Investigating LLM-based tools to support Usability, Accessibility, User Experience in HCI activities: A Systematic Literature Mapping

Simone de Oliveira Santos
Robson de Moraes Alves
Francisco Carlos de Mattos Brito Oliveira
sdo.santos@aluno.uece.br
robson.moraes@aluno.uece.br
fran.oliveira@uece.br
State University of Ceará
Fortaleza, Ceará, Brazil

Francisco Ivanilso Soares Araujo
Atlantico Institute
State University of Ceará
Fortaleza, Ceará, Brazil
ivanilso_soares@atlantico.com.br

ABSTRACT

This study presents a Systematic Literature Mapping (SLM) of recent research on the application of Large Language Models (LLMs) within Human–Computer Interaction (HCI). Drawing from 104 peer-reviewed articles published between 2021 and 2025 in the ACM Digital Library and IEEE Xplore, the mapping investigates the role of LLMs in supporting usability, accessibility, and user experience (UX) activities. LLMs, rooted in Natural Language Processing (NLP), have demonstrated widespread utility across various domains, including healthcare, education, and software engineering. Within HCI, their applications include conversational agents, accessibility enhancements, user journey modeling, prototyping tools, and adaptive user interfaces. The findings highlight how UX professionals have integrated LLM-based solutions into design workflows, leveraging these technologies for improved engagement, personalization, and collaboration. This research addresses 4 guiding questions concerning the domains of application, methodologies, and practitioner adoption of LLMs. The study contributes by (i) offering a comprehensive mapping of HCI-related LLM literature, (ii) identifying areas of active research, and (iii) uncovering emerging trends and open challenges that may inform future investigations.

KEYWORDS

Generative AI, usability, accessibility, UX, HCI activities, systematic mapping

1 INTRODUCTION

Large Language Models (LLMs) perform tasks through Natural Language Processing (NLP) using pre-trained models. Among various types of applications, LLMs have been widely employed across domains such as business and industry, education, politics, healthcare and medicine, science and technology, and many others [5]. Within Human-Computer Interaction (HCI), LLMs have supported a range of activities including usability [25, 36], accessibility [2], and user experience evaluations [87]. The rapid growth of research involving LLMs in this field highlights the need for further investigation and categorization of the impacted domains. A thorough understanding of the potential use of these tools to automate processes may contribute to the advancement of the HCI discipline.

Research involving LLMs has been explored within the broader field of Software Engineering (SE), raising questions regarding their application in SE-related tasks, requirements engineering, data collection methods, tool performance evaluation, and unresolved issues [31, 35]. Discussions held in workshops have addressed the topic, highlighting human-centered aspects, emerging trends, ethical concerns, security, testing, and sustainable SE [71].

Given the ongoing discussions on the topic, it is essential to understand how LLMs are being used to support usability and accessibility activities, as well as how UX professionals are integrating these tools into their workflows. To address this, we aimed to conduct a Systematic Literature Mapping (SLM) using ACM Digital Library and IEEE Xplorer, cataloging 104 articles published in the last five years (2021-2025) that focus on support usability, accessibility, user experience and involve the use of LLMs. Our research questions include: *Where have LLM-based techniques been applied for usability and accessibility? How have researchers used LLMs for user journey analyses? How have LLM-based approaches contributed to user experience? How do UX practitioners use LLM to support their activities?*. We also categorized the papers into domains based on a taxonomy describe by Pang *et al.* [67].

We identified that research supporting usability and accessibility includes the development of chatbots designed to assist various user groups such as older adults, patients, and caregivers, particularly in healthcare and education contexts. Chatbots now rely on natural language processing, improving usability for diverse audiences and facilitating collaboration with clients in product creation and accessibility violation checks. Regarding the user journey, persona simulation and fictional users, capture user intention, have been used for mapping user journey. LLMs enhance the user experience by enabling conversational interfaces, mediating human-AI interaction, and supporting the execution of complex tasks. UX professionals have benefited from prototyping and collaboration tools within design spaces. LLM-based tools are used to evaluate source-code and user interface to improve usability and check for violations of accessibility. Applications and techniques were identified across the domains of Augmented Capabilities, Communication

In: Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia'2025). Rio de Janeiro, Brazil. Porto Alegre: Brazilian Computer Society, 2025.

