

Descrições Automáticas de Imagens por IA: Desafios Éticos e Diretrizes para Inclusão no Contexto do GC6

**Julia M. Rossetti¹, Vih Angel S. Silva¹, Luciana C. L. F. Borges¹, Thiago M. Ventura¹,
Cristiano Maciel¹, Patricia C. de Souza¹, Eunice P. S. Nunes¹**

¹Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)
Cuiabá – MT – Brasil

{juliamariarossetti, vitoria.angel2002}@gmail.com

{lucianafariaborges, eunice.ufmt}@gmail.com

{thiago, patricia}@ic.ufmt.br, cristiano.macie1@ufmt.br

Abstract. *Introduction:* Automatic image description by artificial intelligence systems is essential for accessibility of people with visual impairments, but it may reproduce biases and compromise representativeness. **Objective:** To analyze ethical implications in description generation, considering identity erasure, racial bias, gender stereotypes, and invisibility of disability, using ChatGPT and Gemini. **Methodology:** Qualitative and exploratory study with seven images containing identity markers, described by both models and analyzed through thematic content analysis. **Results:** Although coherent, some descriptions presented omissions, generic language, and stereotypes, indicating the need for guidelines towards more representative and inclusive descriptions. **Keywords** Artificial Intelligence, Digital Accessibility, Automatic Image Description, AI Ethics, Representativeness, Inclusion, GC6.

Resumo. *Introdução:* A descrição automática de imagens por sistemas de inteligência artificial é essencial para a acessibilidade de pessoas com deficiência visual, mas pode reproduzir vieses e comprometer a representatividade. **Objetivo:** Analisar implicações éticas na geração de descrições, considerando apagamento identitário, viés racial, estereótipos de gênero e invisibilização da deficiência, usando ChatGPT e Gemini. **Metodologia:** Estudo qualitativo e exploratório com sete imagens contendo marcadores identitários, descritas pelos dois modelos e analisadas por conteúdo temático. **Resultados:** Algumas descrições, embora coerentes, apresentaram omissões, linguagem genérica e estereótipos, apontando a necessidade de diretrizes para descrições mais representativas e inclusivas.

Palavras-Chave Inteligência Artificial, Acessibilidade Digital, Descrição Automática de Imagens, Ética em IA, Representatividade, Inclusão, GC6.

1. Desafio Abordado

O Grande Desafio abordado é o GranDIHC-BR 06 – Implicações da Inteligência Artificial em Interação Humano-Computador: Uma Discussão sobre Paradigmas, Ética e Diversidade, Equidade e Inclusão [Duarte et al. 2024].

2. Relatos e análises sobre o que foi realizado em 2024–2025

Entre 2024 e 2025, desenvolveu-se um estudo qualitativo e exploratório voltado à análise das descrições automáticas de imagens geradas por sistemas de Inteligência Artificial (IA), com foco na acessibilidade digital e nos desafios éticos associados. A pesquisa busca compreender como ferramentas amplamente difundidas, como ChatGPT [OpenAI 2025] e Gemini [Google AI 2023], lidam com representações de identidade, diversidade e deficiência [Rossetti et al. 2025].

O estudo contemplou a coleta de sete imagens em bancos públicos [Pexels 2024, Pixabay 2024], das quais sete foram selecionadas para análise final, priorizando representatividade de raça, gênero, deficiência e idade. As imagens foram submetidas aos modelos ChatGPT 4.0 e Gemini 1.5 Pro, que geraram descrições automáticas posteriormente analisadas pela técnica de análise de conteúdo temática [Bardin 2016]. Essa abordagem permitiu a identificação de quatro categorias centrais de vieses: apagamento identitário, estereótipos de gênero, invisibilização da deficiência e uso de linguagem genérica ou desumanizante [Birhane e van Dijk 2020, Noble 2018, D'Ignazio e Klein 2020].

