

Acessibilidade em Aplicativos Móveis: Análise de Cobertura e Escopo na Aplicação da Norma ABNT NBR 17060

Arienne Alves Navarro
Universidade Federal de Lavras
Lavras, Minas Gerais
arienne.navarro@estudante.ufla.br

Lívia Della Garza Silva
Universidade Federal de Lavras
Lavras, Minas Gerais
livia.silva18@estudante.ufla.br

Leonardo Filipe da Silva
Universidade Federal de Lavras
Lavras, Minas Gerais
leonardo.silva24@estudante.ufla.br

Monique Scalco Soares Siqueira
Universidade Federal de Lavras
Lavras, Minas Gerais
monique.siqueira1@ufla.br

Eduardo Augusto Ribeiro
Universidade Federal de Lavras
Lavras, Minas Gerais
eduardo.ribeiro3@estudante.ufla.br

Paulo Henrique dos Anjos
Silveira
Universidade Federal de Lavras
Lavras, Minas Gerais
paulo.silveira2@estudante.ufla.br

Gabriel Aguiar A. Silva
Universidade Federal de Lavras
Lavras, Minas Gerais
gabriel.silva53@estudante.ufla.br

Ruan Pablo Gomes Rocha
Universidade Federal de Lavras
Lavras, Minas Gerais
ruan.rocha1@estudante.ufla.br

Ruan Víctor Henrique
Universidade Federal de Lavras
Lavras, Minas Gerais
ruan.henrique@estudante.ufla.br

Saimon Gabriel de Andrade
Universidade Federal de Lavras
Lavras, Minas Gerais
saimon.andrade2@estudante.ufla.br

Paulo Afonso Parreira Júnior
Universidade Federal de Lavras
Lavras, Minas Gerais
pauloa.junior@ufla.br

Silvia Helena Rigatto
Universidade Federal de Lavras
Lavras, Minas Gerais
shrigatto@ufla.br

André Pimenta Freire
Universidade Federal de Lavras
Lavras, Minas Gerais
apfreire@ufla.br

ABSTRACT

The ABNT NBR 17060 standard (Associação Brasileira de Normas Técnicas, Norma Brasileira), launched in 2022, is the first Brazilian technical standard focused on accessibility in mobile applications, based on the Web Content Accessibility Guidelines (WCAG) and aimed at supporting laws such as the Brazilian Inclusion Law. Despite the progress it represents, the standard still presents limitations in its coverage and the scope of its subsections, which can hinder the proper framing and classification of problems in practical accessibility assessments. The objective of this article is to present a diagnosis of the coverage of this standard based on real evaluations. The methodology consisted of a document analysis of accessibility evaluations conducted in an extension project to support the reporting of accessibility problems. Evaluations of five mobile applications (Wise, Gov.br, Nubank, Zé Delivery, and Minas App) were used, identifying cases where the problems found could not be adequately mapped in the planned subsections. The results reveal three main types of limitations in the standard's practical application. The first refers to the complete absence of normative coverage for certain mobile resources and usage situations. The second refers

to partial coverage of recurring problems, requiring registration in related, but technically imprecise, items. The third refers to normative overlap: excessively broad criteria that group heterogeneous problems, of different WCAG principles, under the same subsection, even when nominally covering them. The study concludes that, despite representing a significant advance for digital accessibility in Brazil, there are opportunities for refinement in its scope and specificity so that real problems can be correctly diagnosed and addressed. The findings can assist future revisions of the standard, contributing to guaranteeing the right to digital accessibility in a more precise and comprehensive way, as well as supporting its use in its current version to frame its scope in practical projects.

KEYWORDS

Acessibilidade e Tecnologias Assistivas, Computação Móvel, ABNT 17060

1 INTRODUÇÃO

A crescente transformação digital tem consolidado aplicações Web e móveis como os principais canais de acesso a serviços públicos e privados, abrangendo domínios como governo eletrônico, educação, saúde, comércio eletrônico e serviços financeiros. Nesse cenário, garantir a qualidade dessas aplicações tornou-se um desafio multi-dimensional, envolvendo atributos como desempenho, segurança,

usabilidade e acessibilidade. Em especial, a acessibilidade digital tem assumido papel estratégico ao permitir que pessoas com deficiência utilizem sistemas digitais de forma autônoma, eficiente e em igualdade de oportunidades, independentemente das tecnologias assistivas empregadas [15, 28].

A importância da acessibilidade torna-se ainda mais evidente diante da crescente utilização de dispositivos móveis como principal meio de acesso à Internet. Segundo dados recentes, smartphones representam o dispositivo predominante para consumo de serviços digitais em grande parte do mundo, exigindo que aplicações móveis sejam concebidas considerando diferentes perfis de usuários, contextos de uso e tecnologias assistivas. Entretanto, as particularidades da interação móvel, como gestos, mudanças de orientação da tela, sensores embarcados, leitores de tela móveis e diferentes mecanismos de entrada, introduzem desafios adicionais para o desenvolvimento e a avaliação da acessibilidade quando comparados às aplicações Web tradicionais [30] [8, 12, 15, 17, 20, 24, 28].

Nesse contexto, a avaliação da acessibilidade constitui uma atividade fundamental durante o ciclo de desenvolvimento de software. Seu objetivo não se limita à identificação de barreiras de uso, mas também ao fornecimento de evidências técnicas que subsidiem a correção de problemas, apoiem a comunicação entre equipes multidisciplinares e permitam monitorar a evolução da qualidade da acessibilidade ao longo do tempo. Embora ferramentas automáticas desempenhem papel importante na identificação de determinadas não conformidades [9], diversos estudos demonstram que uma parcela significativa dos problemas de acessibilidade somente pode ser identificada por meio de avaliações manuais conduzidas por especialistas, especialmente aqueles relacionados à interação com tecnologias assistivas e ao comportamento dinâmico das interfaces [15, 27, 29].

Para garantir consistência durante as avaliações de acessibilidade, torna-se necessário utilizar diretrizes e normas técnicas que orientem a identificação, a classificação e a documentação dos problemas encontrados. Nesse contexto, as Web Content Accessibility Guidelines (WCAG) consolidaram-se como o principal referencial internacional para o desenvolvimento e a avaliação da acessibilidade digital, sendo amplamente adotadas por governos, organizações e organismos de padronização como base para legislações, normas e processos de avaliação de conformidade em diferentes países [13, 28], conforme detalhado na Seção 2.1.

No contexto brasileiro, esse movimento resultou na publicação da norma da ABNT, a NBR 17060:2022 [2], primeira norma técnica nacional dedicada especificamente aos requisitos de acessibilidade para aplicativos móveis, detalhada na Seção 2.2.

