



UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO — UNIRIO
CENTRO DE CIÊNCIAS HUMANAS E SOCIAIS — CCH
PROGRAMA DE PÓS-GRADUAÇÃO EM BIBLIOTECONOMIA — PPGB
MESTRADO PROFISSIONAL EM BIBLIOTECONOMIA — MPB

JAQUELINE ANGELA CASAES E SILVA

AUDIODESCRIÇÃO AUTOMATIZADA: INVESTIGAÇÃO DA
VIABILIDADE DA UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL NAS
BIBLIOTECAS

Rio de Janeiro, RJ

2026

JAQUELINE ANGELA CASAES E SILVA

**AUDIODESCRIÇÃO AUTOMATIZADA: INVESTIGAÇÃO DA
VIABILIDADE DA UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL NAS
BIBLIOTECAS**

Dissertação apresentada ao Programa de Pós-Graduação em Biblioteconomia, no Curso de Mestrado Profissional em Biblioteconomia da Universidade Federal do Estado do Rio de Janeiro, como requisito parcial para obtenção do grau de Mestre em Biblioteconomia

Área de Concentração: Biblioteconomia, Cultura e Sociedade.

Linha de Pesquisa: Gestão, Inovação e práticas sustentáveis em Biblioteconomia.

Orientadora: Prof.^a Dr.^a Cládice Nóbile Diniz.

Rio de Janeiro, RJ

2026

Silva, Jaqueline Angela Casaes e

S586 Audiodescrição automatizada: investigação da viabilidade da utilização de inteligência artificial nas bibliotecas / Jaqueline Angela Casaes e Silva. - Rio de Janeiro : UNIRIO, 2026.

117 p. : il. color.

Orientadora: Cládice Nóbile Diniz.

Dissertação (Mestrado) - Universidade Federal do Estado do Rio de Janeiro, Programa de Pós-Graduação em Biblioteconomia, 2026.

1. Acessibilidade informacional. 2. Audiodescrição. 3. Inteligência artificial. 4. Deficiência visual. I. Diniz, Cládice Nóbile, orient.. II. Título.

CDU: 027.7:004.8

JAQUELINE ANGELA CASAES E SILVA

Audiodescrição automatizada: investigação da viabilidade da utilização de inteligência artificial nas bibliotecas

Dissertação apresentada ao Programa de Pós-Graduação em Biblioteconomia, no Curso de Mestrado Profissional em Biblioteconomia da Universidade Federal do Estado do Rio de Janeiro, como requisito parcial para obtenção do grau de Mestre em Biblioteconomia.

Aprovado em: _____

BANCA EXAMINADORA

Dr.^a Cládice Nóbile Diniz — (Orientadora)
Universidade Federal do Estado do Rio de Janeiro — UNIRIO

Dr. Cláudio José Silva Ribeiro — Titular Interno
Universidade Federal do Estado do Rio de Janeiro — UNIRIO

Dr. Hélio Ferreira Orrico — Titular Externo
Universidade Estadual Paulista Júlio de Mesquita Filho — UNESP

Dr. Marcos Miranda — Suplente Interno
Universidade Federal do Estado do Rio de Janeiro — UNIRIO

Dr. Ernani Ribeiro — Suplente Externo
Universidade Federal de Pernambuco — UFPE

Dedico este trabalho a Deus e à
minha família, por todo amor e
apoio ao longo de minha
jornada.

AGRADECIMENTOS

Gostaria de expressar minha profunda gratidão a todos que foram fundamentais para a realização desta pesquisa.

Em primeiro lugar, agradeço a Deus, que me guia, fortalece e sustenta nos momentos mais difíceis, revelando diariamente sua bondade e seu amor. A espiritualidade de luz, meus guias e protetores que são a base da minha vida e me acompanham em cada passo dessa caminhada.

Aos meus pais, Denise e José, minha eterna gratidão por todo o amor, cuidado e sacrifício dedicados ao longo de todos esses anos. Aos meus avós, Neuza Rocha, Aroldo Angelo, Jurea Lima e José Casaes, e ao meu tio, Alexandre Angelo, que seguem me dando forças do céu — saudades eternas.

Ao meu esposo, Júnior, agradeço pelo incentivo constante, pela paciência e por sempre acreditar na minha capacidade. Seu apoio foi essencial para que eu seguisse estudando, me qualificando e conquistando meus objetivos.

Aos meus irmãos e irmãs — Christiane, Felipe, Henrique, Guilherme e Alice — agradeço pelo carinho, pela parceria e pelo apoio ao longo dessa trajetória. Um agradecimento especial à nova integrante da família, minha sobrinha Antonela, que chegou trazendo ainda mais amor, alegria e esperança.

Não poderia deixar de agradecer aos membros peludos da família, que são verdadeiros companheiros de vida: minha cachorrinha Mel e minha gatinha Chitoka. A presença e o afeto de vocês tornam meus dias mais leves, aliviam o estresse e transformam as preocupações em momentos de alegria e diversão.

Aos amigos que fiz ao longo do mestrado, meu muito obrigada. Vocês tornaram essa jornada inesquecível, compartilhando desafios, aprendizados e conquistas. O apoio, a troca e a parceria de vocês foram fundamentais para que eu chegasse até aqui.

Meu agradecimento especial a Wallace Santana, que me incentivou desde o meu ingresso no mestrado, na permanência e na conclusão dessa etapa tão importante, sempre com muito carinho e generosidade.

A todos os professores da UNIRIO, minha gratidão pela contribuição à minha formação acadêmica e profissional. Em especial, à Prof.^a Dr.^a Cládice

Nobile Diniz, minha orientadora, uma pessoa iluminada e inspiradora. Suas contribuições para a área da acessibilidade foram essenciais para o desenvolvimento deste trabalho. Uma mulher excepcional, tanto na docência quanto na vida. Agradeço também ao Professor Marcos Miranda, ser humano extraordinário e profissional dedicado, cuja atuação inspira todos que têm o privilégio de conhecê-lo.

Agradeço os membros da banca avaliadora — Dr. Cláudio José Silva Ribeiro, Dr. Hélio Ferreira Orrico, Dr. Marcos Miranda e Dr. Ernani Ribeiro — pela disponibilidade, pela leitura atenta e pelas contribuições realizadas, que certamente enriqueceram este trabalho.

Este trabalho representa o resultado de uma trajetória construída com esforço, apoio e afeto, e marca uma etapa importante da minha formação acadêmica, que levo comigo com gratidão e responsabilidade.

RESUMO

Investiga a viabilidade do uso de ferramentas de inteligência artificial para a geração automatizada de audiodescrição de capas de livros em bibliotecas universitárias, considerando os princípios da acessibilidade informacional voltados às pessoas com deficiência visual. Adotou-se para o estudo a abordagem qualitativa e o método comparativo com fins descritivos, apoiado em investigação bibliográfica, documental e de campo. Os estudos bibliográfico e documental levaram ao referencial teórico, que aborda a audiodescrição, a acessibilidade informacional, a inclusão no contexto das bibliotecas universitárias e o uso da inteligência artificial na mediação da informação. A pesquisa de campo considerou uma amostra composta por três capas de livros de acervos universitários e gerou audiodescrições por quatro diferentes ferramentas de inteligência artificial, analisando os resultados quanto à estrutura textual, nível de detalhamento, objetividade, uso de linguagem acessível e conformidade normativa, bem como se as audiodescrições produzidas apresentam clareza, detalhamento, precisão e alinhamento às recomendações normativas brasileiras, em especial às diretrizes da Associação Brasileira de Normas Técnicas, no que se refere à neutralidade descritiva e à adequação do conteúdo. Os resultados indicam que as ferramentas de inteligência artificial apresentam potencial significativo para apoiar a produção de audiodescrição em bibliotecas universitárias, especialmente no que se refere à agilidade dos processos e à ampliação do acesso à informação. Entretanto, identificam-se limitações relacionadas à padronização das descrições, à interpretação de elementos visuais complexos e à necessidade de revisão humana para assegurar qualidade, precisão e aderência às normas técnicas. Conclui-se que a inteligência artificial pode atuar como ferramenta complementar no processo de audiodescrição, desde que integrada a práticas profissionais e orientada por critérios técnicos e éticos, contribuindo para o fortalecimento da acessibilidade informacional no ambiente acadêmico.

Palavras-chave: Acessibilidade informacional. Audiodescrição. Inteligência artificial. Deficiência visual.

ABSTRACT

This study investigates the feasibility of using artificial intelligence tools for the automated generation of audio descriptions of book covers in university libraries, considering the principles of informational accessibility for people with visual impairments. A qualitative approach and a comparative method with descriptive purposes were adopted, supported by bibliographic, documentary, and field research. The bibliographic and documentary studies informed the theoretical framework, which addresses audio description, informational accessibility, inclusion in the context of university libraries, and the use of artificial intelligence in information mediation. The field research considered a sample consisting of three book covers from university library collections and generated audio descriptions using four different artificial intelligence tools. The results were analyzed in terms of textual structure, level of detail, objectivity, use of accessible language, and normative compliance, as well as whether the produced audio descriptions demonstrate clarity, detail, accuracy, and alignment with Brazilian normative recommendations, particularly the guidelines of the Brazilian Association of Technical Standards, regarding descriptive neutrality and content adequacy. The findings indicate that artificial intelligence tools have significant potential to support the production of audio description in university libraries, especially in terms of process efficiency and the expansion of access to information. However, limitations were identified regarding the standardization of descriptions, the interpretation of complex visual elements, and the need for human review to ensure quality, accuracy, and adherence to technical standards. It is concluded that artificial intelligence can act as a complementary tool in the audio description process, provided that it is integrated into professional practices and guided by technical and ethical criteria, thereby contributing to the strengthening of informational accessibility in academic environments.

Keywords: Informational accessibility. Audio description. Artificial intelligence. Visual impairment.

LISTA DE ILUSTRAÇÕES

Figura 1 - Livro pelo catálogo <i>on-line</i> da UNIRIO.....	66
Figura 2 - Catálogo da Biblioteca Nacional, obra “[Cabeça de mulher]”.....	67
Figura 3 - Catálogo da Biblioteca Nacional, obra “[Duas figuras e uma árvore]”.	68
Figura 4 - Capa da obra “Manual de Primeiros Socorros”.....	79
Figura 5 - Capa da obra “Informação e controle bibliográfico: um olhar sobre a cibernética”.....	85
Figura 6 - Capa da obra “O Eu profundo e os outros Eus”.....	93

LISTA DE QUADROS

Quadro 1 - Percurso metodológico.....	62
Quadro 2 - Expressões de busca utilizadas nas bases de dados	64

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
API	<i>Application Programming Interface</i> (Interface de Programação de Aplicações)
BDTD	Biblioteca Digital de Teses e Dissertações
CENSUP	Censo da Educação Superior
DL	<i>Deep Learning</i> (aprendizado profundo)
GPT	<i>Generative Pre-trained Transformer</i>
HTML	<i>HyperText Markup Language</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
IA	Inteligência Artificial
IFLA	Federação Internacional de Associações e Instituições Bibliotecárias
LLM	<i>Large Language Models</i> (grandes modelos de linguagem)
ML	<i>Machine Learning</i> (aprendizado de máquina)
NAPNE	Núcleos de Apoio às Pessoas com Necessidades Educacionais Específicas
NBR	Norma Brasileira Registrada
ONU	Organização das Nações Unidas
PLN	Processamento de Linguagem Natural
PNAD	Pesquisa Nacional por Amostra de Domicílios
SAP	<i>Secondary Audio Program</i> (canal de áudio secundário)
UNIRIO	Universidade Federal do Estado do Rio de Janeiro
WCAG	<i>Web Content Accessibility Guidelines</i>
W3C	<i>World Wide Web Consortium</i>

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivo geral e específicos	19
1.2	Justificativa	20
2	REFERENCIAL TEÓRICO	22
2.1	Acessibilidade frente à legislação brasileira.....	22
2.1.1	A Lei Brasileira de Inclusão da Pessoa com Deficiência	22
2.1.2	O Decreto do atendimento prioritário às pessoas com deficiência	25
2.1.3	O processo de acessibilização para implantar a legislação	27
2.2	A semiótica da imagem	32
2.3	A capa de livro como elemento descritivo em bibliotecas	36
2.4	Audiodescrição: princípios, normas e práticas	39
2.5	Inteligência Artificial e audiodescrição automatizada.....	45
3	METODOLOGIA.....	54
3.1	O objeto da pesquisa, seu universo e amostra.....	54
3.2	Descrição da amostra: as ferramentas de IA utilizadas	55
3.2.1	ChatGPT-5.1.....	55
3.2.2	Gemini 2.5 Flash.....	56
3.2.3	Microsoft Copilot	57
3.2.4	DeepSeek-V3.2	58
3.3	Delimitação do objeto e ferramentas	59
3.4	Os dados e seu tratamento por abordagem qualitativa e o percurso metodológico	60
3.5	Caracterização da pesquisa.....	63
3.5.1	As investigações bibliográficas e documentais	63
3.5.2	A investigação de campo	65
3.6	Roteiro técnico de audiodescrição	70
3.7	Elaboração dos <i>prompts</i>.....	73
3.8	CrITÉrios de avaliação das descrições	75
4	RESULTADOS	77
4.1	Conjunto de imagens analisadas	77
4.2	Análises por ferramenta.....	78
4.2.1	Análise da capa do livro Manual de Primeiros Socorros (Fundação	

	Oswaldo Cruz, 2003)	79
4.2.2	Análise da capa do livro Informação e controle bibliográfico: um olhar sobre a cibernética (Machado, 2023).....	85
4.2.3	Análise da capa do livro O Eu profundo e os outros Eus (Pessoa, 1913)	92
5	DISCUSSÕES	101
6	CONSIDERAÇÕES FINAIS	104
	REFERÊNCIAS.....	107
	APÊNDICE A – MANUAL DE PROCEDIMENTOS PARA GERAÇÃO E AVALIAÇÃO DAS AUDIODESCRIÇÕES	113
	ANEXO A – DADOS FORNECIDOS PELO NÚCLEO DE ACESSIBILIDADE E INCLUSÃO (NAI/UNIRIO) SOBRE ESTUDANTES COM DEFICIÊNCIA VISUAL (2025).....	114

1 INTRODUÇÃO

A audiodescrição constitui um recurso indispensável para a inclusão de pessoas com deficiência visual, ao possibilitar o acesso aos conteúdos imagéticos por meio da linguagem verbal. No âmbito das bibliotecas, a audiodescrição assume papel central na garantia da acessibilidade informacional, uma vez que materiais bibliográficos incorporam capas ilustradas, gráficos, fotografias ou elementos visuais que complementam o texto. Sem a devida mediação, esses elementos tornam-se barreiras para usuários que dependem exclusivamente da percepção auditiva ou tátil para acessar a informação.

Com os avanços dos modelos de IA, passaram a existir ferramentas capazes de identificar e gerar descrições automáticas de elementos presentes em uma imagem. Contudo, permanecem em aberto questões sobre a precisão, a neutralidade e a adequação das audiodescrições produzidas por tais tecnologias, sobretudo quando confrontadas com os parâmetros definidos pelas diretrizes normativas que regem a prática da audiodescrição no Brasil.

A presente pesquisa investiga a viabilidade da audiodescrição automatizada, com uso da Inteligência Artificial (IA) no âmbito das bibliotecas, em prol da acessibilidade informacional de usuários com deficiência visual.

Esta dissertação foi elaborada em conformidade com as diretrizes de acessibilidade, visando garantir sua plena utilização por pessoas com deficiência, especialmente usuários de leitores de tela. O documento adota estrutura textual adequada, linguagem clara e recursos que favorecem a leitura assistiva, de modo a assegurar o acesso equitativo à informação e atender aos princípios da acessibilidade informacional.

Para pessoas sem deficiência visual, a descrição de uma imagem auxilia a compreensão do conteúdo textual. No entanto, para aquelas com deficiência visual, a descrição assume papel fundamental: comunicar, por meio de palavras, o conteúdo visual e sua conexão com o texto, tornando acessível o que não pode ser visto. Isso é importante, considerando que *softwares* leitores de tela, como o

Dosvox¹, não interpretam elementos gráficos, somente texto. Logo, para pessoas com deficiência visual em seus diferentes tipos, como baixa visão, cegueira, daltonismo, entre outros, a descrição de imagem é um recurso de acessibilidade fundamental.

No contexto das bibliotecas, a acessibilidade é relevante, pois essas instituições são fundamentais na disseminação do conhecimento e na promoção da igualdade de oportunidades, principalmente durante a vida acadêmica dos estudantes. No entanto, pessoas com deficiência visual enfrentam barreiras ao acessar materiais imagéticos impressos ou digitais nas bibliotecas. Isso inclui a dificuldade em compreender elementos como capas de livros, imagens e gráficos no conteúdo textual. A ausência de descrição desses elementos compromete a experiência do usuário e limita sua autonomia informacional.

O cenário é relevante em bibliotecas universitárias, espaços estratégicos de apoio à aprendizagem, pesquisa e formação, onde a presença de estudantes com deficiência visual reforça a necessidade de integrar práticas de acessibilidade tanto em recursos físicos quanto digitais.

Uma das soluções previstas na legislação brasileira, definida pelo Decreto n.º 5.296/2004, art. 53, § 2º, inciso III é “a descrição e narração em voz de cenas e imagens” (Brasil, 2004), ou seja, audiodescrição, que pode ser feita mediante voz, mas também ser textual e pretende tornar o conteúdo visual acessível, por meio da descrição das imagens. No entanto, a produção de audiodescrições é uma atividade que requer conhecimento e habilidade, ao envolver uma síntese de informações visuais de forma sucinta, para transmitir com eficácia os elementos essenciais de uma cena ou imagem. A audiodescrição também é um recurso para transpor as barreiras nas comunicações e na informação, termo utilizado pela Lei n.º 13.146, de 06 de julho de 2015, que “institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência)”, para se referir a qualquer dificuldade que “impossibilite a expressão ou o recebimento de mensagens e de informações por intermédio de sistemas de comunicação e de tecnologia da informação” (Brasil, 2015).

¹ DOSVOX é um *software* de código aberto desenvolvido pelo Núcleo de Computação Eletrônica (NCE) da Universidade Federal do Rio de Janeiro. Disponível em: <http://www.dosvox.nce.ufrj.br>. Acesso em: 02 fevereiro de 2026.

Com o avanço das tecnologias digitais, novas possibilidades se abrem para apoiar a produção de audiodescrições nas bibliotecas. Ferramentas capazes de interpretar imagens podem atuar como apoio ao trabalho do bibliotecário em etapas como indexação ou representação descritiva, sobretudo em acervos extensos. Apesar da crescente adoção de IA para diferentes tarefas informacionais, ainda há poucas pesquisas que examinem sua eficácia na geração de descrições acessíveis de maneira sistemática no contexto das bibliotecas.

Ainda conforme a Lei n.º 13.146/2015, a deficiência é caracterizada pela interação entre limitações corporais, sensoriais e intelectuais do indivíduo com barreiras que podem dificultar ou impedir sua participação na sociedade (Brasil, 2015). Em consonância, Fernandes (2023) reitera que a deficiência não se limita somente às condições físicas ou sensoriais de uma pessoa, mas também é influenciada pela forma como o ambiente e a sociedade interagem com essas condições.

Consoante os resultados da Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua), realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), a população total do Brasil em 2022 foi estimada em 203,1 milhões de indivíduos. Dentre essa população, 8,9% (18,6 milhões) são pessoas com idades a partir de 2 anos que possuem algum tipo de deficiência. Entre as deficiências investigadas, a dificuldade de enxergar, que os auxílios ópticos não resolvem, afeta 3,1% (6,3 milhões) da população, classificando-se como a segunda deficiência mais comum (IBGE, 2022).

A expressiva diferença na taxa de analfabetismo entre pessoas com e sem deficiência, apontada pela pesquisa, expõe uma grave barreira no acesso à educação e reforça a urgência de políticas inclusivas. Enquanto somente 4,1% das pessoas sem deficiência são analfabetas, essa taxa chega a 19,5% para pessoas com deficiência. Além disso, somente 25,6% das pessoas com deficiência relataram ter concluído pelo menos o Ensino Médio, em comparação com 57,3% das pessoas sem deficiência (IBGE, 2022).

Ao analisarmos os dados por faixa etária, observamos um aumento nas dificuldades de visão à medida que a população envelhece. Para a faixa etária de 40 a 49 anos, 2,9% das pessoas enfrentam dificuldades nesse aspecto. Essa

porcentagem cresce com o avanço da idade, atingindo 4,7% para a faixa de 50 a 59 anos, 7,1% para a faixa de 60 a 69 anos, 10,2% para a faixa de 70 a 79 anos e chegando a 17,9% para a faixa de 80 a 89 anos (IBGE, 2022).

Indivíduos com deficiência visual enfrentam uma variedade de desafios e barreiras em suas atividades diárias, independentemente de sua faixa etária, conforme evidenciado pelo IBGE em 2022. Entre essas dificuldades estão incluídas restrições significativas na capacidade de comunicação e no acesso à informação.

Voltando a considerar a Lei n.º 13.146/2015 e o Decreto n.º 5.296/2004, que estabelece normas e critérios para a promoção da acessibilidade, juntamente com o Decreto n.º 6.949/2009, que promulga a Convenção Internacional sobre os Direitos das Pessoas com Deficiência e reafirma o direito de acesso à informação e à comunicação, se pode afirmar que para garantir a acessibilidade de pessoas com deficiência visual na esfera acadêmica, é imprescindível incluir a audiodescrição entre os recursos educacionais que servem para facilitar e conduzir a aprendizagem, garantindo que os conteúdos sejam acessíveis, independentes de seus suportes, sejam físicos ou digitais (Brasil, 2004, 2009, 2025).

Por outro lado, a observação empírica nas vivências práticas em biblioteca pela pesquisadora, tanto em projetos de leitura acessível quanto nas descrições institucionais de acervos visuais, foi notada a ausência de audiodescrição em materiais bibliográficos, embora a Norma Brasileira Registrada NBR 16452:2016 da Associação Brasileira de Normas Técnicas (ABNT) estabeleça diretrizes para a audiodescrição em publicações, e a ABNT NBR 17225:2025, aborde requisitos para acessibilidade digital, a prática da audiodescrição em catálogos digitais de bibliotecas ainda é incipiente (Associação ..., 2016, 2015).

Nesse contexto, ao investigar a viabilidade do uso das ferramentas de IA para a produção automatizada de descrições de imagens, não foi encontrado na literatura científica significativa quantidade de estudos sobre a eficácia e as limitações das ferramentas de IA na geração de audiodescrições para capas de livros, considerando os fundamentos da linguística e da semiótica para a interpretação da imagem como discurso e construção de sentido, fato que

ensejou a proposição deste estudo, propondo-se à busca de soluções por meio de ferramentas de IA para a produção automatizada de descrições de imagens, para não somente melhorar a acessibilidade para pessoas com deficiência visual, mas também proporcionar uma experiência mais inclusiva para os usuários das bibliotecas.

Limitou-se a pesquisar sobre descrições acessíveis das capas dos livros dado o prazo de investigação estabelecido, porque apesar da relevância da audiodescrição e das normativas que orientam sua aplicação, os catálogos bibliográficos raramente apresentam essas descrições acessíveis; e a um caso, o das bibliotecas da Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Resultando dessa escolha, que a questão a se investigar é a de quais são as potencialidades e limitações do uso de ferramentas de IA na geração automatizada de descrições de capas de livros nas bibliotecas da UNIRIO²?

1.1 Objetivo geral e específicos

O objetivo geral deste trabalho é investigar a viabilidade do uso de ferramentas de IA para a audiodescrição automatizada de capas de livros em bibliotecas universitárias. A análise considera os modelos ChatGPT (OpenAI), Copilot (Microsoft), Gemini (Google) e DeepSeek, utilizando um *prompt* elaborado com base na ABNT NBR 16452:2016 e em consonância com os princípios de acessibilidade da WCAG 2.2.

Os objetivos específicos são:

² A UNIRIO foi selecionada como estudo de caso por um conjunto de fatores institucionais, acadêmicos e metodológicos que a tornam representativa para a investigação sobre acessibilidade informacional em bibliotecas universitárias. É uma universidade pública de reconhecida relevância no cenário do ensino superior, com ampla diversidade de cursos, incluindo formações em áreas da saúde que demandam interação com materiais técnico-científicos. Sua estrutura multicampi e o sistema de bibliotecas dividido em central e setoriais atendem a perfis variados de usuários e combinam serviços presenciais e digitais, permitindo observar a implementação de práticas de acessibilidade em diferentes contextos internos. Foi solicitado ao Núcleo de Acessibilidade e Inclusão (NAI) da UNIRIO um levantamento sobre a presença de estudantes com deficiência visual na universidade. O núcleo encaminhou uma listagem de cursos em que há estudantes cadastrados com cegueira ou baixa visão/visão monocular, indicando a presença desse público em diferentes áreas, como Ciências Biológicas, Biomedicina, Enfermagem, Medicina, Nutrição, Biblioteconomia, Arquivologia, Museologia, Pedagogia, Direito, Serviço Social, Teatro, Música, Administração Pública, entre outras (**Erro! Fonte de referência não encontrada.**). Esses dados reforçam que a UNIRIO abriga um contingente de estudantes com deficiência visual distribuído em diversos cursos e turnos, o que torna ainda mais relevante a análise das condições de acessibilidade informacional em seu sistema de bibliotecas (UNIRIO, 2025).

- Construir um referencial teórico consistente sobre audiodescrição e acessibilidade informacional;
- Analisar a clareza e o nível de detalhamento das audiodescrições geradas pelas ferramentas de IA (ChatGPT, Copilot, Gemini e DeepSeek) para três capas selecionadas no contexto universitário;
- Avaliar a neutralidade e a precisão das descrições, considerando a ABNT NBR 16452:2016 e os princípios de acessibilidade digital da WCAG 2.2;
- Comparar o desempenho dos quatro modelos multimodais, identificando convergências, divergências e inconsistências na produção automatizada das descrições;
- Identificar potencialidades e limitações do uso de IA para audiodescrição no contexto das bibliotecas universitárias;
- Produzir um manual prático com orientações sobre o uso das ferramentas analisadas.

1.2 Justificativa

Assegurar a acessibilidade nas comunicações e no acesso à informação é um princípio fundamental na construção de uma sociedade inclusiva, na qual todas as pessoas tenham acesso equitativo ao conteúdo e ao conhecimento. No contexto acadêmico, torna-se essencial garantir que pessoas com deficiência visual possam compreender e assimilar conteúdos imagéticos, uma vez que imagens são utilizadas para ilustrar conceitos, apresentar dados por meio de gráficos e apoiar a compreensão de fenômenos abstratos, atuando de forma complementar ao texto escrito.

Apesar dos avanços na promoção da acessibilidade, por meio de dispositivos legais que asseguram a educação em condições de igualdade, persistem lacunas importantes. A legislação educacional brasileira estabelece que bibliotecas escolares devem atender ao princípio da acessibilidade, assegurando condições físicas e informacionais adequadas para que se constituam como espaços inclusivos. Entretanto, ainda se observa uma ausência significativa de descrições de imagens em artigos, livros, trabalhos acadêmicos e demais materiais que utilizam elementos visuais.

A norma ABNT NBR 16452:2016, ao tratar de publicações acessíveis, determina que produtos digitais destinados ao público devem disponibilizar audiodescrição para imagens, tabelas, gráficos e demais elementos visuais. Contudo, a observação cotidiana de catálogos digitais de bibliotecas revela que essa diretriz nem sempre é implementada, o que gera lacunas no acesso à informação para usuários com deficiência visual. Embora títulos e legendas ofereçam identificação básica, eles não substituem uma descrição capaz de transmitir o conteúdo visual e seu papel no contexto, sobretudo quando a imagem apresenta múltiplos detalhes relevantes (Associação..., 2016).

Reconhecer o valor das imagens é essencial em um mundo orientado pelo visual. Elas não constituem apenas ilustrações, mas elementos discursivos que participam da construção de sentido. Nesse aspecto, teorias linguísticas e semióticas, como as de Saussure e Peirce, ressaltam que as imagens possuem uma estrutura própria de significação, demandando mediação quando não são acessíveis. Títulos e legendas, por sua brevidade, raramente são suficientes para traduzir essa complexidade. A audiodescrição, assim, não se limita a complementar o texto, mas assume o papel de mediação semiótica capaz de reconstruir, em linguagem verbal, a narrativa visual originalmente acessível apenas pela percepção da imagem.

Nesse cenário, justifica-se investigar a viabilidade do uso de ferramentas de IA como tecnologias assistivas para apoiar a audiodescrição de imagens em bibliotecas universitárias. Ao explorar o potencial dessas ferramentas, este estudo busca não apenas mitigar barreiras informacionais persistentes, mas também promover práticas mais inclusivas e equitativas de acesso à informação, alinhando-se às exigências normativas vigentes e às demandas de uma sociedade que preza pela inclusão, autonomia e participação plena de todos os seus membros.

2 REFERENCIAL TEÓRICO

Este referencial teórico apresenta os fundamentos que sustentam a discussão proposta, abordando conceitos de acessibilidade informacional, princípios da audiodescrição, bases linguísticas e semióticas envolvidas na mediação de imagens, bem como o papel da IA na automatização desse processo. Também são destacadas pesquisas e normativas, incluindo a ABNT NBR 16452:2016, a recente ABNT NBR 17225:2025 e a WCAG 2.2, que estabelecem padrões para a produção de materiais acessíveis e para acessibilidade digital.