Os resultados indicaram que, embora ambos os modelos produzam descrições tecnicamente coerentes, persistem limitações significativas relacionadas à sensibilidade cultural e à representatividade. Em casos como a descrição de pessoas indígenas ou com deficiência, observou-se redução de marcadores identitários a meros ornamentos estéticos, bem como omissão de elementos visuais relevantes, o que reforça preocupações já apontadas pela literatura [Benjamin 2019, Seo et al. 2024]. Além disso, constatou-se que a suposta neutralidade algorítmica pode resultar em exclusão simbólica, uma vez que a ausência de referência a identidades racializadas ou corporificadas reforça padrões normativos [Rossetti et al. 2025].

Esses achados ressaltam que, apesar dos avanços técnicos dos modelos multimodais, a geração automática de descrições de imagens ainda enfrenta obstáculos éticos e políticos, especialmente quando voltada a públicos que dependem da acessibilidade digital para participação social plena. O trabalho realizado em 2024–2025 contribuiu, portanto, para evidenciar lacunas nos sistemas atuais e fundamentar diretrizes para descrições mais representativas e inclusivas no campo da IHC.

3. Reflexões críticas sobre as direções apontadas nos desafios

A descrição de imagens é um recurso fundamental para garantir a acessibilidade de pessoas com deficiência visual, permitindo que informações visuais sejam transmitidas por meio de texto ou áudio de forma compreensível e contextualizada [W3C 2018]. Tradicionalmente, essa prática é realizada por audiodescritores humanos, que interpretam elementos visuais com base em princípios de clareza, objetividade e adequação ao contexto [Almeida e Mazzoni 2020]. Com os avanços da inteligência artificial, especialmente em modelos multimodais, possibilitou-se automatizar parte desse processo, gerando descrições de forma rápida e em larga escala [Adewumi et al. 2024].

Contudo, análises recentes, incluindo [Rossetti et al. 2025], indicam que sistemas como ChatGPT e Gemini, podem apresentar apagamento identitário [Birhane e van Dijk 2020, Prabhu e Birhane 2020], linguagem genérica ou

desumanizante [Noble 2018, D'Ignazio e Klein 2020] e omissão de marcadores culturais e de deficiência [Venkit et al. 2023], afetando a representatividade e a equidade [Benjamin 2019]. Entre as limitações observadas, destaca-se que a neutralidade aparente pode ser excludente. Ao evitar mencionar raça, gênero ou deficiência, sistemas de IA tendem a reforçar padrões normativos e a invisibilizar identidades marginalizadas [Birhane e van Dijk 2020, Benjamin 2019, Noble 2018]. Podem persistir ainda estereótipos, como a associação de mulheres à delicadeza e homens à força, mesmo sem instruções explícitas [D'Ignazio e Klein 2020, West et al. 2019]. Esses padrões podem relevar possível influência dos dados de treinamento e das escolhas de design, que podem perpetuar desigualdades históricas e reforçar hierarquias sociais por meio da linguagem gerada.

Tal cenário evidencia limitações estruturais no treinamento e na validação destes sistemas e a presença de possíveis vieses algorítmicos na forma como sistemas de IA descrevem pessoas e contextos visuais, afetando a precisão, representatividade e a sensibilidade cultural das descrições [Birhane e van Dijk 2020, Venkit et al. 2023], com impactos diretos para usuários com deficiência visual.

Este trabalho busca fomentar um debate crítico sobre a importância de estabelecer critérios claros e métricas representacionais que assegurem descrições automatizadas não apenas corretas, mas também representativas e inclusivas, evidenciando o papel da IA na promoção de uma IHC ética, diversa e inclusiva em consonância com o GC6 [Duarte et al. 2024].

