Journal on Interactive Systems, 2025, 16:1, doi: 10.5753/jis.2025.5983

Analyzing Accessibility, Usability, and User Experience in Mobile Apps Through User Reviews: An Extended Systematic Literature Review

Alberto Dumont Alves Oliveira [**University of São Paulo** | ***albertodumont@alumni.usp.br***]
Marcelo Medeiros Eler [**University of São Paulo** | ***marceloeler@usp.br***]

School of Arts, Sciences and Humanities - University of São Paulo. Rua Arlindo Béttio, 1000, Ermelino Matarazzo, São Paulo, SP, 03828-000, Brazil.

Received: 01 May 2025 • **Accepted:** 15 August 2025 • **Published:** 23 August 2025

Abstract: User reviews have become an essential source for evaluating quality attributes in mobile applications. Although prior research has examined privacy, security, and digital accessibility for people with disabilities through user feedback, comprehensive studies simultaneously analyzing accessibility, usability, and user experience (UX) through this lens remain scarce. This study expands on our earlier work by conducting a systematic literature review (SLR) to explore the connections between these three dimensions through user feedback analysis. From 670 studies published since 2013, we selected 42 for thematic analysis. The selected papers reflect the field's multidisciplinary nature, encompassing computer science, health sciences, and social sciences, with most studies focusing on applications in both physical and mental health domains. The findings highlight accessibility as an ongoing challenge for users with disabilities and those with specific needs. Key results include: (i) a disparity between two positive and six negative accessibility factors; (ii) usability findings revealing six strengths (e.g., performance, design simplicity) but fourteen shortcomings (e.g., navigation difficulties, unintuitive interfaces); (iii) nine core UX elements, with accessibility and usability as central components; and (iv) thirty-three actionable recommendations for future research. This study provides a holistic understanding of how usability and UX improvements can enhance accessibility. The findings aim to assist developers and organizations in creating more inclusive mobile applications while guiding future research directions.

Keywords: Accessibility, Usability, User Experience, Mobile Applications, App Stores, User Reviews, Systematic Literature Review

1 Introduction

Mobile applications have become integral to daily life, offering users a wide range of services across health, education, finance, and entertainment domains. The rapid growth in mobile app usage has been accompanied by increasing user engagement through online platforms, particularly app stores such as Google Play and Apple App Store. These platforms not only serve as distribution channels but also facilitate user feedback through ratings and textual reviews. Such reviews provide valuable insights into user satisfaction, expectations, and challenges encountered during interaction with mobile applications [Oliveira *et al*., 2024; Nakamura *et al*., 2025].

Digital accessibility remains a persistent issue within mobile technologies. While several initiatives and guidelines, such as the Web Content Accessibility Guidelines (WCAG) and the BBC Mobile Accessibility Guidelines, recent studies indicate that many applications continue to lack adequate support for diverse user needs [Eler *et al*., 2019; Mateus *et al*., 2021; Oliveira *et al*., 2023]. Accessibility shortcomings can negatively affect not only users with disabilities but also broader user groups, including older adults, individuals with temporary impairments, and those in low-resource contexts [Santiago and Marques, 2023; Tavares *et al*., 2024].

In parallel, usability and user experience (UX) have gained prominence in the evaluation of interactive systems [Brito *et al*., 2025]. Usability refers to the degree to which a system enables users to perform tasks effectively and efficiently [Nielsen, 2012], while UX encompasses the emotional and subjective responses resulting from interaction [Campos and Zaina, 2024]. These dimensions are interdependent with accessibility: applications that fail to meet accessibility standards often compromise usability and degrade overall UX. Conversely, improvements in usability and UX can enhance accessibility for a wide range of users [Tavares *et al*., 2024].

Despite the growing availability of user reviews in app stores, it remains unclear how these reviews capture and reflect the intertwined dimensions of accessibility, usability, and UX in mobile applications. Most existing studies address these dimensions in isolation, lacking a holistic analysis of how they co-occur and influence one another from the users' perspective [Palomba *et al*., 2015; Di Sorbo *et al*., 2017; Palomba *et al*., 2018; Li *et al*., 2018; Pelloni *et al*., 2018; Silveira *et al*., 2024].

Therefore, in our previous study, we conducted a Sistematic Literature Review (SLR) to provide a comprehensive overview of how user reviews contribute to identifying accessibility issues in mobile applications [Oliveira and Eler, 2024]. The initial search yielded 638 articles, from which we selected 16 studies published between 2013 and 2024. Our analysis reveals: i) various purposes, methodologies, and approaches that researchers have employed to explore digital accessibility; ii) four main strategies for collecting and analyzing accessibility reviews, with commonly referenced stan-

dards including BBC guidelines, Google Material Design, and WCAG; iii) persistent barriers for users with disabilities; and iv) synthesize 31 recommendations for future research directions in areas such as machine learning and the automatic extraction and classification of user reviews.

In this new manuscript, we further investigate mobile application accessibility while significantly extending our previous work. First, rather than focusing solely on digital accessibility for people with disabilities, we broaden the scope to include usability and UX, resulting in Extended SLR. Second, we incorporate an additional year of studies (April 2024 to April 2025), ensuring an Updated SLR. This revised search yielded 670 articles, with 42 ultimately selected for analysis. Third, we introduce new Research Questions (RQs) to examine accessibility in a broader context: i) positive and negative accessibility aspects highlighted in user reviews; ii) usability strengths and shortcomings identified by users; iii) Key UX elements emerging from app reviews; iv) the interplay between accessibility, usability, and UX; and v) challenges for future research.

Accordingly, we framed our investigation in five RQs, of which RQ_2, RQ_3, RQ_4 were introduced in this extended version of our previous work. Notably, RQ_1 and RQ_5, which were partially addressed in the previous work, were substantially refined through contemporary perspectives and advanced analytical techniques.

RQ_1: What positive and negative aspects of accessibility are most frequently highlighted in user reviews of mobile applications? This question focuses on identifying the positive and negative aspects of accessibility most frequently highlighted in user reviews of mobile applications. While this RQ was partially addressed in previous work, which focused primarily on accessibility issues, the current analysis expands the scope by considering a broader range of studies and reveals what these studies report about the positive aspects of accessibility found in user evaluations.

RQ_2: What usability strengths and shortcomings are emphasized in user reviews of mobile applications? This question investigates the usability aspects emphasized in app reviews, focusing on recurring themes such as interface intuitiveness, task efficiency, and error management. By categorizing these observations, we also examined the usability heuristics referenced in the selected studies, such as Nielsen's and ISO 9241 heuristics.

RQ_3: What key elements of UX emerge from app reviews? This question explores the experiential aspects captured in reviews, including emotional responses (e.g., satisfaction or frustration) and subjective perceptions of interaction quality. RQ_3 aims to uncover how users prioritize non-functional attributes like aesthetics, engagement, and emotional resonance in their evaluations.

RQ_4: What synergistic or conflicting patterns between accessibility, usability, and UX are reported in user reviews of mobile applications? This question investigates how user feedback reveals interconnections among the three quality dimensions, focusing on patterns where improvements in one area support or hinder others. The analysis emphasizes both synergy (e.g., accessibility features that enhance overall usability and UX) and tension (e.g., usability solutions that may reduce accessibility), offering a more integrated view of how users perceive system quality across dimensions.

RQ_5: What challenges or gaps in accessibility, usability, and UX are highlighted as areas for future work? This question, present in the previous work, has been expanded with new thematic elements. It examines the gaps and challenges highlighted in reviews, mapping unmet user needs to opportunities for future research or design innovation. This question focuses on persistent issues that transcend individual apps, suggesting broader directions for the HCI community.

This research makes the following key implications and contributions to the field:

- *Comprehensive Research Synthesis:* We provide researchers and practitioners with a systematic overview of how app reviews are currently being utilized to improve digital accessibility including contexts of usability and UX. The findings reveal how accessibility concerns frequently overlap with usability breakdowns and UX dissatisfaction, reinforcing their interdependence in mobile application experiences.
- *User-Centered Evidence from Naturalistic Feedback:* The research highlights the analytical value of spontaneous user reviews as a complementary data source. These reviews provide real-world perspectives that can surface post-deployment issues often missed by traditional evaluation methods, especially among users with specific accessibility needs.
- *Cross-Domain Insights for Inclusive Design:* The selected studies cover applications from health, education, mental well-being, and everyday utilities. This diversity enables the identification of recurring accessibility and UX issues across domains, offering generalizable insights for inclusive mobile design across different application contexts.
- *Practical Implications:* The study identifies common design flaws that persist even after updates, indicating the need for continuous post-release monitoring, and offering actionable findings for mobile application developers, UX designers, and policy makers. These evidence-based recommendations can directly inform accessibility, usability, and UX improvement strategies and user feedback integration processes.
- *Directions for Future Research and Practice:* We propose a forward-looking research agenda that addresses current limitations in the field. Our findings suggest specific avenues for methodological innovation and empirical investigation to advance accessibility research.

This paper is organized as follows. Section 2 presents the fundamental background concepts. Section 3 examines the related work in the field. Section 4 details the research methodology employed for this SLR. Section 5 presents and discusses the findings in relation to our RQs. Section 6 analyzes potential threats to the validity of this study. Finally, Section 7 provides concluding remarks and proposes directions for future research.

2 Background

This section introduces the three interconnected core concepts examined in this extended SLR: accessibility, usability, and UX.

2.1 Accessibility

Accessibility in the context of mobile applications and interactive systems refers to the extent to which a product, device, service, or environment is available to as many users as possible, including people with disabilities [Santiago and Marques, 2023]. This involves addressing barriers for individuals with auditory, cognitive, neurological, physical, and visual impairments, ensuring they can interact with the system effortlessly and that both software and hardware are adapted to their needs [Oliveira *et al*., 2024].

For developers and organizations, ensuring accessibility in mobile applications is critical to delivering higher-quality solutions that cater to a broader audience [Dos Santos *et al*., 2024; Melo *et al*., 2024]. The lack or inadequacy of digital accessibility affects all users, but its challenges disproportionately impact people with disabilities, who represent 16% of the global population [World Health Organization, 2025]. Beyond functionality, accessibility promotes social inclusion for individuals with disabilities, older adults, people in rural areas, and those in developing countries [Abrahao *et al*., 2014].

To guide the implementation of accessibility standards, well-established guidelines exist for web and mobile platforms, such as the WCAG and the BBC Mobile Accessibility Guidelines [Oliveira *et al*., 2023]. Additionally, regional regulations include: Section 508 (United States), EAA/EN 301549 (Europe), ACA/AODA (Canada), and ABNT NBR 17225 (Brazil). Mobile platforms also feature developer specific initiatives, such as Google's Android Accessibility and Apple's iOS Accessibility frameworks [Mateus *et al*., 2021].

2.2 Usability

Usability addresses to the degree to which a system enables users to achieve specified goals effectively, efficiently, and satisfactorily within a given context [Campos *et al*., 2025]. In mobile applications, usability is critical for ensuring intuitive navigation and interaction, directly influencing UX. Common usability feedback in user reviews often highlights design elements such as font size, button placement, and spacing between interactive components [Oliveira *et al*., 2024].

Nielsen [2012], defined usability as a quality attribute that evaluates the ease of use of user interfaces, encompassing methodologies for improving ease-of-use during the design process. The author proposes five key components: learnability, efficiency, memorability, error recovery, and satisfaction. These guidelines are widely adopted for designing and evaluating interfaces, including mobile applications [Campos *et al*., 2025].

Further formalizing this concept, the ISO 9241-11:2018 standard defines usability as the extent to which specified users can employ a product to achieve goals with effectiveness (accuracy), efficiency (resource expenditure), and satisfaction (user comfort) in a particular context [International Organization for Standardization, 2018].

2.3 UX

UX encompasses a person's overall experience when interacting with a product, system, or service. It focuses on user perceptions, emotional responses, and physical/psychological reactions that occur throughout all stages of interaction, including anticipation, active use, and post-use reflection. Evaluating UX quality enables developers to identify improvement opportunities, thereby optimizing development time, effort, and costs [Campos *et al*., 2025].

The ISO 9241-210 standard defines UX as a person's perceptions and responses resulting from the use or anticipated use of a product, system, or service [International Organization for Standardization, 2019]. While ISO 9241-11 (2018) focuses specifically on usability, ISO 9241-210 addresses human-centered design for interactive systems with explicit emphasis on UX [Park *et al*., 2019; Nakamura *et al*., 2025].

Complementing ISO standards, established evaluation methodologies exist for assessing UX and usability in interactive systems. These include standardized instruments such as the System Usability Scale and the User Experience Questionnaire, which provide measurable UX indicators [Campos *et al*., 2025].

3 Related Work

To the best of our knowledge, no existing SLR has comprehensively examined user reviews in the context of mobile app accessibility, usability, and UX collectively. However, separate SLRs have addressed these domains individually. In the domain of usability and UX, Campos *et al*. [2025] presented an updated systematic mapping study on touchable holographic solutions, analyzing 5,429 studies and selecting 65 relevant publications discussing 200 evaluation technologies. Their key findings revealed a continued focus on time efficiency in usability evaluations and the predominance of generic UX metrics, particularly pleasure/fun dimensions.

For instance, Mariano *et al*. [2024] conducted a systematic mapping study on technologies for evaluating hedonic aspects in text-based chatbots, identifying 69 evaluation technologies with General UX being the most prevalent aspect.

Wang *et al*. [2024] analyzed 167 papers (2012-2022) on app review mining, demonstrating an increasing research trend with rating systems as the most frequently examined characteristic. Their study identified Google Play Store as the primary data source and simple random sampling as the dominant strategy, while highlighting domain knowledge, textual content, and sentiment analysis as crucial future research directions.

Focusing on accessibility, Da Costa Nunes *et al*. [2024] reviewed 406 papers from the Brazilian Symposium on Human Factors in Computing Systems (2016-2022), selecting 91 studies that addressed digital accessibility. Their analysis showed visual impairment as the most researched area (33 papers) and revealed limited legal awareness among re-

searchers, with only 21 papers referencing relevant accessibility legislation.