Entretanto, a utilização de uma norma técnica em auditorias de acessibilidade depende não apenas da qualidade de seus requisitos, mas também de sua capacidade de representar, de forma clara e consistente, os problemas identificados durante a avaliação. Em processos de avaliação de conformidade, os problemas encontrados são documentados em relação aos critérios normativos aplicáveis, permitindo registrar evidências de não conformidade, subsidiar ações de correção e produzir resultados comparáveis entre diferentes avaliações [4, 31]. Assim, a organização e a granularidade da norma influenciam diretamente a qualidade dos diagnósticos produzidos. Quando a associação entre problemas e critérios normativos é ambígua ou determinados problemas não encontram

correspondência explícita na norma, o processo de classificação torna-se menos preciso, comprometendo a rastreabilidade dos problemas e a consolidação do conhecimento produzido durante as auditorias [15].

A pesquisa em acessibilidade digital concentra-se predominantemente na avaliação da conformidade de aplicações Web e móveis; investigações sobre a capacidade das próprias normas técnicas em classificar consistentemente os problemas encontrados em auditorias permanecem escassas. No caso da ABNT NBR 17060 [2], publicada em 2022 como primeiro referencial técnico brasileiro para acessibilidade em aplicativos móveis, essa lacuna evidencia a necessidade de investigar empiricamente sua cobertura normativa [7, 20].

Diante desse cenário, este trabalho busca responder à seguinte questão de pesquisa: Quais lacunas de cobertura normativa podem ser identificadas na aplicação da ABNT NBR 17060 durante avaliações manuais de acessibilidade em aplicativos móveis?

Para responder a essa questão, este artigo apresenta um diagnóstico da cobertura normativa da ABNT NBR 17060 a partir da análise de um conjunto de avaliações manuais de acessibilidade realizadas em cinco aplicativos móveis pertencentes a diferentes domínios de aplicação. Diferentemente de estudos que investigam a incidência de problemas de acessibilidade em produtos específicos, este trabalho concentra-se na capacidade da estrutura normativa em representar esses problemas durante o processo de classificação realizado pelos avaliadores. A análise considera avaliações produzidas no âmbito de um projeto de extensão universitária voltado ao apoio à denúncia de problemas de acessibilidade digital, permitindo investigar a norma sob a perspectiva de sua aplicação prática em situações reais.

Como principal contribuição, este artigo apresenta um diagnóstico empírico da cobertura normativa da ABNT NBR 17060 baseado em avaliações reais de acessibilidade, identificando padrões recorrentes de limitações relacionados à granularidade dos requisitos, à cobertura parcial de determinados problemas e à ausência de requisitos específicos para situações observadas durante avaliações manuais. Espera-se que os resultados contribuam para subsidiar futuras revisões da norma, apoiar pesquisadores e profissionais envolvidos na avaliação de acessibilidade e fortalecer os processos de garantia da qualidade de aplicações Web e móveis sob a perspectiva da acessibilidade digital.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção apresenta os dois referenciais normativos centrais para este trabalho: as diretrizes internacionais WCAG e a norma brasileira ABNT NBR 17060, que adapta essas diretrizes para o contexto de aplicativos móveis.

2.1 Web Content Accessibility Guidelines (WCAG)

As Web Content Accessibility Guidelines (WCAG), publicadas pelo World Wide Web Consortium (W3C), consolidaram-se como o principal referencial internacional para o desenvolvimento e a avaliação da acessibilidade digital [28]. As diretrizes organizam-se em torno de quatro princípios: conteúdo deve ser perceptível, operável, compreensível e robusto. Cada princípio se desdobra em diretrizes mais

específicas, e cada diretriz é operacionalizada por meio de critérios de sucesso testáveis, distribuídos em três níveis de conformidade (A, AA e AAA). Essa estrutura hierárquica permite que uma diretriz ampla, como a que trata de nome, função e valor de componentes de interface, seja avaliada por meio de um critério de sucesso objetivo e verificável (WCAG 4.1.2), o que é explorado na análise de colisão apresentada neste artigo (Seção 4.3.3).

Ao longo de suas revisões, de WCAG 2.0 a 2.2, as diretrizes ampliaram gradualmente a cobertura de cenários de uso mais específicos, incluindo recomendações voltadas à interação móvel, como orientação de tela, gestos e tamanho de alvos de toque [30]. Esse esforço de adaptação das WCAG para contextos além do conteúdo Web tradicional também é formalizado por documentos complementares como o WCAG2ICT e o WCAG2Mobile [18, 22]. As WCAG também foram formalmente reconhecidas como norma internacional pela ISO/IEC 40500 [13], o que reforça seu papel como base técnica para normas nacionais, incluindo a ABNT NBR 17060, discutida a seguir.

2.2 ABNT NBR 17060

A ABNT NBR 17060:2022 é a primeira norma técnica brasileira dedicada especificamente aos requisitos de acessibilidade para aplicativos móveis [2], aplicável a aplicativos nativos, baseados em navegador ou híbridos. A norma organiza seus requisitos e recomendações em apenas quatro blocos temáticos: percepção e compreensão (item 5.1.1), controle e interação (item 5.1.2), mídia (item 5.1.3) e requisitos para codificação (item 5.1.4), este último tratado como um único item, sem subdivisões internas. Um Anexo informativo associa cada item da norma a critérios de sucesso da WCAG 2.1 [14] e a outras referências nacionais, mas essa associação é limitada: não há indicação de que um mesmo item possa corresponder a múltiplos critérios WCAG de princípios diferentes, o que sugere uma lacuna de granularidade entre a organização temática da norma e a especificidade de sua referência.

A publicação da norma representa um avanço para a acessibilidade digital no Brasil, alinhado à Lei Brasileira de Inclusão [5] e à Convenção sobre os Direitos das Pessoas com Deficiência [26]. Por ser recente e ainda pouco estudada empiricamente, sua capacidade de representar, de forma consistente, os problemas encontrados em avaliações reais permanece uma questão em aberto, investigada neste artigo a partir de um corpus de avaliações conduzidas em cinco aplicativos móveis.

3 TRABALHOS RELACIONADOS

A análise da cobertura de diretrizes de acessibilidade sobre problemas reais encontrados em avaliações é uma linha com intensa atividade de pesquisa, tanto no contexto da web quanto, mais recentemente, no contexto de aplicativos móveis. Esta seção apresenta trabalhos relacionados ao presente estudo, organizados em dois grupos: (i) estudos que investigaram a cobertura de diretrizes para problemas de acessibilidade em aplicações móveis; e (ii) estudos que analisaram a cobertura de diretrizes no contexto da acessibilidade web.

3.1 Cobertura de Diretrizes em Aplicações Móveis

Um dos estudos mais diretamente relacionados ao presente trabalho é o de Alajarmeh [1] que investigou a cobertura da WCAG 2.1 sobre problemas de acessibilidade encontrados por 16 usuários com deficiência visual ao interagir com dispositivos touchscreen. O estudo identificou 34 categorias principais de problemas e os mapeou para os critérios de sucesso da WCAG 2.1, analisando tanto a existência de critérios aplicáveis quanto a adequação dos níveis de conformidade. Os resultados revelaram que vários problemas não possuíam critérios de sucesso relevantes na WCAG 2.1 ou eram endereçados por critérios com níveis de conformidade insuficientes para mitigar o impacto real sobre usuários com deficiência visual, sinalizando lacunas concretas na cobertura da norma para uso móvel não-visual. Essa abordagem metodológica é análoga à empregada no presente artigo, embora no presente estudo o objeto de análise seja a norma ABNT NBR 17060 [2] em vez da WCAG 2.1, com contribuições importantes para o contexto nacional.