4. Caminhos, estratégias e articulações para os próximos anos

Com base nas discussões apresentadas, nas evidências da literatura, em experiências práticas e alinhando-se às preocupações do GC6, propõem-se ações para tornar a geração automática de descrições de imagens por IA mais ética, sensível e inclusiva. É fundamental a participação ativa de pessoas com deficiência visual em todas as etapas de design, teste e avaliação, garantindo que as soluções reflitam necessidades e experiências reais [Cerejeira e Alves 2022, Santos e Cavalcante 2020]. Essa participação deve ser acompanhada do treinamento de modelos com dados diversos e culturalmente contextualizados, contemplando raça, gênero, idade, expressões culturais e deficiência, de modo a reduzir vieses e promover representatividade [Neves 2022, D'Ignazio e Klein 2020].

Outra medida relevante é a adoção de mecanismos híbridos que combinem IA e revisão humana em contextos críticos, equilibrando eficiência automatizada e sensibilidade interpretativa [Santos e Cavalcante 2020]. Nesse sentido, também se propõe o ajuste narrativo dinâmico, permitindo que o usuário escolha entre descrições mais técnicas ou mais sensíveis/contextualizadas, o que favorece a autonomia interpretativa [Rossetti et al. 2025].

Para assegurar evolução constante, recomenda-se a implementação de canais de feedback contínuo para identificação e correção de omissões e vieses. Por fim, a adoção de diretrizes éticas nacionais, como a Recomendação da UNESCO sobre Ética da Inteligência Artificial, é essencial para incorporar princípios de justiça, não discriminação, transparência e supervisão humana [UNESCO 2021].

5. Lacunas, oportunidades e parcerias

A geração automática de descrições por IA permanece estratégica para a acessibilidade digital, mas enfrenta limitações importantes. Usuários cegos e com baixa visão relatam dificuldades para confiar nas respostas geradas, devido a omissões, imprecisões e descrições genéricas. Esses problemas podem configurar uma falsa neutralidade algorítmica, que tende a reforçar invisibilizações e padrões normativos excludentes [Seo et al. 2024, Venkit et al. 2023].

Como estratégia, usuários com deficiência visual frequentemente recorrem à comparação entre modelos, à validação por conhecimento prévio ou mesmo à aceitação de respostas equivocadas por falta de alternativas. Fatores como clareza na justificativa, estilo da linguagem, tempo de resposta e nível de detalhe também influenciam diretamente na escolha dos modelos e na autonomia interpretativa [Seo et al. 2024]. Neste sentido, ainda faltam métricas representacionais e mecanismos de personalização que considerem o contexto de uso e a experiência subjetiva. Abordagens mais interativas são bem recebidas, mas ainda enfrentam limitações técnicas, como a incapacidade de analisar imagens na totalidade [Chen et al. 2024, Jiang et al. 2025].

A comunidade deve avançar com avaliações participativas, integração de recursos multimodais e critérios que promovam representatividade, sensibilidade cultural e adaptabilidade. Ainda, há oportunidades para soluções híbridas IA mais humanas em contextos críticos [Santos e Cavalcante 2020], alinhadas a diretrizes éticas como as da UNESCO [UNESCO 2021] e aos princípios do GC6 [Duarte et al. 2024].

6. Contribuições e reflexões para o avanço da área

Este trabalho contribui para a área de IHC ao evidenciar, de forma empírica e crítica, os limites atuais da descrição automática de imagens por IA em contextos de acessibilidade digital. A análise de descrições geradas por ChatGPT e Gemini revelou desafios recorrentes, como apagamento identitário, estereótipos de gênero, invisibilização da deficiência e uso de linguagem genérica ou desumanizante [Noble 2018, Benjamin 2019, Venkit et al. 2023]. Essas evidências reforçam a pertinência das preocupações apontadas no GC6 e demonstram que a neutralidade aparente dos modelos algorítmicos pode, na prática, reforçar exclusões simbólicas [Birhane e van Dijk 2020, Seo et al. 2024].