Similarly, Masruroh *et al*. [2022] examined usability and accessibility evaluation methods for mobile applications (2017-2021), identifying 15 studies with seven focus areas and fourteen distinct evaluation methods, none of which emerged as a standard approach. Mateus *et al*. [2021] conducted a systematic mapping of accessibility issues in websites and mobile applications, comparing three evaluation approaches: automated testing, manual inspections, and user assessments. Their study performed a meta-analysis of 38 distinct accessibility problem types identified across 38 selected studies, derived from an initial pool of 304 articles.

Building upon our previous SLR of digital accessibility studies using user reviews, this extended SLR represents the first comprehensive analysis of user review findings at the intersection of accessibility, usability, and UX. Our research identifies key future research directions, including systematic mapping of accessibility issues, and the advancement of inclusive design principles to promote equality and independence for users with disabilities or specific design preferences.

4 Study Design

For this extended SLR, we maintained the same methodological procedures adopted in the previous study [Oliveira and Eler, 2024] including the guidelines proposed by Kitchenham and Charters Kitchenham and Charters [2007]. The methodology was implemented in three main phases: (i) planning, during which the research objectives and protocol were established; (ii) execution, which involved conducting the search, applying inclusion and exclusion criteria, and extracting data from the selected papers; and (iii) synthesis, aimed at understanding the state of the art in the investigated area and presenting the results.

Figure 1 illustrates the phases and steps of our research methodology, with particular emphasis on the execution phase, which consists of four steps: Step 1 involved conducting searches, yielding 670 papers; Step 2 applied selection criteria based on title and abstract reviews, narrowing the count to 37 papers; Step 3 involved a thorough full-text review, further reducing the selection to 27 papers; Step 4 applied forward snowballing, adding 15 more papers and resulting in a final total of 42 papers. The entire process was conducted between March and April 2025, and the supplementary material is publicly available[1].

4.1 Planning

This study aims to explore how user reviews contribute to identifying issues and enhancing digital accessibility, usability and UX principles in mobile applications. The planning phase encompassed formulating RQs, selecting keywords, conducting pilot searches, exploring scientific databases, and defining selection criteria and data extraction fields. These efforts culminated in the development of a research protocol, with its main components outlined in this subsection.

4.1.1 Research questions

To ensure methodological transparency and coherence between the study's objectives and the review protocol, the formulation of the research questions was guided by the PICOC (Population, Intervention, Comparison, Outcome, Context) criteria, which is widely adopted in systematic reviews [Kitchenham and Charters, 2007].

- *Population (P):* Users of mobile applications.
- *Intervention (I):* Use of user reviews to investigate quality attributes of apps.
- *Comparison (C):* No explicit comparison was defined, although implicit aspects among the three dimensions (accessibility, usability, and UX) may emerge during analysis.
- *Outcome (O):* Identification of positive and negative aspects, interdependencies, and research gaps across the three dimensions.
- *Context (C):* Mobile applications.

The following research questions, as outlined in Section 1, were defined based on the PICOC model:

- RQ_1: What positive and negative aspects of accessibility are most frequently highlighted in user reviews of mobile applications?
- RQ_2: What usability strengths and shortcomings are emphasized in user reviews of mobile applications?
- RQ_3: What key elements of UX emerge from app reviews?
- RQ_4: What synergistic or conflicting patterns between accessibility, usability, and UX are reported in user reviews of mobile applications?
- RQ_5: What challenges or gaps in accessibility, usability, and UX are highlighted as areas for future work?

4.1.2 Keywords and search

In the previous study, we conducted an extensive pilot search to assess the state of the art in digital accessibility. During that phase, preliminary searches were performed in the ACM Digital Library and IEEE Xplore databases. The following keywords were used: "accessibility", "digital accessibility", "person with disability", "mobile app", "application", "user reviews" and "app reviews". Some expressions were also tested: "mobile digital accessibility", "user reviews digital accessibility" and "user reviews mobile digital accessibility".

For the new study, we expanded our search strategy by incorporating additional terms such as "usability", "user experience", and "UX". While the previous research focused specifically on accessibility for people with disabilities, this extended investigation adopts a broader perspective encompassing general accessibility, including usability and UX considerations. Our preliminary analysis suggested the need for a broader search strategy to capture relevant results across

[1] `https://doi.org/10.5281/zenodo.15293712`. Access on 20 August 2025.

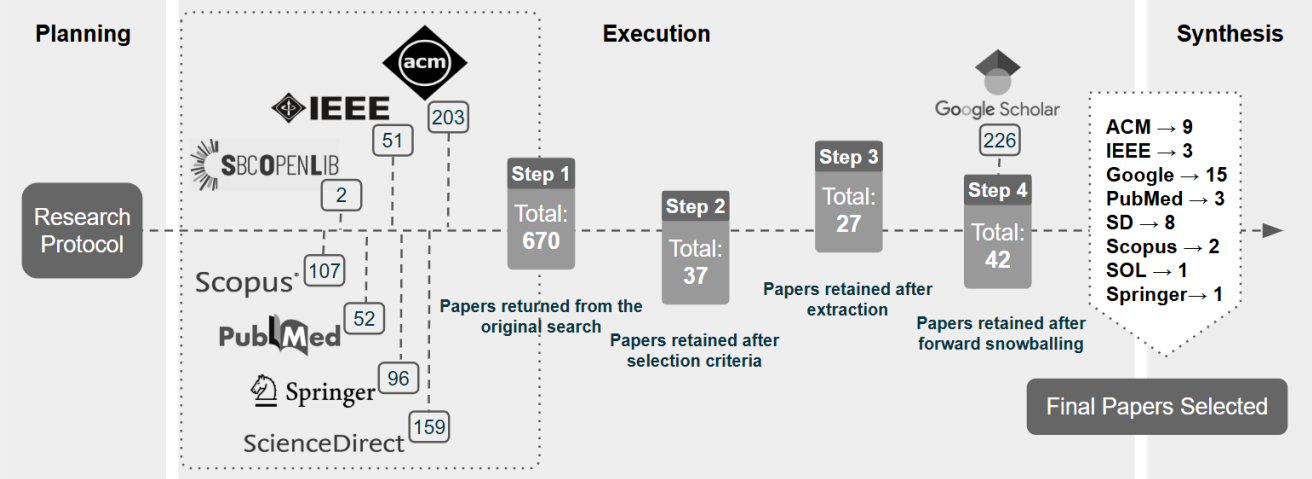

Figure 1. Overview of each phase and step of theSLR process.

multiple themes, which were subsequently refined through systematic filtering.

To conduct a comprehensive search and thereby maximize the possibility of accessing the greatest number of papers, we defined the search string as follows: "("Abstract":accessibility) AND ("Abstract": usability) AND ("Abstract":user experience) AND ("Abstract":reviews*)". Since we opted to use broader and less specific search strings to obtain more results, we also opted to search within the abstracts, as they provide a larger sampling of texts, in addition to titles and keywords. We chose the terms "accessibility", "usability", and "user experience" as they represent the three core themes of this study, while the term "reviews" reflects our primary focus on user evaluations.

The following search engines and digital libraries were selected for this extended SLR: ACM Digital Library[2], IEEE Xplore[3], PubMed[4], Science Direct[5], Scopus[6], SBC Open Lib[7], and Springer Link[8]. Table 1 provides a detailed description and justification for selecting these research engines and databases.

While Web of Science[9] was included in our previous SLR [Oliveira and Eler, 2024], technical constraints prevented its use in the current study. During our search period, the platform's institutional access was limited to author-based queries, with keyword search functionality temporarily unavailable. This restriction would have compromised the reproducibility and methodological consistency of our search strategy. To maintain rigorous comparability across databases, we excluded Web of Science and compensated by expanding queries in other specialized digital libraries. In total, compared to our previous paper, we expanded the search scope from five to seven databases by including PubMed, SBC Open Lib, and Springer Link.

Table 1. Selected search engines and digital libraries with their respective selection criteria for the SLR.

Search Engine / Digital Library	Description and Selection Justification
ACM Digital Library	Official repository of the Association for Computing Machinery, providing comprehensive coverage of computing and human-computer interaction (HCI) literature, essential for technical aspects of mobile app research.
IEEE Xplore	Premier digital library for engineering and technology research, with strong emphasis in accessibility standards and software engineering.
PubMed	Selected for biomedical accessibility studies, as it indexes health-related app usability research.
SBC Open Lib	Institutional repository of the Brazilian Computer Society. Key Brazilian repository for computer science research, ensuring coverage of regional studies in Portuguese.
Science Direct	Elsevier's integrated search platform hosting peer-reviewed journals in computing and social sciences. Complements technical databases with interdisciplinary perspectives.
Scopus	Multidisciplinary citation database offering robust source tracking. Functions as both a bibliometric search engine and curated index.
Springer Link	Publishing platform of Springer Nature. Provides access to high-impact computer science books and conference proceedings.

4.1.3 Selection criteria

The preliminary search phase often yields an extensive collection of articles. To streamline the process and ensure that only pertinent studies are considered, we defined five exclusion criteria (EC01–EC05) and two inclusion criteria (IC01 and IC02) for study selection in this extended SLR. Table 2

[2] `https://dl.acm.org`. Access on 20 August 2025.
[3] `https://ieeexplore.ieee.org/Xplore/home.jsp`. Access on 20 August 2025.
[4] `https://pubmed.ncbi.nlm.nih.gov/`. Access on 20 August 2025.
[5] `https://www.sciencedirect.com/`. Access on 20 August 2025.
[6] `https://www.scopus.com/`. Access on 20 August 2025.
[7] `https://sol.sbc.org.br/`. Access on 20 August 2025.
[8] `https://link.springer.com/`. Access on 20 August 2025.
[9] `https://www.webofscience.com/wos/WOSCC/SMART-search`. Access on 20 August 2025.

outlines these criteria, which are consistent with those used in the prior study.

Table 2. Seleciton criteria.

Criteria	Description
EC01	Papers published before 2013.
EC02	Papers not written in English or Portuguese.
EC03	Papers that were duplicates or unavailable in full text.
EC04	Papers that were not published in conference proceedings or journals
EC05	Papers that fall outside the scope of the research questions.
IC01	Papers that specifically analyzed user reviews as a method for assessing information.
IC02	Papers that provided empirical findings, insights, or recommendations related to mobile app accessibility, usability or UX.

EC01 excludes articles published before 2013, as although mobile applications gained widespread popularity as early as 2008, our latest SLR indicates that the first study addressing user reviews in the context of accessibility appeared in 2013 [Memon *et al*., 2023; Oliveira and Eler, 2024]. EC02 restricts the selection to articles written in English or Portuguese, as these are the languages universally adopted in research and native to the authors. EC03 and EC04 eliminate duplicate or inaccessible studies, as well as those published only as abstracts or posters. EC05 was applied during full-text screening to exclude articles that, while discussing accessibility, usability, UX, or user reviews, were not relevant to our RQs. A study was excluded if it met any of these criteria.

IC01 and IC02 guided the inclusion of studies in this review. IC01 encompassed studies that employed and described the use of user reviews as a methodology for examining digital accessibility, usability, or UX. IC02 included studies presenting empirical findings on the accessibility, usability, or UX of mobile applications. For inclusion, a study had to satisfy both IC01 and IC02, meaning it must not only leverage user reviews as an investigative tool but also provide empirical results or insights related to mobile applications in one of the three dimensions.

4.1.4 Data extraction fields

Once the primary studies were selected, it was essential to extract relevant information from them. Only two data points were collected externally: (i) the "number of citations" (extracted from Google Scholar[10]) and (2) the "article origin" (classified by selection method: original search or forward snowballing).

The primary data extracted from the articles include: authorship details, publication title, year of publication, target platform (iOS / Android), number of applications examined, total volume of user reviews analyzed, terminology applied to user reviews, reviews colleted method, publication venue, data source (libraries / search engines), identified research gaps, study objectives, keywords, abstract content, research methodology, accessibility positive or negative aspects, usability strengths and weaknesses, core UX components, findings, conclusions, future research directions, study limitations, academic contributions, and principal research themes.

[10] `https://scholar.google.com/`. Access on 20 August 2025.

4.2 Execution

To conduct this research, we employed the bibliometric analysis tool StArt[11], which facilitates the organization and systematization of the study selection process Fabbri *et al*. [2016]. The research protocol was configured within StArt to ensure methodological consistency. For data consolidation, we utilized Google Sheets[12], leveraging its filtering capabilities and data visualization features for analysis.

4.2.1 Step 1 - Search

In this step, we applied the search string to the seven predefined databases and obtained the following number of articles, as shown in Figure 1: ACM Digital Library (203), IEEE Xplore (51), PubMed (52), SBC OpenLib (2), ScienceDirect (159), Scopus (107), and Springer Link (96). In total, the original search returned 670 articles. For each search, a BibTex file was generated from each search engine. This format was chosen because it is compatible with the StArt software. The BibTex files allowed for the loading of search results into the software.

4.2.2 Step 2 - Selection criteria

The exclusion criteria (EC01 to EC05) were applied during the title and abstract screening of the 670 articles. After this step, 37 articles remained.

4.2.3 Step 3 - Data extraction

For data extraction, all 37 candidate articles underwent full-text review. During this phase, we reapplied inclusion criteria (IC01 and IC02) and exclusion criterion EC05, which resulted in the removal of 10 articles that were not aligned with our research objectives. The final corpus of 27 articles proceeded to full data extraction, as depicted in Figure 1. While not all extracted data appear explicitly in our SLR results, this comprehensive dataset substantially informed our analytical framework.

4.2.4 Step 4 - Forward snowballing

Following the selection of primary studies, we examined their citations to identify additional relevant literature. Given that this research addresses a relatively underexplored domain, we incorporated forward snowballing methodology to capture recent and potentially significant publications [Wohlin, 2014]. At this stage, we applied exclusion criteria EC1 through EC5 to screen the titles and abstracts of 226 articles cited in Google Scholar by our initial set of 27 selected papers from Step 3. This process yielded 15 additional relevant articles, resulting in a final corpus of 42 studies for our SLR.

[11] `https://www.lapes.ufscar.br/resources/tools-1/start-1`. Access on 20 August 2025.

[12] `https://docs.google.com/spreadsheets/`. Access on 20 August 2025.