Grellmann et al. [11] compararam problemas observados em testes com usuários com afasia com o que uma auditoria baseada nas diretrizes WCAG (versão mobile) conseguia capturar. O estudo evidenciou diferenças expressivas entre os problemas identificados nos testes e aqueles cobertos pela auditoria, discutindo a eficácia limitada das regras WCAG para necessidades cognitivas e linguísticas específicas e propondo recomendações adicionais para aplicações web progressivas. Os resultados reforçam a perspectiva de que diretrizes gerais podem apresentar cobertura insuficiente quando confrontadas com problemas de grupos específicos de usuários.

No campo das heurísticas para mobile, Park et al. [21] propuseram dez heurísticas específicas a partir de observações com usuários com deficiência visual e walkthrough por especialistas, argumentando que as diretrizes vigentes à época deixavam lacunas nas dificuldades reais de touchscreen. Billi et al. [3], de forma complementar, propuseram uma metodologia unificada de acessibilidade e usabilidade mobile, reforçando que diretrizes gerais precisam ser adaptadas ao contexto móvel para não negligenciar questões que heurísticas específicas capturam.

No domínio de aplicações de saúde móvel (*mHealth*), Nasr et al. [19] desenvolveram e validaram o conjunto de diretrizes UGmHA (*Usability Guidelines for mHealth Applications*), comparando-o experimentalmente com outras diretrizes em dois aplicativos. Os resultados mostraram que diretrizes específicas de domínio detectaram mais problemas de usabilidade do que diretrizes genéricas, indicando que a especialização setorial pode ampliar a cobertura sobre problemas reais. Esse achado dialoga diretamente com as limitações identificadas no presente estudo, em que itens normativos excessivamente amplos agrupam falhas heterogêneas, reduzindo a precisão diagnóstica da norma.

3.2 Cobertura de Diretrizes de Acessibilidade Web

No contexto da acessibilidade web, o trabalho seminal de Power et al. [23] constitui referência central para estudos que analisam a cobertura de guidelines sobre problemas reais. Os autores conduziram avaliações baseadas em tarefas com 32 usuários cegos em 16 websites, registrando 1.383 instâncias de problemas e mapeando-as

para os critérios de sucesso da WCAG 2.0. Os resultados revelaram que apenas 50,4% dos problemas encontrados eram cobertos por algum critério de sucesso da WCAG 2.0, e que, mesmo quando a norma aparentava cobrir o problema, as técnicas implementadas nem sempre o resolviam na prática. Esse estudo estabelece um patamar empírico de referência para a avaliação da cobertura normativa e inspira diretamente a abordagem metodológica adotada no presente trabalho, embora o presente trabalho aborde inspeções realizadas por especialistas, e não testes com pessoas com deficiência.

Calvo et al. [6] realizaram inspeções de acessibilidade por sete especialistas em 62 avaliações de websites (móveis e desktop) e aplicações móveis, usando a WCAG 2.0 nível AA como base e registrando explicitamente problemas potenciais não cobertos pela norma. O estudo documentou múltiplas lacunas e recomendou práticas e técnicas adicionais que vão além do escopo da WCAG, contribuindo para a compreensão de quais categorias de problemas ficam sistematicamente fora da cobertura normativa.

Lopes and Carriço [16] desenvolveram um *framework* de análise que diferencia requisitos de usuários dos requisitos técnicos da WCAG, usando testes com usuários com deficiência para identificar partes da norma que não podem ser verificadas automaticamente. O estudo demonstrou que avaliações automáticas cobrem apenas uma fração dos requisitos WCAG e que testes in-vivo com usuários são essenciais para expor problemas que diretrizes e ferramentas automáticas não capturam, evidenciando lacunas práticas entre conformidade técnica e experiência real de uso.

Em conjunto, os trabalhos revisados nesta seção evidenciam que a análise sistemática da cobertura de diretrizes sobre problemas reais é um problema relevante e recorrente na pesquisa em acessibilidade, tanto para a web quanto para o contexto móvel. O presente estudo contribui com essa linha ao aplicar uma abordagem similar à norma brasileira ABNT NBR 17060, avaliando sua cobertura sobre um corpus de 471 problemas coletados em cinco aplicativos móveis e caracterizando os padrões de lacuna identificados.

4 METODOLOGIA

O objetivo deste estudo foi verificar se a norma ABNT NBR 17060 [2] oferece cobertura suficiente para enquadrar os problemas de acessibilidade que aparecem em avaliações reais de aplicativos móveis. Para isso, partiu-se de um conjunto de registros de problemas já produzidos por avaliadores de um projeto de extensão e analisou-se cada um deles à luz dos itens da norma. A Seção 4.1 descreve de onde vêm esses registros e a Seção 4.3 apresenta o protocolo de classificação.

4.1 Origem dos Dados e Aplicativos Avaliados

Os dados vêm de avaliações de acessibilidade realizadas por um núcleo de extensão da Universidade Federal de Lavras que apoia a denúncia formal de problemas em aplicativos móveis e web. Essas avaliações são utilizadas como base para notificações extrajudiciais enviadas às empresas responsáveis e possíveis notificações a órgãos de controle, em uma parceria entre o Departamento de Ciência da Computação e o Departamento de Direito [10, 25].

Foram analisadas avaliações de cinco aplicativos de domínios diferentes: serviços financeiros (Wise e Nubank), governo federal

(Gov.br), entrega por demanda (Zé Delivery) e governo estadual (Minas App). Os aplicativos foram selecionados a partir do conjunto de avaliações já conduzidas pelo núcleo de extensão até o momento da coleta de dados. As avaliações abrangeram tanto a plataforma iOS quanto a Android, utilizando como tecnologias assistivas o VoiceOver e o TalkBack, respectivamente, conforme a plataforma de cada aplicativo avaliado. Cada avaliação foi conduzida por uma dupla ou trio de graduandos com experiência na norma ABNT NBR 17060 e nas diretrizes WCAG 2.1 [14]: Wise (Arienne e Lívia), Gov.br (Lívia e Paulo Henrique), Zé Delivery (Gabriel e Ruan Víctor), Nubank (Gabriel, Ruan Víctor e Ruan Pablo) e Minas App (Gabriel, Paulo Henrique e Ruan Pablo). Não houve treinamento formal estruturado: a familiaridade dos pesquisadores com a norma e com as WCAG foi construída pela prática, ao longo de avaliações sucessivas conduzidas pelo próprio núcleo de extensão; em todas as avaliações foi utilizada como apoio a lista de verificação (*checklist*) derivada da própria ABNT NBR 17060. A classificação foi conduzida integralmente por uma única pesquisadora, sem segunda codificação ou medida formal de concordância entre avaliadores; essa limitação é discutida na Seção 7.