Do ponto de vista científico, a principal contribuição está na proposição de diretrizes que buscam integrar princípios de justiça, diversidade e inclusão à prática da descrição automática, com base em evidências coletadas e em experiências práticas [Rossetti et al. 2025]. Tais diretrizes contemplam desde a participação ativa de pessoas com deficiência visual no desenvolvimento até a adoção de mecanismos híbridos IA-humanos e o treinamento de modelos com dados diversos e culturalmente contextualizados [D'Ignazio e Klein 2020, Cerejeira e Alves 2022].

Por fim, este estudo amplia o debate sobre ética algorítmica na IHC ao propor caminhos que não apenas mitigam riscos, mas também apontam possibilidades de inovação inclusiva. Ao fomentar o uso de métricas representacionais, canais de feedback contínuo e ajuste narrativo dinâmico, busca-se fortalecer a autonomia interpretativa dos usuários e contribuir para o avanço da área em consonância com diretrizes internacionais como a Recomendação da UNESCO [UNESCO 2021] e os princípios estabelecidos pelo GC6 [Duarte et al. 2024].

7. Aspectos éticos envolvidos

A presente proposta foi desenvolvida a partir de revisão de literatura, análise crítica de pesquisas recentes e formulação de diretrizes para práticas inclusivas na descrição automática de imagens por IA. Não houve coleta de dados pessoais nem participação direta de pessoas, de modo que não se aplicam procedimentos de consentimento informado. Assim, o trabalho encontra-se em conformidade com a Lei Geral de Proteção de Dados Pessoais [Brasil 2018].

Além disso, o trabalho de Rossetti, Borges e Souza [2025], também de autoria de integrantes desta proposta, já havia discutido limitações éticas na geração automática de descrições, como apagamento identitário e invisibilização de marcadores de diversidade. A presente reflexão amplia esses achados, reafirmando o compromisso em promover inclusão, diversidade e justiça algorítmica em consonância com os princípios do GC6 [Duarte et al. 2024] e com recomendações internacionais [UNESCO 2021].

8. Agradecimentos

As autoras e autores agradecem à Universidade Federal de Mato Grosso (UFMT) pelo apoio institucional, ao Programa de Pós-Graduação em Ciência da Computação (PPGCOMP) e ao Instituto de Computação pelo espaço de pesquisa e discussão. Reconhecem também o apoio da Pró-Reitoria de Pesquisa da UFMT (PROPESQ) e da Pró-Reitoria de Pós-Graduação da UFMT (PROPG) pelos financiamentos concedidos, que viabilizaram o desenvolvimento deste trabalho.

Este estudo contou ainda com o uso de ferramentas de IA, utilizadas exclusivamente como apoio para a revisão da escrita e organização textual, não desempenhando qualquer papel de autoria intelectual.

Referências

- Adewumi, T., Alkhaled, L., Gurung, N., van Boven, G., e Pagliai, I. (2024). Fairness and bias in multimodal ai: A survey. <https://arxiv.org/abs/2406.19097>. Acesso em: 19 ago. 2025.
- Almeida, N. M. e Mazzoni, A. F. (2020). Design centrado na pessoa com deficiência: práticas inclusivas na criação de tecnologias assistivas. *Revista Brasileira de Design da Informação (InfoDesign)*, 17(2):145–159.
- Bardin, L. (2016). *Análise de conteúdo*. Edições 70, Lisboa.
- Benjamin, R. (2019). *Race after Technology: Abolitionist Tools for the New Jim Code*. Polity Press, Oxford, England.
- Birhane, A. e van Dijk, J. (2020). Robot rights? let's talk about human welfare instead. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, pages 207–213, New York, NY, USA. Association for Computing Machinery.
- Brasil (2018). Lei nº 13.709, de 14 de agosto de 2018. lei geral de proteção de dados pessoais (lgpd). Diário Oficial da União, Brasília, DF, 15 ago. 2018. Acesso em: 26 ago. 2025.
- Cerejeira, T. d. L. T. e Alves, J. F. (2022). Narração na audiodescrição por pessoas com deficiência visual: expandindo o campo profissional. *Revista Educação e Cultura Contemporânea*, 19(59):114–129.