2.1 Acessibilidade frente à legislação brasileira

A acessibilidade constitui um princípio fundamental para a garantia de direitos e para a promoção da inclusão social das pessoas com deficiência. No Brasil, esse princípio encontra respaldo em um conjunto de dispositivos legais que estabelecem diretrizes para a eliminação de barreiras físicas, comunicacionais, informacionais e tecnológicas. A consolidação desse marco normativo ocorreu de forma gradual, articulando legislações específicas, decretos regulamentadores e políticas públicas voltadas à promoção da equidade no acesso a serviços, espaços e informações. Nesse contexto, destacam-se instrumentos legais que orientam a implementação da acessibilidade em diferentes dimensões da vida social, incluindo educação, comunicação, tecnologia e acesso à informação. As subseções a seguir apresentam alguns dos principais dispositivos da legislação brasileira relacionados à acessibilidade.

2.1.1 A Lei Brasileira de Inclusão da Pessoa com Deficiência

A Lei n.º 13.146/2015 pretende assegurar a inclusão social e a igualdade de direitos para pessoas com deficiência no Brasil (Brasil, 2015). Fundamentada na Convenção sobre os Direitos das Pessoas com Deficiência, ratificada pelo Brasil em 2008, estabelece diretrizes para a acessibilidade, autonomia e participação efetiva das pessoas com deficiência na sociedade, definindo a pessoa com deficiência como o indivíduo que possui impedimentos físicos, mentais, intelectuais ou sensoriais de longo prazo que, ao interagir com

barreiras, podem ter dificultada sua participação na sociedade (Brasil, 2015, art. 2º). A avaliação dessa deficiência deve seguir um modelo biopsicossocial, considerando fatores ambientais, psicológicos e estruturais que afetam a vida da pessoa (Brasil, 2015).

Essa Lei reforça a importância da acessibilidade, garantindo condições para que pessoas com deficiência usufruam de espaços, serviços, meios de transporte, comunicação e tecnologias assistivas (Brasil, 2015, art. 3º, I). Além disso, o Desenho Universal é estabelecido como princípio norteador para a criação de produtos e ambientes acessíveis desde sua concepção, sem necessidade de adaptações posteriores (Brasil, 2015, art. 3º, II). Em relação às barreiras, a legislação classifica diversos tipos, incluindo barreiras arquitetônicas, urbanísticas, comunicacionais, tecnológicas e atitudinais, sendo esta última caracterizada por comportamentos e preconceitos que impedem a inclusão das pessoas com deficiência (Brasil, 2015, art. 3º, IV). E define direitos e serviços essenciais, como moradia inclusiva, suporte de profissionais de apoio escolar e garantia de acessibilidade na informação e comunicação, incluindo o uso da Língua Brasileira de Sinais (Libras) e de outras tecnologias assistivas; reforçando o dever do Estado e da sociedade em eliminar barreiras e promover a inclusão (Brasil, 2015).

No capítulo II da citada Lei (2015), é assegurado o direito à igualdade de oportunidades e a não discriminação para todas as pessoas com deficiência no Brasil (Brasil, 2015). O seu artigo 4º define que qualquer forma de distinção, restrição ou exclusão que limite direitos fundamentais da pessoa com deficiência, incluindo a recusa de adaptações razoáveis e o fornecimento de tecnologias assistivas, configura discriminação (Brasil, 2015, art. 4º, § 1º).

Ainda sobre essa Lei, verifica-se que ela também protege as pessoas com deficiência contra qualquer forma de violência, exploração, negligência e tratamento degradante, estabelecendo atenção especial para grupos vulneráveis, como crianças, adolescentes, mulheres e idosos com deficiência (Brasil, 2015, art. 5º). Além disso, esse texto legal reforça que a deficiência não restringe a plena capacidade civil da pessoa, garantindo o direito ao casamento, à constituição de família, à reprodução e ao acesso à informação sobre planejamento familiar, vedando práticas como a esterilização compulsória

(Brasil, 2015, art. 6º). Em seu artigo 8º, estabelece o compromisso do Estado, da sociedade e da família na promoção dos direitos das pessoas com deficiência, abrangendo áreas essenciais como saúde, educação, trabalho, previdência social, acessibilidade, cultura, esporte, lazer, comunicação e avanços científicos e tecnológicos (Brasil, 2015, art. 8º).

No capítulo IV da Lei (2015), é estabelecido que a educação é um direito fundamental da pessoa com deficiência, devendo ser garantido um sistema educacional inclusivo em todos os níveis, promovendo a equidade no acesso, permanência e participação (Brasil, 2015, art. 27). O Estado, a família e a sociedade são responsáveis por assegurar uma educação de qualidade, prevenindo qualquer forma de discriminação ou negligência. O seu artigo 28 detalha as obrigações do poder público na implementação da educação inclusiva, incluindo a adoção de medidas de acessibilidade, a institucionalização do atendimento educacional especializado e o incentivo a práticas pedagógicas inclusivas (Brasil, 2015, art. 28, IV). Esse artigo determina que instituições privadas de ensino não podem cobrar valores adicionais por adaptações voltadas aos estudantes com deficiência, assegurando que a inclusão não seja tratada como um serviço extra (Brasil, 2015, art. 28, §1º).

No ensino superior e técnico, a Lei n.º 13.146/2015, exige a adoção de medidas de acessibilidade nos processos seletivos, garantindo a disponibilização de provas em formatos acessíveis, recursos de tecnologia assistiva e tempo adicional, quando necessário (Brasil, 2015, art. 30).

Essa Lei reafirma que a educação inclusiva deve contemplar atividades recreativas, esportivas e culturais, garantindo a participação equitativa da pessoa com deficiência em todas as esferas da vida acadêmica. Assim, consolida a educação como um direito essencial, exigindo a eliminação de barreiras e a implementação de políticas que assegurem não apenas o acesso, mas a permanência e a aprendizagem dos estudantes com deficiência (Brasil, 2015).

A acessibilidade na informação e comunicação é um dos pilares fundamentais da inclusão. O seu artigo 63 determina que todos os sítios da *internet* mantidos por empresas e órgãos públicos no Brasil devem ser acessíveis para pessoas com deficiência, conforme as melhores práticas

internacionais (Brasil, 2015, art. 63). Isso significa que os *websites* devem permitir navegação por leitores de tela, oferecer audiodescrição de imagens e vídeos, além de possuir *design* responsivo para facilitar o uso por diferentes públicos. Além disso, o poder público deve incentivar a produção e comercialização de livros em formatos acessíveis, garantindo que publicações financiadas com recursos públicos sejam disponibilizadas em versões adaptadas, como arquivos compatíveis com leitores de tela e impressão em Braille (Brasil, 2015, art. 68).

A tecnologia assistiva consiste em um conjunto de ferramentas que promove autonomia e qualidade de vida para pessoas com deficiência, incluindo equipamentos, dispositivos e *softwares* acessíveis. A Lei garante o direito de acesso a esses recursos e determina que o poder público deve criar incentivos para facilitar sua aquisição (Brasil, 2015, art. 74). Entre as medidas previstas, estão a facilitação do crédito especializado, que inclui linhas subsidiadas para a compra de tecnologia assistiva, e a simplificação da importação de dispositivos assistivos, eliminando barreiras burocráticas e alfandegárias. Além disso, a legislação prevê o fomento à pesquisa e produção nacional de tecnologia assistiva, garantindo incentivos fiscais para empresas que desenvolvem produtos acessíveis. Outro ponto importante é a redução ou isenção de tributos sobre tecnologia assistiva, tornando esses produtos mais acessíveis para a população (Brasil, 2015, art. 75). A Lei também prevê a capacitação contínua de tradutores e intérpretes de Libras, guias-intérpretes para surdocegos, profissionais especializados em Braille, audiodescrição e legendagem (Brasil, 2015, art. 73).

Por fim, a Lei n.º 13.146/2015 estabelece diretrizes para a acessibilidade ser tratada não apenas como um direito individual, mas como um dever coletivo do Estado, das empresas e da sociedade. Ao integrar infraestrutura acessível, comunicação inclusiva e incentivo à tecnologia assistiva, promove uma abordagem sistêmica da acessibilidade, garantindo que pessoas com deficiência tenham plena participação na vida social, econômica e digital do país.

2.1.2 O Decreto do atendimento prioritário às pessoas com deficiência

O Decreto n.º 5.296, de 2 de dezembro de 2004, regulamenta as Leis n.º 10.048/2000 (que estabelece atendimento prioritário às pessoas com deficiência) e 10.098/2000 (que estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas com deficiência ou com mobilidade reduzida). O Decreto aborda a acessibilidade em diferentes áreas, incluindo infraestrutura urbana, transporte, comunicação e tecnologia assistiva, garantindo direitos fundamentais para a inclusão e autonomia desses indivíduos. No que diz respeito à acessibilidade na comunicação e informação, o decreto determina que serviços de radiodifusão de sons e imagens, como emissoras de rádio e TV, devem adotar recursos acessíveis para pessoas com deficiência auditiva e visual. Entre as medidas obrigatórias estão subtitulação oculta (*closed caption*), janela de Libras e audiodescrição, garantindo que esses públicos tenham acesso ao conteúdo audiovisual de forma inclusiva (Brasil, 2004, art. 52).

Esse Decreto também reforça a necessidade de acessibilidade digital, estabelecendo que sítios eletrônicos de órgãos públicos devem ser desenvolvidos e mantidos conforme padrões que permitam sua navegação por pessoas com deficiência. Isso inclui compatibilidade com leitores de tela, teclados adaptados e outras tecnologias assistivas, assegurando a equidade no acesso à informação governamental (Brasil, 2004, art. 47).

Outro aspecto relevante do decreto é a exigência de que documentos oficiais, materiais educativos e publicações institucionais sejam disponibilizados em formatos acessíveis, como versões em Braille, caracteres ampliados e arquivos digitais compatíveis com *softwares* de leitura. Essa medida garante que pessoas com deficiência visual possam ter acesso ao conteúdo produzido por órgãos públicos e demais instituições (Brasil, 2004, art. 53). A implementação dessas diretrizes é fundamental para que bibliotecas, arquivos e centros de informação adotem práticas acessíveis. O decreto fortalece a necessidade de adaptação de acervos físicos e digitais, implementação de tecnologias assistivas e disponibilização de materiais audiovisuais acessíveis, tornando esses espaços verdadeiramente inclusivos.

Dessa forma, a acessibilidade informacional constitui um princípio fundamental para a promoção da equidade no acesso ao conhecimento, ao garantir a eliminação de barreiras que dificultam a apropriação da informação

por parte de todos os indivíduos, independentemente de suas limitações físicas, sensoriais ou cognitivas.

2.1.3 O processo de acessibilização para implantar a legislação

Conforme destacam Santos e Araújo (2015), a acessibilidade transcende o conteúdo em si, abrangendo igualmente como a informação é estruturada, disponibilizada e compartilhada.

Com os avanços das tecnologias digitais, ampliaram-se as possibilidades de inclusão informacional, destacando-se o papel dos objetos digitais acessíveis, como arquivos de texto, documentos *HyperText Markup Language* (HTML) e recursos em áudio. Tais formatos são compatíveis com tecnologias assistivas — a exemplo de leitores de tela, ampliadores e sintetizadores de voz — e contribuem para uma mediação mais eficaz entre o conteúdo informacional e os usuários com deficiência (Santos; Araújo, 2015). Assim, a acessibilidade informacional consolida-se como eixo estratégico nas políticas públicas voltadas à cidadania digital.

Ferreira (2023) reforça essa perspectiva ao compreender a informação pública como um bem intangível e um patrimônio social que deve ser acessado de forma equitativa e autônoma por todos os cidadãos. Nesse sentido, a acessibilidade informacional deve integrar-se ao planejamento urbano por meio da articulação com os princípios do Desenho Universal, superando a visão reducionista que associa acessibilidade apenas a rampas ou pisos táteis. A comunicação visual clara, objetiva e compreensível torna-se, portanto, elemento essencial para a inclusão.

Meireles (2024) indica que a acessibilidade informacional não se resume à presença de rampas ou pisos táteis, mas também envolve a comunicação visual clara, objetiva e compreensível. O autor também cita a ABNT NBR 9050:2020. Esta norma recomenda que “as informações devem ser completas, precisas e claras [...], dispostas segundo o princípio dos dois sentidos, perceptíveis e legíveis para todos, inclusive às pessoas com deficiência” (Associação..., 2020, p. 37). No campo educacional, a inclusão de pessoas com deficiência no ensino superior constitui um processo multidimensional, que extrapola o cumprimento da legislação vigente. Ribeiro (2022) aponta que,

apesar do respaldo legal proporcionado pela Lei n.º 13.146/2015, persistem barreiras estruturais, informacionais e atitudinais que dificultam o acesso pleno de estudantes com deficiência visual. A autora destaca que o direito à acessibilidade deve ser compreendido como parte indissociável da qualidade da educação superior, exigindo políticas públicas eficazes e sensibilidade institucional para o enfrentamento das desigualdades (Brasil, 2015).

Sob essa ótica, a acessibilidade informacional deve ser reconhecida como direito fundamental, à semelhança da acessibilidade física. Ribeiro (2022) enfatiza que a comunicação inclusiva demanda a disponibilização de informações em múltiplos formatos acessíveis, como Braille, arquivos digitais compatíveis com leitores de tela e audiolivros. A autora argumenta que a concepção de ambientes informacionais deve prever a diversidade dos usuários desde o início, evitando adaptações excludentes.

Nesse sentido, Sales (2023) reforça que a acessibilidade digital deve ser concebida desde a estruturação dos sistemas informacionais. A autora aponta que os ambientes virtuais devem ser acessíveis, integrando as dimensões da acessibilidade (arquitetônica, comunicacional, informacional, instrumental, atitudinal e programática) e define a acessibilidade informacional como a possibilidade de acesso pleno ao conteúdo, com autonomia e sem prejuízo das singularidades corporais ou sensoriais dos usuários.

Observa-se que a Lei n.º 13.146/2015 determina que a acessibilidade digital seja considerada na concepção e na orientação de projetos que tratem de sistemas de tecnologia da informação e comunicação de uso público ou privado coletivo devem ser orientados pelos princípios do Desenho Universal como regra de caráter geral (Brasil, 2015, § 1º, art. 55).

Braseliano (2024) destaca que o processo de acessibilização ultrapassa o mero fornecimento de acesso formal à informação. Trata-se de garantir a apropriação efetiva do conteúdo, considerando a pluralidade de letramentos e competências cognitivas dos sujeitos. Nesse contexto, a autora propõe uma leitura da acessibilidade como uma “gramática” composta por três pilares: “acesso (acessável), acessibilidade (acessível) e acessibilização (acessibilizável)” (Braseliano, 2024).

Conforme aponta a autora: “não basta conferir acesso e tornar acessível, é preciso haver engajamento para sua apropriação” (Braseliano, 2024), reafirmando que a inclusão pressupõe uma pedagogia comunicacional e informacional sensível às singularidades dos sujeitos.

A acessibilidade informacional, portanto, emerge como um direito transversal, que perpassa diversas áreas da vida social e institucional. A Declaração de Lyon sobre o Acesso à Informação e ao Desenvolvimento, lançada em 2014 durante o Congresso da Federação Internacional de Associações e Instituições Bibliotecárias (IFLA), destaca a importância do acesso livre e equitativo à informação como um direito fundamental para o desenvolvimento social, econômico e cultural das nações (IFLA, 2014). Seu principal objetivo é incentivar políticas públicas que garantam que todas as pessoas, independentemente de suas condições socioeconômicas ou limitações físicas, tenham acesso à informação de forma plena e acessível.

Um dos pilares centrais da Declaração de Lyon é a promoção da acessibilidade informacional, reconhecendo que pessoas com deficiência enfrentam barreiras significativas no acesso ao conhecimento (IFLA, 2014). O documento defende a criação de infraestruturas acessíveis, o uso de tecnologias assistivas, a adaptação de conteúdos digitais e físicos e a capacitação de profissionais da informação para oferecer serviços inclusivos (IFLA, 2014). A Declaração também enfatiza a necessidade de políticas públicas inclusivas, alinhadas com legislações como a Convenção da Organização das Nações Unidas (ONU) sobre os Direitos das Pessoas com Deficiência e a Lei n.º 13.146/2015, garantindo que bibliotecas, arquivos e centros de documentação adotem práticas de acessibilidade. Isso inclui a oferta de recursos em Braille, audiolivros, leitores de tela, sinalização tátil e treinamento para o uso de tecnologias assistivas (IFLA, 2014).

Ao defender a informação como um instrumento essencial para a participação cidadã, a Declaração de Lyon reforça que a acessibilidade deve ser um compromisso global. Sua implementação no contexto brasileiro exige a adoção de estratégias que integrem a acessibilidade digital e física nas políticas de gestão informacional, tornando as bibliotecas verdadeiros espaços de inclusão e democratização do conhecimento (IFLA, 2014).

A Lei n.º 4.084/1962 foi um marco regulatório para a profissão de bibliotecário e a organização dos serviços de biblioteconomia no Brasil. Porém, essa lei não contempla as necessidades das pessoas com deficiência. Com o avanço das legislações sobre direitos das pessoas com deficiência, a inclusão informacional passou a ser um tema central na atuação do profissional bibliotecário. Essa análise encontra respaldo nos pressupostos de Tarapanoff, Suaiden e Oliveira (2002), que atribuem ao bibliotecário um papel social de mediação da cidadania informacional.

Em seu art. 2º, a Lei estabelece que é competência do bibliotecário organizar, dirigir e executar serviços de biblioteca e documentação (Brasil, 1962). Esse dispositivo pode ser interpretado para incluir a responsabilidade do profissional na promoção da acessibilidade informacional, garantindo que acervos e espaços sejam adequados para o atendimento a usuários com diferentes necessidades. O art. 3º, que trata da organização dos serviços de biblioteca para facilitar o estudo e a pesquisa, pode ser expandido para incluir a necessidade de adaptação dos acervos e recursos digitais para pessoas com deficiência visual, auditiva, física ou intelectual (Brasil, 1962). O bibliotecário, como gestor da informação, deve estar preparado para garantir que seu acervo e serviços sejam acessíveis a todos os seus usuários. O conceito de Desenho Universal também deve ser incorporado na gestão de espaços e recursos bibliográficos, garantindo que mobiliário, tecnologias e recursos digitais sejam adaptáveis para diferentes usuários.

A implementação efetiva da acessibilidade informacional enfrenta desafios significativos, como a falta de recursos, a carência de profissionais capacitados e as barreiras atitudinais são obstáculos que dificultam a concretização desse direito. A garantia da acessibilidade informacional requer um esforço contínuo e articulado entre governo, instituições e sociedade civil, visando à construção de uma sociedade verdadeiramente inclusiva.

No âmbito das bibliotecas, a acessibilidade informacional enfrenta diversos desafios. Um dos principais problemas é a falta de acervo acessível, uma vez que muitas bibliotecas universitárias não possuem materiais adaptados para pessoas com deficiência visual, como livros em Braille, audiolivros e formatos digitais acessíveis. Somam-se a isso a escassez de tecnologias

assistivas, como leitores de tela e *softwares* de reconhecimento de voz, o que reduz a autonomia dos usuários, limitando seu aprendizado e desenvolvimento acadêmico. A exclusão digital também se apresenta como um obstáculo significativo, pois a precariedade dos serviços e a ausência de infraestrutura adequada dificultam o acesso informacional.

Para Gonçalves (2024), a biblioteca deve ser concebida como um espaço de acolhimento, aprendizagem e autonomia, garantindo oportunidades equitativas de acesso à informação para todos os usuários. O autor argumenta que a biblioteca, muitas vezes, permanece distante da realidade dos usuários com necessidades educacionais especiais. Sasaki (2006) defende que a superação das barreiras informacionais requer não apenas capacitação contínua, mas também uma mudança cultural institucional que reconheça a diversidade como valor fundamental. Gonçalves (2024) propõe, de forma conclusiva, que as bibliotecas abandonem a filosofia da posse da informação e adotem a filosofia do acesso, por meio do compartilhamento de recursos e da cooperação interinstitucional.

Diniz e Abreu (2024) reforçam esse entendimento ao proporem a utilização dos quesitos do Censo da Educação Superior (CENSUP) como parâmetros diagnósticos para avaliar a acessibilidade em bibliotecas universitárias. A partir de uma abordagem quantitativa aplicada a 23 bibliotecas do Instituto Federal do Maranhão, as autoras demonstram que as lacunas na promoção da acessibilidade não decorrem apenas da ausência de recursos financeiros, mas da inexistência de políticas institucionais sistematizadas. A pesquisa aponta a necessidade de articulação entre bibliotecas e Núcleos de Apoio às Pessoas com Necessidades Educacionais Específicas (NAPNE), reafirmando o papel das bibliotecas como espaços de cidadania, nos quais a acessibilidade deve ser princípio estruturante.

A exclusão social e educacional gerada pela falta de acessibilidade digital resulta em desigualdade de oportunidades, limitando a participação plena desses estudantes na vida acadêmica e social. No aspecto profissional, a restrição ao acesso a cursos *on-line* e plataformas de aprendizado afeta a formação dos usuários, comprometendo sua inserção no mercado de trabalho, onde a proficiência digital é cada vez mais necessária (Fernandes, 2018).

A conscientização sobre a acessibilidade nas bibliotecas tem aumentado, tornando essencial a implantação de dispositivos que garantam a efetivação de políticas públicas amparadas pelo marco legal de inclusão da pessoa com deficiência. Além disso, é essencial direcionar investimentos para a aquisição de tecnologias assistivas, garantindo a disponibilização de *softwares* de leitura de tela e dispositivos adaptados (Fernandes, 2018).

Para mitigar esses desafios, algumas soluções e recomendações devem ser implementadas. As bibliotecas universitárias precisam adotar políticas institucionais que promovam a inclusão informacional, investindo na capacitação de profissionais para atender às necessidades específicas dos estudantes com deficiência.

2.2 A semiótica da imagem

A imagem faz parte da nossa forma de compreender e de nos relacionarmos com o mundo. Ela não é somente aquilo que se vê, mas aquilo que se interpreta, sente e articula com experiências anteriores. Corrêa (2021) destaca que a imagem deve ser entendida como uma linguagem sensível, acessível por múltiplas vias perceptivas, e não restrita ao olhar. Para a autora, a imagem pode ser transposta para outras linguagens — como a verbal — desde que essa mediação preserve as suas funções comunicativas.

Compreender a imagem como linguagem exige retomar fundamentos da semiótica. Saussure (1916) estabelece que o signo resulta da articulação entre significante e significado, relação que só produz sentido dentro de um sistema. Embora tenha privilegiado o signo verbal, suas proposições evidenciam que nenhum elemento comunicativo possui sentido isolado. Isso implica que, quando a imagem não pode ser percebida visualmente, é necessário reconstruir verbalmente o seu significante para preservar o funcionamento do signo. A audiodescrição torna-se, assim, uma forma de tradução semiótica, em que a linguagem verbal procura restituir à pessoa com deficiência visual a estrutura significativa originalmente acessível pela visão. Como sintetiza Santaella (2001, p. 59), “[...] o significado de um signo é outro signo [...] criado na mente pelo signo anterior”, reforçando a ideia de que toda experiência semiótica implica transposição e reconstrução.

Peirce (2005) classifica as imagens como hipoícones da Primeiridade, a categoria das qualidades sensíveis que se manifestam à percepção. Cor, forma, textura e luminosidade, por exemplo, integram esse domínio qualitativo que, embora pareça imediato, é moldado por convenções representacionais. Pinturas, fotografias e ilustrações não são meras reproduções do real: integram códigos culturais de composição e estilo. Quando essas qualidades não são acessíveis pelo olhar, a audiodescrição precisa transpor para a linguagem verbal aquilo que caracteriza a imagem como ícone. Santaella (2001) reforça essa ideia ao afirmar que “uma imagem é um hipoícone porque a qualidade de sua aparência é semelhante à qualidade da aparência do objeto que a imagem representa”. O reconhecimento visual, portanto, não depende apenas da identificação de formas, mas de uma operação semiótica que articula percepção, memória e convenções culturais. A descrição torna-se, nesse sentido, uma mediação entre os modos sensoriais e simbólicos, convertendo qualidades visuais — próprias da Primeiridade³ — em signos verbais capazes de desempenhar função equivalente (Peirce, 2005). Garantir essa equivalência é fundamental para que o usuário com deficiência visual participe de forma autônoma da experiência simbólica que a imagem produz.

Lyons (1987) amplia essa reflexão ao entender a linguagem como sistema simbólico capaz de representar realidades ausentes. Para o autor, a linguagem verbal permite a reconstrução mental de elementos que não estão disponíveis à percepção. Essa concepção é central para a audiodescrição: quando a visão é impossibilitada, o texto descritivo deve funcionar como meio para recompor a cena, não somente enumerando objetos, mas articulando sentidos, relações espaciais e elementos narrativos. Descrever uma imagem, portanto, não consiste em apenas traduzir qualidades visuais para palavras, mas construir um signo verbal que permita ao interlocutor acessar a narrativa visual de forma inteligível e autônoma.

Ao discutir o papel da imagem, Zimmermann (2016) mostra que ela pode funcionar como representação do real, como símbolo ou como signo. Em outras

³ Primeiridade, refere-se à categoria de experiência que antecede qualquer relação ou comparação, representando a qualidade pura e imediata de uma coisa em si mesma, anterior a qualquer mediação ou interpretação (Pierce, 2005).

palavras, a imagem pode tentar mostrar o mundo como ele é, pode representar ideias mais abstratas, ou pode apontar para algo que está fora dela. A imagem também envolve interpretação. Ver não é somente enxergar, é dar sentido ao que está sendo visto, considerando fatores como cultura, memória, vivência e contexto. Uma ideia importante apresentada por esse autor é a diferença entre duas formas de entender a imagem: uma abordagem que considera que é preciso associar a imagem a outros conhecimentos (como a memória, por exemplo), e outra que acredita que o próprio olhar já dá conta de compreender o que está sendo visto. O autor destaca que essas duas formas não são opostas, mas sim complementares (Zimmermann, 2016).

Barthes (1984), ao analisar a fotografia, propõe uma diferença entre o que ele chama de *studium* e *punctum*. O *studium* está relacionado ao conteúdo geral da imagem, aquilo que pode ser explicado culturalmente. Já o *punctum* é aquele detalhe que chama atenção de forma mais emocional, que toca o espectador de maneira pessoal, mesmo que o fotógrafo não tenha planejado isso. Essa ideia mostra como as imagens podem ter sentidos diferentes para cada pessoa, dependendo do que se percebe e de como se interpreta.

A imagem pode sugerir narrativas, ou seja, contar histórias, o que acontece quando ela traz elementos que remetem ao tempo e ao espaço, como personagens, cenários ou acontecimentos (Zimmermann, 2016). Quando pensamos em acessibilidade, é importante considerar que a descrição de uma imagem precisa também buscar transmitir esse sentido narrativo, e não somente listar elementos visuais. Por fim, a imagem não é apenas aquilo que se vê, mas também aquilo que se sente, se lembra e se interpreta. Torná-la acessível não é apenas uma tarefa técnica, mas um compromisso com a inclusão.

Como mostram Corrêa (2021) e os autores discutidos por Zimmermann (2016), descrever uma imagem é também um exercício de linguagem, de sensibilidade e de cuidado com quem vai recebê-la.

A imagem, conforme analisa Moura (2021), deve ser compreendida como um meio simbólico de sobrevivência e continuidade da existência humana diante da finitude biológica. Ela não se limita a uma mera representação visual, mas atua como um verdadeiro “segundo corpo”, que ocupa o espaço simbólico da cultura e prolonga a presença de quem já não está fisicamente. A imagem se

constitui, assim, como extensão da corporeidade, operando como suporte da memória, da identidade e da permanência.

Segundo Moura (2021), essa duplicação simbólica não é uma simples cópia, mas uma forma de persistência: “um corpo que não se apaga por completo, pois resta uma aparência que o simula e o substitui”. A autora compreende a imagem como algo que transcende o suporte material e adentra a esfera ritual, emocional e social, tornando-se uma presença ativa mesmo na ausência do corpo físico.

Essa autora (2021) argumenta que o desejo humano de permanência motiva a produção imagética desde a Antiguidade, e essa motivação permanece na contemporaneidade sob novas formas, se configurando no contexto atual a imagem como um instrumento de resistência à morte, uma tentativa de escapar da assimetria⁴ imposta pela finitude da vida, apontando que esse movimento se ancora em uma dinâmica cultural profunda, na qual a imagem não é somente reflexo da vida, mas estratégia simbólica para contorná-la. Como ela afirma, “há uma tentativa constante de que o corpo não desapareça com a morte, mas se prolongue na imagem” (Moura, 2021).

Contudo, a autora (2021) também reflete criticamente sobre os limites dessa lógica no presente marcado-pela reprodutibilidade técnica, a imagem do corpo passa a ser intensamente estetizada, mediada e mercantilizada. No caso de figuras públicas, como artistas e celebridades, há uma transformação de seus corpos em imagens públicas que sobrevivem à existência biológica, mas esvaziam a individualidade. O corpo, nesse processo, torna-se simulacro: uma presença que se mantém visível, mas desvinculada da organicidade e da experiência vivida. Moura sintetiza: “a imagem encarna a ausência; é corpo que fala sem corpo, é presença de quem já não está” (Moura, 2021).