Em cada avaliação, os problemas foram registrados com título, descrição da falha e o item da norma indicado pelo avaliador como aplicável. Esse registro é a unidade de análise deste estudo: trabalha-se com o que os avaliadores descreveram, sem acrescentar observações próprias sobre os aplicativos. No total, o corpus reúne 471 problemas. Todos foram incluídos na análise, sem nenhum critério de exclusão.

4.2 Como o Protocolo Foi Desenvolvido

O protocolo foi construído de forma iterativa, e não definido a priori. A classificação inicial, baseada apenas no julgamento da pesquisadora, mostrou-se pouco consistente entre problemas semelhantes, o que motivou a criação de perguntas sequenciais e regras explícitas, incluindo uma regra de rejeição para descartar cobertura apenas superficial e critérios de fronteira para separar itens frequentemente confundidos, como o item 5.1.4 (Requisitos para codificação), cuja redação ampla poderia, sem essas travas, ser aplicada a praticamente qualquer problema de interface. O protocolo foi refinado em um estudo-piloto com os dados do Wise (161 problemas, o corpus mais extenso), ao longo de cinco análises completas com versões sucessivas. Esse piloto também revelou uma limitação técnica: versões que fechavam o desfecho de cada problema individualmente deixavam colisões passarem despercebidas, o que motivou dividir a aplicação em duas passagens obrigatórias, detalhadas na Seção 4.3.

4.3 Protocolo de Classificação

O protocolo busca responder duas perguntas para cada problema: a norma ABNT NBR 17060 cobre esse problema de forma adequada? Se não, que tipo de lacuna existe? Para cada problema, são determinados o item da norma que melhor o representa e um critério de sucesso WCAG 2.1 [14]. O resultado final é um dos quatro desfechos descritos na Seção 4.3.3.

4.3.1 Unidade de Análise. Cada problema é classificado como um todo, exatamente como foi descrito pelo avaliador original. Quando um registro reúne mais de um sintoma, o que é comum em avaliações de tela ou elemento, ele não é dividido em falhas menores. A

classificação recai sobre a *causa-raiz dominante* do problema, chamada aqui de *problema central*; os sintomas que são consequência dessa causa são apenas anotados, sem gerar registros extras.

Essa decisão tem duas razões práticas. Primeiro, dividir um problema em partes menores depende fortemente de quem faz a divisão, e diferentes recortes produziram resultados diferentes, o que tornaria a análise pouco reprodutível. Segundo, fragmentar os problemas tenderia a esconder lacunas: cada pedaço isolado tende a caber bem em algum item da norma, fazendo desaparecer exatamente as situações de cobertura incompleta que este estudo quer revelar.

Para identificar o problema central, o protocolo usa um teste simples: se a causa hipotética fosse corrigida, os outros sintomas desapareceriam junto? Os que desaparecem são consequências da causa; os que persistem indicam que o registro mistura mais de uma falha independente. Mesmo nesses casos, o problema não é dividido: classifica-se pelo aspecto que o avaliador focou, e o restante é tratado na medição de cobertura.

4.3.2 Atribuição do Item da Norma e do Critério WCAG. Para cada problema, o protocolo parte do item originalmente indicado pelo avaliador e verifica se ele cobre tudo o que foi descrito. Quando não cobre, busca-se outro item da norma que melhor represente o problema central, com base na correspondência entre o elemento afetado, o tipo de falha e o nível de obrigação que o requisito impõe, e não na semelhança temática. Por exemplo: um elemento que tem rótulo visível mas é anunciado de forma genérica pelo leitor de telas é uma falha de codificação (item 5.1.4, WCAG 4.1.2), não uma ausência de rótulo (item 5.1.1.4, WCAG 3.3.2). Sem critérios explícitos de fronteira como esse, essa distinção tende a ser feita de forma inconsistente.

Depois de identificar o item candidato, aplica-se a *regra de rejeição*: o item é descartado se a cobertura for apenas temática, se exigir interpretação forçada do texto da norma, ou se for uma recomendação quando o problema viola um requisito obrigatório. O item 5.1.4 tem uma trava extra: só é registrado quando o sintoma é diretamente a exposição incorreta ou ausente de nome, função, valor ou estado de um componente. Isso impediu que ele funcionasse como destino fácil para qualquer problema de interface, o que, como os resultados mostram, é justamente o risco que a redação ampla do item cria. Cada problema mantido no estudo recebe um único critério WCAG, correspondente à causa-raiz. Critérios secundários são anotados separadamente.

4.3.3 Desfechos e Definição de Colisão. O desfecho de cada problema é definido por duas perguntas independentes: o item que melhor representa o problema o cobre completamente? E esse item *colide*?

Colisão ocorre quando dois ou mais problemas do mesmo aplicativo têm o mesmo item da norma, mas pertencem a princípios ou diretrizes WCAG diferentes. O escopo é sempre dentro do mesmo aplicativo: o mesmo item em aplicativos diferentes não conta como colisão. Diferenças de critério dentro de uma mesma diretriz (por exemplo, 1.4.3 e 1.4.6) também não são colisão, por representarem gradações do mesmo requisito.

Um ponto importante: colisão tem prioridade sobre cobertura. Um problema pode ser completamente coberto por um item e ainda assim gerar colisão, porque o problema identificado é do item, não

do problema em si: o item é amplo demais, e agrupa falhas de naturezas diferentes. Os quatro desfechos possíveis são:

- **Fora do estudo:** um item cobre o problema completamente e não há colisão. A norma apresentou cobertura adequada para esse caso.
- **Cenário 1:** o item que melhor cobre o problema colide, ele agrupa, sob uma mesma subseção, problemas de princípios ou diretrizes WCAG diferentes. Indica que o item é amplo demais.
- **Cenário 2:** o item cobre o problema só em parte. Algum aspecto descrito pelo avaliador não é alcançado por nenhum item da norma, e o problema acaba registrado em um item correlato, porém impreciso.
- **Cenário 3:** nenhum item da norma cobre o problema, nem parcialmente. A norma simplesmente não tem requisito para aquele tipo de falha.

4.3.4 Procedimento em Duas Passagens. Como a colisão depende de comparar todos os problemas de um mesmo aplicativo entre si, ela não pode ser verificada problema a problema. Por isso, o protocolo é aplicado em duas etapas.

Na **primeira passagem**, atribui-se o item da norma e o critério WCAG a todos os problemas do aplicativo, sem fechar o resultado de nenhum deles ainda. Essa separação foi necessária porque versões anteriores do protocolo, que fechavam o desfecho durante a análise individual, deixavam colisões passarem: alguns problemas saíam como “Fora do estudo” antes que seus pares fossem analisados, e a colisão nunca era detectada. O único desfecho que pode ser fechado nessa etapa é o Cenário 3: sem nenhum item atribuído, não há colisão possível.