4.3 Synthesis

This phase involves analyzing and synthesizing the 42 selected papers included in our SLR, with results presented and discussed in Section 5. Table 3 displays the final corpus of studies, organized by article identifier, authorship and publication year, paper title, citation count, number of mobile applications examined, and quantity of accessibility reviews performed.

4.3.1 Data Analysis

To analyze the data collected in the extended SLR, we employed a qualitative approach grounded in thematic analysis. After extracting relevant information into a structured spreadsheet, we conducted an open coding process to identify recurring themes and concepts. This inductive process allowed us to generate initial codes from the data, which were then organized into higher-order themes.

These themes were grouped into three overarching categories (accessibility, usability, and UX) based on the focus of the studies and the nature of the findings. This categorization enabled us to map and interpret the diverse aspects addressed across the literature, facilitating a structured understanding of the field.

4.3.2 Software Tools

In addition to StArt, used in planning and execution phases, and Google Sheets, applyed for data extraction phase, the following tools were employed to support data analysis:

- Dimensions[13]: A bibliometric platform used to analyze citation diversity metrics, including author gender distribution and geographic representation of cited works, supporting the Citation Diversity Statement.
- Google Sheets: Used for data organization, preliminary analysis, and to generate both the publication timeline visualization (Figure 2) and the journal distribution data (Table 4).
- NotebookLM[14]: An AI-assisted analysis tool used to augment manual content examination of the 42 selected articles, supporting thematic synthesis through semantic search and summary generation of key concepts.
- Voyant Tools[15]: An open-source, web-based text analysis platform used to generate the word cloud visualization (Figure 4), enabling quantitative and qualitative exploration of textual data.
- VOSViewer[16]: A bibliometric software for constructing and visualizing co-author networks (Figure 3) and research trend maps (Figure 5) based on citation data.
- Zotero[17]: A reference management tool utilized to convert bibliographic files from BibTeX format (used by Start) to Research Information Systems format (compatible with VOSViewer).

[13] `https://www.dimensions.ai/`. Access on 20 August 2025.
[14] `https://notebooklm.google.com/`. Access on 20 August 2025.
[15] `https://voyant-tools.org/`. Access on 20 August 2025.
[16] `https://www.vosviewer.com/`. Access on 20 August 2025.
[17] `https://www.zotero.org/`. Access on 20 August 2025.

4.4 Ethical Considerations

Consistent with Apple App Store and Google Play policies, all analyzed user reviews were publicly visible and intended for developer feedback [Apple Inc., 2025; Google LLC, 2025]. The study exclusively utilized anonymized data, excluding any personally identifiable information such as names, photos, or contact details. This practice is aligned with all primary studies in our extended SLR.

The AI tool NotebookLM was applied restricted to processing public article texts for preliminary screening and pattern detection, with mandatory human validation of all analytical outputs.

5 Study Results and Discussion

This section presents and discusses the key findings of our extended SLR. We begin with a bibliometric analysis to outline the research landscape. Subsequently, each of the five RQs is examined in an individual subsection, guided by a thematic synthesis of the selected literature.

The bibliometric analysis focused on four dimensions: publication chronology, author collaborations, dissemination venues, and content relationships via abstracts and keywords. These analyses revealed temporal trends, collaborative networks, and conceptual linkages that grounded our RQs.

Figure 2 compares the annual publication frequency of our two SLRs. The blue line represents findings from the current 2025 extended SLR, while the green line displays results from the prior 2024 study [Oliveira and Eler, 2024]. Trend lines illustrate the overall growth patterns for each review. The 2025 SLR demonstrates greater dynamism, with more pronounced fluctuations and higher publication volume, reflecting its broader scope encompassing usability and UX alongside accessibility. In contrast, the 2024 SLR focused exclusively on accessibility, which appears as a more established and consequently less emergent research domain based on publication trends.

The frequency analysis of publications over the years reveals a significant growth in research output related to the 2025 SLR topic, particularly from 2022 onwards, peaking at 15 publications in 2024. In contrast, the 2024 SLR shows a more stable and modest trend, with fewer publications and a relatively flat growth line. These patterns suggest that the theme of the 2025 SLR has gained increasing academic attention in recent years, whereas the 2024 SLR topic remains steady but less dynamic. This trend suggests a significant surge in scholarly interest regarding the research scope covered by the 2025 SLR.

Figure 3 presents the co-authorship network visualization generated from bibliometric data analyzed using VOSviewer. The analysis identified 138 unique authors, with 19 distinct collaborative connections among them. The visualization identifies four major collaborative clusters among the authors: (i) the red cluster led by Wajdi Aljedaani, (ii) the blue cluster centered on Stephanie Ludi, (iii) the green cluster around Mohamed Wiem Mkaouer, and (iv) a smaller yellow cluster focused on Marcelo M. Eler.

Table 3. Final Selected Scientific Articles in the Extended SLR.

ID	Reference	Title	Cit.	Apps	Reviews
P01	Anam and Yeasin [2013]	Accessibility in smartphone applications: what do we learn from reviews?	16	25	173
P02	Park *et al*. [2019]	Mobile Phone Apps Targeting Medication Adherence: Quality Assessment and Content Analysis of User Reviews	136	704	1,323
P03	Oyebode *et al*. [2020]	Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews	154	104	88,125
P04	Wang *et al*. [2020]	User Perceptions of Virtual Hospital Apps in China: Systematic Search	11	329	2,686
P05	Widnall *et al*. [2020]	User perspectives of mood-monitoring apps available to young people: qualitative content analysis	33	53	1,803
P06	Aljedaani *et al*. [2022]	Automatic Classification of Accessibility User Reviews in Android Apps	24	276	2,663
P07	Bowie-DaBreo *et al*. [2022]	User Perspectives and Ethical Experiences of Apps for Depression: A Qualitative Analysis of User Reviews	29	40	2,217
P08	Du *et al*. [2022]	Listening to Stakeholders Involved in Speech-Language Therapy for Children With Communication Disorders: Content Analysis of Apple App Store Reviews	22	16	721
P09	Haque and Rubya [2022]	For an App Supposed to Make Its Users Feel Better, It Sure is a Joke" – An Analysis of User Reviews of Mobile Mental Health Applications	31	193	4,923
P10	Jang and Park [2022]	Satisfied or not: user experience of mobile augmented reality in using natural language processing techniques on review comments	18	5	8,627
P11	Seon Hong Lee [2022]	Enhancing the Prediction of User Satisfaction with Metaverse Service Through Machine Learning	39	1	4,783,669
P12	Li *et al*. [2022]	An Exploratory Analysis of Human-centric Issues in Parking Solutions Using Surveys and Mobile Parking App Reviews	4	5	1,192
P13	Lieser *et al*. [2022]	The Current State of Mobile Apps Owned by Large Pediatric Hospitals in the United States: Systematic Search and Analysis on Google Play and Apple App Stores	8	11	1,433
P14	Maharjan *et al*. [2022]	What's Up With These Conversational Health Agents? From Users' Critiques to Implications for Design	3	1	2,741
P15	Malik *et al*. [2022]	Evaluating User Feedback for an Artificial Intelligence–Enabled, Cognitive Behavioral Therapy–Based Mental Health App (Wysa): Qualitative Thematic Analysis	47	1	7,929
P16	Reyes Arias *et al*. [2022]	Accessibility Feedback in Mobile Application Reviews: A Dataset of Reviews and Accessibility Guidelines	17	276	1,200
P17	Santiago and Marques [2022]	Are user reviews useful for identifying accessibility issues that autistic users face? an exploratory study	15	8	777
P18	Zhou *et al*. [2022]	A Tale of Two Perspectives: Harvesting System Views and User Views to Understand Patient Portal Engagement	4	1	404
P19	Ahn and Park [2023]	Motivations for user satisfaction of mobile fitness applications: An analysis of user experience based on online review comments	30	–	16,461
P20	Algamdi *et al*. [2023]	Investigating the usability issues in mobile applications reviews using a deep learning model	6	–	3,220
P21	Aljedaani *et al*. [2023]	The state of accessibility in blackboard: Survey and user reviews case study	20	1	15,478
P22	Haque and Rubya [2023]	An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews	177	10	6,235
P23	Jo *et al*. [2023]	Exploring User Perspectives of and Ethical Experiences With Teletherapy Apps: Qualitative Analysis of User Reviews	5	8	3,268
P24	Lobo *et al*. [2023]	Detecting user experience issues from mHealth apps that support stroke caregiver needs: an analysis of user reviews	8	46	13,368
P25	Santiago and Marques [2023]	Exploring user reviews to identify accessibility problems in applications for autistic users	4	8	777
P26	Adaji *et al*. [2024]	Insights from the Review of Apps that Influence Environmental Sustainability	0	70	111,682
P27	Rodrigues Lima De Aguiar *et al*. [2024]	What do Portuguese-Libras automatic translation application users care about? An analysis of user comments to improve application accessibility by the quality assurance team	2	1	584

ID	Reference	Title	Cit.	Apps	Reviews
P28	Alrayani *et al*. [2024]	From Customer's Voice to Decision-Maker Insights: Textual Analysis Framework for Arabic Reviews of Saudi Arabia's Super App	2	–	20,311
P29	Campos and Zaina [2024]	UX-RIVIS: Visualization of information about UX data based on app reviews	0	1	500
P30	Chaudhry and Debi [2024]	User perceptions and experiences of an AI-driven conversational agent for mental health support	7	1	159
P31	Darko *et al*. [2024]	Predicting determinants influencing user satisfaction with mental health app: An explainable machine learning approach based on unstructured data	6	–	17,717
P32	Dos Santos *et al*. [2024]	Evolution may come with a price: analyzing user reviews to understand the impact of updates on mobile apps accessibility	15	340	694
P33	Ganguly and Dasari [2024]	Comparative Study Of AI-Driven Ed-Tech Applications: Insights From Google Play Store Data	0	84	42,000
P34	Guluzade and Sas [2024]	Functionality and User Review Analysis of Mobile Apps for Mindfulness Eating and Eating Disorders	1	27	1,248
P35	Heissler *et al*. [2024]	Can AI Digital Personas for Well–Being Provide Social Support? A Mixed–Method Analysis of User Reviews	0	1	485
P36	Kang and Reynolds [2024]	"This app said I had severe depression, and now I don't know what to do": the unintentional harms of mental health applications	7	36	6,253
P37	Nakamura *et al*. [2024]	UX-MAPPER: A User eXperience Method to Analyze App Store Reviews	5	5	10,000
P38	Nellore *et al*. [2024]	Unveiling User Perspectives: Exploring Themes in Femtech Mobile App Reviews for Enhanced Usability and Privacy	0	46	966,402
P39	Oliveira *et al*. [2024]	Exploring the Influence of Software Evolution on Mobile App Accessibility: Insights from User Reviews	0	340	694
P40	Syukron *et al*. [2024]	A comprehensive study of disaster support mobile apps	2	45	28,161
P41	Nakamura *et al*. [2025]	UX-MAPPER: An automated approach to analyze app store reviews with a focus on UX	0	5	10,000
P42	Okuboyejo *et al*. [2025]	Insights from User Reviews to Improve Suicide Prevention Apps: A Machine Learning and Thematic Analysis-Based Approach	0	39	110,338

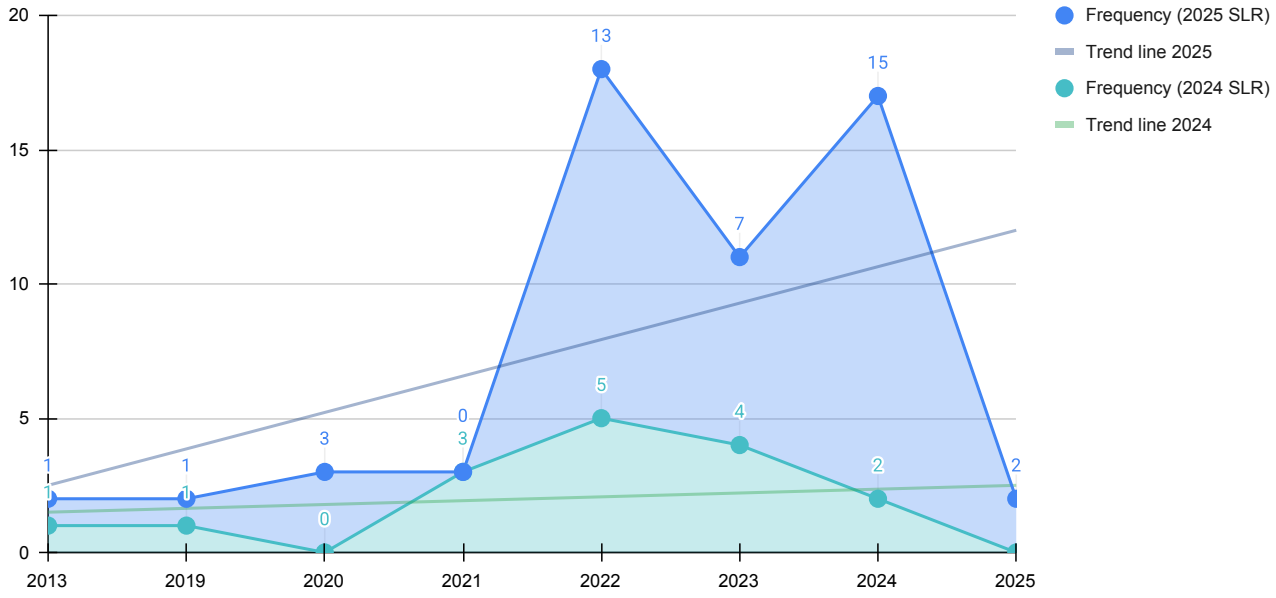

Figure 2. Overview of publications from 2013 to March 2025 and comparison between the 2024 and 2025 SLRs.