A autora ainda observa que, na contemporaneidade, esse desejo de permanência é alimentado por dispositivos tecnológicos e midiáticos que

⁴ A assimetria, designa o desequilíbrio fundamental entre vida e morte: a morte é mais forte que a vida na Primeira Realidade (biológica). Apenas através da Segunda Realidade (simbólica e imagética) é possível superar essa assimetria, permitindo que a imagem funcione como resistência à finitude Moura (2021).

possibilitam não somente a manutenção da imagem, mas sua repetição, multiplicação e circulação (Moura, 2021).

Nessa perspectiva integrada dos autores citados, a audiodescrição se configura como prática que combina teoria da linguagem, semiótica e acessibilidade. Ela não se reduz a uma técnica de descrição, mas constitui mediação simbólica que busca garantir o direito de participação plena na experiência cultural, acadêmica e informacional. É essa compreensão que fundamenta a discussão sobre a necessidade de audiodescrever capas de livros em bibliotecas, no contexto universitário, e orienta a análise proposta nesta pesquisa.

2.3 A capa de livro como elemento descritivo em bibliotecas

A capa de um livro, embora historicamente vinculada ao design gráfico e à função material de proteção do volume impresso, constitui hoje um componente essencial da mediação simbólica entre o leitor e a obra. Autores como Corrêa (2021), Sehn (2009) e Carvalho (2008) demonstram que a capa opera como linguagem visual carregada de sentidos, articulando elementos gráficos e semióticos que produzem expectativas, ativam memórias e orientam interpretações. A imagem da capa, portanto, não se reduz à sua aparência; ela se apresenta como signo cultural, afetivo e narrativo. Sob essa perspectiva, sua acessibilidade se torna elemento central em bibliotecas inclusivas para usuários com deficiência visual. Corrêa (2021) chama atenção para o fato de que a imagem não se reduz ao que é visto, mas envolve camadas interpretativas que precisam ser preservadas quando transpostas para outras linguagens — o que é relevante no caso da audiodescrição para pessoas com deficiência visual.

Nas bibliotecas, esse caráter comunicativo da capa assume importância ainda maior. Sehn (2009) observa que, mesmo em contextos em que o conteúdo do livro é privilegiado — como nas prateleiras onde apenas as lombadas ficam expostas — a capa continua exercendo uma função simbólica relevante em ações de mediação, como exposições, mostras e vitrines temáticas. A autora destaca que o primeiro contato do leitor com o livro é sensorial e que, nesse momento, a imagem da capa atua como vetor de percepção e julgamento, ativando expectativas, afetos e associações cognitivas (Sehn, 2009). Ignorar

essa dimensão em práticas bibliotecárias inclusivas é desconsiderar uma etapa fundamental da experiência leitora.

Carvalho (2008) amplia essa discussão ao demonstrar que a capa evoluiu historicamente de um mero invólucro protetor para um artefato visual dotado de valor simbólico, comercial e cultural. Sua análise revela que a capa condensa elementos visuais — como cor, tipografia, imagem e composição — que atuam conjuntamente para sugerir atmosferas, indicar gêneros e estabelecer uma primeira ponte entre leitor e narrativa. Sobre tal, a autora comenta:

A capa de livro constitui um formato com características muito particulares no contexto do design gráfico, uma vez que reúne em si um conjunto diversificado de funções. Surge como um mecanismo fundamentalmente prático, com o fim de proteger o miolo do livro, e que rapidamente ao longo da história acumula outros propósitos. [...] Enquanto face visível do livro, a capa assume um papel privilegiado na comunicação com o público e, conseqüentemente, constitui um veículo privilegiado de promoção comercial. Comparativamente a outros formatos que lhe são próximos, como a capa de jornal ou de revista, a capa de livro tem uma maior longevidade e presença, fazendo com que muitas se tornem símbolos marcantes, associadas a um texto e a um período histórico (Carvalho, 2008, p. 13).

Essa lógica, quando aplicada ao contexto da acessibilidade, revela uma lacuna importante: a ausência de recursos que permitam ao público cego ou com baixa visão acessar essas camadas visuais de sentido. Nesse ponto, a audiodescrição se impõe não como complemento opcional, mas como condição de equidade comunicativa (Carvalho, 2008).

Silveira (2006) destaca que o primeiro contato do leitor com o livro se dá antes mesmo da leitura do texto — é um contato sensorial, visual e tátil, que inicia a experiência de leitura. Essa primeira percepção é estética e simbólica, sendo a capa um convite visual que estimula a imaginação e antecipa atmosferas do conteúdo, como frisa:

A capa de um livro é a primeira responsável por estabelecer a ligação entre o mundo exterior, do leitor, e o mundo interior, da narrativa. Ela funciona como uma ponte visual que se vale de referências simbólicas para expressar, comunicar e convidar o leitor a adentrar o mundo da leitura e os caminhos daquele texto. Uma boa capa deve sugerir sem

entregar, insinuar sem explicar. Sua função é expressar esteticamente algum ponto central da narrativa ou apenas sugerir uma ambientação que estimule a criatividade do leitor e desperte nele o desejo de partilhar e interagir com o universo de caminhos aberto pelo autor (Silveira, 2006, p. 41).

A autora (2006) enfatiza que o *design* da capa, ao sintetizar aspectos temáticos da narrativa por meio de elementos gráficos e simbólicos, carrega intencionalidades comunicativas que, quando não acessíveis a todos, criam uma barreira na experiência leitora. Assim, no contexto das bibliotecas, onde o acesso democrático ao livro deve ser garantido, a presença da imagem na capa não pode ser negligenciada em sua dimensão comunicacional para leitores com deficiência visual.

Ao considerar o livro em sua dimensão sensorial e comunicacional, Silveira (2006) defende que a leitura de imagens deve ser acessível e significativa. Nesse sentido, a audiodescrição das capas se apresenta como extensão legítima da mediação literária.

Freitas (2014) contribui ao reforçar o papel estratégico da capa entre leitores adolescentes, cuja decisão de leitura muitas vezes se ancora em estímulos visuais. A autora argumenta que a leitura da capa opera em três níveis — sensorial, emocional e racional — e que esses níveis precisam ser considerados na criação de descrições acessíveis. A audiodescrição da capa, portanto, não deve se limitar à enumeração de elementos visuais, mas buscar reconstruir, por meio da linguagem, a narrativa visual que a capa sugere. Em bibliotecas escolares e públicas, isso significa garantir que todos os leitores, independentemente de suas condições sensoriais, tenham acesso ao mesmo convite simbólico à leitura.

No campo da linguagem visual, a cor é um dos elementos mais potentes na criação de sentidos, como discutem Lima e Pereira (2019). Estas autoras demonstram que os matizes presentes nas capas podem induzir interpretações, evocar emoções e sugerir temáticas específicas. No entanto, essas informações são inacessíveis ao leitor com deficiência visual, o que reforça a importância de uma mediação que traduza não somente o “que” está na capa, mas também “como” esses elementos operam simbolicamente (Lima e Pereira, 2019).

A ausência de mediação acessível nas bibliotecas expõe a centralidade da visualidade nos processos de design editorial e curadoria do acervo. Nesse contexto, Araújo (2011) oferece uma perspectiva fundamental ao propor um modelo de catálogo bibliográfico orientado ao usuário. Para o autor, o catálogo deve deixar de ser um artefato fechado, centrado na lógica interna do sistema, e passar a responder às necessidades práticas, cognitivas e sensoriais dos leitores. Isso significa, por exemplo, incorporar descrições acessíveis das capas nos registros bibliográficos — estratégia que vai ao encontro da defesa de uma biblioteca inclusiva e sensível à pluralidade perceptiva dos seus usuários (Araújo, 2011).

Araújo (2011) também defende a adoção de ontologias e modelos conceituais que possibilitem maior plasticidade e interoperabilidade dos sistemas de informação. Essa abordagem amplia o campo de ação da audiodescrição ao reconhecê-la como elemento legítimo de representação da informação. Ao transformar a descrição imagética em conteúdo integrado ao catálogo, as bibliotecas passam a operar com uma lógica de inclusão sensorial, na qual o simbólico não se restringe à visualidade, mas se expande para o verbal, o tátil e o sonoro. Essa transposição exige não apenas recursos técnicos, mas sensibilidade linguística e compromisso ético com o direito à leitura plena (Araújo, 2011).

Em síntese, a orientação ao usuário proposta por Araújo (2011) exige do bibliotecário uma escuta sensível às múltiplas formas de interação com o acervo. Essa postura acolhe também as subjetividades sensoriais, reconhecendo que a leitura se inicia, muitas vezes, pela imagem da capa — mesmo que essa imagem precise ser reconstruída em palavras. Assim, a audiodescrição das capas nos catálogos não é somente um recurso técnico, mas uma mediação simbólica que humaniza o acesso à biblioteca e amplia sua função social.

2.4 Audiodescrição: princípios, normas e práticas

A audiodescrição é um recurso de acessibilidade comunicacional que transforma elementos visuais em linguagem verbal, permitindo que pessoas com deficiência visual compreendam conteúdos baseados em imagens estáticas ou dinâmicas. No Brasil, sua regulamentação é estabelecida principalmente pelas

normas ABNT NBR 16452:2016, ABNT NBR 15290:2005 e, de modo complementar, pela ABNT NBR 9050:2020.

Publicada em 2016, a ABNT NBR 16452 estabelece diretrizes para a elaboração, estruturação e aplicação da audiodescrição em diferentes contextos, com base nos princípios do desenho universal. Seu escopo define que: “Esta Norma fornece diretrizes para a produção da audiodescrição. [...] visam favorecer a percepção, a compreensão e a fruição das informações contidas em imagens dinâmicas ou estáticas, para pessoas impossibilitadas de ver ou com dificuldade para compreender tais imagens” (Associação ..., 2016).

A norma orienta que a descrição deve ser feita de maneira objetiva, concisa e sem interpretações subjetivas, utilizando sempre o tempo presente do modo indicativo. Ela também define responsabilidades técnicas dos profissionais envolvidos, como o audiodescritor roteirista, o audiodescritor narrador e o audiodescritor consultor (geralmente uma pessoa com deficiência visual), assegurando qualidade e validação da acessibilidade.

O roteiro descritivo deve conter informações como a identificação de personagens e elementos visuais relevantes (nomes, papéis, relações), descrição de aparência física, vestuário, expressões, gestos e maneirismos, caracterização do cenário e da ambientação, descrição de sons não identificáveis contextual ou logicamente, sempre com linguagem clara, evitando juízos de valor, metáforas ou adjetivações subjetivas.

A norma ainda define parâmetros técnicos para audiodescrição em diferentes suportes, como televisão, cinema, espetáculos ao vivo, exposições, vídeos, eventos acadêmicos e plataformas digitais.

A ABNT NBR 15290:2005, por sua vez, antecede a 16452 e trata especificamente da acessibilidade na programação televisiva, considerando múltiplos recursos, como audiodescrição, legenda oculta (*closed caption*) e janela de Libras. A norma afirma: “Para ser considerada acessível, a programação televisiva deve atender ao disposto nesta Norma. [...] visa possibilitar às pessoas cegas ou com baixa visão o acesso às mensagens transmitidas de forma essencialmente visual” (Associação ..., 2005).

A norma também estabelece diretrizes para a inserção da audiodescrição no canal de áudio secundário (SAP), permitindo que o usuário o ative conforme sua necessidade. Os elementos descritos devem ser sincronizados com o conteúdo original, sem interferir na trilha sonora ou nos diálogos. Além disso, define o conceito de audiodescrição como “narração descritiva em voz de sons e elementos visuais-chave — movimentos, vestuário, gestos, expressões faciais, mudanças de cena, textos e imagens [...] desapercibidos ou incompreensíveis sem o uso da visão” (Associação ..., 2005).

Embora voltada à acessibilidade física e arquitetônica, a ABNT NBR 9050:2020 reforça a definição de acessibilidade como “possibilidade e condição de alcance, percepção e entendimento para utilização, com segurança e autonomia, de espaços, mobiliários, equipamentos urbanos, edificações, transportes, informação e comunicação” (Associação ..., 2015). Esse conceito embasa todas as práticas normativas voltadas à comunicação acessível, integrando recursos como audiodescrição, Libras, legendagem e sinalização tátil, em consonância com a Lei n.º 13.146/2015 e o Decreto nº 5.296/2004 (Brasil, 2025 e 2004).

A audiodescrição, ao ser normatizada pelas diretrizes da ABNT NBR 16452:2016 e ABNT NBR 15290:2005, consolida-se como uma tecnologia assistiva indispensável à efetivação dos direitos das pessoas com deficiência visual. Sua aplicação deve ser orientada por critérios técnicos, éticos e linguísticos rigorosos, promovendo o acesso equitativo à informação, à cultura e à cidadania. A padronização desses critérios representa um avanço importante para a comunicação inclusiva, ampliando o compromisso de instituições, meios de comunicação e profissionais com a acessibilidade comunicacional de qualidade (Associação ..., 2016 e 2005).

A ABNT NBR 17225:2025, publicada em março de 2025, representa um marco importante na normalização da acessibilidade digital no Brasil, ao estabelecer requisitos específicos para o desenvolvimento de conteúdo e aplicações web acessíveis (Associação ..., 2025). A norma foi elaborada com base na WCAG 2.2, do *World Wide Web Consortium* (W3C), para eliminar ou mitigar barreiras de acesso aos ambientes virtuais por pessoas com deficiência, ou com outras necessidades de acessibilidade, sejam elas situacionais,

temporárias ou permanentes (ABNT, 2025). Nesse sentido, a norma nacional incorpora critérios técnicos alinhados aos níveis de conformidade internacionalmente reconhecidos pelo W3C adaptando-os ao contexto jurídico e social brasileiro, conforme a Lei n.º 13.146/2015 determina (Brasil, 20215).

Além de seguir os princípios fundamentais da WCAG 2.2, a NBR 17225:2025 propõe critérios específicos relacionados à realidade brasileira, como a obrigatoriedade de recursos de audiodescrição para vídeos e a ênfase na linguagem simples como ferramenta de inclusão (Associação ..., 2025). A norma também aborda com profundidade aspectos técnicos da acessibilidade, como a interação por teclado, contraste de cores, uso de cabeçalhos semânticos, navegação consistente e acessibilidade de formulários, demonstrando uma preocupação com diferentes dimensões da experiência do usuário. Com isso, a norma não apenas orienta os desenvolvedores, mas estabelece um novo patamar de qualidade e equidade no acesso à informação digital no Brasil (Associação ..., 2025).

A WCAG 2.2, publicada pelo W3C em 2023 e traduzida oficialmente para o português em 2025, estabelece um conjunto de recomendações técnicas para tornar o conteúdo web mais acessível a pessoas com diferentes tipos de deficiência. Essa diretriz é estruturada em quatro princípios fundamentais: perceptível, operável, compreensível e robusto, os quais orientam a criação de conteúdos digitais que possam ser utilizados com autonomia e segurança por todos os usuários (W3C, 2023). A WCAG 2.2 organiza seus requisitos em três níveis de conformidade: o nível A, que representa os requisitos mínimos para acessibilidade; o nível AA, mais exigente e geralmente adotado por legislações e políticas públicas; e o nível AAA, que contempla o mais alto grau de acessibilidade, oferecendo uma experiência mais completa para usuários com múltiplas deficiências (W3C, 2023).

Entre os novos critérios introduzidos pela WCAG 2.2, destacam-se diretrizes como Entrada Redundante (3.3.7), Ajuda Consistente (3.2.6), Localização do Foco (2.4.11), Aparência do Foco (2.4.13), Movimentos de Arrastar (2.5.7) e Tamanho do Alvo (Mínimo) (2.5.8) (W3C, 2023). Essas adições evidenciam o esforço do W3C em ampliar a abrangência das diretrizes para

grupos que anteriormente encontravam maiores barreiras no ambiente digital, sobretudo em interfaces complexas e formulários.

No entanto, a adoção e implementação das WCAG 2.2 ainda enfrentam desafios significativos. Um dos principais obstáculos está na tendência de empresas e instituições públicas focarem apenas nos requisitos de nível AA, considerados o “mínimo legal”, em detrimento das diretrizes de nível AAA, que poderiam assegurar uma experiência digital realmente inclusiva. Além disso, o próprio W3C reconhece que as diretrizes ainda são limitadas no que diz respeito à acessibilidade de pessoas com deficiências cognitivas, de linguagem e de aprendizagem, indicando a necessidade de futuras revisões para contemplar esses públicos de forma mais eficaz (W3C, 2023). Dessa forma, embora a WCAG 2.2 represente um avanço técnico importante, ela não substitui a necessidade de normas nacionais contextualizadas, como a ABNT NBR 17225:2025, que traduzem esses princípios em exigências mais aplicáveis à realidade local.

A audiodescrição é tratada de forma central na WCAG 2.2, na Diretriz 1.2 — Mídias com base em tempo, que estabelece requisitos específicos para tornar conteúdos audiovisuais acessíveis. A norma define audiodescrição como um recurso que fornece, em forma de narração, uma descrição das informações visuais importantes de um vídeo, permitindo que pessoas com deficiência visual compreendam o conteúdo de maneira equivalente. O objetivo principal é garantir a equivalência perceptiva, ou seja, que elementos visuais como ações, expressões, cenários e textos na tela sejam compreensíveis também por meio auditivo.

No nível A de conformidade, a WCAG 2.2 estabelece, por meio do Critério de Sucesso 1.2.3 – Audiodescrição ou Mídia Alternativa (Pré-gravada), que conteúdos de vídeos pré-gravados devem oferecer uma das seguintes opções: audiodescrição sincronizada ou uma mídia alternativa (como transcrições detalhadas), garantindo o acesso ao conteúdo visual por pessoas cegas ou com baixa visão. Já no nível AA, o critério 1.2.5 – Audiodescrição (Pré-gravada) torna obrigatória a presença de audiodescrição sincronizada, elevando o padrão de acessibilidade ao incorporar esse recurso ao fluxo audiovisual. Isso demonstra

um avanço técnico importante na promoção de uma web mais inclusiva, para conteúdos educacionais, institucionais e governamentais.

No nível AAA, de mais alto grau de exigência, os critérios da WCAG 2.2 vão ainda mais longe ao exigir audiodescrição estendida (1.2.7), destinada a vídeos cujo ritmo ou quantidade de informação visual não permite que a descrição seja incluída nos espaços de pausa natural do áudio original. Nesses casos, a narrativa precisa ser pausada para inserir as descrições detalhadas, assegurando a plena compreensão do conteúdo. Essa abordagem revela uma preocupação com a complexidade semântica dos vídeos e com a experiência do usuário na totalidade, embora sua adoção prática ainda seja rara fora de contextos institucionais específicos. A própria norma reconhece que nem todos os conteúdos acessíveis estão ao alcance de todas as deficiências, o que reforça a importância de aplicar múltiplas estratégias de acessibilidade para atingir diferentes públicos (W3C, 2023).

Corrêa (2021) enfatiza que a audiodescrição deve ser estruturada de forma clara e objetiva, respeitando a experiência do usuário e evitando sobreposição com diálogos ou elementos sonoros essenciais.

No contexto educacional, a audiodescrição tem sido cada vez mais reconhecida como um instrumento pedagógico para promover a aprendizagem de estudantes com deficiência visual. Ribeiro (2022) aponta que o uso de audiodescrição em livros didáticos, vídeos educativos e apresentações acadêmicas é essencial para garantir a equiparação de oportunidades no ambiente escolar. A ausência desse recurso em materiais pedagógicos compromete a autonomia dos estudantes e reforça barreiras no acesso ao conhecimento.

Além disso, a audiodescrição não se restringe somente ao ambiente educacional, sendo aplicada em produções audiovisuais, espetáculos teatrais, exposições e museus. A Lei n.º 13.146/2015 determina a obrigatoriedade desse recurso em conteúdos televisivos e cinematográficos, contribuindo para a democratização do acesso à cultura e ao entretenimento (Brasil, 2015). Fernandes (2023) destaca que, apesar dos avanços legais, a implementação da audiodescrição ainda enfrenta desafios, como a falta de profissionais capacitados e a resistência de algumas produtoras em adotar o recurso.

A audiodescrição também se beneficia dos avanços tecnológicos no campo da IA. Corrêa (2021) discute como ferramentas de *computer vision* e aprendizado de máquina desenvolveram-se para automatizar a audiodescrição, permitindo que algoritmos analisem imagens e gerem descrições automaticamente. No entanto, a qualidade da audiodescrição automatizada ainda não se equipara à produção humana, que exige sensibilidade e interpretação contextual, como comenta Correa (2021, p. 17):

Pessoas com deficiência visual utilizam screen readers para escutar o conteúdo de uma tela e essas tecnologias assistivas podem dar às pessoas acesso à web que de outra forma não teriam sido capazes de se envolver. Mas eles só podem auxiliar as pessoas até certo ponto. [...] o conteúdo visual dessas plataformas somente será totalmente acessível com descrições de imagem e *alt-text* apropriados.

A formação de audiodescritores é um aspecto essencial para garantir a qualidade desse serviço. A ABNT NBR 15290:2016 estabelece diretrizes sobre a clareza da narração, fidelidade ao conteúdo original, uso de terminologia apropriada e técnicas de sincronização para que a audiodescrição seja eficiente e compreensível para o público-alvo.

Por fim, a audiodescrição é um recurso fundamental para garantir a acessibilidade comunicacional e a inclusão de pessoas com deficiência visual na sociedade. Seu uso em diferentes contextos, desde a educação até o entretenimento, contribui para a construção de um ambiente mais inclusivo e equitativo. No entanto, desafios como a resistência de algumas instituições, a falta de investimentos e a necessidade de maior fiscalização do cumprimento das normas ainda precisam ser superados para que a audiodescrição se torne uma realidade efetiva em todas as esferas sociais.

2.5 Inteligência Artificial e audiodescrição automatizada

A IA constitui um campo multidisciplinar cuja proposta central é o desenvolvimento de sistemas capazes de simular a inteligência humana. Tais sistemas são projetados para aprender com experiências, adaptar-se a novos estímulos e executar tarefas que tradicionalmente exigem raciocínio humano (Dourado, 2023). Embora o termo "inteligência artificial" tenha sido utilizado

durante a Conferência de Dartmouth, em 1956, sua conceituação permanece em constante evolução (Bringsjord e Govindarajulu, 2022). Russell e Norvig (2013) definem IA como "a ciência e a engenharia de produzir máquinas inteligentes", destacando que o campo busca compreender e construir agentes que percebam seu ambiente e realizem ações de forma autônoma e racional (Russell e Norvig, 2013).

Um marco histórico importante no delineamento da IA foi o trabalho de Alan Turing (1950), que propôs o "Teste de Turing" como critério operacional para avaliar a inteligência de uma máquina (Dourado, 2023). Segundo Turing, uma máquina pode ser considerada inteligente se for capaz de manter uma conversa indistinguível da de um ser humano. A proposta implicava, portanto, que um sistema computacional deveria dominar habilidades como o processamento de linguagem natural, raciocínio automatizado, representação de conhecimento e aprendizado de máquina (Russell e Norvig, 2013).

Apesar de sua consolidação como área de pesquisa e desenvolvimento, não há consenso universal sobre a definição de IA, o que representa um desafio considerável para sua regulamentação e para a formulação de políticas públicas voltadas à tecnologia (Renda e Engler, 2023). A classificação da IA pode seguir distintos critérios, incluindo a comparação com capacidades humanas, o que distingue a chamada IA Estreita (ou fraca), voltada a tarefas específicas, da IA Geral, que simularia habilidades cognitivas humanas amplas, ou ainda conforme sua base funcional, sendo dividida em IA Simbólica, que utiliza regras explícitas de representação lógica, e IA Conexionista, que se baseia em redes neurais artificiais (Dourado, 2023).

Entre os avanços mais significativos da IA nas últimas décadas está o aprendizado de máquina (*machine learning*, ML), que Russell e Norvig (2013) tratam como um dos motores fundamentais da IA moderna. Ao contrário da programação tradicional, baseada em regras fixas, o ML permite que sistemas aprendam a partir de dados, encontrando padrões e fazendo previsões sem que todas as regras sejam explicitamente programadas (Dourado, 2023). Para os autores, "o papel da aprendizagem como uma extensão do alcance do projetista em ambientes desconhecidos [...] favorece a representação explícita do conhecimento e do raciocínio" (Russell e Norvig, 2013).

O processo de generalização, conceituado pela capacidade de aplicar o conhecimento aprendido a novos dados, é uma característica central do ML. Essa abordagem se mostra eficaz em contextos nos quais há grande volume de dados e necessidade de adaptação constante (Dourado, 2023). Nesse sentido, o aprendizado de máquina transforma dados e experiências passadas em modelos matemáticos, possibilitando a criação de sistemas mais flexíveis e autônomos (Russell e Norvig, 2013).

Um subtipo do aprendizado de máquina é o aprendizado profundo (*deep learning*, DL), cuja principal característica é o uso de redes neurais artificiais compostas por múltiplas camadas (Dourado, 2023). Tais redes imitam, de maneira simplificada, o funcionamento das conexões neurais do cérebro humano, possibilitando o processamento paralelo e hierárquico de informações (Dourado, 2023). Russell e Norvig (2013) destacam que redes neurais artificiais consistem em modelos computacionais capazes de realizar generalizações a partir de dados, sendo eficazes em tarefas como reconhecimento de padrões, classificação e previsão. O DL levou essa abordagem a um novo patamar ao permitir o processamento de grandes volumes de dados com mínima intervenção humana, sobretudo por meio de técnicas como a retropropagação (*backpropagation*), que ajusta os pesos da rede com base no erro da saída prevista (Russell e Norvig, 2013).

Um dos campos de aplicação mais expressivos do aprendizado profundo é o Processamento de Linguagem Natural (*Natural Language Processing*, PLN), área que busca permitir a comunicação entre seres humanos e máquinas por meio da linguagem natural (Dourado, 2023). A compreensão computacional da linguagem humana requer não apenas a análise sintática das frases, mas também a consideração do contexto, da semântica e das intenções subjacentes aos enunciados. Conforme afirmam Russell e Norvig (2013), "a linguagem é um sistema extremamente complexo e ambíguo, cuja interpretação requer tanto conhecimento linguístico quanto extralinguístico" (Russell e Norvig, 2013). Técnicas de PLN são fundamentais para viabilizar a comunicação fluida entre usuários e sistemas baseados em IA, e sua aplicação tem sido cada vez mais comum em assistentes virtuais, mecanismos de busca, tradução automática e, mais recentemente, em modelos de linguagem generativos (Dourado, 2023).

A ascensão dos chamados modelos fundacionais (*foundation models*), como a família GPT (*Generative Pre-trained Transformer*), representa um novo paradigma no desenvolvimento de sistemas baseados em IA (Dourado, 2023). Esses modelos são treinados em grandes conjuntos de dados não estruturados e são capazes de realizar múltiplas tarefas sem a necessidade de reconfigurações específicas, graças à sua habilidade de compreender e gerar linguagem com alto grau de coerência. Sua escalabilidade e versatilidade decorrem do uso intensivo de aprendizado profundo e da adoção de arquiteturas uniformizadas, como os transformadores (*transformers*), que passaram a constituir o núcleo técnico de sistemas como ChatGPT, Google Gemini e Microsoft Copilot (Dourado, 2023).

A "emergência" observada nesses modelos refere-se à sua capacidade de generalizar para tarefas inéditas com base apenas em instruções expressas em linguagem natural. Esse comportamento emergente, não necessariamente programado de forma explícita, deriva da complexidade dos padrões capturados durante o treinamento em larga escala. Além disso, a "homogeneização" no desenvolvimento de modelos fundacionais indica uma convergência de técnicas, arquiteturas e objetivos, o que permite que diversas tarefas sejam realizadas a partir de uma única base modelar, otimizando tempo e recursos computacionais.

Ferramentas como ChatGPT, Google Gemini e Microsoft Copilot exemplificam aplicações comerciais desses modelos, conhecidas como grandes modelos de linguagem (*Large Language Models*, LLM). Esses sistemas não somente compreendem e geram respostas em linguagem natural, como também são capazes de sintetizar, traduzir, explicar e até auxiliar na tomada de decisões. Sua utilização em ambientes clínicos e educacionais tem se expandido, sobretudo por sua aptidão para produzir conteúdo acessível e adaptado ao público-alvo, fomentando, assim, práticas mais inclusivas (Russell e Norvig, 2013).

A arquitetura dos modelos fundacionais representa uma evolução dos modelos de linguagem probabilísticos, que originalmente operavam com base em n-gramas (unigrama, bigrama, trigrama), estruturas estatísticas mais simples e limitadas (Russell e Norvig, 2013). Enquanto esses modelos mais antigos eram eficazes apenas em tarefas localizadas e com vocabulários restritos, os modelos

generativos atuais conseguem inferir significados, construir respostas coerentes e até propor conteúdos inéditos. Nesse sentido, o caráter generativo dos LLM aproxima-se da cognição humana, na medida em que mobiliza associações entre dados prévios para produzir saídas novas, adaptativas e contextualmente relevantes (Russell e Norvig, 2013).