Na **segunda passagem**, todos os problemas com item atribuído são revistos em conjunto, a colisão é verificada e o desfecho final de cada um é registrado.

A Figura 1 sintetiza o procedimento em duas passagens.

4.4 Considerações éticas

Este trabalho não envolveu o recrutamento de participantes para sua realização. Os dados utilizados foram gerados a partir de inspeções de acessibilidade realizadas pelos próprios pesquisadores.

5 RESULTADOS

Esta seção apresenta os resultados da aplicação do protocolo aos 471 problemas do corpus. A Seção 5.1 mostra a distribuição geral dos desfechos; as Seções 5.2, 5.3 e 5.4 detalham cada um dos três cenários de limitação identificados, com exemplos concretos extraídos das avaliações.

5.1 Distribuição Geral dos Desfechos

A Figura 2 mostra a distribuição proporcional dos quatro desfechos. Em 296 casos (62,8%), a norma cobriu adequadamente o problema, sem colisão (Fora do estudo). Os 175 restantes (37,2%) evidenciaram limitações na aplicação prática da norma, divididas em três tipos: ausência total de cobertura normativa (Cenário 3, 12 casos, 2,5%), cobertura parcial de aspectos específicos do problema (Cenário 2, 40 casos, 8,5%) e sobreposição normativa, em que itens excessivamente amplos agrupam, sob a mesma subseção, falhas associadas a

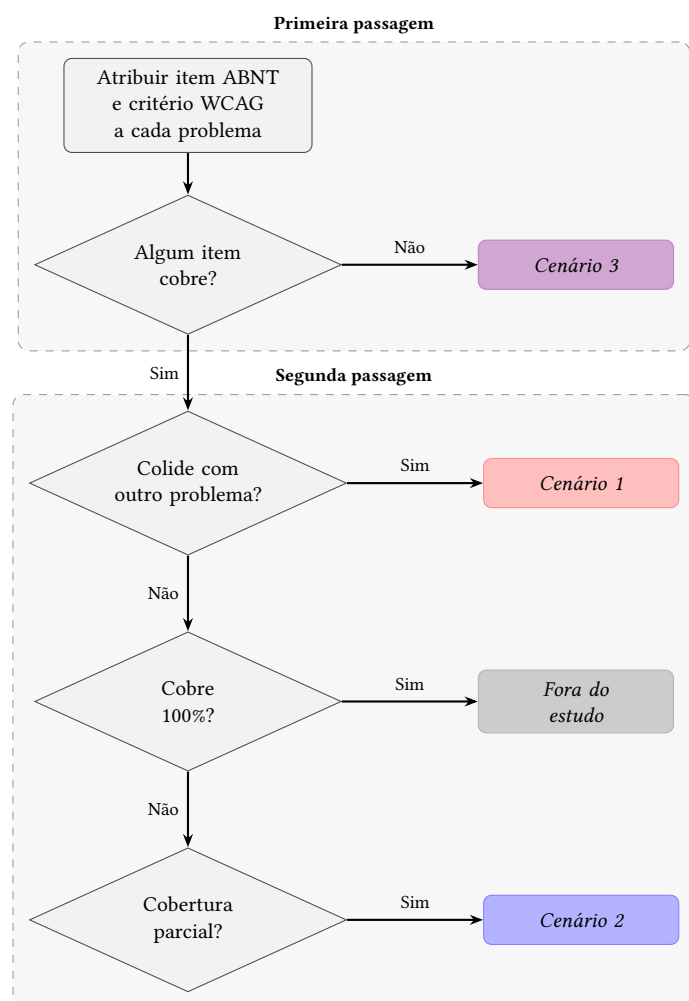


Figura 1: Fluxo esquemático do protocolo em duas passagens. A primeira atribui item e critério WCAG (Cenário 3 pode fechar aqui); a segunda verifica colisão e encerra os desfechos restantes.

critérios WCAG de naturezas distintas (Cenário 1, 123 casos, 26,1%). É importante notar que, no Cenário 1, o item da norma cobre o problema de forma completa; a limitação não está na ausência de cobertura, mas na falta de granularidade do requisito.

Separando granularidade de cobertura, 419 dos 471 problemas (89,0%) tiveram cobertura completa por algum item da norma (independentemente de colisão), e apenas 52 (11,0%) não tiveram cobertura completa (Cenários 2 e 3). O Cenário 1 (26,1%) não indica ausência de cobertura, mas falta de granularidade do item responsável.

A Figura 3 detalha a distribuição por aplicativo. As lacunas concentram-se em Wise e Gov.br, que respondem juntos por 116 dos 123 casos de Cenário 1 (94,3%) e por 28 dos 40 de Cenário 2 (70%). Nubank e Minas App tiveram proporção maior de cobertura

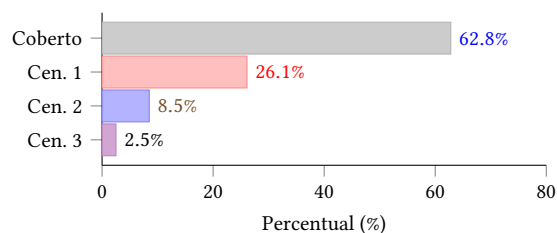


Figura 2: Distribuição proporcional dos 471 problemas pelos quatro desfechos. Em 62,8% dos casos a norma cobriu adequadamente o problema; os 37,2% restantes evidenciaram limitações, seja por ausência, cobertura parcial ou sobreposição normativa.

adequada, o que pode refletir tanto o perfil de acessibilidade desses aplicativos quanto o escopo das avaliações.

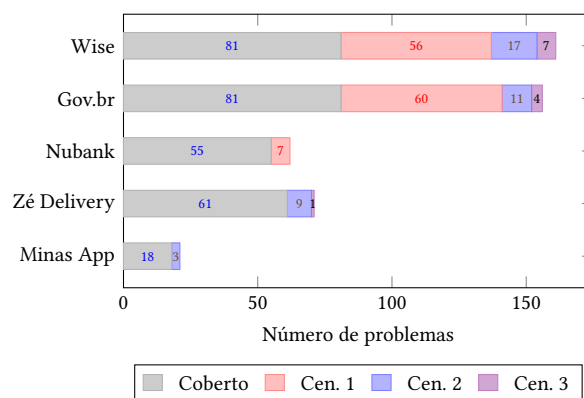


Figura 3: Distribuição dos 471 problemas por aplicativo e desfecho. Wise e Gov.br concentram a grande maioria dos casos de Cenário 1 e Cenário 2.

Como o Wise serviu de piloto para o desenvolvimento do protocolo, recalculou-se a distribuição excluindo-o do corpus (310 problemas): a cobertura adequada sobe de 62,8% para 69,4%, e o Cenário 1 cai de 26,1% para 21,6%, mantendo a mesma ordem relativa entre os desfechos, indício de que o padrão central não é artefato exclusivo do piloto.

Os problemas se distribuíram por 36 itens distintos da norma, associados a 30 critérios de sucesso WCAG 2.1.