The size of each node is proportional to the number of documents authored or co-authored, while the thickness of the links represents the strength of the collaboration between authors. The red cluster shows a dense internal collaboration pattern, while the blue and green clusters exhibit more selective partnerships. The yellow cluster, although smaller, sug-

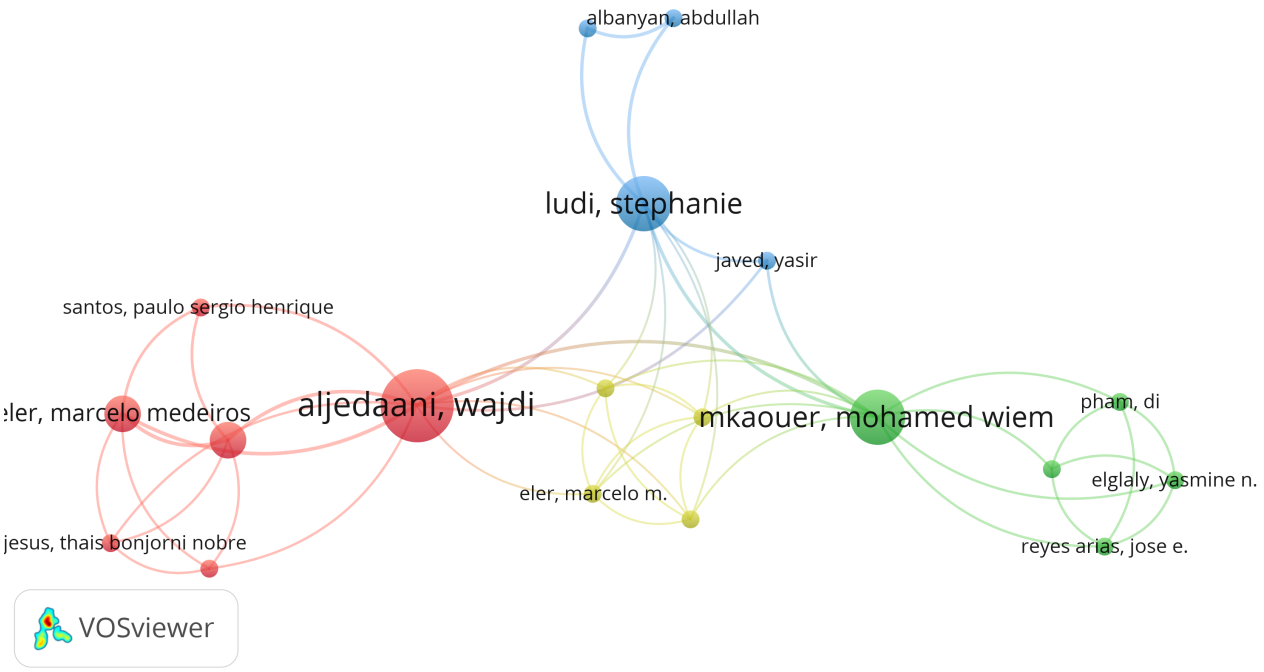

Figure 3. Co-authorship network visualization generated using VOSviewer.

gests emerging research activity or bridging roles between larger groups. Cross-cluster connections indicate interactions across different research teams, highlighting the multidisciplinary nature of studies in the three dimensions.

The bibliometric analysis of publication venues reveals a pronounced predominance of Computer Science research (59.5% – 25 articles), followed by contributions from Health Sciences (26.2% – 11 articles) and Human and Social Sciences (14.3% – 6 articles). This distribution highlights both the continued centrality of technological approaches in addressing the three dimensions challenges, and the field's growing interdisciplinary character.

The 42 selected publications appeared across 31 distinct venues. Journal publications predominated (61.9% – 26 articles), significantly outnumbering conference proceedings (38.1% – 16 articles), reflecting a strong preference for peer-reviewed dissemination channels that typically offer greater methodological rigor and scientific impact.

Table 4 details the major publication venue in the first column, followed by its disciplinary field, number of published articles, and corresponding percentage. Among these venues, JMIR mHealth and uHealth (Health Sciences) and the Proceedings of the Brazilian Symposium on Human Factors in Computing Systems (Computer Science) emerged as the most frequent publication outlets, each accounting for 9.5% (4 articles) of the total corpus.

The comparative analysis of the 2024 and 2025 SLRs reveals significant evolution in the research landscape. The 2024 SLR, exclusively focused on digital accessibility, comprised 16 articles solely from Computer Science. In contrast, the extended 2025 SLR (N=42 articles) demonstrates broader interdisciplinary engagement, incorporating Health Sciences and Human and Social Sciences. Notably, the 2025 SLR exhibits a marked shift toward journal publications (61.9%), diverging from the conference-dominated dissemination pattern of the 2024 SLR. These findings suggest both expanding multidisciplinary interest and advancing research maturity in the three dimensions.

Figure 4 presents a word cloud generated from the abstracts of the 42 analyzed articles, revealing key thematic priorities in the three dimensions. Dominant terms including "apps", "users", "reviews", "health", "mobile", "usability", and "accessibility" reflect a strong research focus on mobile application evaluation, particularly in healthcare and support services. Recurrent references to "mental health", "support", "patient care", and commercial "app stores" further underscore the intersection between technological development and healthcare needs. Methodological terms (e.g., "analysis", "study", "themes") and concepts like "ethical" considerations and "design" highlight the multidisciplinary, user-centered orientation of current studies.

Figure 5 presents a keyword co-occurrence network visualization derived from our extended SLR. The analysis reveals four primary thematic clusters: (i) a green cluster, predominantly featuring "user reviews" and "mHealth", which encapsulates research on user feedback analysis in mobile health applications; (ii) a red cluster, anchored by "machine learning" and "sentiment analysis", underscoring the application of computational methods in UX assessment; (iii) a blue cluster, focused on "accessibility," "privacy", and "usability", representing critical concerns in digital interface design; and (iv) a purple cluster, centered on "mobile applications" and their usability evaluation.

Additionally, while nascent research areas (e.g., artificial

Table 4. Overview of Major Publication Venues.

Conferences and Jornals	Field	Articles	%
JMIR Mhealth and Uhealth	Health	4	9.5%
Proceedings of the Brazilian Symposium on Human Factors in Computing Systems	Computer Science	4	9.5%
JMIR Pediatrics and Parenting	Health	3	7.1%
Journal on Interactive Systems	Computer Science	2	4.8%
Proceedings of the ACM on Human-Computer Interaction	Computer Science	2	4.8%
Proceedings of the CHI Conference on Human Factors in Computing Systems	Computer Science	2	4.8%
Journal of the Brazilian Computer Society	Computer Science	1	2.4%
International Conference on Data Science and Machine Learning Applications	Computer Science	1	2.4%
Human Behavior and Emerging Technologies	Human and Social Sciences	1	2.4%
IEEE International Conference on Healthcare Informatics	Health	1	2.4%

Figure 4. Word Cloud from Abstract Analysis of Scientific Articles in the Extended SLR.

intelligence integration, ethical implications, privacy preservation, and chatbot-mediated mental health interventions) are discernible within the network, their current representation remains comparatively limited relative to established research streams. The node size and edge density within the network reflect both the scholarly prominence and interconnectedness of these research domains.

Figures 4 and 5 were generated to provide a preliminary overview of the selected studies. The word cloud is based on the abstracts of the 42 articles, while the keyword co-occurrence map was created from the author-defined keywords. Although terms like "WCAG" are not among the keywords, they appear three times in the abstracts of two articles [Reyes Arias *et al*., 2022; Oliveira *et al*., 2024]. Conversely, terms such as "ethics" occur less frequently in the abstracts but are included as keywords in two articles [Bowie-DaBreo *et al*., 2022; Jo *et al*., 2023], which explains their presence in the network visualization.

While these visualizations offer an initial conceptual mapping of the corpus, they are inherently limited by the scope of abstracts and predefined keywords. Therefore, the identification of critical areas and the responses to our RQs were based on a full-text analysis of all 42 selected studies. This in-depth review allowed us to uncover nuanced findings, thematic connections, and underrepresented issues that would not be evident from metadata alone.

To elucidate the conceptualization of user reviews in scholarly discourse, we conducted a comprehensive terminological analysis across the 42 articles comprising this extended SLR. As presented in Table 5, our analysis identified seven novel terms beyond those previously cataloged in the 2024 SLR. These emergent terminologies include: "app reviews analysis", "consumer reviews", "online review", "user intent", "user online reviews", "user reviews analysis", "and user satisfaction".

This terminological expansion demonstrates an increasing lexical diversity in academic discussions of user-generated content within mobile app research. The findings show how scholarly approaches to user feedback are evolving in accessibility, usability, and UX research, reflecting methodological specialization and interdisciplinary influences.

5.1 RQ_1: Accessibility Aspects Highlighted in User Reviews

RQ_1 aims at identifying the most frequently mentioned positive and negative aspects related to accessibility in user

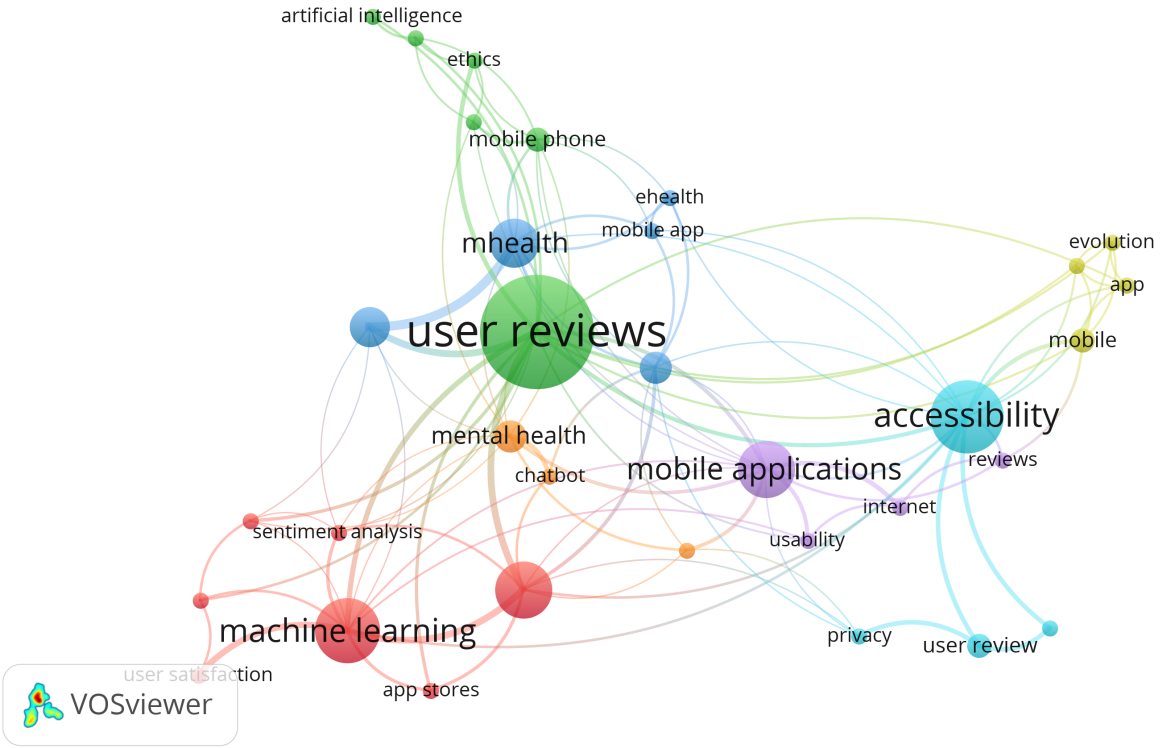

Figure 5. Keyword Network Visualization genereated from the Extended SLR.

Table 5. Extended User Review Terminology Found in the Literature.

accessibility feedback	feedback	**user intent***	user reviews feedback
accessibility reviews	Google App Store reviews	user interaction reviews	**user satisfaction***
accessibility update reviews	mobile app reviews	user feedback	user's feedback
accessibility user reviews	**online review***	**user online reviews***	users' feedback
app review	reviews	user review	
app reviews analysis*	user comments	user reviews	
consumer reviews*	user comments analysis	**user reviews analysis***	

reviews of mobile applications. In addressing this question, we examined how users express accessibility-related experiences, highlighting both well-received features and recurring limitations. Although users often describe these aspects using informal or non-technical language, their feedback reflects practical accessibility concerns that frequently align with established guidelines and inclusive design principles.

Positive aspects related to accessibility in user reviews include:

- *Praise for existing accessibility:* Some users express satisfaction when they encounter apps they consider accessible [Santiago and Marques, 2023].
- *Recognition of accessibility improvements:* Users notice and appreciate accessibility enhancements implemented in app updates, such as color scheme adjustments, increased text size, and interface customization options [Dos Santos *et al*., 2024; Oliveira *et al*., 2024].

We present three user review examples from the analyzed studies that demonstrate positive accessibility evaluations in users' perceptions.

> *"The first and by far the best feature to me on this update is the color labels. Being color blind is now easier to deal with on Google calendar."* (Google Calendar) [Dos Santos *et al*., 2024]

> *"I absolutely LOVE this app. It's easy to use, it has EVERYTHING and it gets constantly updated. I installed it to track my mood and my triggers, but now I use it for everything: I track my sleep, what I eat, how much I drink, how much I exercise. It also has great podcasts and advices. I use A LOT videos of yoga, pil."* (App name not mentioned) [Okuboyejo *et al*., 2025]

> *"I am a blind person. In previous updates the app never supported with my TalkBack option but from last two updates.. The app is very well performing with my talk back.. today 1st time in my life I ordered a food dish without taking any help of others"* (Zomato - Online Food Delivery & RestaurantReviews) [Oliveira *et al*., 2024]

Prevalent negative aspects in user reviews regarding accessibility include:

- *Bugs affecting accessibility:* Reviews highlight how the presence of bugs can hinder effective app use for individuals with accessibility needs [Guluzade and Sas, 2024].
- *Incompatibility with assistive technologies:* Limitations in supporting assistive technologies, such as screen readers commonly used by blind or visually impaired individuals, are also a significant reported issue [Dos Santos *et al*., 2024; Oliveira *et al*., 2024].
- *Incompatibility with device accessibility settings:* Problems arise when apps fail to respect or function properly with the operating system's accessibility settings [Dos Santos *et al*., 2024; Oliveira *et al*., 2024].
- *Lack of specific accessibility features:* A common critique is the absence of essential functionalities such as adjustable text size, customizable color schemes (including dark mode), proper audio controls, and video captions [Widnall *et al*., 2020; Reyes Arias *et al*., 2022; Aljedaani *et al*., 2023; Dos Santos *et al*., 2024].
- *Removal of accessibility features in updates:* Users often express dissatisfaction when app updates remove previously available accessibility features, making apps less usable [Dos Santos *et al*., 2024; Oliveira *et al*., 2024].
- *Visibility and feedback issues for users with Autism Spectrum Disorder:* In apps designed for this user group, reviews may highlight problems related to system state visibility and inadequate feedback during interactions [Santiago and Marques, 2022, 2023].