A evolução da IA nas últimas décadas tem sido marcada por avanços não apenas tecnológicos, mas também metodológicos. Um dos marcos mais importantes para a consolidação da área foi a adoção mais rigorosa do método científico, com foco em experimentos comparativos, replicabilidade e benchmarks consistentes. Esse movimento contribuiu para que os resultados obtidos em IA se tornassem mais previsíveis e comparáveis, permitindo avaliar com maior clareza a eficácia de diferentes abordagens (Russell e Norvig, 2013).

Outro fator determinante para essa evolução foi a disponibilidade, a partir dos anos 2000, de grandes volumes de dados coletados da web em forma de textos, imagens e interações humanas. A partir de 2001, tornou-se evidente que o sucesso de muitas aplicações de IA não dependia exclusivamente da sofisticação algorítmica, mas também da escala dos dados disponíveis para treinamento. Como observam Russell e Norvig (2013), “muitas vezes, um algoritmo simples com acesso a um volume massivo de dados supera um algoritmo sofisticado limitado por conjuntos de dados menores”. Esse princípio passou a guiar a formação de modelos fundacionais, baseados em aprendizado profundo, treinados com trilhões de palavras e bilhões de imagens.

A versatilidade desses sistemas se reflete em seus usos principais, entre os quais se destacam a geração de conteúdo, a análise de dados, e a compreensão e tradução de linguagem natural. No campo da geração de texto e imagens, os modelos de DL têm sido empregados para produzir conteúdo inédito com base em dados fornecidos (Loch, 2021). Essa capacidade não se limita a uma função instrumental, mas abre espaço para uma contribuição cognitiva, criativa e até estética, como no caso de aplicações artísticas baseadas em IA generativa. A arquitetura dos modelos fundacionais viabiliza a construção de textos coesos e imagens estilizadas, com considerável grau de originalidade (Loch, 2021).

Em relação à análise de dados, os modelos de aprendizado de máquina e aprendizado profundo têm se consolidado como ferramentas essenciais para a análise preditiva. Essa técnica tem sido aplicada, por exemplo, na área da saúde, com o objetivo de prever desfechos clínicos, identificar fatores de risco e auxiliar na gestão de recursos. O uso de dados em tempo real e históricos, coletados a partir de registros eletrônicos de saúde, permite o desenvolvimento de sistemas que reconhecem padrões e geram insights úteis para profissionais da saúde, gestores e pacientes. Em muitos desses casos, o PLN é empregado para a análise semântica de prontuários clínicos, oferecendo uma interpretação automatizada do conteúdo textual desses documentos.

No que diz respeito à compreensão da linguagem, os modelos de linguagem treinados em larga escala apresentam desempenho elevado em tarefas como identificação de idioma, categorização textual e análise de sentimentos (Russell e Norvig, 2013). Técnicas baseadas em aprendizado supervisionado, combinadas com vetores de características derivados de n-gramas, são utilizadas em classificações automáticas, como a detecção de spam ou a triagem de conteúdo em redes sociais. A precisão alcançada por esses sistemas é notável, chegando a superar 99% em tarefas de identificação de idioma, mesmo com textos curtos. Além disso, a tradução automática, que inicialmente se baseava em modelos estatísticos simples, passou a se beneficiar de modelos de DL que integram contextos linguísticos mais amplos e são capazes de produzir traduções mais naturais e semanticamente adequadas (Russell e Norvig, 2013).

No contexto específico dos grandes modelos de linguagem utilizados nesta pesquisa, dois conceitos técnicos são fundamentais: *tokens* e janela de contexto. Os modelos não processam o texto como palavras isoladas, mas como sequências de *tokens*, unidades mínimas de processamento que podem corresponder a palavras inteiras, partes de palavras ou sinais de pontuação, dependendo do esquema de segmentação adotado. Tanto o texto inserido pela pessoa usuária quanto a resposta gerada pela ferramenta são convertidos em tokens, e cada modelo possui um limite máximo de tokens por interação, que define a chamada janela de contexto. Esse limite interfere na dimensão dos prompts que podem ser utilizados, na quantidade de histórico de conversa que

pode ser mantida e na extensão das audiodescrições geradas, razão pela qual o conceito de token é mencionado na metodologia.

Outro elemento técnico necessário para compreender as possibilidades de aplicação futura desses modelos em bibliotecas é a noção de API (*Application Programming Interface*). Uma API é um conjunto de regras e *endpoints* que permite que sistemas distintos “conversem” entre si de forma automatizada. No caso da IA generativa, a API de um modelo de linguagem possibilita que plataformas de catálogo, repositórios digitais ou sistemas de gerenciamento de bibliotecas enviem capas de livros e *prompts* ao modelo e recebam as descrições produzidas, sem intervenção manual. Embora esta pesquisa utilize somente as interfaces gratuitas acessadas via navegador, a menção às API é importante porque elas constituem o caminho técnico para uma futura integração sistêmica das audiodescrições aos catálogos bibliográficos, o que amplia o potencial de escalabilidade das soluções aqui discutidas.

Portanto, os usos contemporâneos da IA refletem sua progressiva maturidade técnica e metodológica. Desde a geração de linguagem até a previsão de desfechos clínicos, passando pela organização semântica de grandes volumes de dados, a IA se tornou uma ferramenta central em múltiplas áreas do conhecimento, com impactos crescentes em ambientes acadêmicos, culturais e assistenciais (Loch, 2021).

As aplicações da IA em contextos educacionais e culturais têm ganhado crescente relevância em relação à promoção da acessibilidade e da inclusão social (Loch, 2021). A convergência entre IA e tecnologias digitais tem potencializado a criação de ambientes mais equitativos, permitindo o acesso ampliado à informação, à educação e à cultura por pessoas com deficiência ou mobilidade reduzida. Essa abordagem integra os princípios do desenho universal, que visam o desenvolvimento de produtos, espaços e serviços utilizáveis pelo maior número possível de pessoas, sem a necessidade de adaptação posterior (Loch, 2021).

A IA tem ampliado o alcance das tecnologias assistivas, sobretudo com a introdução de sistemas baseados em aprendizado de máquina e processamento de linguagem natural, que permitem a interpretação de contextos linguísticos e a personalização da experiência do usuário (Meireles, 2024). A capacidade

desses sistemas de aprender com dados reais e adaptar-se a diferentes perfis de uso os torna eficazes em ambientes educacionais, bibliotecas e museus. Em ambientes digitais, por exemplo, ferramentas como o VLibras, mantidas pelo Governo Federal brasileiro, traduzem conteúdos sonoros e textuais para a Língua Brasileira de Sinais, promovendo a inclusão de pessoas surdas na vida acadêmica, profissional e cultural (Meireles, 2024).

De acordo com a legislação brasileira, cabe ao Estado fomentar a pesquisa científica, o desenvolvimento tecnológico e a inovação voltados à inclusão das pessoas com deficiência. Isso inclui a destinação de recursos para programas de apoio à pesquisa em tecnologias assistivas, bem como a capacitação de profissionais da educação para o uso de Libras, Braille e outras ferramentas acessíveis (Brasil, 2015). A presença de diretrizes curriculares que contemplem o desenho universal no ensino superior e na formação técnica reforça a perspectiva de inclusão desde os níveis mais iniciais da formação docente (Brasil, 2015).

Em bibliotecas e acervos digitais, a acessibilidade dos conteúdos é um imperativo ético e legal. A disponibilização de obras em múltiplos formatos acessíveis, inclusive em extensões abertas e não proprietárias, favorece a inclusão de estudantes com deficiência visual, auditiva ou cognitiva (Brasil, 2015). Além disso, o uso da IA para organizar e interpretar acervos favorece a descoberta de conteúdos relevantes de maneira personalizada, contribuindo para uma experiência mais significativa e autônoma (Brasil, 2015).

Em espaços culturais e de preservação histórica, como museus e sítios arqueológicos, as tecnologias baseadas em IA podem desempenhar funções como mediação interativa, personalização da visita e atualização automatizada de informações (Meireles, 2024). Projetos como a “Rota Afro” em Jundiá, por exemplo, utilizam sinalização acessível, intervenções artísticas e roteiros digitais com foco na educação patrimonial, tornando visíveis narrativas historicamente marginalizadas (Meireles, 2024). Essas iniciativas demonstram que a IA, quando utilizada com intencionalidade pedagógica e social, pode fortalecer o acesso à cultura e à memória coletiva.

A IA também é utilizada na produção artística, expandindo os limites da criação estética. Técnicas de DL vêm sendo utilizadas para gerar obras visuais

e sonoras originais, como exemplificam os projetos “*Edmond de Belamy*” e “*The Butcher’s Son*”, que ilustram a colaboração entre artistas e algoritmos (Loch, 2021). No entanto, o debate contemporâneo sobre a arte produzida por IA ressalta que sua legitimidade está associada não somente à capacidade técnica da máquina, mas à construção de narrativas simbólicas, críticas ou sociais que permitam interlocução com os públicos e o contexto cultural. A arte gerada por IA, nesse sentido, pode contribuir para a democratização da cultura, ao desafiar hierarquias tradicionais de autoria e ampliar os meios de criação.

Cabe ressaltar que, apesar dos avanços recentes da IA, sua capacidade de interpretar imagens ainda enfrenta limitações técnicas significativas, sobretudo na identificação precisa de cores, formas, tipografias, sombras e detalhes visuais sutis. Como destaca a Recomendação sobre a Ética da Inteligência Artificial da UNESCO (2021), sistemas automatizados operam a partir de bases de dados e processos algorítmicos que não são transparentes e que podem apresentar lacunas importantes na capacidade de representação. No caso da audiodescrição, tais limitações podem resultar em alucinações, omissões e descrições que não correspondem ao conteúdo visual real, comprometendo a confiabilidade da informação acessível e a própria finalidade da descrição de imagens para usuários cegos ou com baixa visão.

Do ponto de vista ético, a UNESCO (2021) reforça que sistemas de IA não devem comprometer direitos fundamentais, incluindo o acesso equitativo à informação. Quando audiodescrições automatizadas são utilizadas sem revisão humana, há risco de que erros de interpretação visual prejudiquem a compreensão da imagem por pessoas cegas ou com baixa visão, afetando a autonomia e a experiência informacional desses usuários. A UNESCO também enfatiza a necessidade de supervisão humana significativa e de transparência nos processos de produção algorítmica, o que torna indispensável o papel mediador do profissional da informação. Dessa forma, a adoção de IA nas bibliotecas requer políticas de uso responsável e diretrizes que assegurem qualidade, confiabilidade e equidade na descrição e disponibilização de conteúdos visuais.

3 METODOLOGIA

A presente seção descreve a metodologia adotada para o desenvolvimento da pesquisa, explicitando os procedimentos que orientaram desde a definição do objeto até a análise dos resultados, considerando que o objetivo central consiste em investigar a viabilidade do uso de ferramentas de IA para a audiodescrição automatizada de capas de livros em bibliotecas universitárias, a partir da análise dos modelos ChatGPT (OpenAI), Copilot (Microsoft), Gemini (Google) e DeepSeek, com base em um *prompt* elaborado segundo a ABNT NBR 16452:2016 e em consonância com os princípios de acessibilidade da WCAG 2.2

3.1 O objeto da pesquisa, seu universo e amostra

A pesquisa foi elaborada por abordagem qualitativa. O universo da pesquisa é o conjunto de dados relativos às audiodescrições da imagem de capas de livros obtidas por ferramentas de IA. A amostra selecionada compõe-se dos dados das audiodescrições de três capas realizadas pelas quatro ferramentas de AI, conforme a seguir se detalha.

As quatro ferramentas de IA utilizadas são gratuitas, com suporte multimodal, ou seja, capazes de analisar imagens e produzir descrições textuais a partir de comandos em linguagem natural. Embora apresentem interfaces orientadas ao usuário, essas ferramentas são construídas sobre arquiteturas complexas de modelos de linguagem e visão que influenciam o modo como interpretam e descrevem elementos visuais.

Optou-se por utilizar as versões gratuitas das ferramentas e seus modelos disponíveis em novembro de 2025, por escolha metodológica reflete uma perspectiva realista do uso em bibliotecas universitárias, que operam sem assinaturas ditas “*premium*” e dependem de recursos abertos. Assim, permitiram avaliar o potencial da IA em cenários cotidianos e suas limitações práticas, como restrições de processamento, tamanho máximo de entrada e filtros de segurança.

3.2 Descrição da amostra: as ferramentas de IA utilizadas

A amostra desta pesquisa foi composta por ferramentas de inteligência artificial capazes de processar imagens e gerar descrições textuais a partir de seu conteúdo visual. Foram selecionados modelos multimodais amplamente acessíveis ao público, priorizando plataformas que disponibilizam versões gratuitas com suporte ao envio e interpretação de imagens. A escolha dessas ferramentas buscou contemplar diferentes ecossistemas tecnológicos e abordagens de desenvolvimento de modelos de linguagem de grande porte, permitindo observar variações na forma como cada sistema interpreta elementos visuais e os traduz em linguagem natural. Nesse contexto, foram analisadas quatro ferramentas: ChatGPT-5.1 (OpenAI), Gemini 2.5 Flash (Google), Microsoft Copilot e DeepSeek-V3.2. As subseções a seguir apresentam uma descrição técnica de cada modelo, destacando suas características multimodais e sua relevância para a produção automatizada de audiodescrições.

3.2.1 ChatGPT-5.1

Desenvolvido pela OpenAI, o ChatGPT-5.1 é apresentado como um modelo multimodal capaz de processar texto e imagens de forma integrada. Conforme o anúncio técnico oficial da organização, a série GPT-5 trouxe avanços notáveis em compreensão contextual, raciocínio estruturado e capacidades multimodais, utilizando uma arquitetura *transformer* de grande escala treinada em diversas modalidades de dados (OpenAI, 2024).

Este modelo funciona como um sistema unificado de linguagem e visão, onde a análise de imagens ocorre sem a necessidade de módulos externos. Ao receber uma imagem, o modelo converte seus elementos em *embeddings* visuais, processados pelo mesmo mecanismo de atenção empregado para o texto. Essa integração permite ao sistema interpretar a composição gráfica, estruturas visuais, padrões cromáticos, elementos tipográficos e objetos na imagem, sintetizando essas informações em linguagem natural de forma coesa (OpenAI, 2024).

Para esta pesquisa, utilizamos a versão gratuita do ChatGPT-5.1, acessível publicamente pela interface web. Embora esta versão apresente

algumas limitações em volume de contexto e personalização avançada, ela mantém a essência da capacidade multimodal do GPT-5.1. Isso nos permite carregar imagens e obter descrições fundamentadas no conteúdo visual, justificando sua escolha pela acessibilidade e por representar um modelo utilizado por estudantes, bibliotecas e o público em geral (OpenAI, 2024).

Metodologicamente, o ChatGPT-5.1 serve como um importante ponto de referência entre as ferramentas gratuitas comparadas, sendo o modelo mais avançado em nosso escopo. Ele nos oferece uma base sólida para avaliar a precisão e a qualidade estrutural da audiodescrição automatizada.

3.2.2 Gemini 2.5 Flash

O Gemini 2.5 Flash, modelo multimodal de grande porte desenvolvido pelo Google, integra a família Gemini. Sua arquitetura de transformadores foi concebida para operar de maneira integrada com múltiplas modalidades de entrada – texto, imagem, código e áudio – em um sistema unificado. A documentação técnica do Google (2024) descreve o Gemini 2.5 como um modelo que realiza análise, síntese e raciocínio a partir de dados visuais e textuais em um fluxo operacional contínuo (Google, 2024).

O *design* do Gemini 2.5 Flash prioriza o desempenho em larga escala, focando em eficiência computacional, baixa latência e responsividade. Essas características o estabelecem como um "modelo de trabalho" versátil dentro da linha Gemini. As versões públicas oferecem suporte a janelas de contexto extensas, o que é relevante para o processamento de documentos longos, sequências imagéticas complexas e interações contínuas sem comprometer a coerência semântica (Google, 2024).

No processamento de imagens, o Gemini 2.5 Flash integra módulos de visão computacional ao seu núcleo, superando a separação entre codificadores visuais e textuais de arquiteturas mais antigas. As imagens fornecidas são convertidas em *embeddings* visuais, que são então manipulados pelo mesmo mecanismo de atenção usado para o texto. Esse processo permite ao modelo estabelecer relações ricas entre elementos gráficos, composicionais e textuais, facilitando tarefas como legendagem automática, análise estrutural, extração de informações visuais e síntese descritiva, incluindo a produção de

audiodescrições automatizadas, que é o foco central deste estudo (Google, 2024).

Para a pesquisa, utilizamos a versão gratuita, disponível na interface pública do Google Gemini. Esta versão permite o envio direto de imagens, garantindo que o modelo processe o conteúdo visual internamente e gere descrições textuais conforme os parâmetros do *prompt*. Apesar de operar com recursos mais limitados em comparação às versões corporativas ou de API, a versão gratuita mantém a base multimodal do Gemini 2.5 Flash, o que assegura a consistência metodológica para nossa análise comparativa.

3.2.3 Microsoft Copilot

O Microsoft Copilot funciona como um assistente multimodal, integrado ao ecossistema Microsoft e construído sobre os LLM da família GPT, utilizando a infraestrutura Azure. A documentação técnica da Microsoft (2024) indica que o Copilot emprega modelos otimizados para ambientes de produtividade digital, com especial atenção à interpretação de imagens, síntese de informações e processamento de linguagem natural (Microsoft, 2024).

Sua arquitetura segue o paradigma *transformer*, suportando entradas multimodais, como texto e imagem, em um único *pipeline* de análise. Ao receber conteúdo visual, o sistema o converte em representações vetoriais por meio de um módulo interno de visão computacional. Esses *embeddings* são então integrados ao modelo de linguagem principal, capacitando o Copilot a realizar operações diversas, como descrição visual, extração de texto, identificação de objetos, leitura de elementos gráficos e síntese de conteúdos imagéticos (Microsoft, 2024).

Nesta pesquisa, a versão gratuita do Copilot, acessível via navegador, foi a escolhida, sem as integrações avançadas com ferramentas corporativas do Microsoft 365. Esta versão, contudo, oferece recursos multimodais que preservam o núcleo da arquitetura baseada em GPT, permitindo o envio direto de imagens para interpretação. Embora possa ter limitações no tamanho da janela de contexto e na profundidade das análises visuais, ela se mostra adequada para nossos fins experimentais, mantendo a capacidade essencial de transformar estímulos visuais em saídas textuais estruturadas (Microsoft, 2024).

A inclusão do Copilot nesta dissertação é estratégica para a análise comparativa. Ele nos permite verificar o desempenho de um modelo multimodal voltado ao uso cotidiano e à produtividade, diferenciando-se de plataformas que priorizam aplicações de pesquisa. Assim, podemos avaliar como distintos ecossistemas técnicos abordam a interpretação visual na produção automatizada de descrições textuais de capas de livros.

3.2.4 DeepSeek-V3.2

O DeepSeek-V3.2 é um modelo de linguagem multimodal desenvolvido pela DeepSeek AI. De acordo com a documentação oficial, a linha DeepSeek foi concebida para equilibrar desempenho computacional e precisão textual, integrando módulos de linguagem e visão em um sistema unificado, orientado por arquiteturas *transformer* (DeepSeek AI, 2024).

No processamento de imagens, o DeepSeek adota o princípio da codificação conjunta: a imagem fornecida é convertida em *embeddings* visuais por um codificador especializado, que são integrados ao modelo principal por meio de camadas de atenção cruzada. Essa arquitetura permite ao sistema identificar elementos gráficos, ler trechos textuais na imagem, reconhecer padrões estruturais e gerar descrições coerentes em um único fluxo de saída textual (DeepSeek AI, 2024).

Para a pesquisa, utilizamos a versão gratuita, disponível na interface pública do DeepSeek, o que viabiliza a análise visual mesmo sem acesso à API. Apesar de suas limitações em profundidade de análise ou na manipulação de instruções complexas, esta versão é suficiente para nossos propósitos de experimentação acadêmica. Sua inclusão no estudo nos permite observar o desempenho de um modelo emergente, cujas características de eficiência e síntese textual oferecem uma perspectiva diferente das abordagens adotadas por empresas mais estabelecidas como OpenAI e Google.

Comparativamente, o DeepSeek-V3.2 representa uma arquitetura multimodal alternativa, com um foco notável na eficiência de inferência e na rapidez de resposta. Essas características são valorizadas em aplicações de uso geral e em contextos educacionais, tornando-o um objeto de estudo relevante para nossa análise.

3.3 Delimitação do objeto e ferramentas

O corpus da pesquisa foi composto por três capas de livros selecionadas no Portal Domínio Público, ambiente que reúne obras com direitos autorais liberados para reprodução e uso acadêmico. A definição desse conjunto buscou preservar a diversidade visual necessária para observar o comportamento das ferramentas de IA diante de diferentes demandas descritivas, sem ampliar o corpus a ponto de comprometer a profundidade analítica. Em pesquisas qualitativas, a escolha de um número reduzido de casos é coerente com o objetivo de examinar o fenômeno em profundidade. Flick (2009) enfatiza que a variedade interna do material é mais relevante do que sua quantidade, e que a análise depende da intensidade com que cada caso é investigado. Gibbs (2009) observa que a análise qualitativa depende de exame detalhado e repetido dos dados, exigindo um corpus que seja gerenciável e que permita atenção minuciosa às variações internas, padrões e possíveis contradições.

A seleção das três capas obedeceu ao critério de variedade temática e editorial, contemplando (1) uma obra técnica na área de ciências da saúde; (2) um título literário; e (3) uma obra técnica na área de biblioteconomia. A escolha por gêneros distintos segue o princípio da amostragem intencional, discutido por Bauer e Aarts (2002), no qual os casos são selecionados por seu potencial de ilustrar contrastes significativos dentro do fenômeno investigado. As capas selecionadas serão apresentadas integralmente na seção **Erro! Fonte de referência não encontrada..** Além disso, a literatura sobre análise da imagem aponta que estudos aprofundados exigem atenção às particularidades semióticas de cada objeto visual. De forma complementar, Santaella (1999) argumenta que diferentes categorias de imagens operam com regimes próprios de produção de sentido, tornando necessário priorizar a diversidade interna do corpus, e não sua extensão numérica.

A pesquisa adota como estudo de caso o contexto das bibliotecas universitárias da UNIRIO, com foco na ausência de audiodescrição das capas no catálogo *on-line* (sistema Sophia). A escolha da UNIRIO fundamenta-se em sua função formadora no ensino superior, na presença de estudantes com deficiência visual ou baixa visão, e na necessidade institucional de recursos de acessibilidade informacional. Assim, o estudo não trata da descrição genérica de

capas em bibliotecas, mas das demandas concretas observadas no contexto da universidade.

A análise incluiu quatro modelos multimodais: ChatGPT, Google Gemini, Microsoft Copilot e DeepSeek. A escolha dessas ferramentas decorre da capacidade que possuem de interpretar imagens e gerar descrições textuais por meio de comandos em linguagem natural, característica indispensável para a etapa experimental. Para garantir comparabilidade, todas as ferramentas receberam o mesmo *prompt*, elaborado com base nos critérios da ABNT NBR 16452:2016, que define orientações para a estrutura da audiodescrição humana, e nas diretrizes da WCAG 2.2, que estabelecem fundamentos de acessibilidade digital em relação à equivalência perceptiva de conteúdos visuais.

As três capas foram submetidas às quatro ferramentas. O mesmo comando, nas mesmas condições de entrada, foi utilizado em todas as execuções. Essa padronização metodológica assegura que as diferenças observadas nas descrições decorram das características internas dos modelos, e não de variações no procedimento. As respostas foram registradas e organizadas em quadros comparativos, garantindo rastreabilidade e transparência.

3.4 Os dados e seu tratamento por abordagem qualitativa e o percurso metodológico

Os dados foram coletados por meio de pesquisa bibliográfica, documental e de campo, sendo tratados segundo uma abordagem qualitativa.

Na pesquisa, adotou-se o método comparativo, pois “procede pela investigação de indivíduos, classes, fenômenos ou fatos, com vistas a ressaltar as diferenças e similaridades entre eles” (Gil, 2010, p.16) com fins descritivos, pois aborda a “descrição das características de determinada população ou fenômeno ou, então, o estabelecimento de relações entre variáveis” (Gil, 2010, p. 28).

A abordagem qualitativa tem o “ambiente como fonte direta dos dados” (Prodanov; Freitas, 2013). A escolha deve-se ao interesse em compreender como os modelos de IA produzem descrições, quais características preservam

ou distorcem, e em que medida suas respostas se aproximam dos parâmetros técnicos definidos pela ABNT NBR 16452:2016 e pelas diretrizes de acessibilidade digital da WCAG 2.2. Esse tipo de abordagem permite observar nuances, incoerências, padrões e limitações (Prodanov; Freitas, 2013).

A análise qualitativa baseou-se em seis eixos, definidos a partir das normas e da literatura especializada: precisão, detalhamento, clareza, relevância, ordem lógica e neutralidade.

Esses critérios seguem o que Flick (2009) descreve como alinhamento entre procedimento e objeto, garantindo consistência entre método e finalidade interpretativa.

A coleta de dados ocorreu em duas fases. Na primeira, foram selecionadas as três capas no Portal Domínio Público, observando a variedade visual e a pertinência temática ao contexto universitário. Na segunda, aplicou-se o *prompt* padronizado aos quatro modelos multimodais. As respostas geradas constituem o corpus principal de análise e foram examinadas de acordo com o conjunto de critérios definidos.

O percurso metodológico foi sintetizado no Quadro 1, que organiza as etapas e seus objetivos. Cada fase foi planejada para articular teoria e prática, permitindo compreender o potencial das ferramentas de IA na audiodescrição das capas e seu uso no ambiente da UNIRIO.

Quadro 1 - Percurso metodológico.

Etapas	Descrição	Instrumento/Fonte	Objetivo
Levantamento teórico	Revisão da literatura sobre imagem, semiótica, audiodescrição, acessibilidade e IA em bibliotecas universitárias.	Livros, artigos científicos, dissertações, normas ABNT, WCAG 2.2.	Fundamentar teoricamente a pesquisa.
Definição do contexto empírico	Identificação das necessidades de acessibilidade informacional na UNIRIO, com foco na ausência de audiodescrição das capas no catálogo Sophia.	Documentos institucionais, site da UNIRIO, análise exploratória do catálogo.	Delimitar o estudo de caso e justificar a relevância institucional.
Seleção do corpus	Escolha de três capas do Portal Domínio Público, representando diversidade visual e editorial: uma obra técnica da área da saúde, uma obra literária e uma obra relacionada à Biblioteconomia ou campos afins, assegurando variedade gráfica para a análise.	Portal Domínio Público.	Variar estilos gráficos sem comprometer a profundidade da análise qualitativa.
Elaboração do prompt	Construção de um comando único baseado na ABNT NBR 16452:2016 e nos princípios de equivalência perceptiva da WCAG 2.2.	Norma ABNT NBR 16452:2016; WCAG 2.2.	Garantir consistência, comparabilidade e rigor descritivo.
Geração das audiodescrições	Aplicação do <i>prompt</i> às quatro ferramentas multimodais (ChatGPT, Copilot, Gemini e DeepSeek) nas mesmas condições de entrada.	Ferramentas de IA.	Produzir as descrições automatizadas que compõem o corpus principal.
Avaliação comparativa	Análise qualitativa baseada em precisão, detalhamento, clareza, relevância, ordem lógica e neutralidade.	Matriz avaliativa construída a partir das normas e da literatura.	Identificar padrões, diferenças, limitações e potencialidades de cada IA.
Redação da dissertação	Organização e escrita dos capítulos, integrando fundamentação teórica, metodologia, análise e discussão.	Documento final.	Apresentar os resultados com consistência acadêmica.
Elaboração do produto	Produção de um manual prático com orientações para bibliotecas universitárias que desejam utilizar IA na audiodescrição das capas.	Material técnico com exemplos, protocolos e recomendações.	Oferecer aplicabilidade prática e contribuir para a acessibilidade informacional da UNIRIO.

Fonte: elaboração própria.

Ao final, será produzido um manual prático com orientações objetivas para bibliotecas universitárias, contendo: sugestões de *prompts*; recomendações de uso das ferramentas; exemplos comentados; diretrizes para integrar descrições

aos catálogos; e limites observados na geração automática. Esse produto atende à dimensão aplicada da pesquisa, ao oferecer instrumentos concretos para aprimorar a acessibilidade no âmbito da biblioteconomia.

Os procedimentos metodológicos foram estruturados para garantir rigor, transparência e possibilidade de reprodutibilidade. Para isso, a metodologia organiza-se em cinco eixos principais: (a) definição do enquadramento teórico-metodológico; (b) seleção de três capas escolhidas em função de sua pertinência ao contexto universitário; (c) elaboração do roteiro de audiodescrição e do *prompt* com base na ABNT NBR 16452:2016 e nas diretrizes da WCAG 2.2; (d) aplicação das ferramentas de IA selecionadas (ChatGPT, Microsoft Copilot, Google Gemini e DeepSeek); e (e) estabelecimento dos critérios de análise comparativa das audiodescrições geradas.

3.5 Caracterização da pesquisa

Esta seção apresenta a caracterização da pesquisa a partir de suas frentes de investigação e do campo empírico adotado. Para isso, articulam-se os procedimentos bibliográficos, documentais e de campo, de modo a sustentar tanto a fundamentação teórica quanto a delimitação do contexto analisado. Essa caracterização permite explicitar as bases que orientam a produção, a coleta e a interpretação dos dados no desenvolvimento do estudo.

