em: 1 out. 2026.