To illustrate how users may express accessibility concerns, we present representative reviews where users identify specific interface barriers and suggest potential improvements.

> *"Could you make it possible to have the English translation? Or at least make it possible to copy the text so that I can translate it."* (App name not mentioned) [Syukron *et al*., 2024]

> *"overpowering [which] hinders the user experience. The blend of colors makes focusing on content difcult, especially with multiple symptoms, resulting in a chaotic graph"* (App name not mentioned) [Guluzade and Sas, 2024]

> *"I like that the buttons on the original page can't be changed/moved around like in other apps. It helps the child memorize the placement of the buttons, so less frustration."* (LAMP WFL) [Du *et al*., 2022]

While user reviews contain significant accessibility perceptions, it is important to note that negative reviews tend to be more prevalent than positive ones in this domain. This may be because users are more likely to report problems than to praise aspects that work well [Widnall *et al*., 2020; Dos Santos *et al*., 2024; Oliveira *et al*., 2024].

Table 6 synthesizes key accessibility perceptions in user reviews, organized into two complementary perspectives: (a) disability types and (b) app contexts. In both cases, characteristics are grouped into positive and negative feature mentions. The classification of disability types follows general accessibility literature and was adapted to reflect the way user needs were expressed in reviews, while the app contexts illustrate how these needs emerge in specific domains, such as mental health and stroke caregiving. This organization clarifies how user needs manifest across conditions and application scenarios, even when not explicitly named by users.

For auditory impairments, few reviews explicitly address accessibility, though issues like the lack of captions or audio controls are noted. In contrast, users with cognitive or neurodivergent profiles frequently mention confusing interfaces, insufficient guidance, and feedback issues, while positively valuing simple layouts and predictable interactions.

Reviews related to visual impairments emphasize both the importance of screen reader compatibility, text scaling, and contrast, as well as frustrations with unreadable fonts or missing dark modes. Reports touching on neurological or physical conditions highlight navigation barriers and control sensitivity, particularly where fine motor skills are involved.

In summary, generalized accessibility concerns, relevant to multiple user profiles, reveal broader challenges such as instability, removal of accessible features in updates, or incompatibility with assistive settings. At the same time, inclusive design principles and customization options are praised for enhancing overall accessibility. These findings reinforce the need for adaptable interfaces and continuity in accessibility support, particularly during application updates. They also point to gaps that future research and inclusive development practices should address more systematically.

5.2 RQ_2: Usability Strengths and Shortcomings

RQ_2 investigates how usability is portrayed in user reviews of mobile applications, focusing on both strengths and shortcomings reported by users. These reviews provide valuable feedback for developers to understand users' experiences and expectations relative to usability.

The following items presents key usability strengths identified in user reviews:

- Good performance with fast and efficient apps is frequently praised in reviews and efficiency are often mentioned as positive aspects [Park *et al*., 2019; Widnall *et al*., 2020].
- Simple, straightforward design and ease of use are frequently valued and praised by users [Bowie-DaBreo *et al*., 2022; Santiago and Marques, 2022].
- Users appreciate intuitive and easy-to-navigate applications [Widnall *et al*., 2020].
- Stable and well-functioning applications are considered easy to use by many users. Application stability and proper functionality are crucial for good usability [Bowie-DaBreo *et al*., 2022; Guluzade and Sas, 2024].
- In applications for specific needs (e.g., medication adherence), users value features like customizable medication details and personalized reminders [Park *et al*., 2019].
- In mental health applications, an aesthetically pleasing interface is mentioned as a positive aspect [Oyebode *et al*., 2020].

Table 6. Principal Accessibility Perceptions in User Reviews by (a) Disabilities / Specific Needs and (b) App Contexts, with Examples of Positive and Negative Features.

Category	Positive Features / Perceptions	Negative Features / Perceptions
(a) Disabilities / Specific Need		
Auditory Impairment	Automatic translation. Availability of captions when present.	Lack of proper audio controls. Absence of video captions. Difficulty understanding multimedia content.
Cognitive / Neurodivergent (e.g., Autism)	Simple andstraightforward design. Memorization of button placement for Autism Spectrum Disorder. Easy-to-navigate applications. Personalized reminders and adaptable routines.	Visibility and feedback issues for Autism Spectrum Disorder. Confusing interfaces. Lack of proper guidance on app usage. Difficulty navigating features. Abstract or non-intuitive elements (e.g., many buttons).
Multiple / General Needs	Inclusive design principles for all users. Customization options (e.g., personalized reminders). Stable and well-functioning applications.	Bugs affecting general app use. Incompatibility with device accessibility settings. Removal of accessibility features in updates. Login and registration process difficulties. Outdated and difficult-to-use interfaces. App failures, slowness, and UI problems.
Neurological / Physical (e.g., Epilepsy, Reduced Mobility)	Apps that respect operating system accessibility settings and avoid sensory overload.	UI negatively impacts people with epilepsy. Bugs affecting effective app use (e.g., button sensitivity leading to accidental presses). Difficulty with navigation requiring fine motor control. Incompatibility with assistive technologies.
Visual Impairment	Adjustable color schemes (e.g., color labels for color blind). Increased text size. Support for screen readers (e.g., TalkBack). Interface customization options.	Unfriendly font and small font. Incompatibility with screen readers. Lack of dark mode and poor color contrast hindering content focus.
(b) App Contexts		
Mental Health Apps (e.g., Therapy, Mood Tracking)	Intuitive and aesthetically pleasant interfaces supporting adherence. Customizable reminders for medication or mood tracking.	Frequent instability (crashes, freezes, errors). Slow or unresponsive apps. Lack of ongoing support. Excessively complex design in apps expected to be simple and welcoming.
Stroke Caregiver Apps	Remote monitoring and notification features valued when working properly.	Login and registration issues. Inconsistent notifications. Faulty location tracking. Slow performance. Outdated and hard-to-use interfaces.

To illustrate this scenario, we present two review examples where users highlight usability strengths. In the first, the user emphasizes the app's effectiveness, while in the second, the user praises its simplicity while suggesting improvements.

> *"I recently had a medical issue that required me to be on several edications.I've never taken meds regularly and this app helped me make sure I took everything when I was supposed to. Highly recommend."* (App name not mentioned) [Park *et al*., 2019]

> *"Even though it is very simple, I found it really good. Could you create a way for users to develop personalized routines that fit each phase of autistic children, teenagers, or adults? I loved it!"* (Jade Autismo) [Santiago and Marques, 2023]

Conversely, the following prevalent usability shortcomings were identified in user reviews:

- Many users report navigation difficulties, confusing interfaces, and slow or inefficient app performance leading to usability issues [Widnall *et al*., 2020; Du *et al*., 2022; Algamdi *et al*., 2023; Guluzade and Sas, 2024].
- Technical issues such as crashes, slowness, and confusing navigation are common sources of usability shortcomings [Park *et al*., 2019; Bowie-DaBreo *et al*., 2022; Lobo *et al*., 2023; Guluzade and Sas, 2024; Okuboyejo *et al*., 2025]
- Complex or unintuitive interfaces hinder application use [Park *et al*., 2019; Bowie-DaBreo *et al*., 2022; Guluzade and Sas, 2024].
- Difficulty navigating the app and its features is a recurring problem [Park *et al*., 2019; Bowie-DaBreo *et al*., 2022].
- Lack of proper guidance on how to use the application can lead to frustration [Bowie-DaBreo *et al*., 2022; Santiago and Marques, 2023].
- Problems with the sign-up, registration, or login process

are mentioned as usability difficulties [Bowie-DaBreo *et al*., 2022; Guluzade and Sas, 2024].
- Malfunctioning app elements (e.g., sounds, notifications, or data recording) negatively affect usability and UX. Issues with alerts and notifications are particularly noted [Bowie-DaBreo *et al*., 2022; Lobo *et al*., 2023; Aljedaani *et al*., 2023; Syukron *et al*., 2024].
- Compatibility issues across different mobile platforms are also reported [Bowie-DaBreo *et al*., 2022].
- In specific-needs applications (e.g., medication adherence), inflexible reminder settings are a negative point [Park *et al*., 2019].
- In mental health applications, stability and functionality issues are frequently highlighted, including crashes upon opening, errors, freezes, and unresponsiveness. Some users also find certain applications overly complex [Bowie-DaBreo *et al*., 2022; Guluzade and Sas, 2024; Syukron *et al*., 2024; Okuboyejo *et al*., 2025].
- Updates that introduce new bugs or remove useful features lead to negative experiences [Dos Santos *et al*., 2024; Oliveira *et al*., 2024].
- In applications for stroke victim caregivers, login issues, notifications, location tracking, and slow performance are common. The outdated and difficult-to-use interface is also criticized [Lobo *et al*., 2023].
- In mindful eating and eating disorder applications, app instability, errors, and slow performance are predominant negative aspects [Guluzade and Sas, 2024].
- In disaster management applications, app failures, slow performance, and UI problems (e.g., inadequate color schemes) are reported [Syukron *et al*., 2024].

Two user reviews were selected to illustrate usability limitations. In the first example, the user states the app "is not useful" as it fails to meet its intended purpose, while the second mentions an issue involving disabled buttons.

> *"This app is limited and not as useful as it should be because I installed it as I wanted the location of street parking, not just for high rise parking spots."* (Parkopedia Park) [Li *et al*., 2022]

> *"I click enable and the text button goes gray. 5s later the enable comes back frombeing grayed out. i press enable Skill again, same thing. will NOT ENABLE"* (Alexa Health & fitness) [Maharjan *et al*., 2022]

Although not consistently applied across studies, some authors referenced established usability heuristics, such as Nielsen's [Reyes Arias *et al*., 2022; Maharjan *et al*., 2022; Algamdi *et al*., 2023] and ISO 9241 [Oyebode *et al*., 2020; Algamdi *et al*., 2023; Ahn and Park, 2023]. These references were often used to present literature review or to justify design recommendations.

For example, Algamdi *et al*. [2023] cited Nielsen to define usability in the related work section and proposed future directions to evaluate application software reviews based on Nielsen's heuristics. Maharjan *et al*. [2022] identified a related study that applied extended Nielsen's heuristics, recommending increased personification of conversational agents. Algamdi *et al*. [2023] also referred to ISO 9241 to define usability, while Ahn and Park [2023] mentioned ISO 9241 in the literature review section to provide a definition of UX.

User reviews demonstrate that usability is a key determinant of satisfaction in mobile applications. These user-generated evaluations provide rich and actionable insights that support the identification of improvement areas, such as feature prioritization, strategic bug resolution, and evidence-based interface optimization, while also suggesting effective pathways to enhance the overall user experience.

5.3 RQ_3: UX Perceptions

RQ_3 explores how UX is perceived and expressed in user reviews of mobile applications. Through the selected studies, we identified nine key UX elements frequently mentioned by users: accessibility, design, emotional experience, functionality, performance, reliability, support, updates, and usability. These aspects influence user satisfaction, dissatisfaction, and overall impressions of app quality. The following items briefly describe each of these UX elements:

- *Accessibility:* Ease of use for all users, including those with disabilities [Reyes Arias *et al*., 2022]. Accessibility emerges as an important aspect, with users valuing apps featuring simple, straightforward designs [Widnall *et al*., 2020]. Incompatibility with device accessibility settings and lack of specific accessibility features are noted as shortcomings [Reyes Arias *et al*., 2022]. App inaccessibility can lead to feelings of negative emotions, such as frustration, disappointment, exclusion and even perceived discrimination among users with disabilities [Widnall *et al*., 2020; Dos Santos *et al*., 2024; Oliveira *et al*., 2024].
- *Design:* Appealing and well-organized interface [Oyebode *et al*., 2020]. Design and User Interface (UI) play a significant role in UX. An aesthetically pleasing interface is mentioned as a positive aspect for some apps. However, negative reviews may cite interface problems leading to lower ratings [Oyebode *et al*., 2020; Nakamura *et al*., 2024].
- *Emotional Experience:* User feelings and reactions during app use [Malik *et al*., 2022]. Emotional and affective experience also proves relevant. Positive feelings stem from useful and enjoyable apps [Bowie-DaBreo *et al*., 2022]. Negative experiences may arise from complex interfaces, errors, and lack of support [Guluzade and Sas, 2024; Okuboyejo *et al*., 2025]. Perceived helpfulness and non-critical nature of some apps contribute to good UX [Malik *et al*., 2022].
- *Functionality:* Feature effectiveness, presence of desired features, customization [Okuboyejo *et al*., 2025]. Functionality and practical aspects are often praised when working properly. Users appreciate features that help them achieve goals effectively [Okuboyejo *et al*., 2025]. Customization options are also valued Widnall *et al*. [2020]. Conversely, missing desired features or malfunctioning existing features lead to criticism [Kang and Reynolds, 2024]. Inflexibility in feature configuration (e.g., reminders) can be a source of dissatisfaction

[Park *et al*., 2019].