3.5.1 As investigações bibliográficas e documentais

A abordagem bibliográfica foi empregada por basear-se em materiais já publicados, tais como monografias, dissertações, teses, periódicos científicos e livros. Já a dimensão documental refere-se à utilização de documentos oficiais, como leis, decretos e atas, que fornecem suporte legal e normativo ao estudo (Prodanov; Freitas, 2013). Essa combinação de fontes permite uma análise abrangente e fundamentada do tema investigado.

Para identificar materiais bibliográficos e documentos relevantes, foram definidos termos e expressões como acessibilidade, deficiência visual, audiodescrição, biblioteca e tecnologia assistiva, combinados com o termo IA. Para refinar a busca, termos compostos foram colocados entre aspas, e foram aplicados filtros de idioma (português e inglês) e período (2014 a 2024). A

escolha dos idiomas se deve à importância de publicações em português, que refletem o contexto nacional, e em inglês, por ser o idioma predominante na produção acadêmica internacional. Já o recorte temporal foi adotado para garantir que os resultados estivessem alinhados com o conhecimento mais atual e relevante para o momento da pesquisa. O operador booleano AND foi utilizado para garantir que todos os termos estivessem presentes nos resultados.

As buscas foram realizadas em bases de dados reconhecidas. O Google Acadêmico foi escolhido por sua diversidade de fontes, incluindo artigos, teses, dissertações e livros. A BDTD (Biblioteca Digital de Teses e Dissertações) foi utilizada por oferecer um acervo robusto e centralizado de trabalhos acadêmicos em várias áreas do conhecimento. O Catálogo de Teses e Dissertações da Capes também foi consultado por ser uma referência confiável e usada no Brasil. Todas essas bases foram escolhidas por sua acessibilidade, interfaces intuitivas e ferramentas que facilitam o refinamento da pesquisa, contribuindo para a obtenção de resultados mais precisos e relevantes.

O Quadro 2 apresenta as expressões de busca utilizadas nas bases de dados, juntamente com a quantidade de documentos recuperados.

Quadro 2 - Expressões de busca utilizadas nas bases de dados

Base de dados	Termos e operadores usados	Documentos recuperados
BDTD	'tecnologia assistiva AND deficiência visual'	239
	'deficiência visual AND audiodescrição'	96
	'deficiência visual AND inteligência artificial'	5
	'acessibilidade AND inteligência artificial'	46
	'biblioteca AND audiodescrição'	36
Google Acadêmico	acessibilidade AND "inteligência artificial" AND audiodescrição	485
	biblioteca AND "tecnologia assistiva" AND "inteligência artificial"	640
	biblioteca AND audiodescrição AND "inteligência artificial"	375
Catálogo de Teses e Dissertações Capes	biblioteca AND acessibilidade AND "deficiência visual"	229
	acessibilidade AND "inteligência artificial"	34
	"deficiência visual" AND "inteligência artificial"	17

Fonte: elaboração própria.

O levantamento bibliográfico foi conduzido no período entre outubro e dezembro de 2024. Durante o processo de busca e seleção de materiais nas

bases de dados escolhidas, foram analisadas as cinco primeiras páginas de resultados, uma vez que os trabalhos mais relevantes ou mais atualizados tendem a aparecer nas primeiras posições.

Conforme destacado por Cendón (2001), os mecanismos de busca das bases de dados empregam algoritmos que classificam os resultados com base em critérios como a localização e a frequência das palavras-chave em artigos, teses, livros e outros documentos. Publicações nas quais os termos pesquisados aparecem em posições estratégicas, como títulos e cabeçalhos, são consideradas mais pertinentes. Além disso, outros fatores que influenciam a posição dos resultados incluem a quantidade de termos da busca presentes no texto, a proximidade entre eles e o tamanho do documento (Cendón, 2001). Essa abordagem proporciona uma visão inicial do que está disponível sobre o tema em estudo.

Cada artigo foi avaliado quanto à sua relevância por meio da análise de elementos como título, resumo e palavras-chave. A leitura desses componentes permitiu uma compreensão preliminar do conteúdo de cada trabalho, facilitando a seleção daqueles que seriam lidos na íntegra. Essa etapa foi essencial para garantir que somente os materiais mais alinhados aos objetivos da pesquisa fossem considerados para análise detalhada.

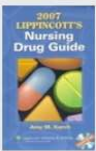
3.5.2 A investigação de campo

Para o estudo de campo, o recorte empírico da pesquisa foi definido a partir das bibliotecas da UNIRIO, adotadas como estudo de caso por se constituírem como bibliotecas universitárias que atendem um público diversificado, incluindo estudantes com deficiência visual ou baixa visão. Para complementar a caracterização do campo empírico, realizou-se uma solicitação formal ao NAI da UNIRIO, por meio de e-mail institucional, requisitando informações sobre a presença de estudantes com deficiência visual matriculados na instituição. Em resposta, o NAI disponibilizou uma planilha contendo curso, turno e tipologia da deficiência visual (cegueira ou baixa visão/visão monocular) dos estudantes cadastrados, sem qualquer identificação nominal ou dado pessoal sensível, assegurando, portanto, o anonimato dos discentes.

A escolha da UNIRIO como campo de estudo responde a dois fatores centrais: (1) sua relevância institucional no contexto da educação superior pública, como espaço de formação e produção de conhecimento em distintas áreas; e (2) a ausência de audiodescrição das capas no catálogo *on-line* Sophia, elemento que evidencia uma lacuna concreta de acessibilidade informacional e reforça a pertinência da investigação proposta.

No caso da UNIRIO, observa-se que o catálogo *on-line*, operado pelo gerenciador Sophia, apresenta majoritariamente informações textuais convencionais relacionadas ao item bibliográfico (autor, número de chamada, ISBN, dados de publicação, descrição física e assuntos), conforme apresentado na Figura 1.

Figura 1 - Livro pelo catálogo *on-line* da UNIRIO.

Detalhes	MARC tags	Dublin Core
Detalhes da obra		
	Inf. publicação	E-book - Inglês
	ISBN	1582556296 (USA)
	Número de chamada	
	Classificação	615
	Notação	L765
	Título	2007 Lippincott's nursing drug guide [recurso eletrônico] / editor, Amy M. Karch.
	Outros títulos	[Lippincott's nursing drug guide]
	Imprensa	Philadelphia : Lippincott Williams & Wilkins, 2007.
	Desc. física	1 v.
	Sites relacionados	
	Site	http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=booktext&NEWS=N&DF=bookdb&AN=00146940/2007_Edition/3&XPATH=/PG(0)
	Assuntos	1. Medicamentos - Administração - Manuais 2. Enfermagem - Manuais, guias, etc.
	Ent. sec.	I. Karch, Amy Morrison, 1949- (Ed.)
	Link do título	https://sophia-web.unirio.br/TerminalWeb/index.asp?codigo_sophia=92377

Fonte: UNIRIO (2024).

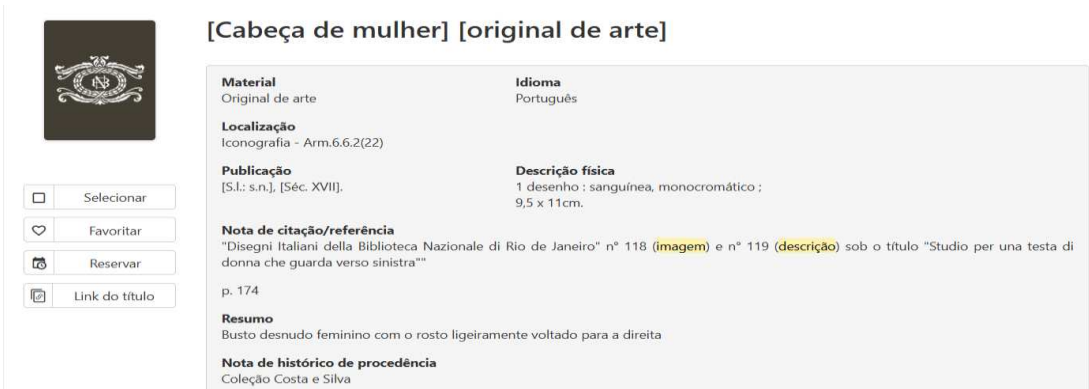
Audiodescrição: A imagem mostra a tela do catálogo *on-line* da UNIRIO, no sistema Sophia, exibindo o registro do livro *2007 Lippincott's Nursing Drug Guide*, editado por Amy M. Karch. À esquerda, aparece a capa do livro. A capa tem fundo em tons de azul e apresenta três formas ovais em destaque, semelhantes a cápsulas estilizadas, nas cores verde, laranja e azul-claro, alinhadas no centro. O título "*2007 Lippincott's Nursing Drug Guide*" está posicionado na parte superior da capa, em letras maiúsculas amarelas. O nome da editora responsável, "Amy M. Karch", aparece na parte inferior, em fonte branca. No canto inferior direito da capa, há o logotipo da editora Lippincott Williams & Wilkins, composto por pequenas formas geométricas coloridas. Abaixo da capa, na coluna esquerda, há quatro botões dispostos verticalmente: "Selecionar", "Favoritar", "Reservar" e "Link do título". À direita, encontram-se os dados do registro: tipo de material (E-book em inglês), ISBN (1582556296, EUA), classificação 615, notação L765 e o título completo "*2007 Lippincott's nursing drug guide [recurso eletrônico] / editor, Amy M. Karch*". Também são exibidos outros títulos relacionados, a informação da editora (Philadelphia: Lippincott Williams & Wilkins, 2007), a descrição física ("1 v."), o link de acesso eletrônico e dois assuntos listados: "Medicamentos –

Administração – Manuais” e “Enfermagem – Manuais, guias, etc.”. Na parte superior direita, estão as abas “Detalhes”, “MARC tags” e “Dublin Core”, destacadas sobre fundo cinza-claro. O restante da tela apresenta fundo branco e texto em fonte preta, organizado em tabela de duas colunas.

Nota-se que não há qualquer descrição textual da capa, apenas a exibição da imagem. Para usuários com deficiência visual, esse tipo de apresentação impõe barreiras significativas: o leitor de tela identifica a capa como mero arquivo gráfico não descrito, inviabilizando o acesso ao conteúdo imagético. Embora o registro apresente descrição física do item (como número de páginas e presença de ilustrações), essa informação não substitui a audiodescrição da capa, que permanece ausente.

Ao analisar a Biblioteca Nacional, nota-se que, embora não utilize o termo “audiodescrição”, a instituição apresenta um tipo de descrição sucinta da imagem em registros de obras iconográficas, sobretudo em materiais de arte, conforme Figura 2 e Figura 3. **Erro! Fonte de referência não encontrada.** Trata-se, contudo, de uma descrição resumida, que não atende aos padrões de detalhamento, neutralidade e informatividade previstos na ABNT NBR 16452:2016, mas que já representa um esforço institucional de transpor parte do conteúdo visual para a linguagem verbal.

Figura 2 - Catálogo da Biblioteca Nacional, obra “[Cabeça de mulher]”.



The image shows a catalog entry from the Biblioteca Nacional. On the left, there is a small, dark, square thumbnail of the artwork. To its right, the title "[Cabeça de mulher] [original de arte]" is displayed. Below the title, there are several fields: "Material" (Original de arte), "Idioma" (Português), "Localização" (Iconografia - Arm.6.6.2(22)), "Publicação" ([S.l.: s.n.], [Séc. XVII]), and "Descrição física" (1 desenho : sanguínea, monocromático; 9,5 x 11cm). There is also a "Nota de citação/referência" section with a long text snippet, a "Resumo" section with a brief description, and a "Nota de histórico de procedência" section. On the far left, there are four buttons: "Selecionar", "Favoritar", "Reservar", and "Link do título".

Fonte: Biblioteca Nacional (2025).

Audiodescrição: A imagem apresenta o registro do catálogo da Biblioteca Nacional para o item *[Cabeça de mulher] [original de arte]*. À esquerda, há uma miniatura da imagem, em tom escuro, no formato de um brasão estilizado, sem detalhes visuais nítidos da obra. À direita, o registro apresenta os campos textuais. O material é “Original de arte”. O idioma é “Português”. A localização está indicada como “Iconografia – Arm.6.6.2(22)”. A publicação aparece como “[S.l.: s.n.], [Século XVIII]”. A descrição física indica “1 desenho: sanguínea, monocromático; 9,5 x 11 cm”. No campo “Nota de citação/referência”, aparece referência bibliográfica que inclui as expressões “(imagem)” e “(descrição)”. Abaixo, o resumo apresenta: “Busto desnudo feminino com o rosto ligeiramente voltado para a direita”. O campo final, “Nota de histórico de procedência”,

indica “Coleção Costa e Silva”. O layout segue padrão de fundo claro com texto em preto distribuído em blocos organizados.

Figura 3 - Catálogo da Biblioteca Nacional, obra “[Duas figuras e uma árvore]”.



☐ Selecionar

 Favoritar

 Reservar

 Link do título

[Duas figuras e uma árvore] [original de arte]

Material Original de arte	Idioma Português
Localização Iconografia - Arm.6.7.1(4)	
Publicação [S.l.: s.n.], [Séc.XVIII].	Descrição física 1 desenho : bico de pena, monocromático ; 12,8 x 18,3cm.
Nota de citação/referência "Disegni Italiani della Biblioteca Nazionale di Rio de Janeiro" n° 168 (imagem) n°172 (descrição) sob o título "Due figure e un albero" e atribui autoria a falsário de Guercino do Séc. XVIII p. 223	
Resumo Grande árvore a esquerda com duas figuras humanas próximas	
Nota de histórico de procedência Coleção Costa e Silva	

Fonte: Biblioteca Nacional (2025).

Audiodescrição: A imagem mostra a tela de um registro do catálogo da Biblioteca Nacional referente ao item *[Duas figuras e uma árvore] [original de arte]*. À esquerda, aparece uma miniatura da obra: uma imagem monocromática com contornos escuros em formato de brasão ornamental, sem detalhes visíveis da ilustração original. À direita, surgem campos textuais. O material é indicado como “Original de arte”. O idioma é “Português”. A localização aparece como “Iconografia – Arm.6.7.1(4)”. A publicação informa “[S.l.: s.n.], [Século XVIII]”. A descrição física indica “1 desenho: bico de pena, monocromático; 12,8 x 18,3 cm”. No campo “Nota de citação/referência”, há um trecho que menciona a referência do item e inclui, entre parênteses, as palavras “(imagem)” e “(descrição)”, sinalizando que há uma descrição resumida associada. Logo abaixo, o resumo apresenta: “Grande árvore à esquerda com duas figuras humanas próximas”. No fim do registro, há o campo “Nota de histórico de procedência”, que indica “Coleção Costa e Silva”. O fundo da página é claro, com texto em fonte preta organizado em blocos.

Esses exemplos demonstram que, mesmo em instituições de referência nacional, a audiodescrição ainda aparece de modo fragmentado, irregular e pouco aprofundado, predominando resumos minimalistas voltados para identificação iconográfica, mas não uma audiodescrição conforme padrões de acessibilidade. Em outras palavras, as descrições encontradas na Biblioteca Nacional não constituem audiodescrição em sentido pleno, mas sim notas iconográficas condensadas, geralmente inseridas para fins de catalogação.

A presença das imagens, quando não acompanhada de descrição verbal adequada, cria uma barreira de acesso que compromete a experiência. Nesse contexto, torna-se pertinente investigar soluções baseadas em IA, capazes de apoiar bibliotecários na produção automatizada de audiodescrições que ampliem

a acessibilidade e promovam o acesso equitativo às obras.

Outro ponto relevante diz respeito aos sistemas de gerenciamento de bibliotecas, que ainda apresentam infraestrutura técnica limitada para atender às demandas contemporâneas de acessibilidade, em relação à representação de elementos imagéticos. Embora capas e outros recursos visuais componham os registros bibliográficos, grande parte desses sistemas não dispõe de campos específicos para a inserção de audiodescrição, o que inviabiliza a descrição adequada das imagens ou restringe sua inserção a soluções improvisadas. Na prática, o que se encontra são metadados que podem ser adaptados, mas não concebidos para essa finalidade, revelando a ausência de diretrizes ou funcionalidades que incentivem ou operacionalizem a audiodescrição no ambiente catalográfico.

Essa ausência reflete a permanência de modelos de catalogação rigidamente orientados por padrões tradicionais, nos quais a imagem é tratada como mero elemento ilustrativo, e não como objeto informacional que também exige representação descritiva. Como observa Araújo (2010), os sistemas de informação em bibliotecas tendem a reproduzir estruturas técnico-metodológicas convencionais e pouco responsivas às transformações sociais, tecnológicas e culturais que atravessam o trabalho informacional.

A dificuldade de incorporar campos para audiodescrição ou outros recursos de acessibilidade nos gerenciadores bibliográficos também está relacionada a uma cultura de gestão que privilegia práticas já consolidadas e demonstradas como “seguras”, em detrimento da experimentação ou da adoção de tecnologias emergentes. Araújo (2010) argumenta que grande parte das rotinas de trabalho em bibliotecas permanece ancorada em procedimentos padronizados, herdeiros de uma lógica operacional que prioriza estabilidade e reprodutibilidade, mesmo quando tais padrões não respondem adequadamente às necessidades informacionais contemporâneas. Nesse cenário, propostas de inovação — como a inclusão de descrições de imagens, metadados acessíveis ou funcionalidades específicas para usuários com deficiência visual encontram barreiras institucionais, seja pela rigidez dos sistemas, seja por decisões administrativas que evitam alterar fluxos consolidados. Assim, a ausência de campos para audiodescrição nos gerenciadores de bibliotecas não constitui

apenas uma limitação técnica, mas também uma expressão das restrições impostas por modelos de gestão que valorizam a manutenção do já estabelecido em detrimento da evolução das práticas.

3.6 Roteiro técnico de audiodescrição

A audiodescrição, enquanto recurso de acessibilidade comunicacional, deve ser construída com base em critérios técnicos, linguísticos e éticos que garantam à pessoa com deficiência visual acesso claro, autônomo e equivalente ao conteúdo imagético. No Brasil, o principal documento normativo que orienta esse processo é a ABNT NBR 16452:2016, que estabelece diretrizes para a tradução de imagens em palavras e define parâmetros como linguagem objetiva, neutralidade descritiva, ordem lógica e relevância informacional dos elementos descritos (Associação ..., 2016).

A WCAG 2.2, adotada internacionalmente para orientar práticas de acessibilidade digital, complementa a NBR 16452:2016 ao reforçar princípios que atravessam também a audiodescrição, como clareza textual, compreensão da informação e oferta de alternativas textuais adequadas para imagens. Esses princípios são relevantes em ambientes digitais, incluindo catálogos *on-line* de bibliotecas, onde a experiência de navegação precisa se manter acessível mesmo quando o conteúdo envolve mídias visuais (Associação ..., 2016).

A NBR 17225:2025, voltada à acessibilidade em conteúdo e aplicações web, é utilizada aqui como documento auxiliar. Embora não trate da elaboração de audiodescrição, suas diretrizes para apresentação textual, organização da informação, alternativas para imagens e estruturação de textos contribuem para definir critérios de qualidade esperados das descrições utilizadas em ambientes digitais. Dessa forma, a norma funciona como apoio complementar, reforçando a coerência entre práticas de audiodescrição e requisitos de acessibilidade web (Associação ..., 2025).

Segundo Ribeiro (2011), a tradução verbal da imagem exige frases claras, coerentes e organizadas de modo a preservar os elementos essenciais da composição. O autor destaca que características humanas observáveis, como etnicidade, gênero, traços faciais, expressões, postura corporal e vestimentas, devem ser descritas apenas quando pertinentes para a compreensão da cena e

sempre com neutralidade. Essa orientação está alinhada às exigências da NBR 16452:2016, que veda inferências subjetivas, julgamentos ou interpretações psicológicas não sustentadas pela materialidade visual (Associação ..., 2016).

A ABNT NBR 17225:2025 reforça que a construção do roteiro deve considerar o efeito de sentido pretendido, o tempo de exposição da imagem e a função informativa da descrição, sem recorrer a interpretações ou avaliações subjetivas.

A seguir, apresenta-se um roteiro técnico baseado nas diretrizes combinadas das dessas normas, organizado por etapas para orientar a produção e a análise da audiodescrição de capas de livros e outras imagens estáticas:

a) Estrutura descritiva: do geral ao específico

A descrição inicia com uma visão panorâmica da imagem, situando a pessoa usuária no contexto geral. Em seguida, são apresentados os elementos específicos da composição, de modo a favorecer a construção mental da cena. Nessa etapa, são identificados: o tipo de imagem (fotografia, ilustração, arte digital, colagem), a paleta cromática predominante e eventuais contrastes relevantes, bem como os principais elementos em primeiro plano e plano de fundo, com indicação de sua localização espacial.

b) Elementos essenciais a serem descritos

Em consonância com a NBR 16452:2016, a descrição deve responder, de forma explícita ou implícita, aos seguintes eixos informativos:

- Quem? (personagens, figuras humanas ou animais, autores representados, quando houver);
- O quê? (objetos, elementos gráficos, títulos, símbolos, logotipos);
- Como? (estilo visual, disposição dos elementos, gestos, posições corporais, texturas e técnicas aparentes);
- Onde? (cenário, ambiente, fundo, relação espacial entre os elementos);
- Quando? (época ou período sugerido, caso a imagem ofereça pistas temporais).

Características fenotípicas, expressões corporais e vestimentas são descritas apenas quando apresentam função comunicacional e podem ser observadas na imagem (Ribeiro, 2011), evitando-se extrapolações quanto a emoções, intenções ou identidades não explicitamente representadas.

c) Clareza e objetividade linguística

A NBR 16452:2016 recomenda o uso de frases curtas, diretas e no tempo presente do indicativo. O vocabulário deve ser preciso e acessível, evitando ambiguidades, metáforas, regionalismos e tecnicismos desnecessários. Em diálogo com a NBR 17225:2025 e com a WCAG 2.2, o texto da audiodescrição deve ser coeso, segmentado de forma lógica e adequado à escuta por leitores de tela e outros recursos de tecnologia assistiva.

d) Relevância informacional

A descrição deve privilegiar informações pertinentes à compreensão do conteúdo imagético, evitando repetições, redundâncias e minúcias que não contribuam para a construção significativa da cena. A seleção dos elementos descritos leva em conta o público-alvo, os objetivos informacionais da obra e o contexto de uso – no caso desta pesquisa, o catálogo de uma biblioteca universitária.

e) Ordem lógica e organização textual

A sequência da audiodescrição acompanha a organização visual da imagem, seguindo uma ordem que favorece a apreensão auditiva. Adota-se, como referência geral:

1. Visão geral da imagem;
2. Elementos principais em destaque;
3. Detalhes complementares relevantes;
4. Leitura do texto inserido na imagem (título, autor, editora, selos, prêmios);
5. Encerramento da descrição, quando necessário.

A ABNT NBR 17225:2025 orienta que textos em ambientes digitais sejam apresentados em unidades organizadas, o que também se aplica ao roteiro de audiodescrição, facilitando sua leitura e posterior reutilização em catálogos.

f) Neutralidade descritiva

A linguagem empregada deve permanecer isenta de adjetivações subjetivas, juízos de valor, estigmas ou inferências não apoiadas na imagem. A imparcialidade é condição para que a pessoa com deficiência visual forme sua própria interpretação a partir das informações fornecidas. Termos avaliativos são evitados, substituídos por descrições observáveis (por exemplo, “cores em tons escuros” em vez de “cores tristes”).

g) Leitura de textos contidos na imagem

Em capas de livros, a leitura dos elementos textuais é parte obrigatória da audiodescrição. Devem ser lidos, sempre que presentes:

- Título da obra;
- Nome do autor ou autora;
- Subtítulo, edição, coleção, prêmios ou selos;
- Nome da editora ou logotipo identificável;
- Particularidades da tipografia quando cumprem função expressiva (por exemplo, letras manuscritas, letras em caixa alta ou títulos dispostos de forma inclinada).

Esse roteiro técnico constituiu a base para a elaboração do *prompt* aplicado às ferramentas de IA para a avaliação das audiodescrições produzidas, garantindo alinhamento entre o procedimento metodológico e os parâmetros normativos de acessibilidade

3.7 Elaboração dos *prompts*

Para padronizar a produção automatizada de audiodescrição por ferramentas baseadas em IA, os *prompts* foram elaborados com base nos parâmetros definidos pela ABNT NBR 16452:2016, norma central para a construção da audiodescrição no Brasil, e nos princípios de acessibilidade digital da WCAG 2.2, que orientam a apresentação clara, perceptível e compreensível das informações em ambientes digitais. A NBR 17225:2025, voltada à acessibilidade em conteúdo web, é utilizada aqui como documento

complementar, pois reforça critérios de organização textual, legibilidade e oferta de alternativas para conteúdo visual.

Combinadas, essas diretrizes oferecem os fundamentos técnicos e linguísticos necessários para garantir que as descrições geradas pelas ferramentas de IA sejam claras, precisas, objetivas, relevantes e organizadas de forma lógica, permitindo seu uso em catálogos bibliográficos e outros ambientes informacionais.

Com base nessas normas e princípios, o seguinte *prompt* foi desenvolvido:

“Aja como um especialista em audiodescrição, seguindo estritamente as diretrizes da ABNT NBR 16452:2016 e os princípios de acessibilidade da WCAG 2.2. Sua tarefa é criar uma audiodescrição detalhada para a capa do livro enviada em anexo em parágrafo único de forma clara, objetiva e estruturada, utilizando frases no tempo presente do indicativo e mantendo neutralidade, precisão e relevância informacional. A audiodescrição deve iniciar com uma visão geral da imagem e, em seguida, apresentar seus elementos de maneira organizada, indo do geral ao específico. Sempre que houver textos presentes na imagem, como título, nome do autor, subtítulo, editora, selos ou outras inscrições, leia e descreva esses elementos no momento em que aparecerem na composição visual. Identifique figuras humanas ou personagens, quando existirem, descrevendo características observáveis como postura, vestimentas e expressões, sem inferências subjetivas. Descreva também objetos, cores, texturas, formas, disposição espacial e demais componentes visuais, bem como o cenário ou ambiente retratado. Inclua indicações temporais somente quando estiverem visíveis na imagem. A descrição deve seguir uma ordem lógica compatível com a percepção visual, garantindo clareza comunicacional.”

Esse *prompt* pode ser adaptado conforme o tipo de imagem — fotografia, ilustração, arte gráfica — ou de acordo com o contexto de uso, como ambientes educacionais, culturais ou institucionais, desde que preservados os parâmetros estabelecidos pela ABNT NBR 16452:2016 e pelos princípios de acessibilidade previstos na WCAG 2.2.

3.8 Critérios de avaliação das descrições

A avaliação das audiodescrições produzidas pelas ferramentas de IA adotadas nesta pesquisa foi conduzida com base em critérios derivados da ABNT NBR 16452:2016, norma central para a elaboração de audiodescrições no Brasil, e da ABNT NBR 17225:2025, utilizada aqui como documento complementar para reforçar princípios de organização textual, legibilidade e coerência informacional em ambientes digitais. Esses critérios foram selecionados de modo a contemplar tanto os aspectos linguísticos quanto os aspectos estruturais envolvidos na qualidade da descrição, garantindo que o material gerado possa ser utilizado por pessoas com deficiência visual em contextos informacionais como bibliotecas e catálogos digitais.

Os critérios adotados pela ABNT NBR 16452:2016 são:

1. **Precisão:** Avalia se a descrição corresponde ao conteúdo visual da imagem, sem erros de identificação, omissões ou inferências indevidas.
2. **Detalhamento:** Verifica se a descrição contempla os elementos essenciais recomendados pela ABNT NBR 16452:2016, incluindo *o que*, *quem*, *como*, *onde* e *quando*, quando aplicáveis.
3. **Clareza:** Considera se a linguagem utilizada é objetiva, direta, coesa e redigida no tempo presente do indicativo, conforme orientações normativas.
4. **Relevância:** Analisa se as informações apresentadas são pertinentes para a compreensão da cena, evitando redundâncias ou detalhes irrelevantes.
5. **Ordem lógica:** Examina se a descrição segue a estrutura sequencial recomendada pela ABNT NBR 16452:2016 (do geral ao específico), favorecendo a construção mental da imagem pelo usuário.
6. **Neutralidade:** Verifica se a descrição evita julgamentos de valor, interpretações subjetivas, especulações ou qualquer inferência não confirmada visualmente.

Para sistematizar essa avaliação qualitativa, desenvolveu-se uma escala descritiva de quatro níveis, que permite graduar o desempenho das ferramentas

em relação a cada critério técnico, facilitando a identificação de padrões e variações. A pontuação é atribuída da seguinte forma:

- **Excelente (4):** A descrição atende a todos os critérios estabelecidos, sem apresentar falhas ou omissões significativas.
- **Bom (3):** A descrição cumpre a maioria dos critérios, com pequenas limitações ou a necessidade de ajustes pontuais.
- **Regular (2):** A descrição atende parcialmente aos critérios, exibindo falhas perceptíveis ou informações incompletas que comprometem a qualidade.
- **Fraco (1):** A descrição não atende aos critérios fundamentais, apresentando omissões graves ou erros que a tornam inadequada para o propósito.