5.2 Cenário 1: Itens Muito Amplos

O Cenário 1 ocorre quando um único item da norma acaba agrupando, na mesma subseção, problemas com causas e correções tecnicamente distintas. Diferentemente dos Cenários 2 e 3, o Cenário 1 não envolve ausência de cobertura: o item da norma cobre o problema por completo, mas de forma excessivamente ampla, reunindo sob a mesma subseção falhas associadas a critérios WCAG

de naturezas distintas. A Tabela 1 mostra os três itens que geraram colisão no corpus e quais critérios WCAG diferentes foram encontrados sob cada um.

Tabela 1: Itens da norma que geraram colisão (Cenário 1), com os critérios WCAG heterogêneos reunidos sob cada um.

Item	Critérios WCAG reunidos	<i>n</i>	Cen. 1
5.1.4	4.1.2; 2.1.1; 2.4.3	120	107
5.1.1.20	1.4.2; 4.1.3	20	12
5.1.2.14	2.1.1; 2.1.2; 2.4.3	7	4

O achado mais expressivo do Cenário 1 é a heterogeneidade técnica escondida sob um único item da norma. Foram atribuídos 120 problemas ao item 5.1.4 (Requisitos para codificação), e des- ses, 107 geraram colisão, 87,0% de todos os casos de Cenário 1. O que torna esse número revelador não é o tamanho do item, mas a diversidade dos problemas que ele agrupa: 97 referem-se a falhas de Nome, Função e Valor (WCAG 4.1.2, princípio Robusto), como ícones lidos de forma genérica pelo leitor de telas ou botões sem semântica adequada; 14 são problemas de operação por tecnologia assistiva (WCAG 2.1.1, princípio Operável), como elementos que não podem ser acionados pelo leitor de telas; e 9 envolvem ordem ou posicionamento de foco (WCAG 2.4.3, princípio Operável). Cada um desses grupos exige uma correção diferente: ajustar a semântica de um componente no código não é a mesma coisa que torná-lo acionável por tecnologia assistiva, e nenhum dos dois se confunde com reordenar a sequência de foco. Ainda assim, os três tipos de falha recaem sobre o mesmo item da norma, o item 5.1.4.

Um caso concreto ilustra essa colisão: no Wise, o logotipo da tela de bloqueio é anunciado corretamente pelo leitor de telas apenas no primeiro foco, passando a ser lido apenas como “Imagem” nas interações seguintes, uma falha de nome acessível (WCAG 4.1.2). Esse problema foi registrado no mesmo item 5.1.4 que outros problemas do Wise cuja causa é a impossibilidade de acionamento de elementos via leitor de telas (WCAG 2.1.1), evidenciando como o item agrupa, sob a mesma subseção, correções tecnicamente independentes.

Vale notar que os próprios avaliadores originais tendiam a evitar o item 5.1.4, possivelmente por perceberem sua amplitude e por quererem ser mais precisos na indicação do problema. Isso significa que parte das colisões só ficou visível após a reatribuição feita neste estudo, quando problemas catalogados sob outros itens foram realocados para o que melhor representava sua causa. Os itens 5.1.1.20 (retorno fornecido pela aplicação) e 5.1.2.14 (navegação com tecnologias assistivas) seguem o mesmo padrão em menor escala.

5.3 Cenário 2: Cobertura Parcial

O Cenário 2 reúne os 40 casos em que a norma cobre o aspecto principal do problema, mas deixa de fora algum aspecto específico que o avaliador descreveu. Nesses casos, o problema foi alocado no item mais próximo disponível, mas a parte não coberta ficou registrada como lacuna.

Um padrão recorrente envolve o comportamento dinâmico do indicador de foco. Nas avaliações do Wise, aparecem casos em

que o indicador de foco existe e está visível, mas não acompanha o scroll horizontal de um carrossel de cartões: o usuário perde o rastro visual do foco enquanto navega. O item 5.1.1.15 (Requisitos para indicar o foco de navegação) exige que o foco esteja visível e bem posicionado, mas não diz nada sobre seu comportamento em situações dinâmicas como essa, por isso o problema é registrado nesse item, ainda que de forma incompleta.

Padrão parecido aparece nos requisitos de tempo. Em mais de uma avaliação, o sistema encerra a sessão automaticamente sem avisar o usuário com antecedência, nem oferecer a opção de estendê-la nesse momento. O item 5.1.2.6 exige que o aplicativo ofereça mecanismos de ajuste de tempo, mas não obriga esse aviso prévio. Ambas as situações foram registradas no item correlato disponível, sem correspondência direta na norma para o aspecto específico que o avaliador descreveu.

5.4 Cenário 3: Sem Cobertura

O Cenário 3 compreende os 12 problemas para os quais nenhum item da norma oferece cobertura, mesmo parcial. Todos se concentram em aspectos de interação por leitor de telas que a norma simplesmente não menciona.

Dois padrões se repetem. O primeiro é a *ordem de leitura diferente da ordem visual*: o leitor de telas anuncia elementos, como botões de fechar de janelas modais ou itens de uma lista, em uma sequência diferente da que aparece na tela. O item 5.1.1.10, que trata do posicionamento visual dos elementos, não cobre a sequência de leitura por tecnologia assistiva; usá-lo nesses casos seria exatamente o tipo de cobertura superficial que a regra de rejeição descarta. O segundo é a *falta de gestão de foco em transições*: ao abrir uma nova tela ou modal, o foco do leitor de telas deveria ir automaticamente para o novo conteúdo, mas isso não acontece, e a norma não estabelece nenhum requisito sobre isso.

Aparece também o caso de elementos distintos lidos como um bloco único: conjuntos de controles ou trechos de texto que o leitor de telas anuncia juntos, sem separação. Esses problemas foram mais frequentes nas avaliações do Wise e do Gov.br, que concentram o maior número de registros de interação por leitor de telas no corpus.

6 DISCUSSÃO

Os resultados indicam que a maior parte das limitações não decorre da ausência de requisitos, mas de como os existentes foram redigidos: o Cenário 1 (26,1% do corpus), concentrado quase inteiramente no item 5.1.4, é por definição um problema de granularidade, não de cobertura, sua redação ampla absorve falhas tecnicamente distintas (semântica de componente, operabilidade por tecnologia assistiva, ordenação de foco) sob o mesmo rótulo normativo. Ainda que essa agregação possa ter sido deliberada por parte dos redatores da norma, a ausência de subcritérios que distingam esses princípios WCAG dentro do mesmo item compromete a rastreabilidade e a comparabilidade dos diagnósticos entre avaliações. Esse padrão é consistente com o risco, descrito na literatura de qualidade de software, de que requisitos pouco específicos reduzam a comparabilidade entre avaliações independentes [4]. O fato de os próprios

avaliadores originais evitarem o item 5.1.4 sempre que possível sugere que essa ambiguidade já era percebida na prática, reforçando a utilidade da regra de rejeição empregada neste estudo.