- *Performance:* Stability, speed, absence of crashes and errors [Guluzade and Sas, 2024; Syukron *et al*., 2024]. Performance and Reliability are other key elements. Issues like crashes, slowness, and errors are frequent sources of dissatisfaction. Users express frustration with apps that fail to open, freeze, or become unresponsive. Bugs preventing registration or login are also reported as usability obstacles [Guluzade and Sas, 2024; Syukron *et al*., 2024]. Fast and well-functioning applications contribute to a positive UX [Widnall *et al*., 2020].
- *Reliability:* App trustworthiness, data security [Oyebode *et al*., 2020]. Users express frustration with unstable, slow, or buggy applications [Park *et al*., 2019; Bowie-DaBreo *et al*., 2022; Guluzade and Sas, 2024; Syukron *et al*., 2024].
- *Support:* User assistance when needed [Oyebode *et al*., 2020]. Tutorials or help options may be desirable [Lobo *et al*., 2023].
- *Updates:* Impact of new versions on functionality and stability [Nakamura *et al*., 2025]. App updates can significantly impact UX. While improvements and bug fixes are generally well-received, updates introducing new bugs or removing useful features lead to negative experiences [Dos Santos *et al*., 2024; Oliveira *et al*., 2024; Nakamura *et al*., 2025].
- *Usability:* Ease of use, intuitive design, clear navigation, efficiency [Widnall *et al*., 2020; Nakamura *et al*., 2024, 2025]. Usability is a crucial element frequently mentioned in reviews. Users value simplicity and straightforward, easy-to-use design [Widnall *et al*., 2020]. Apps considered fast and efficient receive positive mentions, while inefficient or slow apps generate frustration [Widnall *et al*., 2020]. Ease of use significantly contributes to positive app evaluations [Malik *et al*., 2022; Nakamura *et al*., 2024]. Conversely, confusing navigation is a common source of negative experiences [Guluzade and Sas, 2024]. App stability and proper functioning are considered essential for usability Oyebode *et al*. [2020]. Reviews frequently express frustration with sluggish or poorly performing apps, as well as interaction challenges [Widnall *et al*., 2020; Algamdi *et al*., 2023; Dos Santos *et al*., 2024].

To illustrate these elements, we selected user reviews from the analyzed studies. The first three reviews express frustration, annoyance, and predominantly negative experiences with app usage. Conversely, the final two examples demonstrate user satisfaction and positive experiences with the applications.

> *"Very frustrated. . .How much of my private information do you actually need? I don't think I should need a separate login at all, but certainly after I've given my Amazon and Facebook, you should not be asking for more!"* (Alexa Health & fitness) [Maharjan *et al*., 2022]

> *"I basically have had issues finding a park on the app even when I'm at venue or when I type it in. So annoying!"* (Wilson Parking) [Li *et al*., 2022]

> *"Warning to anyone with epilepsy or neurological conditions - may not have desired effect! Don't have epilepsy but do have RSD and now in a lot more pain."* (Ihunda Binaural Beat) [Reyes Arias *et al*., 2022]

> *"This is an amazing app that will help you focus more on yourself and your own happiness. Starting with the fact that it lets you rate your mood and write about your day and ending at letting you set tasks for yourself and motivating you. It has really helped me to love myself and learn how to have a better lifestyle...I recommend this app to anyone and everyone, because writing your feelings down and learning to love yourself really will help you gain more confidence and trust in yourself. You will feel more secure."* (App name not mentioned) [Bowie-DaBreo *et al*., 2022]

> *"BOT helps me to change my manner of thinking within minutes. From being pessimistic and hopeless to feeling confident and hopeful. I check in a few times a day depending on when I feel myself slipping back to old ways. I believe that over the long term this app will help me in all aspects of my life exponentially. It's like you're talking to a friend who really knows how you're feeling and cares."* (App name not mentioned) [Haque and Rubya, 2022]

User reviews reveal that UX is influenced by technical, functional, and emotional factors, with accessibility, usability, and performance as key drivers of satisfaction or frustration. Positive features increase perceived value, while critical shortcomings may impair UX, especially for users with specific needs. Recognizing these factors is essential to guide more inclusive and effective design.

5.4 RQ_4: Cross-Dimensional Patterns in Accessibility, Usability, and UX

RQ_4 investigates the presence of synergistic and conflicting patterns among accessibility, usability, and UX as reflected in user reviews of mobile applications. These patterns highlight how improvements in one dimension may reinforce or, conversely, compromise another. By examining these interactions, we aim to better understand the multidimensional nature of app quality from the user's perspective. In what follows, we first outline cross-dimensional associations observed in user reviews for each theme individually:

- *Accessibility as a component of usability and UX:* Reviews frequently demonstrate that accessibility is perceived as a fundamental aspect of usability. An inaccessible app is inherently difficult to use for certain user groups, negatively impacting their overall experience [Reyes Arias *et al*., 2022]. For instance, issues with font size or color contrast are often mentioned in reviews as

usability challenges [Reyes Arias *et al*., 2022; Syukron *et al*., 2024].

> *"really really bad new font extremely unfriendly for those with low vision or dyslexia especially since you seem to not allow a phone's UI to change it to the user's preferred font. i would give it a zero if icould." (Twitter)* [Oliveira *et al*., 2024]

- *Usability as a core factor for good UX, influenced by accessibility:* Usability, referring to ease and effectiveness of use, is a common theme in reviews and directly affects UX [Bowie-DaBreo *et al*., 2022]. Usability problems such as confusing navigation or malfunctioning features lead to negative experiences [Bowie-DaBreo *et al*., 2022; Guluzade and Sas, 2024]. Accessibility influences usability, as apps with accessibility barriers become less usable for some users, impairing their UX [Reyes Arias *et al*., 2022].

> *"Too many buttons. They don't make sense Aesthetics (e.g., 'little'/mouse picture goes to page with about a hundred flowers???? What is that about?!!!!!"* (Speak for yourself) [Du *et al*., 2022]

- *Mutual impact between accessibility and UX:* Reviews show that lack of accessibility can cause frustration and negative emotions, reducing overall satisfaction with the app. Conversely, apps demonstrating accessibility considerations are often praised, contributing to a positive and inclusive UX [Santiago and Marques, 2022, 2023].

> *"The sensitivity on this app is extremely annoying. Just scrolling will cause a button to be pushed at random. It's quite annoying when you're trying to use this and random buttons are getting pushed not by you intentionally."* (Proloquo2Go) [Du *et al*., 2022]

We then present the most salient synergistic patterns, such as inclusive design, customization features, and app stability and performance, which span across the three dimensions.

- *Inclusive Design:* Evidence suggests that inclusive design principles treat accessibility as a requirement that enhances usability for all users, with or without disabilities [Reyes Arias *et al*., 2022]. Reviews praising simple, straightforward, and easy-to-use designs may reflect apps that incorporated accessibility principles, resulting in better overall usability and, consequently, improved UX for everyone [Anam and Yeasin, 2013; Reyes Arias *et al*., 2022].

> *"I gave it five stars because it is really good, but to make it better, it could have audio options like drag the letter a and then say the name of the letter and then say what was formed with the letters."* (App name not mentioned) [Santiago and Marques, 2023]

- *Customization Features:* The ability to personalize app functionalities, such as configuring medication reminders or accessing information through audio, text, and images is valued by users. This improves both accessibility and usability, positively contributing to UX [Anam and Yeasin, 2013; Widnall *et al*., 2020; Santiago and Marques, 2023].

> *"This used to be a good app. After the last update they made the text so small that it's almost impossible to read for people with vision problems. There's also no way to make the text bigger."* (Amazon Prime Video) [Dos Santos *et al*., 2024]

- *Stability and Performance:* Stable and well-functioning apps are considered easier to use [Anam and Yeasin, 2013]. Addressing technical issues affecting stability and performance benefits both general usability and accessibility (e.g., ensuring screen reader compatibility), leading to better UX [Guluzade and Sas, 2024].

> *"... Crashes when tapping onmyexisting vehicle in settings. And when booking, I select "today" and it selects "tomorrow" instead and locks it so that I can no longer change the day. It also shows 2 "today"s in the list."* (Wilson Parking) [Li *et al*., 2022]

Finally, we identify two potentially conflicting patterns that may negatively affect the three dimensions when not properly addressed.

- *Incluse Design Conflict:* While not explicitly highlighted as direct conflicts in reviews, certain approaches to improving usability for one group may inadvertently create accessibility barriers for another (and vice versa). However, sources primarily emphasize synergy through inclusive design [Reyes Arias *et al*., 2022].

> *"Please consider making this app accessible to all, not just sighted users. Blind people need DBT, too. Thank you."* (DBT Coach) [Kang and Reynolds, 2024]

- *Software Update:* App updates aimed at improving usability or adding new features may sometimes introduce new accessibility bugs or remove functionalities useful for users with disabilities, resulting in negative UX for these users [Dos Santos *et al*., 2024]. Reviews may express frustration when updates make an app less accessible than previous versions [Oliveira and Eler, 2024].

> *"Does not work on the latest android system update!! Cannot reinstall!! I need this app to park at uni, please fix!"* (Cellopark Australia) [Li *et al*., 2022]

To consolidate the observed interdependencies across dimensions, Table 7 presents a comparative synthesis of how key app characteristics were reflected in user reviews related to accessibility, usability, and UX. The selected characteristics were derived from the thematic synthesis in this subsection and reflect recurring elements discussed across the three dimensions.

Table 7. Overview of Key Characteristics Across Accessibility, Usability, and UX in Mobile Applications.

Characteristic	Accessibility	Usability	UX
Customization / Personalization	Customizable color schemes and adjustable text size are desired. Personalization of app functionalities (e.g., audio, text, images, medication reminders) is valued.	Customizable medication details and personalized reminders are valued. Inflexible reminder settings are a negative point.	Users prioritize non-functional attributes like engagement. Customization options are valued.
Design Simplicity / Clarity	Users value simple, straightforward designs. Inclusive design principles treat accessibility as a requirement.	Ease of use is frequently valued and praised. Intuitive and easy-to-navigate applications are appreciated.	Aesthetic pleasing interface is a positive aspect. Simplicity contributes to positive app evaluations.
Emotional Response	App inaccessibility can lead to feelings of negative emotions (frustration, disappointment, exclusion, discrimination).	Usability problems (e.g., confusing navigation or malfunctioning features) lead to negative experiences.	Encompasses emotional responses (satisfaction, frustration) and subjective perceptions. Positive feelings stem from useful and enjoyable apps.
Feedback / Guidance	Feedback issues for users with Autism Spectrum Disorder are most highlighted.	Lack of proper guidance on how to use the application can lead to frustration.	Emotional responses (e.g., satisfaction or frustration) result from interaction.
Functionality Effectiveness	Absence of essential functionalities (e.g., video captions) is a critique.	App enables users to perform tasks effectively and efficiently. Malfunctioning app elements negatively affect usability.	Feature effectiveness and presence of desired features influence perception. Missing desired features or malfunctioning features lead to criticism.
Navigation	Impact on cognitive or neurodivergent users, especially those with autism spectrum disorder.	Intuitive and easy-to-navigate applications are appreciated. Navigation difficulties and confusing interfaces lead to usability issues.	Confusing navigation is a common source of negative experiences.
Performance / Speed	Addressing technical issues affecting stability and performance benefits accessibility (e.g., ensuring screen reader compatibility).	Fast and efficient apps are frequently praised. Slowness and inefficient app performance lead to usability issues.	Fast and well-functioning applications contribute to a positive user experience. Sluggish or poorly performing apps generate frustration.
Problem Reporting / Identification	User reviews often indicate specific violations of accessibility guidelines. Reviews highlight how bugs hinder app use for individuals with accessibility needs.	Reviews provide valuable feedback for developers to understand users' experiences and expectations.	Evaluating UX quality enables developers to identify improvement opportunities. Users report challenges encountered during interaction.
Stability / Reliability	Addressing technical issues affecting stability benefits accessibility. Bugs hinder effective app use for individuals with accessibility needs.	Stable and well-functioning applications are considered easy to use. Technical issues such as crashes, slowness, and confusing navigation are common sources of shortcomings.	Users express frustration with apps that fail to open, freeze, or become unresponsive. App trustworthiness and data security are elements.
Updates Impact	App updates may remove previously available accessibility features. Updates can introduce new accessibility bugs.	Updates that introduce new bugs or remove useful features lead to negative experiences.	App updates can significantly impact UX; well-received if improvements, negative if new bugs/removed features.

Customization and personalization, for example, are appreciated for enabling visual adjustments (accessibility), tailored reminders (usability), and more engaging experiences (UX). Design simplicity supports clarity for users with cognitive or sensory needs, facilitates navigation, and enhances satisfaction through visual appeal. Emotional responses, such as frustration or satisfaction, are consistently linked to the success or failure of features across all three dimensions.

Functionality, navigation, and performance are critical in all dimensions. Bugs or lack of responsiveness hinder accessibility reduce usability, and lead to negative user experiences. Similarly, app updates can improve or degrade multiple aspects simultaneously, depending on whether they fix issues or introduce new barriers.

Lastly, the theme of user feedback reinforces the role of reviews in identifying barriers, guiding improvements, and informing inclusive design practices. This comparative view emphasizes how changes in one area, whether positive or negative, often cascade across the others, reinforcing the need for holistic and inclusive development strategies.

5.5 RQ_5: Challenges and Unmet Needs

RQ_5 explores the main challenges, gaps, and unmet needs reported across studies on accessibility, usability, and UX in mobile applications. While several studies explicitly point to areas requiring further research or design improvements, others only imply these needs through limitations or overlooked aspects. In this subsection, we synthesize and categorize the identified directions for future research according to the three dimensions analyzed in this extended SLR.

Several studies in our corpus employed manual thematic analysis or coding strategies to interpret user reviews [Reyes Arias *et al*., 2022; Santiago and Marques, 2023; Dos Santos *et al*., 2024], a pattern also observed in our prior SLR focused exclusively on accessibility [Oliveira and Eler, 2024]. Other studies applied computational techniques such as sentiment analysis, topic modeling, or automated extraction [Wang *et al*., 2020; Aljedaani *et al*., 2022; Memon *et al*., 2023].

Although the adoption of advanced methods such as Natural Language Pprocessing and Machine Learning remains limited, these approaches offer promising opportunities to scale and refine the analysis of user feedback. Future research should explore such techniques to enhance reproducibility, analytical depth, and automation in systematic evaluations.