Essa gradação permite uma análise matizada do desempenho de cada ferramenta, fornecendo uma base robusta para as discussões subsequentes.

4 RESULTADOS

A análise das audiodescrições geradas pelos quatro modelos — ChatGPT, Gemini, Copilot e DeepSeek, todos em suas versões gratuitas — revelou padrões de desempenho distintos na interpretação das capas de livros selecionadas. Cada ferramenta demonstrou abordagens específicas no processamento de imagens, influenciando a completude, a neutralidade e o detalhamento das descrições produzidas. Essas variações não se restringem à qualidade das imagens, mas refletem as arquiteturas multimodais inerentes a cada sistema. Os resultados, portanto, ilustram como a lógica interna de cada modelo impacta a aderência às normas técnicas (ABNT NBR 16452:2016 e princípios da WCAG 2.2), delineando as potencialidades e os limites da inteligência artificial na audiodescrição para bibliotecas universitárias.

4.1 Conjunto de imagens analisadas

A seleção das obras para análise foi planejada para oferecer um panorama diversificado dos desafios que as ferramentas de IA enfrentam na audiodescrição.

A primeira obra, *Manual de Primeiros Socorros*, da Fundação Oswaldo Cruz (2003), é um material técnico da área de ciências da saúde. Sua capa, de cunho funcional e institucional, exhibe elementos gráficos característicos de publicações governamentais e de saúde pública, como paletas cromáticas padronizadas e uma organização visual que prioriza a clareza. A descrição desse tipo de imagem exige precisão, dada a combinação de elementos simbólicos da área com componentes textuais e informativos.

Em seguida, *Informação e Controle Bibliográfico: um olhar sobre a cibernética*, de Ana Maria Nogueira Machado (2023), representa o campo da Biblioteconomia. Sua inclusão visa observar como a IA processa elementos visuais inerentes ao universo informacional. Capas dessa natureza exibem diagramações mais conceituais, com foco em abstrações gráficas, símbolos e composições tipográficas que aludem ao campo científico em questão. Isso demanda das ferramentas de IA a capacidade de interpretar estruturas visuais menos figurativas.

Por fim, a obra *O Eu profundo e os outros Eus*, de Fernando Pessoa (1913) e disponível no Portal Domínio Público, foi selecionada para representar o universo literário. Capas de obras dessa natureza demonstram vasta diversidade em estilo e composição, mobilizando elementos estilísticos, metáforas visuais ou composições simbólicas. Isso desafia os modelos de IA a reconhecer camadas visuais mais subjetivas, sem comprometer a neutralidade imposta pela NBR 16452:2016.

Dessa forma, os três títulos constituem um conjunto visual heterogêneo e metodologicamente pertinente para analisar a performance das ferramentas de IA na geração de audiodescrições automatizadas. Cada capa, ao apresentar desafios distintos — do técnico-institucional ao conceitual e ao literário —, permite comparar a interpretação de estruturas visuais diversas pelos modelos e a tradução desses elementos em linguagem verbal, em conformidade com os parâmetros normativos da pesquisa.

4.2 Análises por ferramenta

Nesta seção, apresentamos as avaliações detalhadas de cada ferramenta de Inteligência Artificial, seguindo os critérios estabelecidos na metodologia. Para a análise de cada audiodescrição, consideraram-se a precisão, o detalhamento, a clareza, a relevância informacional, a ordem lógica e a neutralidade descritiva, conforme orientado pelas normas ABNT NBR 16452:2016 e as diretrizes complementares da WCAG 2.2 e da NBR 17225:2025, visando uma avaliação rigorosa do desempenho de cada modelo.

As três capas das publicações estão apresentadas nesta ordem:

- 1) Manual de Primeiros Socorros (Fundação Oswaldo Cruz, 2003) (Figura 4);
- 2) Informação e controle bibliográfico: um olhar sobre a cibernética (Machado, 2023) (Figura 5);
- 3) *O Eu profundo e os outros Eus* (Pessoa, 1913) (Figura 6).

4.2.1 Análise da capa do livro Manual de Primeiros Socorros (Fundação Oswaldo Cruz, 2003)

A capa considerada está apresentada na Figura 4, a qual segue as quatro respectivas audiodescrições pelas ferramentas selecionadas.

Figura 4 - Capa da obra “Manual de Primeiros Socorros”.



Fonte: Fundação Oswaldo Cruz (2003).

1) Audiodescrição gerada pelo ChatGPT-5.1:

A imagem apresenta a capa de um livro com design gráfico em alto contraste, predominando as cores branco, vermelho e preto. Em plano geral, o fundo é totalmente branco, e sobre ele se distribuem elementos textuais e ilustrativos. Na parte superior, em letras grandes e espessas, aparece o título “manual DE primeiros socorros”, disposto em três linhas, com as palavras “manual” e “socorros” em vermelho, enquanto “DE” e “primeiros” estão em preto. À esquerda, há uma forma gráfica estilizada em vermelho que lembra um contorno curvo, semelhante a um símbolo abstrato ou ao traço de uma espiral aberta. Abaixo e à direita do título, visível no quadrante inferior direito da capa, há a ilustração de um estojo de primeiros socorros em forma de caixa retangular com cantos arredondados. O estojo é branco com bordas e detalhes em vermelho, incluindo uma cruz vermelha sobre um círculo branco na tampa. A tampa também exibe o texto “MANUAL DE PRIMEIROS SOCORROS” em letras vermelhas. Na base da capa, centralizado na porção inferior, encontra-se novamente a palavra “socorros” em vermelho, escrita em fonte grande e espessa. Logo abaixo, à

direita, aparece o logotipo “NUBIO” acompanhado do texto “Núcleo de Biossegurança Fundação Oswaldo Cruz”, ambos em preto. Não há figuras humanas ou elementos de cenário; todos os componentes são gráficos e tipográficos, organizados de forma simétrica e com forte contraste visual.

Critérios de avaliação:

- a) Precisão - 3 (Bom):** A descrição corresponde aos elementos presentes na capa. Contudo, apresenta um erro relevante ao afirmar que a palavra “primeiros” está inteiramente em vermelho, quando na imagem ela aparece dividida cromaticamente, com “prim” em preto e “eios” em vermelho. Não há inferências sobre o símbolo gráfico ou sobre o estojo além do que é verificável visualmente.
- b) Detalhamento - 3 (Bom):** Os elementos principais e secundários da imagem foram descritos: título, cores predominantes, símbolo gráfico, estojo ilustrado, textos complementares e logotipo institucional. Entretanto, seria possível ampliar ligeiramente o detalhamento das proporções relativas, posição exata dos elementos em relação às margens e dimensões visuais aproximadas do símbolo à esquerda.
- c) Clareza - 4 (Excelente):** O texto está redigido de forma fluida, objetiva e coesa, no tempo presente do indicativo. As frases mantêm boa legibilidade e organização, facilitando o processamento auditivo de usuários com deficiência visual.
- d) Relevância - 4 (Excelente):** Todas as informações fornecidas são pertinentes para a compreensão da capa; não há excesso descritivo, especulação ou elementos irrelevantes.
- e) Ordem lógica - 4 (Excelente):** A organização geral segue o movimento recomendado do geral para o específico.
- f) Neutralidade - 4 (Excelente):** Não apresenta julgamentos de valor ou interpretações subjetivas. O símbolo gráfico é descrito de forma neutra (“forma gráfica estilizada”), sem atribuição de significados não evidenciáveis.
- g) Comentário geral -** A audiodescrição apresenta alta qualidade de acordo com os critérios estabelecidos.

2) **Audiodescrição gerada pelo Gemini 2.5 Flash:**

A capa do livro apresenta um fundo branco e uma composição gráfica em vermelho e preto, dominada pelo título principal, que se distribui verticalmente e em cores alternadas: o termo MANUAL está em vermelho na parte superior, seguido por DE em preto, e então PRIMEIROS em vermelho. A palavra SOCORROS, também em preto e com letras grandes e em negrito, finaliza o título na parte inferior da composição. À esquerda, uma forma abstrata e curvilínea em vermelho, semelhante a uma fita, ocupa parte do espaço, enquanto uma linha curva e fina em preto atravessa o centro da capa. No lado direito, repousa a ilustração de uma caixa de primeiros socorros branca, de formato retangular e cantos arredondados, que exibe uma cruz vermelha no centro de sua tampa e traz a inscrição ESTOJO DE PRIMEIROS SOCORROS em letras vermelhas pequenas. Na porção inferior da capa, centralizados, estão o selo ou logo do NUBIO em preto estilizado e, à direita dele, a identificação institucional em letras menores e pretas: Núcleo de Biossegurança e, abaixo, Fundação Oswaldo Cruz.

Critérios de avaliação:

- a) **Precisão** - 3 (Bom): A descrição identifica corretamente a maior parte dos elementos presentes na capa, incluindo cores predominantes, disposição geral dos componentes, forma gráfica à esquerda e características do estojo de primeiros socorros. Contudo, apresenta um erro relevante ao afirmar que a palavra “primeiros” está inteiramente em vermelho, quando na imagem ela aparece dividida cromaticamente, com “prim” em preto e “eiros” em vermelho. Além disso, descreve o título como organizado verticalmente, quando na realidade ele está distribuído em linhas horizontais escalonadas.
- b) **Detalhamento** - 3 (Bom): O texto detalha adequadamente os elementos centrais da capa, incluindo tipografia, formas gráficas, ilustração do estojo e identificação institucional. Porém não acerta a disposição exata do título em três linhas e o destaque para a repetição da palavra “socorros”. A descrição também poderia situar com maior precisão o traço curvo preto e explicitar relações proporcionais entre os elementos.
- c) **Clareza** - 4 (Excelente): A redação é fluida, objetiva e coesa, com uso do tempo presente do indicativo. As frases mantêm boa legibilidade e ritmo, favorecendo a compreensão auditiva por usuários com deficiência visual.
- d) **Relevância** - 4 (Excelente): As informações incluídas são pertinentes e direcionadas ao entendimento da composição visual. O texto mantém

foco nos elementos que efetivamente impactam a reconstrução mental da imagem.

- e) **Ordem lógica** - 3 (Bom): A organização geral segue o movimento recomendado do geral para o específico, mas apresenta pequenas inconsistências ao caracterizar a organização vertical do título, o que interfere no fluxo perceptivo.
- f) **Neutralidade** - 4 (Excelente): Não apresenta julgamentos de valor ou interpretações subjetivas.
- g) **Comentário geral** - A audiodescrição demonstra boa qualidade técnica, com clareza linguística, neutralidade adequada e seleção de informações pertinentes. Entretanto, a precisão sofre impacto devido a erros relacionados à tipografia e à disposição gráfica do título, elementos centrais nesta capa específica.

3) **Audiodescrição gerada pelo Copilot:**

A imagem apresenta a capa de um manual em formato vertical, com fundo branco e elementos gráficos em vermelho e preto. No centro superior, em letras maiúsculas e negritadas, está o título “manual DE primeiros socorros”, com a palavra “manual” em vermelho e as demais em preto. Abaixo do título, à esquerda, há um símbolo gráfico composto por uma cruz vermelha estilizada, com bordas arredondadas e preenchimento sólido, representando o símbolo universal da saúde. Sobreposta à cruz, há a ilustração de uma maleta de primeiros socorros em perspectiva frontal, na cor branca com contorno preto, contendo uma cruz vermelha centralizada. Na parte inferior da imagem, centralizado, aparece o logotipo da instituição responsável: a palavra “NUBIO” em letras pretas maiúsculas, seguida da inscrição “Núcleo de Biossegurança Fundação Oswaldo Cruz” em letras menores, também pretas, distribuídas em duas linhas. Não há figuras humanas, cenário ou indicação temporal visível. A composição visual é simétrica e organizada, com contraste adequado entre texto e fundo, facilitando a leitura.

Crerérios de avaliação:

- a) **Precisão** - 1 (Fraco): O título não está representado corretamente: o termo “primeiros” não aparece inteiramente em preto, pois na capa real há divisão cromática evidente, com “prim” em preto e “eios” em vermelho. A descrição inclui ainda uma cruz vermelha estilizada abaixo do título, elemento que não existe na capa. O sistema descreve uma cruz abstrata com bordas arredondadas, preenchimento sólido e com uma maleta sobreposta; no entanto, na imagem real não há cruz vermelha solta na

composição e sim fazendo parte da maleta. A maleta é descrita com contorno preto, mas na capa o contorno é vermelho. Além disso, omite o grande grafismo curvo em vermelho à esquerda e o traço preto sinuoso que organiza visualmente a composição, dois elementos centrais da página.

- b) Detalhamento** - 2 (Regular): Embora descreva alguns componentes textuais e identifique parte da área inferior, o deixa de mencionar elementos fundamentais para a compreensão integral da capa, como o grafismo vermelho de grande porte, a divisão tipográfica da palavra “primeiros”, a presença da palavra “socorros” em tipografia destacada na parte inferior da capa e o traço curvo preto que atravessa a composição. Também não diferencia a repetição do termo “socorros”, nem aborda a disposição em três linhas do título.
- c) Clareza** - 4 (Excelente): Apesar dos problemas factuais, a narrativa é linguisticamente clara, estruturada com frases bem-organizadas, verbos no tempo presente e fluxo compreensível.
- d) Relevância** - 2 (Regular): A audiodescrição mistura elementos realmente presentes com outros inexistentes, o que reduz a relevância das informações.
- e) Ordem lógica** - 3 (Bom): Segue uma sequência relativamente ordenada, começando por uma visão geral e passando a elementos específicos. Porém há uma quebra da fidelidade do fluxo perceptivo.
- f) Neutralidade** - 4 (Excelente): O texto mantém postura descritiva, sem julgamentos ou interpretações subjetivas.
- g) Comentário geral** - A audiodescrição produzida demonstra clareza textual e neutralidade, mas falha no aspecto da fidelidade visual. A descrição dos elementos, altera a posição e estrutura gráfica do título. Em termos acadêmicos, o desempenho seria avaliado como fraco em precisão e apenas regular no conjunto geral, ainda que a redação seja clara e formalmente correta.

4) Audiodescrição gerada pelo DeepSeek V-3.2:

A imagem apresenta a capa de um livro com fundo branco, organizada visualmente de cima para baixo. No terço superior, centralizado, encontra-se o título do livro em letras grandes, na cor verde-escura: "MANUAL DE PERFOS".

Logo abaixo do título, aparece o subtítulo em letras menores, também na cor verde-escura: "SOCOTROS". Segue-se uma linha horizontal fina na cor verde, que separa o título do restante da capa. Abaixo dessa linha, no centro, está o logotipo do Núcleo de Biossegurança, representado por uma figura circular que combina elementos gráficos de folhas verdes dispostas simetricamente, sugerindo um símbolo relacionado à biossegurança e natureza. Sob o logotipo, em duas linhas, consta a inscrição "NÚCLEO de Biossegurança". Na parte inferior da capa, centralizada, está o nome da instituição em letras grandes e na cor azul-escura: "Fundação Oswaldo Cruz". Não há figuras humanas, cenários ou outros elementos ilustrativos além do logotipo descrito. A composição é limpa, estruturada e utiliza um esquema de cores restrito ao verde-escuro, azul-escuro e branco.

Crítérios de avaliação:

- a) Precisão - 1 (Fraco):** A audiodescrição é imprecisa e descreve uma capa que não corresponde de forma alguma à imagem real. O sistema identifica cores inexistentes, afirmando que o título e subtítulo estão em verde-escuro, quando na capa verdadeira o título aparece em vermelho e preto. Além disso, inventa completamente palavras que não existem na capa, como "MANUAL DE PERFOS" e "SOCOTROS", que não estão presentes de forma alguma. Também descreve um logotipo circular com folhas verdes, o que não corresponde ao logotipo quadrado e tipográfico do NUBIO. A menção à instituição "Fundação Oswaldo Cruz" aparece na cor azul-escura, quando na imagem real o texto está em preto e com fonte muito menor. O sistema ignora todos os elementos centrais da capa verdadeira e não há qualquer relação visual entre a descrição fornecida e a imagem real, configurando falha completa no critério de precisão.
- b) Detalhamento - 1 (Fraco):** Embora o texto apresente um nível de detalhamento interno coerente com a imagem que o modelo imagina, ele não detalha absolutamente nada que esteja presente na capa real.
- c) Clareza - 3 (Bom):** A redação é clara, organizada e utiliza frases bem estruturadas em tempo presente. Linguisticamente, o texto é compreensível, fluido e coeso, atendendo às recomendações estilísticas.
- d) Relevância - 1 (Fraco):** A descrição não apresenta informações relevantes sobre a capa real, pois se baseia em elementos inexistentes.
- e) Ordem lógica - 3 (Bom):** Apesar de descrever uma imagem inexistente, o texto segue uma sequência lógica adequada, iniciando pelo geral,

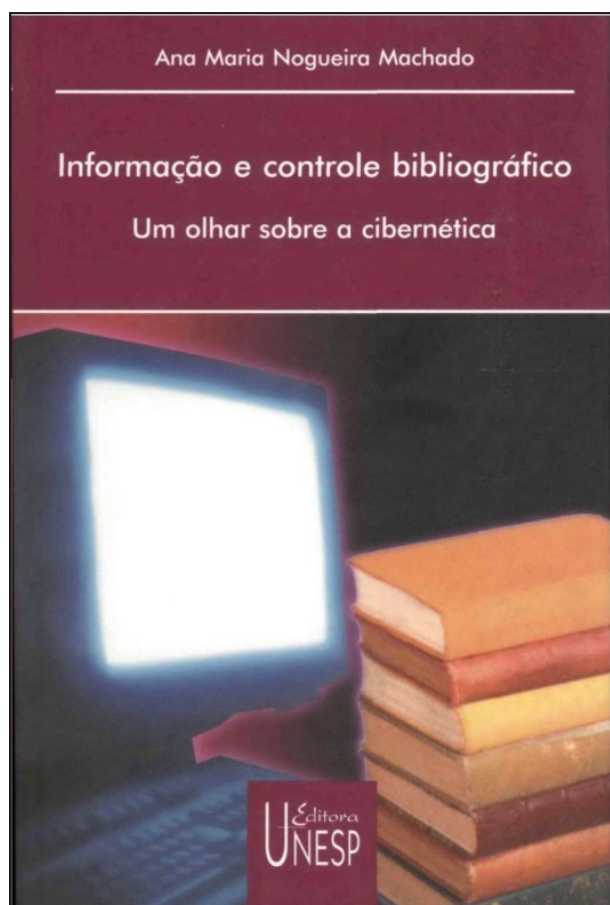
passando para elementos centrais, subtítulo, linha divisória, logotipo e textos institucionais.

- f) **Neutralidade** - 4 (Excelente): O texto mantém neutralidade descritiva, sem adjetivos subjetivos, julgamentos ou inferências interpretativas, mesmo ao descrever componentes inexistentes.
- g) **Comentário geral** - A audiodescrição é completamente incompatível com a capa real do livro. O modelo cria uma imagem inexistente, com título alterado, elementos gráficos inventados, cores incorretas, logotipo fictício e estrutura visual que não corresponde ao objeto analisado.

4.2.2 Análise da capa do livro Informação e controle bibliográfico: um olhar sobre a cibernética (Machado, 2023)

A análise da segunda imagem, correspondente à capa da obra Informação e controle bibliográfico: um olhar sobre a cibernética (Figura 5).

Figura 5 - Capa da obra “Informação e controle bibliográfico: um olhar sobre a cibernética”.



Fonte: Machado (2023).

1) **Audiodescrição gerada pelo ChatGPT-5.1:**

A imagem apresenta a capa de um livro com predominância das cores bordô, preto e tons quentes de amarelo e laranja, organizada em dois blocos principais: a faixa superior com textos e a área inferior com imagens ilustrativas. No topo da capa, centralizado sobre um fundo bordô uniforme, aparece o nome da autora “Ana Maria Nogueira Machado” em letras brancas. Abaixo desse nome, uma linha horizontal branca separa o cabeçalho do bloco com o título. Logo em seguida, também centralizado e em letras brancas, está o título “Informação e controle bibliográfico”; abaixo dele, em fonte menor e igualmente branca, surge o subtítulo “Um olhar sobre a cibernética”. A metade inferior da capa apresenta uma composição gráfica com predominância de tons escuros. À esquerda, ocupa boa parte do espaço a imagem de um monitor de computador antigo, visto de frente, com a tela retangular emitindo uma luz branca intensa que cria um halo luminoso sobre o fundo. Abaixo da tela, apenas parcialmente visível, há um teclado escuro desfocado. À direita do monitor, vê-se uma pilha de seis livros dispostos verticalmente, empilhados de forma alinhada; as capas apresentam tons terrosos como amarelo, laranja, marrom e vermelho, com lombadas lisas e sem textos visíveis. Na porção inferior central, sobre a divisão entre as imagens e o fundo bordô, aparece o logotipo da “Editora UNESP” em letras brancas, inscritas em um retângulo bordô. Não há presença de figuras humanas na composição. A capa conjuga elementos analógicos e digitais para representar visualmente a temática da obra, mantendo contraste suficiente entre texto e fundo para leitura adequada.

Crerios de avaliaão:

- a) Precisoão - 4 (Excelente):** A audiodescrião corresponde aos elementos visíveis na capa. Os textos sã transcritos corretamente, incluindo nome da autora, título, subtítulo e editora. As cores predominantes sã descritas de modo compatível com a imagem, e os objetos — monitor de computador, teclado e pilha de livros — sã representados de acordo com suas formas, posioões e característias observáveis. Apenas a contagem dos livros empilhados, a audiodescrião indica seis, porém sã sete. Nã há inserão de elementos inexistentes, nem omissão de elementos centrais da composioão, garantindo alta fidelidade à imagem real.
- b) Detalhamento - 4 (Excelente):** O nível de detalhamento é adequado, contemplando forma, cor, posioão, textura sugerida, organizaão espacial e relaões entre os objetos. O monitor antigo é descrito com precisoão, incluindo o brilho da tela; a pilha de livros é quantificada e caracterizada cromaticamente; os blocos superior e inferior da capa sã descritos de maneira completa; o logotipo da editora é identificado corretamente. O

detalhamento é suficiente para permitir reconstrução mental da imagem, sem extrapolações.

- c) **Clareza** - 4 (Excelente): O texto está redigido com objetividade, clareza e fluidez, utilizando o tempo presente do indicativo, conforme recomendam as normas de acessibilidade.
- d) **Relevância** - 4 (Excelente): Todas as informações apresentadas são pertinentes e contribuem para a compreensão da capa. Não há excesso descritivo nem detalhes irrelevantes.
- e) **Ordem lógica** - 4 (Excelente): A organização segue rigorosamente o movimento recomendado: visão geral da composição, seguida da descrição de elementos textuais superiores e, em seguida, dos elementos gráficos inferiores, passando do geral ao específico e mantendo coerência espacial. A sequência é natural, fluida e reproduz o percurso típico do olhar.
- f) **Neutralidade** - 4 (Excelente): Não há julgamentos de valor, inferências subjetivas, interpretações simbólicas ou atribuições de significado não verificáveis. Os elementos são descritos de forma objetiva e factual, respeitando integralmente o princípio da neutralidade.
- g) **Comentário geral** - A audiodescrição apresenta excelente qualidade técnica, atendendo de forma consistente a todos os critérios avaliativos e às orientações normativas.

2) **Audiodescrição gerada pelo Gemini 2.5 Flash:**

A capa do livro, disposta verticalmente, apresenta uma paleta de cores dominada por tons de vinho e marrom, e está dividida entre o cabeçalho e uma ilustração principal, estabelecendo um contraste entre texto e imagem. Na parte superior, sobre um fundo cor de vinho, estão inscritos em branco o nome da autora, Ana Maria Nogueira Machado, e abaixo, o título principal, Informação e controle bibliográfico, seguido pelo subtítulo, Um olhar sobre a cibernética. A seção inferior é ocupada por uma ilustração em tons escuros de azul e preto, onde se destacam dois elementos centrais: à esquerda, um monitor de computador antigo de tubo, com a tela branca-azulada intensamente iluminada, emitindo um forte brilho que é a principal fonte de luz do ambiente; à direita, adjacente e ligeiramente à frente do monitor, uma pilha vertical de cinco livros de capa dura em tons de laranja e marrom repousa. Sobre a junção dos elementos da ilustração, na parte inferior central, encontra-se o selo da editora, um retângulo cor de vinho com a inscrição em branco, Editora UNESP, com a letra 'U' em destaque.

Cr terios de avalia  o:

- a) Preci  o - 4 (Excelente):** O Gemini identifica corretamente a paleta de cores (tons de vinho, marrom, azul e preto), o posicionamento do nome da autora, do t tulo e do subt tulo, bem como a divis  o superior-inferior da composi   o. Tamb  m descreve adequadamente o monitor de tubo com luminosidade intensa e a pilha de livros   direita, incluindo a predomin  ncia de tons quentes. O selo da Editora UNESP   identificado de maneira correta, incluindo o ret  ngulo vinho e o destaque da letra “U”. A maior parte dos elementos visuais est  representada com precisi  o, sem inser   o de conte  do inexistente ou omiss  es significativas. A  nica varia   o m  nima — n  mero de livros (cinco na descri   o, sete na imagem) — n  o compromete a compreens  o geral e se enquadra como um deslize de baixa gravidade.
- b) Detalhamento - 4 (Excelente):** O n  vel de detalhamento atende  s exig  ncias das normas: os elementos textuais s  o mencionados no momento em que aparecem na composi   o, a paleta crom  tica   explicitada com precisi  o, a natureza do monitor   corretamente descrita como “antigo” e de tubo, e a pilha de livros recebe caracteriza   o crom  tica e posicional coerente. Tamb  m   descrito rela    o espacial entre monitor e livros, al  m do papel da tela como fonte de luz, o que contribui para a reconstru   o mental da cena sem extrapolar o que   observ  vel. O detalhamento   completo e n  o apresenta excessos.
- c) Clareza - 4 (Excelente):** A reda   o   fluida, coesa e objetiva, com uso adequado do tempo presente do indicativo.
- d) Relev  ncia - 4 (Excelente):** Todas as informa    es s  o pertinentes para compreender a composi   o visual e n  o h   elementos desnecess  rios nem interpreta    es simb  licas. O foco permanece nos componentes estruturais da capa — textos, imagem, cores, disposi   o espacial e elementos editoriais — sem digress   es ou adjetiva   o subjetiva. A descri   o n  o omite informa    es essenciais e n  o apresenta redund  ncia.
- e) Ordem l  gica - 4 (Excelente):** A sequ  ncia da descri   o segue com precisi  o o percurso visual esperado: inicia com uma vis  o geral da

disposição vertical e da paleta, avança para a parte superior com os textos, segue para a área inferior com os elementos ilustrados e finaliza com o selo da editora.

- f) Neutralidade - 4 (Excelente):** A descrição é inteiramente objetiva. O Gemini utiliza expressões descritivas (“monitor de computador antigo de tubo”, “forte brilho”, “tons de laranja e marrom”) sem atribuir julgamentos ou sentidos metafóricos.
- g) Comentário geral -** A audiodescrição apresenta desempenho técnico de alto nível, com adesão consistente às normas de audiodescrição e às boas práticas de acessibilidade informacional.

3) Audiodescrição gerada pelo Copilot:

A imagem apresenta a capa de um livro em orientação vertical, com fundo predominantemente escuro e elementos visuais contrastantes em tons claros. No canto superior esquerdo, aparece o nome da autora “Ana Maria Nogueira Machado” em letras brancas. Abaixo, centralizado e em fonte maior, está o título “Informação e controle bibliográfico”, seguido do subtítulo “Um olhar sobre a cibernética”, ambos também em letras brancas. À esquerda da composição, há a imagem de um monitor de computador com tela iluminada em tons azulados, emitindo um brilho que se destaca contra o fundo escuro. À direita, vê-se uma pilha de livros dispostos verticalmente, com lombadas voltadas para o observador, representando o universo da informação tradicional. Os livros têm cores neutras e textura fosca, contrastando com o brilho da tela digital. Na parte inferior central da capa, está o logotipo da editora “Editora UNESP”, em letras brancas. Não há figuras humanas, personagens ou indicações temporais visíveis. A disposição dos elementos sugere uma relação entre o mundo digital e o conhecimento impresso, com composição equilibrada e contraste visual que favorece a leitura.

Critérios de avaliação:

- a) Precisão - 3 (Bom):** A audiodescrição corresponde de forma ampla à imagem real, identificando corretamente o nome da autora, o título, o subtítulo, a localização geral dos textos, a presença do monitor iluminado e a pilha de livros à direita. Entretanto, há duas imprecisões importantes: a descrição afirma que o fundo da capa é “predominantemente escuro”, quando na verdade somente a área inferior possui tons escuros; a parte superior, onde estão os textos, é claramente bordô, não preta ou neutra. Além disso, a descrição menciona que os livros possuem “cores neutras”, o que não corresponde exatamente à imagem, já que eles apresentam tons quentes e variados (amarelo, laranja, vermelho, marrom). Apesar

dessas inconsistências cromáticas, a maior parte dos elementos está representada de forma correta, justificando a nota bom.