O Cenário 2 (8,5%) sugere um problema diferente: não de amplitude excessiva, mas de escopo insuficiente dentro de um requisito já existente. Os dois exemplos discutidos, o comportamento do indicador de foco durante rolagem horizontal e a ausência de aviso prévio ao encerramento automático de sessão, têm em comum o fato de tratarem de comportamento dinâmico da interface, algo que as WCAG abordam de forma mais explícita do que a NBR 17060 em sua adaptação para o contexto móvel [28, 30]. Isso indica que, ao traduzir os critérios da WCAG para o formato da norma brasileira, parte da granularidade original pode ter se perdido justamente nos aspectos mais específicos da interação móvel, como gestos, rolagem e temporização de sessão.

Já o Cenário 3, embora pequeno em proporção (2,5%), é o que mais diretamente evidencia uma lacuna estrutural: os problemas que o compõem giram em torno da experiência de navegação por leitor de telas, como ordem de leitura, gestão de foco em transições e agrupamento de elementos, área em que a interação móvel se distancia mais da interação Web tradicional [12, 30]. A ausência completa de requisitos para esses casos sugere que a adaptação das WCAG para dispositivos móveis, embora tenha incorporado aspectos como gestos e orientação de tela, não acompanhou da mesma forma as particularidades do uso de tecnologias assistivas nesse contexto.

Tomados em conjunto, os três cenários sugerem uma leitura prática para dois públicos. Para avaliadores que utilizam a norma em auditorias, o item 5.1.4 merece atenção redobrada: sua redação ampla o torna um destino cômodo para problemas diversos, mas essa comodidade mascara heterogeneidade técnica que dificulta tanto a priorização de correções quanto a comparação entre avaliações de diferentes aplicativos. Para os responsáveis pela norma, os padrões aqui identificados apontam uma agenda concreta de revisão: subdividir o item 5.1.4 por tipo de falha de codificação; detalhar os critérios de tempo aplicáveis a sessões e processos cronometrados; e criar requisitos específicos para comportamento dinâmico do foco, gestão de foco em transições de contexto e ordem de leitura por tecnologia assistiva.

A concentração das lacunas em Wise e Gov.br também merece interpretação cuidadosa. Como esses dois aplicativos respondem por 94,3% dos casos de Cenário 1 e 70% dos de Cenário 2, é possível que parte do efeito observado decorra do volume e da profundidade das avaliações realizadas nesses aplicativos, e não necessariamente de uma diferença real na qualidade de acessibilidade entre os produtos. Essa possibilidade é discutida com mais detalhe como limitação do estudo na conclusão.

7 CONCLUSÃO

Este artigo investigou a cobertura da norma ABNT NBR 17060 [2] sobre um corpus de 471 problemas de acessibilidade coletados em cinco aplicativos móveis de domínios distintos. O objetivo não era avaliar os aplicativos em si, mas examinar, de forma sistemática, em que medida a norma oferece instrumentos normativos precisos para enquadrar os problemas que avaliadores encontram na prática.

Os resultados mostram que, em 62,8 problema relatado. Os 37,2 padrões apontam tanto cuidados práticos para avaliadores quanto uma agenda concreta para futuras revisões da norma.

Uma primeira limitação do estudo diz respeito à validação da classificação: todo o processo foi conduzido por uma única pesquisadora, sem segunda codificação independente nem medida formal de concordância entre avaliadores. Embora o protocolo tenha sido desenvolvido justamente para reduzir a subjetividade na atribuição de itens e desfechos, por meio de regras explícitas de fronteira e rejeição (Seção 4.2), a ausência de verificação por um segundo avaliador não permite quantificar o grau de concordância que outros pesquisadores teriam ao aplicar o mesmo protocolo. Trabalhos futuros devem envolver múltiplos avaliadores independentes e reportar uma medida de concordância, fortalecendo a confiabilidade dos desfechos aqui apresentados.

Uma segunda limitação diz respeito à abrangência do corpus: os cinco aplicativos avaliados concentram grande parte das colisões em Wise e Gov.br, que respondem conjuntamente por 94,3 Cenário 2. Aplicativos com menor densidade de problemas de interação por leitor de telas, como Nubank e Minas App, apresentaram proporção elevada de cobertura adequada, o que pode refletir tanto a maturidade de acessibilidade desses produtos quanto o escopo das avaliações realizadas. Como trabalhos futuros, propõe-se a replicação do protocolo com corpora mais amplos, a investigação de domínios não contemplados, como saúde, educação e transporte, e a elaboração de um mapeamento formal entre os itens da ABNT NBR 17060 e os critérios WCAG 2.1, que possa subsidiar tanto revisões da norma quanto o desenvolvimento de instrumentos de avaliação mais precisos.

AGRADECIMENTOS

Os autores agradecem a FAPEMIG, CNPq e CAPES pelo apoio financeiro. As ferramentas de Inteligência Artificial Claude Sonnet 4.6 e Gemini 3.5 Flash foram utilizadas para apoio na revisão e organização do texto. Os autores atestam responsabilidade intelectual por todo o texto do artigo.

REFERÊNCIAS

- [1] Nancy Alajarmeh. 2021. The extent of mobile accessibility coverage in WCAG 2.1: sufficiency of success criteria and appropriateness of relevant conformance levels pertaining to accessibility problems encountered by users who are visually impaired. *Universal Access in the Information Society* 21, 3 (2021), 681–701. doi:10.1007/S10209-020-00785-W
- [2] Associação Brasileira de Normas Técnicas. 2022. *NBR 17060: Acessibilidade — Requisitos de acessibilidade para aplicativos móveis*. Norma Brasileira. ABNT, Rio de Janeiro.
- [3] Marco Billi, Laura Burzagli, Tiziana Catarci, Giuseppe Santucci, and Enrico Bertini. 2010. A unified methodology for the evaluation of accessibility and usability of mobile applications. *Universal Access in the Information Society* 9, 4 (2010), 361–374. doi:10.1007/S10209-009-0180-1
- [4] Giorgio Brajnik, Yeliz Yesilada, and Simon Harper. 2012. Is accessibility conformance an elusive property? A study of validity and reliability of WCAG 2.0. *ACM Trans. Access. Comput.* 4, 2, Article 8 (March 2012), 28 pages. doi:10.1145/2141943.2141946
- [5] Brasil. 2015. Lei n.º 13.146, de 6 de julho de 2015 — Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Diário Oficial da União. https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm
- [6] Rocío Calvo, Faezeh Seyedarabi, and Andreas Savva. 2016. Beyond Web Content Accessibility Guidelines: Expert Accessibility Reviews. In *Proceedings of the 7th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (DSAI '16)*. ACM, 77–84. doi:10.1145/3019943.3019955