Future work identified in the field of Accessibility:

- Develop improved accessibility testing methods beyond manual approaches [Aljedaani *et al*., 2022].
- Combine accessibility and usability guidelines to benefit both disabled and non-disabled users [Reyes Arias *et al*., 2022].
- Prioritize identifying and communicating high-risk accessibility errors to designers and developers [Reyes Arias *et al*., 2022].
- Investigate how end-users discuss specific accessibility features using non-technical terminology distinct from guideline taxonomies [Reyes Arias *et al*., 2022].
- Reconsider accessibility in interface design for marginalized populations (elderly, visually impaired) regarding font size, color contrast, and voice control. Accessibility must also address technological literacy and generational habits (e.g., ability to print information, understanding e-bills) [Zhou *et al*., 2022].
- Resolve the most common accessibility issues reported by users of applications like Blackboard [Aljedaani *et al*., 2023].
- Optimize automated classification of accessibility-related user reviews [Santiago and Marques, 2023].
- Consider implementing dark mode and allowing font size/color scheme customization to enhance accessibility [Kang and Reynolds, 2024].
- Conduct deeper analysis of accessibility update reviews, potentially comparing them with general accessibility reviews [Oliveira *et al*., 2024].
- Investigate update impacts on accessibility reviews, including identification of common update terminology, user characteristics, WCAG principles, accessibility issues/improvements, user outcomes, and demands [Dos Santos *et al*., 2024; Oliveira *et al*., 2024].
- Enhance content organization and in-app navigation for greater accessibility and ease of use [Chaudhry and Debi, 2024].
- Expand inclusion of diverse user populations in app development and evaluation, addressing platform-specific app shortages [Okuboyejo *et al*., 2025].

Future work identified in the field of Usability:

- Extend review analysis approaches to applications in other domains to identify strengths/weaknesses and provide design suggestions [Oyebode *et al*., 2020].
- Research should consider app stores and mobile user reviews as effective channels for understanding human-centered needs [Li *et al*., 2022].
- Improve conversational agents regarding functionality, reliability, usability, and enjoyment. Develop deeper understanding of user needs and experiences with emerging technologies like conversational agents for health and wellbeing [Maharjan *et al*., 2022].
- Continue exploring critical review analysis to understand technology usage [Maharjan *et al*., 2022].
- Propose future directions to evaluate application software reviews based on Nielsen's heuristics [Algamdi *et al*., 2023].
- Examine impact of app features (gamification, offline capabilities, UI design) on user engagement in Ed-Tech applications [Ganguly and Dasari, 2024].
- Study effects of app updates, content quality, and engagement features (notifications) on app performance [Ganguly and Dasari, 2024].
- Address significant issues with app stability and limited technical support [Guluzade and Sas, 2024].
- UX-MAPPER tool could be enhanced with temporal review analysis, improving usability and mobile compatibility [Nakamura *et al*., 2023, 2025].

Future work identified in the field of UX:

- Develop empirical understanding of mental health app limitations through user-centered perspectives [Haque and Rubya, 2022].
- Research and design implications for improving transparency, usability, and UX of mental health apps [Haque and Rubya, 2022].
- Future work should account for domain-specific characteristics (beyond games/non-games) when investigating UX, and should address limitations of current UX studies [Jang and Park, 2022].
- Study user goals and success/failure scenarios in telehealth [Zhou *et al*., 2022].
- Holistic understanding of user needs/experiences is required to identify new design opportunities for conversational agents [Maharjan *et al*., 2022].
- Continued adoption and future exploration of review analysis can enhance ability to generate insights about user technology experiences [Maharjan *et al*., 2022].
- UX-MAPPER tool could be improved to extract more useful features and could include app/category comparisons [Nakamura *et al*., 2023].
- Understand the relative/comparative influence of determinants (app design, content, usability, interaction) on user satisfaction with mental health apps [Darko *et al*., 2024].
- Further research should focus on using NLP advances to improve personalized interactions with AI chatbots [Chaudhry and Debi, 2024].
- Longitudinal studies are needed to understand personalization's impact on tool effectiveness [Chaudhry and Debi, 2024].
- Future work could use machine learning to predict user satisfaction determinants [Darko *et al*., 2024].
- Future plans include incorporating walkthrough evaluations, surveys, and interviews to obtain deeper data about user preferences, concerns, experiences, and perspectives from femtech industry stakeholders [Nellore *et al*., 2024].
- Future research could identify cultural/gender influences in reviews [Nakamura *et al*., 2025].

The reviewed literature reveals a wide range of future research opportunities across the three dimensions. In accessibility, studies emphasize the need for more inclusive design practices, automation in accessibility testing, and deeper understanding of user feedback terminology. Usability-focused research highlights the importance of review-based analyses, improvements to conversational agents, and investigation into engagement mechanisms. In the UX domain, future directions include personalized and context-aware design, longitudinal studies, and methodological enhancements for analyzing user feedback. Collectively, these directions reflect a growing interest in refining user-centered approaches and addressing gaps that hinder the inclusivity, functionality, and overall effectiveness of digital applications.

6 Threats to Validity

Assessing Threats to Validity is crucial for ensuring the scientific value of SLR in Software Engineering Zhou *et al*. [2016]. We discuss these threats across four dimensions: search strategy, selection process, data extraction, and interpretative synthesis.

Search strategy validity may be affected by the formulation of the search strings and the selection of digital libraries. Although we expanded our query to include broader terms such as "usability" and "user experience" in addition to "accessibility" and "reviews," relevant studies may have been missed due to variations in terminology or indexing practices. Furthermore, institutional constraints prevented the use of the Web of Science database in this review, which may have limited the comprehensiveness of the search.

Selection bias is a potential threat due to the reliance on inclusion and exclusion criteria based on abstracts and titles in the initial screening. Although these criteria were rigorously defined and consistently applied, subjective judgment during the screening and full-text review stages may have influenced the selection of studies. To mitigate this, we employed a structured protocol and performed forward snowballing to identify additional relevant publications.

Data extraction and coding validity could be impacted by inconsistencies in how key constructs were reported across studies. Given the heterogeneity in terminology, review types, and methodological details, some degree of interpretation was required during data extraction and thematic coding. While efforts were made to ensure consistency, there remains the possibility of misclassification or partial capture of certain findings.

The findings presented in the synthesis phase are base on secondary data reported in the reviewed studies and may be influenced by the original authors' framing and scope. In addition, while user reviews provide valuable perspectives, they often lack contextual information such as demographic characteristics or technical environments, which may limit the generalizability of findings derived from them. Despite these limitations, the study applied established guidelines for SLR and adopted complementary strategies such as snowballing, qualitative coding, and cross-domain analysis to strengthen methodological rigor.

7 Conclusion

This study presented an extended SLR that examined how user reviews of mobile applications reflect and interconnect the dimensions of accessibility, usability, and UX. Building upon previous work focused solely on digital accessibility, this new review incorporated a broader analytical scope and synthesized findings from 42 articles across computer science, health sciences, and human and social sciences.

The analysis discloses that user reviews offer insights into mobile application quality, particularly in identifying barriers and enablers for diverse user populations. Three key findings emerge: First, accessibility challenges persist, with negative reports substantially outnumbering positive evaluations. Second, usability perceptions are praised for attributes like speed, intuitiveness, and stability, but criticized when navigation, feedback mechanisms, or core functionality prove inadequate. Third, UX perceptions encompass emotional, functional, and technical dimensions. Crucially,

the SLR demonstrates significant thematic overlap and interdependence among the three dimensions, suggesting that enhancements in one domain may cascade to others.

The findings demonstrate that analysis of user reviews can effectively identify and prioritize the three dimensions concerns, providing understandings for mobile development processes. These results advocate for incorporating review mining as a standard complementary practice to benefit developers, researchers, and stakeholders in creating more inclusive applications.

Future research efforts should focus on three key directions: (i) advancing automated analysis of user feedback through natural language processing and machine learning techniques, (ii) developing domain-specific classification frameworks for the three dimensions concerns, and (ii) integrating review mining with complementary methods like usability testing and behavioral analytics. Additionally, expanding data sources beyond app stores to incorporate social media discourse would enable more holistic assessment of UX across mobile platforms.

Declarations

Acknowledgements

We extend our appreciation to the users whose app store reviews provided the foundational data for this study, as well as to the researchers whose prior work enabled our systematic analysis. Their contributions collectively advance understanding of accessibility, usability, and UX in mobile applications.

Funding

This study was financed in part by the São Paulo Research Foundation (FAPESP - grant 2022/10910-3).

Authors' Contributions

Conceptualization: A.D.A. Oliveira (ADAO) conceived the original study design; M.M. Eler (MME) extended the framework for this work. **Methodology:** ADAO and MME designed the research approach. **Investigation:** ADAO conducted the SLR, and both ADAO and MME performed the analysis. **Writing:** ADAO prepared the original manuscript draft. **Supervision:** MME provided critical revisions and editorial guidance. **Approval:** All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests, financial or otherwise, that could influence this work.

Availability of data and materials

The complete research materials for this study, including: (i) the StArt project file, (ii) SLR process spreadsheet, (iii) BibTeX file of the 42 selected studies, and (iv) RIS file of the 42 selected studies, are publicly available in the Supplementary Materials section at https://doi.org/10.5281/zenodo.15293712. These resources enable full transparency and facilitate reproducibility of our SLR methodology.

Citation Diversity Statement

While our extended SLR prioritized comprehensive coverage of accessibility research, we recognize that citation practices may reflect broader academic biases. Of the 42 cited articles, we observed that 35.7% (15 articles) were authored by women and 24% (10 articles) originated from institutions in the Global South.

References

Abrahao, A. L., Cavalcanti, A., Pereira, L., and Roque, L. (2014). Accessibility study of touch and gesture interaction with seniors. *Journal on Interactive Systems*, 5(2). DOI: https://doi.org/10.5753/jis.2014.647.

Adaji, I., Idoko, P., and Adisa, M. O. (2024). Insights from the review of apps that influence environmental sustainability. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*, UMAP Adjunct '24, page 154–159, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3631700.3664878.

Ahn, H. and Park, E. (2023). Motivations for user satisfaction of mobile fitness applications: An analysis of user experience based on online review comments. *Humanities and Social Sciences Communications*, 10(1):3. DOI: https://doi.org/10.1057/s41599-022-01452-6.

Algamdi, S., Albanyan, A., and Ludi, S. (2023). Investigating the usability issues in mobile applications reviews using a deep learning model. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0108–0113. DOI: https://doi.org/10.1109/CCWC57344.2023.10099350.

Aljedaani, W., Alkahtani, M., Ludi, S., Mkaouer, M. W., Eler, M. M., Kessentini, M., and Ouni, A. (2023). The state of accessibility in blackboard: Survey and user reviews case study. In *Proceedings of the 20th International Web for All Conference*, W4A '23, page 84–95, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3587281.3587291.

Aljedaani, W., Mkaouer, M. W., Ludi, S., and Javed, Y. (2022). Automatic classification of accessibility user reviews in android apps. In *2022 7th International Conference on Data Science and Machine Learning Applications (CDMA)*, pages 133–138. DOI: https://doi.org/10.1109/CDMA54072.2022.00027.

Alrayani, B., Kalkatawi, M., Abulkhair, M., and Abukhodair, F. (2024). From customer's voice to decision-maker insights: Textual analysis framework for arabic reviews of saudi arabia's super app. *Applied Sciences*, 14(16):6952. DOI: https://doi.org/10.3390/app14166952.

Anam, A. I. and Yeasin, M. (2013). Accessibility in smartphone applications: what do we learn from reviews? In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '13, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/2513383.2513421.

Apple Inc. (2025). App store review guidelines. https://developer.apple.com/app-store/review/guidelines. Access on 27 April 2025.

Bowie-DaBreo, D., Sas, C., Iles-Smith, H., and Sünram-Lea, S. (2022). User perspectives and ethical experiences of apps for depression: A qualitative analysis of user reviews. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3491102.3517498.

Brito, L., Barbosa, I., Silva, M., Drumond, C., Melo, R., Guedes, D., Simões, J., and Sousa, L. (2025). Usability in ipc price collection application: Evaluation of interface and user experience. In *Anais do XXI Simpósio Brasileiro de Sistemas de Informação*, pages 683–692, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/sbsi.2025.246616.

Campos, A. E. S. and Zaina, L. (2024). Ux-rivis: Visualization of information about ux data based on app reviews. In *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems*, IHC '24, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3702038.3702087.

Campos, T., Castello, M., Damasceno, E., and Valentim, N. (2025). An updated systematic mapping study on usability and user experience evaluation of touchable holographic solutions. *Journal on Interactive Systems*, 16(1):172–198. DOI: https://doi.org/10.5753/jis.2025.4694.

Chaudhry, B. M. and Debi, H. R. (2024). User perceptions and experiences of an ai-driven conversational agent for mental health support. *mHealth*, 10(0). DOI: https://doi.org/10.21037/mhealth-23-55.

Da Costa Nunes, E. H., Ribeiro, G. V., Monteiro, I. T., and Gonçalves, E. (2024). Digital accessibility at the brazilian symposium on human factors in computing systems (ihc): An updated systematic literature review. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems*, IHC '23, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3638067.3638075.

Darko, A. P., Antwi, C. O., Adjei, K., Zhang, B., and Ren, J. (2024). Predicting determinants influencing user satisfaction with mental health app: An explainable machine learning approach based on unstructured data. *Expert Syst. Appl.*, 249(PB). DOI: https://doi.org/10.1016/j.eswa.2024.123647.

Di Sorbo, A., Panichella, S., Alexandru, C. V., Visaggio, C. A., and Canfora, G. (2017). Surf: Summarizer of user reviews feedback. In *2017 IEEE/ACM 39th International Conference on Software Engineering Companion (ICSE-C)*, pages 55–58. DOI: https://doi.org/10.1109/ICSE-C.2017.5.

Dos Santos, P. S. H., Oliveira, A. D. A., De Jesus, T. B. N., Aljedaani, W., and Eler, M. M. (2024). Evolution may come with a price: analyzing user reviews to understand the impact of updates on mobile apps accessibility. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems*, IHC '23, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3638067.3638081.

Du, Y., Choe, S., Vega, J., Liu, Y., and Trujillo, A. (2022). Listening to stakeholders involved in speech-language therapy for children with communication disorders: Content analysis of apple app store reviews. *JMIR Pediatr Parent*, 5(1):e28661. DOI: https://doi.org/10.2196/28661.

Eler, M. M., Orlandin, L., and Oliveira, A. D. A. (2019). Do android app users care about accessibility? an analysis of user reviews on the google play store. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*, IHC '19, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3357155.3358477.

Fabbri, S., Silva, C., Hernandes, E., Octaviano, F., Di Thommazo, A., and Belgamo, A. (2016). Improvements in the start tool to better support the systematic review process. In *Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering*, EASE '16, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/2915970.2916013.