- b) Detalhamento** - 3 (Bom): A descrição identifica o monitor, a tela iluminada, a pilha de livros e os elementos textuais, mas apresenta detalhamento apenas moderado. Não informa o número aproximado de livros, não descreve a paleta cromática com precisão, não cita o contraste bordô/branco da seção superior e não distingue adequadamente o retângulo vinho do logotipo da editora. O detalhamento espacial também é discreto, sem referência à separação nítida entre a faixa superior textual e a área inferior ilustrada. Apesar disso, o essencial está presente, justificando a avaliação como bom.
- c) Clareza** - 4 (Excelente): O texto é claro, direto e fácil de acompanhar. A redação utiliza frases bem estruturadas e o tempo verbal adequado, sem ambiguidades.
- d) Relevância** - 4 (Excelente): Todas as informações apresentadas dizem respeito à capa e contribuem para a compreensão da composição visual. Não há excesso de dados nem inclusão de interpretações fantasiosas. A descrição é objetiva e foca nos elementos principais da cena.
- e) Ordem lógica** - 4 (Excelente): A sequência segue adequadamente a estrutura recomendada: inicia com visão geral, depois apresenta o bloco superior textual, seguida da área inferior ilustrada, encerrando com o logotipo. A progressão do geral ao específico e do topo para a base da capa é mantida de forma coerente e fluida.
- f) Neutralidade** - 3 (Bom): A descrição é neutra, mas um trecho compromete parcialmente esse critério: “representando o universo da informação tradicional”, o que constitui uma inferência interpretativa e não algo observável. Também “sugere uma relação entre o mundo digital e o conhecimento impresso”, o que extrapola a função da audiodescrição ao propor uma leitura simbólica da imagem.
- g) Comentário geral** - A audiodescrição é clara, coesa e abrange os elementos essenciais da capa. Apresenta bom desempenho em precisão, com algumas inconsistências cromáticas e duas inferências interpretativas que reduzem a neutralidade. Ainda assim, a informação

visual principal é corretamente transmitida e a estrutura respeita as normas de audiodescrição.

4) Audiodescrição gerada pelo DeepSeek V-3.2:

A imagem apresenta a capa de um livro com fundo de cor laranja-claro, organizada visualmente em blocos de texto centralizados. Na parte superior, em letras grandes e na cor azul-escura, consta o nome da autora: "ANA MARIA NOGUEIRA MACHADO". Abaixo, centralizado e em fonte de tamanho maior, está o título principal em letras maiúsculas na cor azul-escura: "INFORMAÇÃO E CONTROLE BIBLIOGRÁFICO". Segue-se, na linha seguinte, o subtítulo em letras menores, também em azul-escuro: "Um olhar sobre a cibernética". Na parte inferior da capa, centralizada, aparece o nome da editora em letras grandes, em caixa alta e na cor amarela: "UNESP". Não há figuras humanas, ilustrações, logotipos ou outros elementos gráficos visíveis, somente o texto descrito sobre o fundo liso laranja-claro. A disposição é simples e linear, com alinhamento central.

CrITÉRIOS de avaliação:

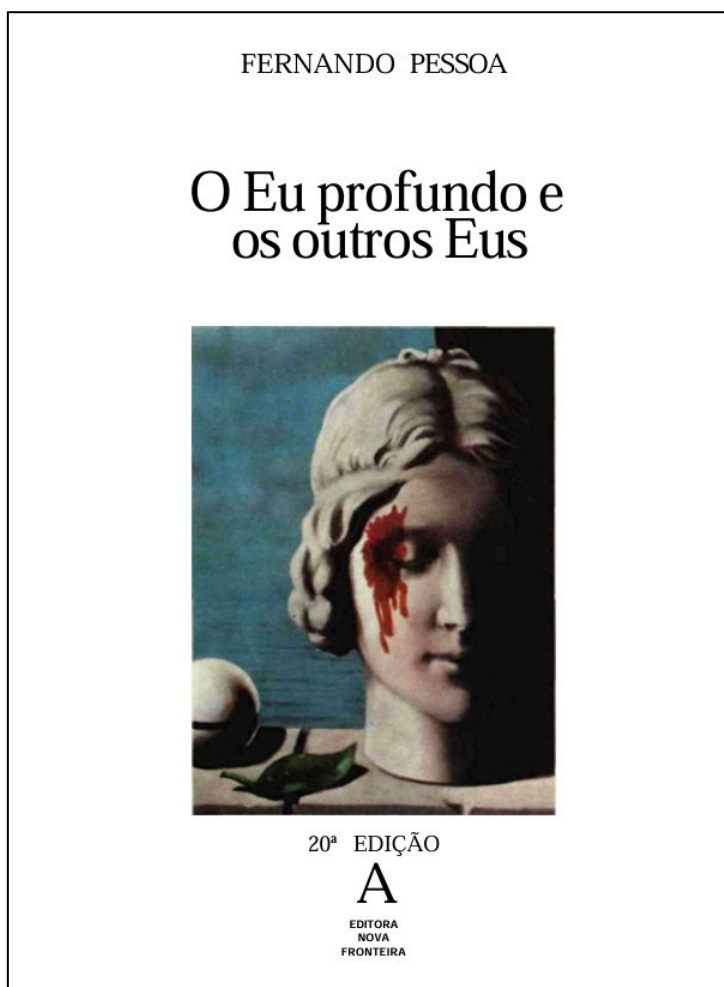
- a) Precisão - 1 (Fraco):** A descrição não corresponde ao conteúdo visual da capa real. Ela apresenta um fundo laranja-claro inexistente, quando a capa verdadeira possui um bloco superior bordô e uma área inferior escura com uma ilustração. Também afirma que não há imagens, ilustrações ou logotipos, quando a capa real contém um monitor de computador iluminado, uma pilha de livros e o logotipo da “Editora UNESP” em destaque. Além disso, as cores do texto foram descritas de maneira incorreta — o DeepSeek afirma que são azul-escuras e amarelas, enquanto na imagem são brancas. A estrutura visual também é representada de forma equivocada: a capa não é minimalista, nem composta somente por texto centralizado. A descrição, portanto, não possui correspondência com a imagem real, justificando a nota mínima.
- b) Detalhamento - 1 (Fraco):** O detalhamento, embora organizado internamente, descreve uma capa completamente diferente, composta apenas por texto sobre fundo laranja. Não há qualquer menção aos elementos essenciais presentes na capa verdadeira: monitor, brilho da tela, teclado, pilha de livros, mistura de tons escuros e bordô, divisão entre áreas ilustrada e textual, nem ao formato do logotipo. Como os elementos reais não são descritos, o detalhamento é inexistente em relação à imagem analisada.

- c) **Clareza** - 4 (Excelente): Apesar de descrever uma capa incorreta, o texto apresenta redação clara, frases bem estruturadas, tempo verbal adequado e vocabulário simples e funcional. Linguisticamente, a audiodescrição é coesa e compreensível, o que justifica nota máxima neste critério.
- d) **Relevância** - 1 (Fraco): A descrição não oferece informações relevantes sobre a capa real, pois se baseia inteiramente em elementos inexistentes. Assim, mesmo com estrutura correta, o conteúdo não é útil para um usuário que depende da audiodescrição para compreender a imagem. A ausência dos elementos centrais da capa compromete totalmente o critério de relevância.
- e) **Ordem lógica** - 3 (Bom): A ordem interna do texto é coerente: inicia com visão geral, segue para o nome da autora, título, subtítulo e editora. A linearidade do fluxo está tecnicamente adequada. Entretanto, como a descrição não corresponde à imagem real, a eficácia dessa organização é limitada, ainda que a estrutura narrativa seja formalmente correta.
- f) **Neutralidade** - 4 (Excelente): O texto é neutro, objetivo e não apresenta julgamentos, interpretações subjetivas ou atribuições metafóricas. A linguagem é descritiva e técnica, sem inferências interpretativas, mantendo aderência formal ao critério de neutralidade.
- g) **Comentário geral** - A audiodescrição apresenta problemas graves de precisão e relevância, pois descreve uma capa totalmente diferente da imagem real. Embora a redação seja clara e respeite princípios de organização textual, o conteúdo não corresponde a nenhum elemento visual da capa analisada, o que compromete sua utilidade para fins de acessibilidade.

4.2.3 Análise da capa do livro O Eu profundo e os outros Eus (Pessoa, 1913)

A análise da terceira imagem, correspondente à capa da obra O Eu profundo e os outros Eus (Figura 6).

Figura 6 - Capa da obra “O Eu profundo e os outros Eus”.



Fonte: Pessoa (1913).

1) Audiodescrição gerada pelo ChatGPT-5.1:

A imagem apresenta a capa de um livro disposta em orientação vertical, com fundo branco e elementos textuais centralizados na parte superior. No topo, aparece o nome do autor “FERNANDO PESSOA” em letras maiúsculas e pretas. Abaixo, também centralizado e em letras pretas de maior tamanho, está o título “O Eu profundo e os outros Eus”. Na região central da capa, há uma imagem retangular que ocupa cerca de um terço do espaço total. Essa imagem mostra a representação de uma escultura clássica em forma de busto feminino, com traços simétricos e cabelos ondulados, vista frontalmente. A escultura é dividida pela iluminação em duas metades, sendo a esquerda iluminada e a direita em sombra. Sobre o lado esquerdo do rosto, abaixo do olho, há um escorrimento de tinta vermelha que desce em direção à bochecha. O fundo da imagem apresenta céu azul e objetos dispostos sobre uma superfície horizontal: à esquerda, há uma esfera clara parcialmente visível; à frente, próximo à base da escultura, encontra-se uma folha verde apoiada sobre a superfície ocre. Abaixo da imagem ilustrada, centralizado, lê-se “20ª EDIÇÃO”. Logo abaixo, aparece a letra “A” em destaque, seguida das linhas “EDITORA NOVA FRONTEIRA”, todas em letras pretas e centralizadas. Não há figuras humanas reais, apenas a representação

escultórica mencionada. A composição é simétrica e apresenta forte contraste entre o fundo branco e os elementos gráficos e textuais.

CrITÉRIOS de avaliação:

- a) Precisão** - 4 (Excelente): A descrição corresponde à capa apresentada. Todos os textos — nome do autor, título, indicação de edição, letra “A” e nome da editora — foram transcritos corretamente, sem erros de leitura. Os elementos visuais centrais, como o busto clássico, a divisão de luz e sombra no rosto, o escorrimento vermelho abaixo do olho, a esfera clara ao fundo, a folha verde sobre a superfície e o céu azul, foram identificados com precisão e sem inferências indevidas.
- b) Detalhamento** - 4 (Excelente): O nível de detalhamento é adequado, a descrição informa a posição dos textos, o tipo de fundo, as características da escultura, a divisão luminosa entre as metades do rosto, os objetos presentes na composição, a disposição espacial, as cores predominantes e a estrutura geral da capa. Os elementos de fundo e plano foram descritos sem exagero, mantendo equilíbrio entre completude e objetividade.
- c) Clareza** - 4 (Excelente): O texto é claro, direto e organizado, com frases bem estruturadas e predomínio do tempo presente do indicativo, como orienta a norma. A linguagem é acessível e fluida, permitindo compreensão auditiva sem esforço excessivo. Não há ambiguidades nem sobrecarga informacional.
- d) Relevância** - 4 (Excelente): Todas as informações fornecidas são pertinentes para a compreensão da capa. Nenhum elemento secundário foi omitido, e não há informações irrelevantes ou ornamentações desnecessárias.
- e) Ordem lógica** - 4 (Excelente): A sequência descritiva segue o percurso típico da percepção visual: apresenta primeiro uma visão geral, depois os textos superiores, em seguida a imagem central, e por fim os textos inferiores. A ordem do geral para o específico e do alto para baixo mantém coerência com as orientações.
- f) Neutralidade** - 4 (Excelente): A descrição mantém postura objetiva e isenta, sem julgamentos, análises simbólicas ou interpretações sobre

intenções artísticas. Elementos potencialmente sensíveis, como o escorrimento vermelho no rosto da escultura, foram descritos somente como forma e cor, sem inferências metafóricas.

- g) Comentário geral** - A audiodescrição apresenta excelente aderência às normas brasileiras e às boas práticas de acessibilidade. Demonstra precisão, organização lógica, clareza textual e detalhamento adequado, permitindo a compreensão plena e fiel da capa.

2) Audiodescrição gerada pelo Gemini 2.5 Flash:

A capa do livro, de formato retangular vertical, é predominantemente branca e concentra seus elementos em uma faixa central, de cima para baixo. No topo, centralizado e em letras pretas, lê-se o nome do autor: FERNANDO PESSOA. Abaixo, em fonte maior, está o título da obra: O Eu profundo e os outros Eus. O centro da capa é ocupado por uma imagem retangular emoldurada, que retrata a cabeça de uma estátua feminina, possivelmente de mármore, com um fundo azul-turquesa na metade superior, que simula um céu, e um horizonte escuro. A estátua está voltada para a direita, com a metade direita do rosto na sombra e a metade esquerda iluminada; do olho esquerdo da estátua, que está virado para o observador, escorrem manchas de cor vermelho-sangue pela face e parte do pescoço. No canto inferior esquerdo da imagem, repousam um objeto esférico branco opaco e uma folha verde de formato oval, sobre o horizonte escuro. Abaixo da imagem, alinhado ao centro, está a inscrição 20ª EDIÇÃO, seguida por uma grande letra A. Por fim, na base da capa, também centralizada, lê-se o nome da editora em duas linhas: EDITORA e NOVA FRONTEIRA.

Critérios de avaliação:

- a) Precisão** - 4 (Excelente): A audiodescrição corresponde à capa real do livro. Todos os textos são transcritos corretamente, incluindo o nome do autor, o título, a indicação da edição, a letra “A” e o nome da editora. Os elementos visuais centrais — o busto feminino clássico, a divisão de luz e sombra no rosto, o escorrimento vermelho abaixo do olho, o fundo azul, a esfera branca e a folha verde — foram descritos de forma compatível com o que está visível na imagem. A direção do rosto da estátua, a iluminação e a composição espacial são representadas sem erros perceptíveis. Não há inserção de elementos inexistentes nem omissões significativas, o que garante alta precisão.
- b) Detalhamento** - 4 (Excelente): A descrição traz detalhes pertinentes: formato vertical da capa, predominância de fundo branco, posição dos textos, estrutura verticalizada da composição, cores predominantes da

ilustração, direção e iluminação no busto, textura sugerida da estátua, descrição da esfera e da folha, além da relação espacial entre os objetos. O detalhamento é suficiente e equilibrado, permitindo ao usuário com deficiência visual reconstruir mentalmente a cena, sem extrapolar o conteúdo visível ou incluir interpretações indevidas.

- c) **Clareza** - 4 (Excelente): O texto está bem-organizado, fluido e estruturado com frases claras, objetivas e coesas. O uso do tempo presente do indicativo está consistente e favorece a compreensão auditiva. Não há ambiguidades, excessos sintáticos ou termos que prejudiquem a acessibilidade.
- d) **Relevância** - 4 (Excelente): Todas as informações são úteis para a compreensão visual da capa. Não há inserção de elementos irrelevantes nem supressão de conteúdos importantes. A descrição concentra-se no essencial — textos, imagem central, elementos gráficos e composição espacial — de modo funcional e direto.
- e) **Ordem lógica** - 4 (Excelente): A sequência segue a estrutura recomendada pela NBR 16452: inicia com uma visão geral, localiza os elementos textuais superiores, descreve o bloco ilustrativo central, e finaliza com os textos inferiores. A progressão do geral para o específico e do alto para baixo é precisa e organizada, facilitando o entendimento.
- f) **Neutralidade** - 4 (Excelente): A audiodescrição mantém total neutralidade. Embora descreva um elemento visual impactante (o escorrimento vermelho da estátua), a redação o trata somente como forma, cor e posição, sem atribuir simbolismos. O uso da expressão “possivelmente de mármore” é aceitável, pois não implica interpretação subjetiva, mas uma observação baseada em características visuais reconhecíveis em esculturas clássicas. Não há julgamentos de valor ou inferências emocionais.
- g) **Comentário geral** - A audiodescrição apresenta desempenho excelente em todos os critérios, demonstrando domínio técnico, estrutura organizada e aderência às normas brasileiras de audiodescrição. O texto equilibra precisão visual, objetividade e clareza comunicacional, proporcionando uma compreensão completa e fiel da capa para usuários

com deficiência visual. Trata-se de uma descrição bem construída e totalmente adequada para fins acadêmicos e de acessibilidade.

3) **Audiodescrição gerada pelo Copilot:**

A imagem apresenta a capa de um livro em orientação vertical, com fundo predominantemente azul e composição visual de estilo surrealista. No centro da capa, destaca-se um busto clássico de mármore representando uma cabeça feminina, voltada ligeiramente para a esquerda, com olhos fechados e expressão serena. Do canto externo do olho direito escorre uma linha vermelha, semelhante a sangue, que contrasta com o tom claro da escultura. Ao fundo, à direita, vê-se um céu azul uniforme e uma esfera branca, com aparência de ovo, posicionada sobre uma superfície plana. À frente da esfera, no canto inferior direito, há um objeto verde em forma de folha ou pétala. No topo da imagem, em letras maiúsculas pretas, aparece o nome do autor “FERNANDO PESSOA”. Abaixo, centralizado, está o título “O Eu profundo e os outros Eus”, em letras pretas e caixa alta e baixa. Na parte inferior esquerda, em letras pequenas pretas, lê-se “20ª EDIÇÃO”. No canto inferior direito, está o logotipo da editora “EDITORIA NOVA FRONTEIRA”, distribuído em quatro linhas, com todas as palavras em letras maiúsculas pretas. Não há figuras humanas, cenário definido ou indicação temporal visível. A disposição dos elementos é simétrica e o contraste entre cores claras e escuras favorece a leitura e a percepção visual.

Critérios de avaliação:

- a) **Precisão** - 3 (Bom): A maioria dos elementos da capa foi descrita corretamente, incluindo o busto feminino, o escorrimento vermelho, a esfera branca, a folha verde, o nome do autor, o título e o crédito editorial. No entanto, a descrição apresenta duas imprecisões relevantes. A primeira é a afirmação de que o busto está voltado “ligeiramente para a esquerda”, quando na imagem o rosto está levemente voltado para a direita do observador (ombro esquerdo da escultura recuado e queixo orientado para a direita). A segunda é a indicação de que determinados textos estão localizados em “parte inferior esquerda” e “canto inferior direito”, quando, na capa original, o “20ª EDIÇÃO” está centralizado logo abaixo da imagem e o logotipo da editora ocupa a região inferior central, não um canto. Apesar dessas imprecisões, a descrição geral corresponde ao conteúdo visual, justificando nota bom.
- b) **Detalhamento** - 3 (Bom): O detalhamento cobre elementos essenciais da composição, como a orientação vertical, o estilo surrealista, a presença do busto, os objetos no plano inferior e a estrutura textual. Entretanto, faltam detalhes importantes para aderir à NBR 16452:2016, como a clara

divisão entre a metade iluminada e a metade sombreada do rosto, o cenário de horizonte azul-turquesa na metade superior da imagem, e o fato de que o busto ocupa somente a imagem central retangular inserida na capa, e não toda a capa. Tais detalhes não inviabilizam a compreensão, mas reduzem o nível de minúcia esperado para uma audiodescrição acadêmica.

- c) **Clareza** - 4 (Excelente): A redação é clara, coesa e estruturada, com frases bem construídas, uso consistente do tempo presente e vocabulário adequado. A descrição é facilmente compreensível em formato sonoro, favorecendo o processamento auditivo pelo usuário. Não há ambiguidades textuais significativas.
- d) **Relevância** - 4 (Excelente): A descrição inclui apenas informações que contribuem para a compreensão da capa, sem elementos supérfluos ou desvios interpretativos extensos. Os detalhes mencionados são adequados e contribuem para a construção mental da imagem por parte do usuário.
- e) **Ordem lógica** - 4 (Excelente): O texto segue a sequência recomendada pela NBR 16452: visão geral inicial, apresentação dos textos superiores, descrição da imagem central e, por fim, os textos inferiores. A ordem espacial é coerente e facilita a compreensão global, apesar de alguns erros localizados apontados no critério de precisão.
- f) **Neutralidade** - 4 (Excelente): A descrição mantém postura objetiva e não atribui significados simbólicos ou emocionais à imagem. Mesmo ao mencionar elementos que poderiam sugerir interpretações subjetivas — como o escorrimento vermelho — o faz de forma descritiva, sem inferências interpretativas. A neutralidade, portanto, é preservada.
- g) **Comentário geral** - A audiodescrição apresenta boa qualidade e atende à maior parte dos requisitos normativos, com clareza comunicacional e foco adequado nos elementos essenciais. No entanto, algumas imprecisões de direção, localização espacial e descrição cromática reduzem levemente sua aderência plena ao modelo de excelência. Trata-se de uma descrição funcional e tecnicamente coerente, mas que requer ajustes para alcançar rigor máximo no contexto acadêmico de avaliação de audiodescrições.

4) **Audiodescrição gerada pelo DeepSeek V-3.2:**

A imagem apresenta a capa de um livro com fundo de cor bege-claro, organizada visualmente em blocos de texto centralizados. Na parte superior, em letras grandes, maiúsculas e na cor azul-escura, está o nome do autor: "FERNANDO PESSOA". Abaixo, centralizado e em fonte de tamanho muito maior e na cor vermelha, encontra-se o título: "O Eu profundo e os outros Eus". Entre o título e as informações seguintes, há uma linha horizontal decorativa formada por uma sequência de pequenos círculos azuis. Na parte inferior da capa, centralizado, aparece o indicativo "20ª EDIÇÃO" em letras azuis. Logo abaixo, a letra "A", também em azul, isolada em sua linha. Na linha final, em letras grandes e na cor azul-escura, está o nome da editora: "A EDITORA NOVA FRONTEIRA". Não há figuras humanas, ilustrações, retratos ou outros elementos gráficos visíveis, apenas o texto e a linha decorativa descritos sobre o fundo liso bege-claro. A disposição é simples, linear e com alinhamento central.

Critérios de avaliação:

- a) **Precisão** - 1 (Fraco): A descrição não corresponde à capa real do livro. O fundo não é bege-claro, mas branco com uma grande imagem central emoldurada. O título não está em vermelho; está em letras pretas. Não existe qualquer linha horizontal decorativa de círculos azuis. A capa real contém uma ilustração central de um busto clássico com um escorrimento vermelho abaixo do olho, um céu azul e objetos sobre uma superfície — elementos completamente ausentes da descrição do DeepSeek. Além disso, o texto da editora não aparece como "A EDITORA NOVA FRONTEIRA" em azul, e sim como "EDITORIA NOVA FRONTEIRA" em preto, sob a letra "A" centralizada. Como a descrição ignora todos os elementos visuais essenciais e inventa características inexistentes, a precisão é mínima.
- b) **Detalhamento** - 1 (Fraco): Ainda que a estrutura interna seja organizada, o detalhamento não se aplica ao conteúdo real da capa. A audiodescrição descreve somente textos genéricos e um suposto fundo bege, omitindo completamente a imagem central, que é o elemento mais importante da capa, e alterando as cores, fontes e disposições. Assim, não existe detalhamento real da imagem analisada, apenas de uma capa imaginada, o que justifica a nota mínima.
- c) **Clareza** - 4 (Excelente): O texto é claro, direto e bem escrito, com frases estruturadas de modo coerente e de fácil compreensão. A redação segue o tempo presente do indicativo e apresenta organização sintática

adequada. A clareza linguística, entretanto, não compensa as falhas de fidelidade visual.

- d) Relevância** - 1 (Fraco): A relevância é comprometida porque as informações apresentadas não correspondem à imagem real. Mesmo sendo claras, elas não ajudam o usuário com deficiência visual a compreender a capa. Por tratar-se de uma audiodescrição completamente dissociada do conteúdo verdadeiro, o critério de relevância não é atendido.
- e) Ordem lógica** - 3 (Bom): A ordem geral da descrição (visão ampla, nome do autor, título, linha decorativa, elementos inferiores) está bem estruturada e segue um fluxo lógico. Contudo, essa estrutura não contribui para o entendimento da capa real, ainda que, narrativamente, esteja organizada. Por isso, recebe nota intermediária.
- f) Neutralidade** - 4 (Excelente): A descrição mantém neutralidade, sem julgamentos subjetivos, interpretações simbólicas ou inferências sobre sentimentos ou significados. Ela descreve somente aquilo que acredita ver, com linguagem objetiva — mesmo que a interpretação visual esteja incorreta.
- g) Comentário geral** - A audiodescrição é linguisticamente clara e formalmente organizada, mas falha completamente em seu propósito principal: representar a capa real do livro. O modelo cria uma capa inexistente, alterando cores, removendo a imagem central, modificando textos e adicionando elementos decorativos imaginários. Assim, embora seja tecnicamente bem escrita, é totalmente inadequada para fins de acessibilidade. Em termos acadêmicos e segundo a ABNT NBR 16452:2016, configura uma audiodescrição inválida, com desempenho insuficiente nos critérios fundamentais de precisão, detalhamento e relevância.

5 DISCUSSÕES

A análise comparativa das audiodescrições produzidas por quatro modelos de IA (ChatGPT-5.1, Gemini 2.5 Flash, Copilot e DeepSeek V-3.2) para as três capas avaliadas — Manual de Primeiros Socorros, Informação e controle bibliográfico: um olhar sobre a cibernética e O Eu profundo e os outros Eus — evidencia um padrão relativamente estável de desempenho, com três grupos bem definidos: modelos com alta aderência às normas de audiodescrição (ChatGPT e Gemini), um modelo intermediário (Copilot) e um modelo com desempenho insuficiente do ponto de vista da acessibilidade (DeepSeek).

De modo geral, ChatGPT-5.1 e Gemini 2.5 Flash apresentaram os melhores resultados de forma consistente nas três imagens. Em praticamente todos os casos, ambos alcançaram conceito “Excelente” para clareza, relevância, ordem lógica e neutralidade. As audiodescrições produzidas por esses modelos adotaram linguagem objetiva, uso adequado do tempo presente do indicativo, boa coesão frasal e organização sequencial compatível com o percurso visual recomendado pela ABNT NBR 16452:2016 (visão geral, textos superiores, elementos gráficos centrais e textos inferiores). Do ponto de vista da precisão e do detalhamento, ChatGPT e Gemini oscilaram entre “Bom” e “Excelente”. Nas capas em que houve pequenas falhas — por exemplo, número exato de livros empilhados ou divisão cromática da palavra “primeiros” na capa do Manual de Primeiros Socorros — os erros foram pontuais, sem comprometer a compreensão global da imagem. Em síntese, esses dois modelos foram capazes de gerar audiodescrições tecnicamente robustas, com boa correspondência visual e suficiente riqueza descritiva para permitir a construção mental das capas por pessoas com deficiência visual.

O Copilot apresentou um desempenho intermediário. Em todas as capas, o modelo manteve alta clareza textual, relevância global das informações, boa ordem lógica e neutralidade adequada na maior parte dos trechos, o que resultou em avaliações elevadas nesses critérios. No entanto, a dimensão visual foi mais problemática: a precisão variou de “Bom” a “Fraco”, com erros na representação de cores, na localização espacial de elementos e, em alguns casos, na estrutura gráfica do título e de ícones. Em Informação e controle bibliográfico, por exemplo, o Copilot descreveu o fundo como “predominantemente escuro”,

desconsiderando a faixa bordô superior, e classificou as cores dos livros como neutras, quando na realidade predominam tons quentes. Em O Eu profundo e os outros Eus, houve equívocos na direção do rosto da escultura e na localização do texto “20ª EDIÇÃO” e do logotipo da editora. Além disso, em uma das descrições surgiram interpretações simbólicas (“representando o universo da informação tradicional”, “sugere uma relação entre o mundo digital e o conhecimento impresso”), o que fere parcialmente o princípio de neutralidade exigido pela NBR 16452:2016. Ainda assim, o Copilot conseguiu descrever os elementos centrais de forma globalmente coerente, configurando, no conjunto, um desempenho entre regular e bom, sobretudo útil se acompanhado de revisão humana especializada.

O DeepSeek V-3.2, por sua vez, apresentou o pior desempenho entre os modelos analisados. Em todas as capas, suas audiodescrições foram classificadas como “Fraco” em precisão, detalhamento e relevância. O padrão observado foi de criação de capas praticamente fictícias: fundos em cores inexistentes (laranja, bege), títulos alterados (“MANUAL DE PERFOS”, “SOCOTROS”), supressão completa das imagens centrais (monitores, livros e busto escultórico) e adição de elementos decorativos que não constam nas capas reais (linhas de círculos azuis, logotipos estilizados com folhas verdes). Em outras palavras, não somente omitem componentes essenciais, como substituem a cena por outra composição, o que as torna inviáveis para fins de acessibilidade. Paradoxalmente, do ponto de vista estritamente linguístico, o texto do DeepSeek tende a ser claro, coeso e neutro, recebendo conceitos altos em clareza, ordem lógica e neutralidade. No entanto, essa correção formal não compensa a ausência de aderência visual, que é o núcleo da audiodescrição.

Quando se observa o conjunto dos critérios, emerge um resultado importante para a discussão metodológica: a dimensão linguística (clareza, coesão, neutralidade e ordem lógica) é relativamente bem atendida por todos os modelos, inclusive pelos que falham gravemente na precisão. Os maiores problemas concentram-se na capacidade de “ver” corretamente a imagem, ou seja, de identificar cores, formas, textos, disposição espacial e elementos relevantes sem alucinações visuais ou inferências equivocadas. Isso indica que, no contexto das ferramentas avaliadas, o principal desafio para a acessibilidade

não está na formulação textual em si, mas na etapa de interpretação visual automática.