- [7] Erik Henrique Da Costa Nunes and Ingrid Teixeira Monteiro. 2024. Exploring the Accessibility Legal Landscape: Accessibility requirements in mobile applications according to ABNT NBR 17060. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems* (Maceió, Brazil) (IHC '23). Association for Computing Machinery, New York, NY, USA, Article 18, 11 pages. doi:10.1145/3638067.3638121
- [8] Matheus Gomes de Andrade, Daniel Mesquita Feijó Rabelo, Ribamar de Souza Martins, and Windson Viana. 2024. Investigating the accessibility of popular mobile Android apps: a prevalence, category, and language study. In *Brazilian Symposium on Multimedia and the Web (WebMedia)*. SBC, 400–404.
- [9] Thomas Fischer, Björn Lundell, and Jonas Gamalielsson. 2025. Coverage of web accessibility guidelines provided by automated checking tools. *Universal Access in the Information Society* 24, 4 (2025), 3615–3637.
- [10] André Pimenta Freire, Paulo Afonso Parreira Junior, Silvia Helena Rigatto, Tiago Silva da Silva, Bruna Poletto Salton, Lael Nervis, Eduardo Augusto Ribeiro, Gabriel Aguiar Alves e Silva, Jamilly Danielle Antonio, João Victor de Oliveira Rosa, et al. 2024. Accessibility Inspections Using the Web Content Accessibility Guidelines by Novice Evaluators: an Experience Report. In *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems*. 1–10.
- [11] Brian Grellmann, Timothy Neate, Abi Roper, Stephanie Wilson, and Jane Marshall. 2018. Investigating Mobile Accessibility Guidance for People with Aphasia. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '18)*. ACM, 338–340. doi:10.1145/3234695.3241011
- [12] International Organization for Standardization. 2008. ISO 9241-171:2008 – Ergonomics of Human-System Interaction – Part 171: Guidance on Software Accessibility. International Standard.
- [13] International Organization for Standardization. 2025. *ISO/IEC 40500:2025 – Information Technology – W3C Web Content Accessibility Guidelines (WCAG) 2.2*. International Standard. International Organization for Standardization.
- [14] Andrew Kirkpatrick, Joshue O'Connor, Alastair Campbell, and Michael Cooper. 2018. *Web Content Accessibility Guidelines (WCAG) 2.1*. W3C Recommendation. World Wide Web Consortium (W3C). <https://www.w3.org/TR/WCAG21/>
- [15] Jonathan Lazar, Daniel F. Goldstein, and Anne Taylor. 2015. *Ensuring Digital Accessibility through Process and Policy*. Morgan Kaufmann, Waltham, MA, USA.
- [16] Rui Lopes and Luis Carriço. 2009. On the Gap between Automated and In-Vivo Evaluations of Web Accessibility. In *Proceedings of the International Conference on Computers for Handicapped Persons (ICCHP 2009)*. Springer, 788–795. doi:10.1007/978-3-642-02713-0_78
- [17] Renan Lopes, Agebson Rocha Façanha, and Windson Viana. 2022. I can't pay! Accessibility analysis of mobile banking apps. In *Proceedings of the Brazilian Symposium on Multimedia and the Web*. 253–257.
- [18] Mary Jo Mueller, Phil Day, and Daniel Montalvo. 2025. Guidance on Applying WCAG 2 to Non-Web Information and Communications Technologies (WCAG2ICT). <https://www.w3.org/TR/wcag2ict-22/>. W3C Group Note, 11 December 2025. Accessed: 01 Jun. 2026.
- [19] Eman Nasr, Wafaa Alsaggaf, and Doaa Sinnari. 2023. Developing Usability Guidelines for mHealth Applications (UGmHA). *Multimodal Technologies and Interaction* 7, 3 (2023), 26. doi:10.3390/mti7030026
- [20] Alberto Dumont Alves Oliveira and Marcelo Medeiros Eler. 2024. Exploring Accessibility of Mobile Applications Through User Feedback: Insights from App Reviews in a Systematic Literature Review. In *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems (IHC '24)*. Association for Computing Machinery, New York, NY, USA, Article 56, 15 pages. doi:10.1145/3702038.3702094
- [21] Kyudong Park, Taedong Goh, and Hyo-Jeong So. 2014. Toward accessible mobile application design: developing mobile application accessibility guidelines for people with visual impairment. In *Proceedings of HCI Korea*. 31–38. doi:10.5555/2729485.2729491
- [22] Kim Patch, Jeanne Spellman, and Kathy Wahlbin. 2015. Mobile Accessibility: How WCAG 2.0 and Other W3C/WAI Guidelines Apply to Mobile. <https://www.w3.org/TR/mobile-accessibility-mapping/>. W3C Working Draft, 26 February 2015. Accessed: 01 Jun. 2026.
- [23] Christopher Power, André Pimenta Freire, Helen Petrie, and David Swallow. 2012. Guidelines are only half of the story: accessibility problems encountered by blind users on the web. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, 433–442. doi:10.1145/2207676.2207736
- [24] Daniel Mesquita Feijó Rabelo, Ribamar de Souza Martins, Isaac Santos, Paulo Henrique G Da Silva, Kiev Gama, and Windson Viana. 2025. Breaking barriers in mobile accessibility: A study of llm-generated native android interfaces. In *2025 IEEE/ACM 12th International Conference on Mobile Software Engineering and Systems (MOBILESoft)*. IEEE, 20–31.
- [25] Gabriel Aguiar A Silva, Paulo Henrique dos Anjos Silveira, Ruan Pablo Gomes Rocha, Felipe Braz, Paulo Afonso Parreira Junior, Monique Scalco Soares Siqueira, André Pimenta Freire, and Silvia Helena Rigatto. 2025. Avaliação da Acessibilidade de um Sistema de Agência de Fomento e Processo de Denúncia Com a Vivência de um Pesquisador com Deficiência Visual. In *Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC)*. SBC, 259–278.
- [26] United Nations. 2006. Convention on the Rights of Persons with Disabilities. Adopted by the United Nations General Assembly.
- [27] World Wide Web Consortium (W3C). 2014. Website Accessibility Conformance Evaluation Methodology (WCAG-EM) 1.0. <https://www.w3.org/TR/WCAG-EM/>. W3C Working Group Note. Accessed: 01 Jun. 2026.
- [28] World Wide Web Consortium (W3C). 2023. Web Content Accessibility Guidelines (WCAG) 2.2. <https://www.w3.org/TR/WCAG22/>. W3C Recommendation. Accessed: June 1, 2026.
- [29] World Wide Web Consortium (W3C). 2024. How People with Disabilities Use the Web. <https://www.w3.org/WAI/people-use-web/>. W3C Web Accessibility Initiative. Accessed: 01 Jun. 2026.
- [30] World Wide Web Consortium (W3C). 2024. Mobile Accessibility. <https://www.w3.org/WAI/mobile/>. W3C Web Accessibility Initiative. Accessed: 01 Jun. 2026.
- [31] World Wide Web Consortium (W3C). 2026. W3C Accessibility Guidelines Evaluation Methodology (WCAG-EM) 2.0. <https://www.w3.org/TR/wcag-em-2/>. W3C Accessibility Guidelines Working Group Draft Note, 5 February 2026. Accessed: 28 Jun. 2026.