Ganguly, R. and Dasari, N. (2024). Comparative study of ai-driven ed-tech applications: Insights from google play store data. In *2024 International Conference on TVET Excellence Development (ICTeD)*, pages 237–243. DOI: https://doi.org/10.1109/ICTeD62334.2024.10844660.

Google LLC (2025). Google play user content and comment posting policy. https://play.google.com/about/comment-posting-policy. Access on 27 April 2025.

Guluzade, L. and Sas, C. (2024). Functionality and user review analysis of mobile apps for mindfulness eating and eating disorders. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, DIS '24, page 1350–1371, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3643834.3661521.

Haque, M. D. R. and Rubya, S. (2023). An overview of chatbot-based mobile mental health apps: Insights from app description and user reviews. *JMIR Mhealth Uhealth*, 11:e44838. DOI: https://doi.org/10.2196/44838.

Haque, M. R. and Rubya, S. (2022). "for an app supposed to make its users feel better, it sure is a joke" – an analysis of user reviews of mobile mental health applications. DOI: https://doi.org/10.48550/arXiv.2209.07796.

Heissler, R., Jonáš, J., Carre, N., Mostovoy, K., and Bunge, E. L. (2024). Can ai digital personas for well-being provide social support? a mixed-method analysis of user reviews. *Human Behavior and Emerging Technologies*, 2024(1):6738001. DOI: https://doi.org/10.1155/2024/6738001.

International Organization for Standardization (2018). Iso 9241-11:2018 - ergonomics of human-system interaction – part 11: Usability: Definitions and concepts. https://www.iso.org/standard/63500.html. Access on 23 April 2025.

International Organization for Standardization (2019). Iso 9241-210:2019 - ergonomics of human-system interaction – part 210: Human-centred design for interactive systems. https://www.iso.org/standard/77520.html. Access on 23

April 2025.

Jang, Y. and Park, E. (2022). Satisfied or not: user experience of mobile augmented reality in using natural language processing techniques on review comments. *Virtual Reality*, 26(3):839–848. DOI: https://doi.org/10.1007/s10055-021-00599-y.

Jo, E., Kouaho, W.-J., Schueller, S. M., and Epstein, D. A. (2023). Exploring user perspectives of and ethical experiences with teletherapy apps: Qualitative analysis of user reviews. *JMIR Ment Health*, 10:e49684. DOI: https://doi.org/10.2196/49684.

Kang, R. M. and Reynolds, T. L. (2024). "this app said i had severe depression, and now i don't know what to do": the unintentional harms of mental health applications. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3613904.3642178.

Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical report, Technical report, EBSE Technical Report EBSE-2007-01. https://docs.edtechhub.org/lib/EDAG684W. Access on 21 August 2025.

Li, R., Obie, H. O., and Khalajzadeh, H. (2022). An exploratory analysis of human-centric issues in parking solutions using surveys and mobile parking app reviews. In *Proceedings of the 33rd Australian Conference on Human-Computer Interaction*, OzCHI '21, page 26–37, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3520495.3520496.

Li, X., Zhang, Z., and Stefanidis, K. (2018). Mobile app evolution analysis based on user reviews. In *SoMeT*. DOI: https://doi.org/10.3233/978-1-61499-900-3-773.

Lieser, T., Huang, Y., and Sezgin, E. (2022). The current state of mobile apps owned by large pediatric hospitals in the united states: Systematic search and analysis on google play and apple app stores. *JMIR Pediatr Parent*, 5(4):e38940. DOI: https://doi.org/10.2196/38940.

Lobo, E. H., Abdelrazek, M., Frølich, A., Rasmussen, L. J., Livingston, P. M., Islam, S. M. S., Kensing, F., and Grundy, J. (2023). Detecting user experience issues from mhealth apps that support stroke caregiver needs: an analysis of user reviews. *Frontiers in Public Health*, 11. DOI: https://doi.org/10.3389/fpubh.2023.1027667.

Maharjan, R., Doherty, K., Rohani, D. A., Bækgaard, P., and Bardram, J. E. (2022). What's up with these conversational health agents? from users' critiques to implications for design. *Frontiers in Digital Health*, Volume 4 - 2022. DOI: https://doi.org/10.3389/fdgth.2022.840232.

Malik, T., Ambrose, A. J., and Sinha, C. (2022). Evaluating user feedback for an artificial intelligence–enabled, cognitive behavioral therapy–based mental health app (wysa): Qualitative thematic analysis. *JMIR Hum Factors*, 9(2):e35668. DOI: https://doi.org/10.2196/35668.

Mariano, P. A. d. L., Souza, A. C. R. d., Guerino, G. C., Chaves, A. P., and Valentim, N. M. C. (2024). A systematic mapping study about technologies for hedonic aspects evaluation in text-based chatbots. *Journal on Interactive Systems*, 15(1):875–896. DOI: https://doi.org/10.5753/jis.2024.4350.

Masruroh, S. U., Rizqy Vitalaya, N. A., Sukmana, H. T., Subchi, I., Khairani, D., and Durachman, Y. (2022). Evaluation of usability and accessibility of mobile application for people with disability: Systematic literature review. In *2022 International Conference on Science and Technology (ICOSTECH)*, pages 1–7. DOI: https://doi.org/10.1109/ICOSTECH54296.2022.9829134.

Mateus, D. A., Silva, C. A., de Oliveira, A. F. B. A., Costa, H., and Freire, A. P. (2021). A systematic mapping of accessibility problems encountered on websites and mobile apps: A comparison between automated tests, manual inspections and user evaluations. *Journal on Interactive Systems*, 12(1):145–171. DOI: https://doi.org/10.5753/jis.2021.1778.

Melo, G. L. N., Menezes, N. d. S., Maciel, A. C. d. S., Teran, L. A., Rocha, T. A. d., Souza, C. R. B. d., and Mota, M. P. (2024). Accessibility issues in establishing awareness on remote collaborative software development. *Journal on Interactive Systems*, 15(1):294–310. DOI: https://doi.org/10.5753/jis.2024.3802.

Memon, Z. A., Munawar, N., and Kamal, M. (2023). App store mining for feature extraction: analyzing user reviews. *Acta Scientiarum. Technology*, 46(1):e62867. DOI: https://doi.org/10.4025/actascitechnol.v46i1.62867.

Nakamura, W., Oliveira, E., Redmiles, D., Oliveira, E., and Conte, T. (2023). O que afeta a ux em aplicações móveis? um resumo de um mapeamento sistemático da literatura sobre fatores que afetam a ux em reviews de lojas de aplicativos. In *Anais Estendidos do XXII Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais*, pages 261–265, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/ihc_estendido.2023.233724.

Nakamura, W. T., C. de Oliveira, E. C., H. T. de Oliveira, E., and Conte, T. (2024). Ux-mapper: A user experience method to analyze app store reviews. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems*, IHC '23, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3638067.3638109.

Nakamura, W. T., Oliveira, E. C. C. d., Oliveira, E. H. T. d., and Conte, T. (2025). Ux-mapper: An automated approach to analyze app store reviews with a focus on ux. *Journal on Interactive Systems*, 16(1):54–74. DOI: https://doi.org/10.5753/jis.2025.4099.

Nellore, N., Mishra, T., and Zimmer, M. (2024). Unveiling user perspectives: Exploring themes in femtech mobile app reviews for enhanced usability and privacy. *Proc. ACM Hum.-Comput. Interact.*, 8(MHCI). DOI: https://doi.org/10.1145/3676530.

Nielsen, J. (2012). Usability 101: Introduction to usability. https://www.nngroup.com/articles/usability-101-introduction-to-usability. Access on 23 April 2025.

Okuboyejo, S., Han, S., Jha, S., Eneja, C., and and, R. O. (2025). Insights from user reviews to improve suicide prevention apps: A machine learning and thematic analysis-based approach. *International Journal of Human–Computer Interaction*, 0(0):1–21. DOI:

https://doi.org/10.1080/10447318.2025.2476711.

Oliveira, A. and Eler, M. (2024). Exploring accessibility of mobile applications through user feedback: Insights from app reviews in a systematic literature review. In *Anais do XXIII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais*, pages 633–647, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.1145/3702038.3702094.

Oliveira, A. D. A., Dos Santos, P. S. H., Marcílio Júnior, W. E., Aljedaani, W. M., Eler, D. M., and Eler, M. M. (2023). Analyzing accessibility reviews associated with visual disabilities or eye conditions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3544548.3581315.

Oliveira, A. D. A., Santos, P. S. H. d., Aljedaani, W., and Eler, M. M. (2024). Exploring the influence of software evolution on mobile app accessibility: Insights from user reviews. *Journal of the Brazilian Computer Society*, 30(1):584–607. DOI: https://doi.org/10.5753/jbcs.2024.4321.

Oyebode, O., Alqahtani, F., and Orji, R. (2020). Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. *IEEE Access*, 8:111141–111158. DOI: https://doi.org/10.1109/ACCESS.2020.3002176.

Palomba, F., Linares-Vásquez, M., Bavota, G., Oliveto, R., Di Penta, M., Poshyvanyk, D., and De Lucia, A. (2015). User reviews matter! tracking crowdsourced reviews to support evolution of successful apps. In *2015 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pages 291–300. DOI: https://doi.org/10.1109/ICSM.2015.7332475.

Palomba, F., Linares-Vásquez, M., Bavota, G., Oliveto, R., Penta, M. D., Poshyvanyk, D., and Lucia, A. D. (2018). Crowdsourcing user reviews to support the evolution of mobile apps. *Journal of Systems and Software*, 137:143 – 162. DOI: https://doi.org/10.1016/j.jss.2017.11.043.

Park, J. Y. E., Li, J., Howren, A., Tsao, N. W., and De Vera, M. (2019). Mobile phone apps targeting medication adherence: Quality assessment and content analysis of user reviews. *JMIR Mhealth Uhealth*, 7(1):e11919. DOI: https://doi.org/10.2196/11919.

Pelloni, L., Grano, G., Ciurumelea, A., Panichella, S., Palomba, F., and Gall, H. C. (2018). Becloma: Augmenting stack traces with user review information. In *2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pages 522–526. DOI: https://doi.org/10.1109/SANER.2018.8330252.

Reyes Arias, J. E., Kurtzhall, K., Pham, D., Mkaouer, M. W., and Elglaly, Y. N. (2022). Accessibility feedback in mobile application reviews: A dataset of reviews and accessibility guidelines. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI EA '22, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3491101.3519625.

Rodrigues Lima De Aguiar, A. C., Melo Alves, L. E., Melo Feliciano De Sá, F., and Costa Aguiar, Y. P. (2024). What do portuguese-libras automatic translation application users care about? an analysis of user comments to improve application accessibility by the quality assurance team. In *Proceedings of the XI Latin American Conference on Human Computer Interaction*, CLIHC '23, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3630970.3631069.

Santiago, M. T. and Marques, A. B. (2022). Are user reviews useful for identifying accessibility issues that autistic users face? an exploratory study. In *Proceedings of the 21st Brazilian Symposium on Human Factors in Computing Systems*, IHC '22, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3554364.3559114.

Santiago, M. T. and Marques, A. B. (2023). Exploring user reviews to identify accessibility problems in applications for autistic users. *Journal on Interactive Systems*, 14(1):317–330. DOI: https://doi.org/10.5753/jis.2023.3238.

Seon Hong Lee, Haein Lee, J. H. K. (2022). Enhancing the prediction of user satisfaction with metaverse service through machine learning. *Computers, Materials & Continua*, 72(3):4983–4997. DOI: https://doi.org/10.32604/cmc.2022.027943.

Silveira, L. A. d., Aljedaani, W., and Eler, M. M. (2024). Strategic insights into integrating accessibility in software development: Motivations, barriers, commitments, and business value. In *Proceedings of the Brazilian Symposium on Human Factors (IHC)*, pages 1–14. ACM. DOI: https://doi.org/10.1145/3702038.3702108.

Syukron, M., Madugalla, A., Shahin, M., and Grundy, J. (2024). A comprehensive study of disaster support mobile apps. DOI: https://doi.org/10.48550/arXiv.2407.08145.

Tavares, D., Lobato, S., Rodrigues, K., Sacramento, C., Faria, M., and Gouveia, Y. (2024). Acessibilidade nas mídias sociais para pessoas com deficiência visual. In *Minicursos do IHC 2024*. Sociedade Brasileira de Computação. DOI: https://doi.org/10.5753/sbc.16123.0.1.

Wang, X., Zhang, T., Tan, Y., Shang, W., and Li, Y. (2024). How to effectively mine app reviews concerning software ecosystem? a survey of review characteristics. *Journal of Systems and Software*, 213:112040. DOI: https://doi.org/10.1016/j.jss.2024.112040.

Wang, Y., Liu, Y., Shi, Y., Yu, Y., and Yang, J. (2020). User perceptions of virtual hospital apps in china: Systematic search. *JMIR Mhealth Uhealth*, 8(8):e19487. DOI: https://doi.org/10.2196/19487.

Widnall, E., Grant, C. E., Wang, T., Cross, L., Velupillai, S., Roberts, A., Stewart, R., Simonoff, E., and Downs, J. (2020). User perspectives of mood-monitoring apps available to young people: Qualitative content analysis. *JMIR Mhealth Uhealth*, 8(10):e18140. DOI: https://doi.org/10.2196/18140.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, EASE '14, New York, NY,

USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/2601248.2601268.

World Health Organization (2025). Global report on health equity for persons with disabilities. https://www.who.int/teams/noncommunicable-diseases/sensory-functions-disability-and-rehabilitation/global-report-on-health-equity-for-persons-with-disabilities. Access on 20 August 2025.

Zhou, J., Arriaga, R. I., Liu, H., and Huang, M. (2022). A tale of two perspectives: Harvesting system views and user views to understand patient portal engagement. In *2022 IEEE 10th International Conference on Healthcare Informatics (ICHI)*, pages 373–383. DOI: https://doi.org/10.1109/ICHI54592.2022.00059.

Zhou, X., Jin, Y., Zhang, H., Li, S., and Huang, X. (2016). A map of threats to validity of systematic literature reviews in software engineering. In *2016 23rd Asia-Pacific Software Engineering Conference (APSEC)*, pages 153–160. DOI: https://doi.org/10.1109/APSEC.2016.031.