Esses achados revelam limitações estruturais das ferramentas de IA generativa no que diz respeito à audiodescrição de imagens, sobretudo em ambientes informacionais que exigem alta precisão visual. Embora apresentem competência linguística avançada, os modelos ainda demonstram dificuldade em reconhecer elementos gráficos com fidelidade, o que resulta em alucinações visuais, erros cromáticos, imprecisões tipográficas e inferências não sustentadas pela imagem. Além disso, aspectos como disposição espacial, direção de iluminação, leitura correta de textos incorporados à imagem e identificação de objetos pequenos ou parcialmente ocultos ainda constituem desafios significativos. Tais limitações reforçam que, apesar do potencial dessas tecnologias, sua utilização isolada não garante conformidade com normas de acessibilidade, exigindo sempre validação humana especializada para assegurar confiabilidade e aderência às diretrizes da ABNT NBR 16452:2016.

Em síntese, os resultados demonstram que, entre os modelos analisados, somente ChatGPT-5.1 e Gemini 2.5 Flash produzem audiodescrições que se aproximam de um padrão aceitável para uso em bibliotecas, catálogos digitais e outros ambientes informacionais comprometidos com as diretrizes da ABNT NBR 16452:2016 e com os princípios da WCAG 2.2. O Copilot apresenta potencial, mas requer supervisão e correção especializada, enquanto o DeepSeek V-3.2, tal como se comportou neste estudo, não atende aos requisitos mínimos de fidelidade visual, sendo inadequado para uso direto em contextos de acessibilidade informacional. Esses achados reforçam a necessidade de mediação humana qualificada na validação de audiodescrições automatizadas e apontam para a importância de critérios normativos claros na avaliação de sistemas de IA aplicados à descrição de imagens em bibliotecas.

6 CONSIDERAÇÕES FINAIS

O objetivo central desta pesquisa foi investigar a viabilidade do uso de ferramentas de IA para a audiodescrição automatizada de capas de livros em bibliotecas universitárias. Este objetivo foi alcançado, na medida em que o estudo produziu uma avaliação sistemática e fundamentada sobre as capacidades e limitações das ferramentas analisadas, permitindo estabelecer conclusões acerca de sua aplicabilidade prática em ambientes informacionais acadêmicos.

Os resultados obtidos neste estudo evidenciam que as ferramentas de inteligência artificial avaliadas apresentam desempenhos distintos na tarefa de gerar audiodescrições de capas de livros, revelando avanços importantes, mas também limitações significativas quando observadas à luz das normas brasileiras de acessibilidade ABNT NBR 16452:2016. As análises demonstraram que ChatGPT-5.1 e Gemini 2.5 Flash são, atualmente, os modelos que mais se aproximam de uma prática descritiva tecnicamente adequada para ambientes informacionais, produzindo textos com boa clareza, ordenação lógica, neutralidade e aderência geral aos elementos visuais. Ainda assim, mesmo nesses modelos, surgem deslizos pontuais na contagem de elementos, na descrição precisa de cores e na leitura de detalhes tipográficos, o que evidencia que, embora promissoras, tais ferramentas ainda não alcançam, de forma autônoma, um desempenho plenamente confiável.

O comportamento do Copilot mostra que boa qualidade linguística não implica necessariamente precisão visual. A ferramenta demonstrou capacidade de organizar a descrição de forma coerente e de apresentar linguagem acessível, mas suas dificuldades recorrentes de identificação cromática, espacial e tipográfica revelam que a interpretação visual automática ainda é um desafio central para a IA. O caso do DeepSeek V-3.2 aprofunda essa constatação ao evidenciar que, em alguns modelos, a descrição textual é dissociada da imagem real, resultando na criação de capas fictícias e impossibilitando sua utilização em contextos de acessibilidade. Esses achados confirmam que a etapa de “ver” a imagem continua sendo o ponto mais frágil das ferramentas atuais, mesmo quando a etapa de “descrever” já alcança boa fluência, objetividade e coesão.

A pesquisa também revelou que a IA tende a apresentar um desempenho linguístico muito superior ao seu desempenho perceptual. Isso significa que erros de leitura visual coexistem com textos claros, bem escritos e aparentemente confiáveis, o que exige cautela por parte de profissionais responsáveis por bibliotecas, repositórios e plataformas digitais. Em contextos de acessibilidade, o risco não está apenas na ausência de descrição, mas também na presença de descrições incorretas, que podem induzir o usuário a concepções equivocadas sobre o conteúdo visual. Assim, a mediação humana permanece indispensável para garantir precisão, fidedignidade e conformidade com as normas vigentes.

Os resultados também evidenciam os riscos associados à dependência acrítica de ferramentas de IA para a produção de audiodescrições. A presença de textos linguisticamente impecáveis, mas visualmente imprecisos, pode gerar uma falsa sensação de confiabilidade, levando instituições e profissionais a adotarem modelos que não garantem aderência às normas de acessibilidade. Em ambientes biblioteconômicos, essa discrepância pode resultar na disseminação de descrições equivocadas, prejudicando a experiência do usuário, comprometendo assim a qualidade informacional dos acervos digitais.

Os achados reforçam, ainda, a necessidade de que bibliotecas e instituições públicas e privadas desenvolvam diretrizes próprias para a adoção de ferramentas de IA na geração de audiodescrições, incorporando etapas de revisão, validação e teste com usuários reais. O presente estudo mostra que a IA pode ser uma aliada significativa na promoção da acessibilidade informacional, mas somente quando utilizada em conjunto com avaliação humana especializada e com parâmetros normativos claros. A automatização integral da audiodescrição, neste momento, implicaria riscos consideráveis para a qualidade da informação disponibilizada.

Do ponto de vista da Biblioteconomia, este estudo reforça o papel central e insubstituível do profissional da informação na mediação da acessibilidade. Mesmo em um cenário em que sistemas de IA possam apoiar a produção de audiodescrições, é o olhar especializado do bibliotecário — orientado pelas normas, pelos princípios éticos e pelas necessidades reais dos usuários — que assegura a fidelidade, a clareza e a utilidade dessas descrições. Assim, a

atuação humana permanece essencial para garantir que a tecnologia seja integrada de forma responsável, transformando o bibliotecário em um mediador crítico e cuidadoso na construção de ambientes informacionais verdadeiramente acessíveis.

Por fim, a pesquisa evidencia o potencial de investigação futura envolvendo modelos multimodais, testes empíricos com pessoas cegas e com baixa visão, ampliação dos tipos de imagens analisadas e desenvolvimento de instrumentos avaliativos mais refinados. A consolidação de políticas de acessibilidade em ambientes digitais requer tanto avanço tecnológico quanto rigor metodológico. A partir dos resultados aqui apresentados, fica evidente que o caminho mais seguro para a adoção de IA em audiodescrição passa pela integração entre tecnologia e conhecimento biblioteconômico, garantindo que o direito ao acesso pleno e equitativo à informação seja efetivamente assegurado.

REFERÊNCIAS

- Associação Brasileira de Normas Técnicas (ABNT). **ABNT NBR 15290:2005** - Acessibilidade em comunicação na televisão. Rio de Janeiro: ABNT, 2005. Disponível em: <https://www.abntcolecao.com.br/mpf/norma.aspx?ID=365121>. Acesso em: 20 ago. 2024.
- Associação Brasileira de Normas Técnicas (ABNT). **ABNT NBR 16452:2016** - Acessibilidade na comunicação: audiodescrição. Rio de Janeiro: ABNT, 2016. Disponível em: <https://drive.prefeitura.sp.gov.br/cidade/secretarias/upload/ABNT%20-%20Acessibilidade.pdf>. Acesso em: 20 ago. 2023.
- Associação Brasileira de Normas Técnicas (ABNT). **ABNT NBR 9050:2020** - Acessibilidade a edificações, mobiliário, espaços e equipamentos urbanos. Rio de Janeiro: ABNT, 2020. Disponível em: https://drive.prefeitura.sp.gov.br/cidade/secretarias/upload/NBR9050_20.pdf. Acesso em: 20 ago. 2023.
- Associação Brasileira de Normas Técnicas (ABNT). **ABNT NBR 17225:2025** – Acessibilidade digital: requisitos para conteúdo e aplicações web. Rio de Janeiro: ABNT, 2025.
- ARAÚJO, Aníbal Perea. Catálogo da biblioteca: o objeto orientado ao usuário. **Perspectivas em Ciência da Informação**, v. 16, n. 2, p. 17-28, abr./jun. 2011. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/133>. Acesso em: 20 ago. 2024.
- BARTHES, Roland. **A câmara clara**. Lisboa: Edições 70, 1989.
- BAUER, Martin; AARTS, Bas. A construção do corpus: um princípio para coleta, organização e análise de dados. In: BAUER, Martin; GASKELL, George (org.). **Pesquisa qualitativa com texto, imagem e som: um manual prático**. Petrópolis: Vozes, 2002. Disponível em: <https://repositorio.ufsc.br/xmlui/handle/123456789/1582>. Acesso em: 27 nov. 2025.
- BRASELIANO, L. A. de A. **Acessibilidade informacional e comunicacional em editais de ingresso da educação superior: um estudo de caso**. 2024. Dissertação (Mestrado em Educação Especial) – Universidade Federal de São Carlos, São Carlos, 2024. Disponível em: <https://repositorio.ufscar.br/handle/20.500.14289/20015>. Acesso em: 31 mar. 2025.
- BRASIL. Lei n.º 13.146, de 6 de julho de 2015. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 2015. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 14 maio 2024.
- BRASIL. Lei n.º 14.126, de 22 de março de 2021. Classifica a visão monocular como deficiência sensorial, do tipo visual. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14126.htm. Acesso em: 14 maio 2024.

BRASIL. Lei n.º 14.837, de 8 de abril de 2024. Altera a Lei n.º 12.244, de 24 de maio de 2010, que “dispõe sobre a universalização das bibliotecas nas instituições de ensino do País”, para modificar a definição de biblioteca escolar e criar o Sistema Nacional de Bibliotecas Escolares (SNBE). **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 2024. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2023-2026/2024/Lei/L14837.htm. Acesso em: 31 mar. 2025.

BRASIL. Lei n.º 4.084, de 30 de junho de 1962. Dispõe sobre a profissão de bibliotecário e regula seu exercício. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 1962a. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/1950-1969/l4084.htm. Acesso em: 14 maio 2024.

BRASIL. Decreto n.º 5.296, de 2 de dezembro de 2004. Regulamenta as leis n.º 10.048, de 8 de novembro de 2000, que dá prioridade de atendimento às pessoas que especifica, e 10.098, de 19 de dezembro de 2000, que estabelece normas gerais e critérios básicos para a promoção da acessibilidade [...]. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 2004. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/decreto/d5296.htm. Acesso em: 14 maio 2024.

BRASIL. Decreto n.º 6.949, de 25 de agosto de 2009. Promulga a Convenção Internacional sobre os Direitos das Pessoas com Deficiência e seu Protocolo Facultativo, assinados em Nova York, em 30 de março de 2007. Brasília, DF: **Congresso Nacional**, 2009. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/decreto/d6949.htm. Acesso em: 5 ago. 2024.

BRINGSJORD, Selmer; GOVINDARAJULU, Naveen Sundar. Artificial Intelligence. In: ZALTA, Edward N.; NODELMAN, Uri (orgs.). **The Stanford Encyclopedia of Philosophy**. Fall 2022 Edition. Stanford, CA: Metaphysics Research Lab, Stanford University, 2022. Disponível em: <https://plato.stanford.edu/archives/fall2022/entries/artificial-intelligence/>. Acesso em: 13 abr. 2025.

CARVALHO, Ana Isabel Silva. **A capa de livro: o objecto, o contexto, o processo**. 2008. Dissertação (Mestrado em Design da Imagem) – Faculdade de Belas Artes, Universidade do Porto, Porto, 2008. Disponível em: <https://hdl.handle.net/10216/67404>. Acesso em: 20 ago. 2024.

CORRÊA, A. C. B. **Imagem invisível: estudo sobre a imagem, acessibilidade e inteligência artificial**. 2021. 123 f. Dissertação (Mestrado em Design de Comunicação) - Universidade de Lisboa, Faculdade de Belas Artes, 2021. Disponível em: <https://repositorio.ul.pt/handle/10451/50078>. Acesso em: 09 maio 2024.

DEEPSEEK. **DeepSeek API Documentation**. 2024. Disponível em: <https://api-docs.deepseek.com/>. Acesso em: 26 fev. 2025.

DINIZ, C. N.; BIANQUE DO NASCIMENTO ABREU, J. Quesitos do Censo da Educação Superior como parâmetros para políticas de acessibilidade e inclusão em bibliotecas: estudo de caso do Instituto Federal do Maranhão. **Revista Brasileira de Biblioteconomia e Documentação**, [S. l.], v. 20, p. 1–17, 2024. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/2052>.

Acesso em: 31 mar. 2025.

DOURADO, Daniel de Araujo. **Regulação da inteligência artificial na saúde**. 2023. Tese (Doutorado em Ciências) – Faculdade de Medicina, Universidade de São Paulo, São Paulo, 2023. Disponível em:

<https://www.teses.usp.br/teses/disponiveis/5/5137/tde-23042024-111255/pt-br.php>. Acesso em: 13 abr. 2025.

FERNANDES, H. R. **A audiodescrição, os textos alternativos e as tecnologias da informação e comunicação**: Um estudo acerca da escolarização das pessoas com deficiência visual. 2023. 171 f. Dissertação (Mestrado em Estudos Linguísticos) - Universidade Federal de Uberlândia, Uberlândia, 2023. Disponível em:

<https://repositorio.ufu.br/handle/123456789/38820>. Acesso em: 15 set. 2024.

FERNANDES, J. D. P. B. **Diagnóstico da acessibilidade informacional na biblioteconomia brasileira**. 2018. 120 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal do Ceará, Centro de Humanidades, Programa de Pós-graduação em Ciência da Informação, Fortaleza, 2018.

Disponível em: <http://www.repositorio.ufc.br/handle/riufc/33425>. Acesso em: 31 mar. 2025.

FERREIRA, R. M. **Projeto de acessibilidade informacional (PAI): boas práticas para a cultura de acessibilidade digital em informação pública**. 2023. 334 f. Tese (Doutorado Profissional em Bens Culturais e Projetos Sociais) - Fundação Getulio Vargas, Rio de Janeiro, 2023. Disponível em:

<https://repositorio.fgv.br/items/12728016-cbca-492e-919a-29abebf46bb2>.

Acesso em: 31 mar. 2025.

FLICK, Uwe. **Introdução à pesquisa qualitativa**. Tradução Joice Elias Costa – 3ª ed. – Porto Alegre: Artmed, 2009. 405 p. Disponível em:

<https://periodicos.sbu.unicamp.br/ojs/index.php/ideias/article/download/8649420/15975/27891>. Acesso em: 27 nov. 2025.

FREITAS, Janaína Vieira de. **Julgando o livro pela capa: a influência da capa do livro para a escolha do leitor**. 2014. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014. Disponível em:

<https://lume.ufrgs.br/bitstream/handle/10183/102311/000935003.pdf;sequence=1>. Acesso em: 31 mar. 2025.

FUNDAÇÃO OSWALDO CRUZ. **Manual de primeiros socorros**. 2003:

Fundação Oswaldo Cruz, 2003. Disponível em:

<https://www.dominiopublico.gov.br/download/texto/fi000007.pdf>. Acesso em: 04 dez. 2025.

GIBBS, Graham. **Analyzing qualitative data**. London: SAGE, 2009. Disponível em:

<https://methods.sagepub.com/book/mono/preview/analyzing-qualitative-data-2e.pdf#> . Acesso em: 27 nov. 2025.

GONÇALVES, A. C. da C. **Biblioteca inclusiva: acessibilidade informacional baseada no desenho universal aos usuários do Instituto Federal da Paraíba (IFPB) – Campus Guarabira**. 2024. Dissertação (Mestrado Profissional em Educação Profissional e Tecnológica) – Instituto Federal da Paraíba, Campus

João Pessoa, João Pessoa, 2024. Disponível em: <https://repositorio.ifpb.edu.br/handle/177683/3709>. Acesso em: 31 mar. 2025.

GOOGLE. **Introducing Gemini 2.5: a new era of AI performance.** 2024. Disponível em: <https://blog.google/technology/google-deepmind/introducing-gemini-2-5/>. Acesso em: 26 fev. 2025.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social.** 6. ed. São Paulo: Atlas, 2010.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico.** 2. ed. Novo Hamburgo: Feevale, 2013. 277 p.

IBGE. Instituto Brasileiro de Geografia e Estatística. **Pesquisa nacional por amostra de domicílios contínua: divulgação dos resultados gerais.** Rio de Janeiro: IBGE, 2022. Disponível em: https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/0a9afaed04d7_9830f73a16136dba23b9.pdf. Acesso em: 10 out. 2024.

IFLA. International Federation of Library Associations and Institutions. **Lyon Declaration on Access to Information and Development.** Lyon, 2014. Disponível em: <https://www.ifla.org/publications/node/8852>. Acesso em: 3 mar. 2025.

JOLY, Martine. **Introdução à análise da imagem.** São Paulo: Papirus, 2005. Disponível em: <https://www.unijales.edu.br/library/downebook/id:63>. Acesso em: 27 nov. 2025.

LIMA, Yasmine; PEREIRA, Carla. A capa do livro e suas cores: o papel do matiz na interpretação da mensagem. In: **Anais do 9º CIDI e 9º CONGIC.** Belo Horizonte: Sociedade Brasileira de Design da Informação – SBDI, 2019. Disponível em: <https://pdf.blucher.com.br/designproceedings/9cidi/1.0091.pdf>. Acesso em: 31 mar. 2025.

LOCH, Carolina Valentim. **A obra de arte na era da inteligência artificial.** 2021. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Tecnológica Federal do Paraná, Curitiba, 2021. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/26309>. Acesso em: 13 abr. 2025.

LYONS, John. **Linguagem e linguística: uma introdução.** São Paulo: Ática, 1987.

MACHADO, Ana Maria Nogueira. **Informação e controle bibliográfico: um olhar sobre a cibernética.** São Paulo: Fundação Editora da Unesp, 2023. Disponível em: <https://www.dominiopublico.gov.br/download/texto/up000023.pdf>. Acesso em: 04 dez. 2025.

MEIRELES, P. M. **Wayfinding e o design informacional no espaço público urbano: fluxos visuais conectando pessoas e lugares.** 2024. 212 f. Dissertação (mestrado) - Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia Civil, Arquitetura e Urbanismo, Campinas, SP. Disponível em: <https://hdl.handle.net/20.500.12733/18573>. Acesso em: 31 mar. 2025.

MICROSOFT. **Visão geral do Microsoft Copilot.** 2024. Disponível em: <https://learn.microsoft.com/pt-br/copilot/overview>. Acesso em: 26 fev. 2025.

- MOURA, Lidia Zuin de. **Homo imago – imagem como sobrevivência e segundo corpo**. 2021. 236 f. Tese (Doutorado em Artes Visuais) – Universidade Estadual de Campinas, Instituto de Artes, Campinas, 2021. Disponível em: https://bdtd.ibict.br/vufind/Record/UNICAMP-30_1336180d388f072089b0a55370ac96c3. Acesso em: 31 mar. 2025.
- OPENAI. **Introducing GPT-5**. 2024. Disponível em: <https://openai.com/pt-BR/index/introducing-gpt-5/>. Acesso em: 26 fev. 2025.
- PEIRCE, Charles Sanders. **Semiótica**. São Paulo: Perspectiva, 2005.
- PESSOA, Fernando. *O Eu profundo e os outros Eus*. [S. l.]: [s. e.], 1913. Disponível em: <https://www.dominiopublico.gov.br/download/texto/vo000009.pdf>. Acesso em: 04 dez. 2025.
- RENDA, Andrea; ENGLER, Alex. **What's in a name? Getting the definition of Artificial Intelligence right in the EU's AI Act**. Brussels: CEPS – Centre for European Policy Studies, fev. 2023. Disponível em: <https://www.ceps.eu/ceps-publications/whats-in-a-name/>. Acesso em: 13 abr. 2025.
- RIBEIRO, E. N. **A imagem na relação de expressão com o texto escrito: contribuições da áudio-descrição para a aprendizagem de educandos surdos**. 2011. Dissertação (Mestrado em Educação) – Universidade Federal de Pernambuco, Recife, 2011.
- RIBEIRO, L. V. M. **Tecnologias assistivas e digitais no atendimento de usuários com deficiência visual na Biblioteca Central da UFMA: o direito à acessibilidade**. 2022. Dissertação (Mestrado Interdisciplinar em Cultura e Sociedade) – Universidade Federal do Maranhão, São Luís, 2022. Disponível em: <https://tedebc.ufma.br/jspui/handle/tede/3559>. Acesso em: 31 mar. 2025.
- RUSSELL, Stuart; NORVIG, Peter. **Inteligência artificial**. 3. ed. São Paulo: Pearson Education do Brasil, 2013.
- SALES, G. A. O. **Acessibilidade para busca de informações: um estudo de caso no Núcleo de Acessibilidade e Educação Inclusiva (NAEDI) para a formação da página web**. 2023. Dissertação (Mestrado Profissional em Gestão da Informação e do Conhecimento) – Universidade Federal de Sergipe, São Cristóvão, SE, 2023. Disponível em: https://ri.ufs.br/bitstream/riufs/18777/2/GLICIA_ANGELICA_OLIVEIRA_SALES.pdf. Acesso em: 20 ago. 2024.
- SANTAELLA, Lúcia. **O que é semiótica**. São Paulo: Brasiliense, 1999. Disponível em: https://projetoex-votosdobrasil.net/wp-content/uploads/2023/01/146282759-o-que-e-semiotica_1.pdf. Acesso em: 27 nov. 2025.
- SANTOS, C. G. dos; ARAÚJO, W. J. de. **Acessibilidade informacional: um estudo sobre configurações de segurança em objetos digitais acessíveis segundo análise de aceitação por pessoas com deficiência visual**. Pesquisa Brasileira em Ciência da Informação e Biblioteconomia, João Pessoa, v. 10, n. 2, p. 209-222, 2015. Disponível em: <https://brapci.inf.br/index.php/res/v/25901>. Acesso em: 25 ago. 2024.
- SASSAKI, R. K. **Inclusão: acessibilidade no lazer, trabalho e educação**.

Revista Nacional de Reabilitação (Reação), São Paulo, Ano XII, mar./abr. 2009, p. 10-16. Disponível em: [https://files.cercomp.ufg.br/weby/up/211/o/SASSAKI - Acessibilidade.pdf](https://files.cercomp.ufg.br/weby/up/211/o/SASSAKI_-_Acessibilidade.pdf). Acesso em: 20 ago. 2024.

SAUSSURE, Ferdinand de. **Curso de linguística geral**. São Paulo: Cultrix, 1916.

SEHN, Thaís Cristina Martino. **Objeto desejo**. 2009. Trabalho de Conclusão de Curso (Bacharelado em Design Gráfico) – Instituto de Artes e Design, Universidade Federal de Pelotas, Pelotas, 2009. Disponível em: <https://www.escriitoriodolivro.com.br/leitura/objeto%20desejo.pdf>. Acesso em: 31 mar. 2025.

SILVEIRA, Mariana Newlands. **Bibliomania no sistema literário**. 2006. Tese (Doutorado em Letras) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2006. Orientadora: Heidrun Friedel Krieger Olinto de Oliveira. Disponível em: https://www.maxwell.vrac.puc-rio.br/9120/9120_5.PDF. Acesso em: 31 mar. 2025.

TARAPANOFF, K.; SUAIDEN, E.; OLIVEIRA C. L. Funções sociais e oportunidades para profissionais da informação. **DataGramaZero: Revista de Ciência da Informação**, [S. l.], v. 3, n. 5, 2002. Disponível em: <https://brapci.inf.br/index.php/res/v/5401>. Acesso em: 5 maio 2025.

UNESCO. **Recomendação sobre a Ética da Inteligência Artificial**. Paris: Organização das Nações Unidas para a Educação, a Ciência e a Cultura, 2021. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>. Acesso em: 03 nov. 2025.

W3C Recommendation. **Diretrizes de Acessibilidade para Conteúdo Web (WCAG) 2.2**. A. Campbell, C. Adams, R. B. Montgomery, M. Cooper e A. Kirkpatrick, eds. Tradução de Reinaldo Ferraz. Ceweb.br. Recomendação W3C de 05 out. 2023. Disponível em: <https://www.w3.org/Translations/WCAG22-pt-BR/>. Acesso em: 13 abr. 2025.

ZIMMERMANN, Jian Marcel. **Da representação da imagem à imagem como representação em O pintor de retratos e Satolep**. 2016. 221 f. Tese (Doutorado em Letras) – Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Letras, Porto Alegre, 2016. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/141253/000992396.pdf?sequence=1>. Acesso em: 13 abr. 2025.

**APÊNDICE A – MANUAL DE PROCEDIMENTOS PARA GERAÇÃO E
AVALIAÇÃO DAS AUDIODESCRIÇÕES**

**ANEXO A – DADOS FORNECIDOS PELO NÚCLEO DE ACESSIBILIDADE E
INCLUSÃO (NAI/UNIRIO) SOBRE ESTUDANTES COM DEFICIÊNCIA
VISUAL (2025)**

Tabela 1: Dados fornecidos pelo Núcleo de Acessibilidade e Inclusão (NAI/UNIRIO) sobre estudantes com deficiência visual (2025).

NOME_CURSO	DESCR_DEFICIENCIA
Ciências Biológicas - Bacharelado - Turno Integral (M/V)	8 - Cegueira
Ciências Biológicas - Bacharelado - Turno Integral (M/V)	8 - Cegueira
Biomedicina - Bacharelado - Turno Integral (M/V)	8 - Cegueira
Ciências Ambientais - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Ciências Ambientais - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Enfermagem - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Enfermagem - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Medicina - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Medicina - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Medicina - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Medicina - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Medicina - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Medicina - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Medicina - Bacharelado - Turno Integral (M/V)	8 - Cegueira
Nutrição - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Nutrição - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Sistemas de Informação - Bacharelado - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Matemática - Licenciatura - Turno Integral (Vespertino/Noturno)	6 - Baixa visão ou visão monocular
História - Bacharelado - Turno Matutino	8 - Cegueira
História - Licenciatura - Turno Matutino - código e-MEC 320064	8 - Cegueira
História - Licenciatura - Turno Matutino - código e-MEC 320064	8 - Cegueira
História - Licenciatura - Turno Matutino - código e-MEC 320064	6 - Baixa visão ou visão monocular
Ciências Sociais - Licenciatura - Turno Matutino	6 - Baixa visão ou visão monocular
Ciências Sociais - Licenciatura - Turno Matutino	6 - Baixa visão ou visão monocular
Arquivologia - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Biblioteconomia - Bacharelado - Turno Matutino	6 - Baixa visão ou visão monocular
Museologia - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular

Museologia - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Museologia - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Museologia - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Museologia - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Museologia - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Pedagogia - Licenciatura - Turno Noturno	8 - Cegueira
Pedagogia - Licenciatura - Turno Noturno	6 - Baixa visão ou visão monocular
Pedagogia - Licenciatura - Turno Noturno	6 - Baixa visão ou visão monocular
Pedagogia - Licenciatura - Turno Noturno	8 - Cegueira
Pedagogia - Licenciatura - Turno Vespertino	6 - Baixa visão ou visão monocular
Pedagogia - Licenciatura - Turno Vespertino	6 - Baixa visão ou visão monocular
Pedagogia - Licenciatura - Turno Vespertino	6 - Baixa visão ou visão monocular
Pedagogia - Licenciatura - Turno Vespertino	6 - Baixa visão ou visão monocular
Direito - Bacharelado - Turno Noturno	8 - Cegueira
Direito - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Turismo - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Filosofia - Bacharelado - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Filosofia - Licenciatura - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Serviço Social - Bacharelado - Turno Noturno	6 - Baixa visão ou visão monocular
Serviço Social - Bacharelado - Turno Noturno	8 - Cegueira
Serviço Social - Bacharelado - Turno Noturno	8 - Cegueira
Estética e Teoria do Teatro - Bacharelado - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Estética e Teoria do Teatro - Bacharelado - Turno Integral (V/N)	8 - Cegueira
Estética e Teoria do Teatro - Bacharelado - Turno Integral (V/N)	8 - Cegueira
Estética e Teoria do Teatro - Bacharelado - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Cenografia e Indumentária - Bacharelado - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Atuação Cênica - Bacharelado - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Música - Licenciatura - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Música - Bacharelado - Hab. Instrumento - Piano	6 - Baixa visão ou visão monocular
Música - Bac. - Hab. MPB - Arranjo Musical - Turno Integral (V/N)	6 - Baixa visão ou visão monocular
Letras - Licenciatura - Turno Noturno	6 - Baixa visão ou visão monocular

Administração Pública - Bacharelado - Turno Integral (M/V)	8 - Cegueira
Administração Pública - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Administração Pública - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Ciência Política - Bacharelado - Turno Integral (M/V)	6 - Baixa visão ou visão monocular
Pedagogia	8 - Cegueira

Fonte: NAI/UNIRIO (2025).