- Chen, M., Cao, Y., Zhang, Y., e Lu, C. (2024). Quantifying and mitigating unimodal biases in multimodal large language models: A causal perspective. <https://arxiv.org/abs/2403.18346>. Acesso em: 19 ago. 2025.
- D'Ignazio, C. e Klein, L. (2020). *Data Feminism*. The MIT Press, Cambridge, MA, USA.
- Duarte, E. F., Alves, D. D., Carneiro, N., Matos, E. S., Baranauskas, M. C. C., e Mendoza, Y. L. L. (2024). GranDIHC-BR 2025–2035 – GC6: Implications of Artificial Intelligence in HCI: A Discussion on Paradigms, Ethics, and Diversity, Equity and Inclusion. In *Anais do Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC)*, pages 1027–1045, Brasília, DF, Brasil. Sociedade Brasileira de Computação.
- Google AI (2023). Gemini: A multimodal ai system. <https://deepmind.google/technologies/gemini/>. Acesso em: 12 maio 2025.
- Jiang, L., Zhu, A., e Oppegaard, B. (2025). Audio description automatons: Exploring perspectives on personas for generative ai description writing assistants. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, New York, NY, USA. Association for Computing Machinery.
- Neves, J. (2022). Translation and accessibility... the translation of everyday things. In Zanettin, F. e Rundle, C., editors, *Routledge Handbook of Translation and Methodology*, pages 441–456. Routledge / Taylor and Francis, London and New York.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press, New York.
- OpenAI (2025). Chatgpt (versão de maio. 2025). <https://chat.openai.com>. Acesso em: 14 maio 2025.
- Pexels (2024). Sobre o pexels. <https://www.pexels.com/pt-br/sobre/>. Acesso em: 29 mar. 2025.
- Pixabay (2024). Termos de serviço. <https://pixabay.com/service/terms/>. Acesso em: 29 mar. 2025.
- Prabhu, V. U. e Birhane, A. (2020). Large image datasets: A pyrrhic win for computer vision? <https://arxiv.org/abs/2006.16923>. Acesso em: 19 ago. 2025.
- Rossetti, J. M. R., Borges, L. C. L. F., e Souza, P. C. (2025). Descrição automática de imagens por ia: uma leitura ética e representativa. In *Anais do XXIV Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC 2025)*, Belo Horizonte, MG, Brasil. Sociedade Brasileira de Computação.
- Santos, S. e Cavalcante, T. (2020). Acessibilidade e audiodescrição: Um olhar para a aprendizagem dos estudantes com deficiência visual. *Educação: Teoria e Prática*, 30:1–19.
- Seo, J., Kamath, S. S., Zeidieh, A., Venkatesh, S., e McCurry, S. (2024). Maidr meets ai: exploring multimodal llm-based data visualization interpretation by and with blind and low-vision users. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–31, St. John's, NL, Canada. Association for Computing Machinery.

UNESCO (2021). Recommendation on the ethics of artificial intelligence. Technical report, UNESCO, Paris. Aprovada em nov. 2021. Acesso em: 19 ago. 2025.

Venkit, P. N., Srinath, M., e Wilson, S. (2023). Automated ableism: An exploration of explicit disability biases in sentiment and toxicity analysis models. <https://arxiv.org/abs/2307.09209>. Acesso em: 19 ago. 2025.

W3C (2018). Accessibility requirements for people with low vision. <https://www.w3.org/TR/low-vision-needs/>. Acesso em: 29 mar. 2025.

West, S. M., Whittaker, M., e Crawford, K. (2019). Discriminating systems: Gender, race and power in ai. Technical report, AI Now Institute, New York. Acesso em: 19 ago. 2025.