ISSN 2966-2753

& Writing, Accessibility, User Experience, and Health. The most extensively researched domains are those focused on improving communication between users and systems, particularly through advancements in chatbots, now often referred to as agents, and efforts aimed at enhancing designer productivity and overall user experience.

In general, this study presents an investigation into the domains in which LLMs are being applied within HCI activities, serving as a guide for the community to identify gaps and explore new research areas. In summary, our contributions are:

- A systematic mapping of 104 papers related to the use of LLMs in usability, accessibility, and user experience activities, based on articles retrieved from the ACM Digital Library and IEEE Xplore (2021-2025);
- A distribution of papers into domains related to HCI activities. The list of all papers accepted for this mapping are available in https://doi.org/10.5281/zenodo.16809386;
- Identification of emerging trends and potential research gaps.

2 BACKGROUND AND RELATED WORK

Usability refers to the ease with which users can learn and operate an interface, as well as the satisfaction derived from its use [64]. It is also considered a quality criterion for systems and, in some cases, this quality, especially when linked to users' feelings and emotions, is referred to as User Experience (UX) [77]. Accessibility, on the other hand, is concerned with eliminating barriers that prevent users from accessing and interacting with a system's interface. This means enabling more people to use the system, regardless of whether they have a disability or not [7].

Evaluation in HCI is one of the components that comprise the design process. The usability and accessibility of a system can be assessed either through expert inspection or by observing end users [7]. Both approaches have their respective costs, advantages, and disadvantages, and may be conducted manually or in an automated manner. The use of techniques and tools to support usability and accessibility has been a recurring subject of research, and, similar to other areas of computing. LLMs and others techniques of Artificial Intelligence (AI) have been employed by HCI researchers to support their activities. LLMs have been applied to automate heuristic evaluations of user interfaces [25, 36], produce accessible video description [52], generate accessible code [2], and also to create chatbots to improve user experience [19].

Due to the rapid growth and advancement of LLMs, their forms of use have also expanded. As a result, the variety of applications of this technology has increased. Several systematic literature reviews have documented and categorized how LLMs are being used in research. For example, the study conducted by Kursat *et al.* [31] listed LLM usage across various phases of Software Engineering (SE). In the field of Human-Computer Interaction (HCI), a systematic literature review by Pang *et al.* [67], which inspired our research, analyzed 153 articles employing LLMs in diverse HCI activities. The review covered specifically the CHI proceedings from 2020 to 2024 and primarily investigated where LLMs were applied across HCI practices, as well as concerns regarding their usage. In our study, the mapping focuses only on usability, accessibility, and User Experience (UX) activities, consulting specialized databases ACM and IEEE. The goal is to conduct a quantitative analysis to categorize the areas in which LLMs are currently being used to support usability, accessibility, and UX tasks, and to identify how UX professionals are leveraging LLM-based tools in their workflows.

3 RESEARCH METHOD

The methodology used to construct this mapping study was based on the protocol defined by Kitchenham *et al.* and Petersen *et al.* [44, 70]. The main objective of this mapping is to investigate how LLMs have been employed to support usability, accessibility, user experience, user journey in interactive systems, and also to support the UX practitioner activities. The research is quantitative as it seeks to list and categorize in which aspects LLMs are being used by UX professionals.

Figure 1 presents an overview of the methodology employed, including the Planning, Conduction, and Reporting stages. Section 3.1 is part of the Planning phase, Sections 3.2 and 3.3 are part of the Conduction phase, and Sections 3.4 and 4 are part of the Report phase and provide the detailed results of the mapping study.

3.1 Research questions

Based on the objectives of the mapping study, its requirements were identified, seeking to address gaps found through a literature review. From this objective, the following Research Questions (RQs) were defined:

RQ1: Where have LLM-based techniques been applied for usability and accessibility?
RQ2: How have researchers used LLMs for user journey analyses?
RQ3: How have LLM-based approaches contributed to user experience?
RQ4: How do UX practitioners use LLM to support their activities?

RQ1 aims to explore the approaches used to improve system usability and accessibility, as well as methods for evaluating interfaces either through screen analysis or source code inspection. RQ2 seeks to identify strategies for analyzing the user journey. RQ3 focuses on understanding how LLMs can contribute to enhancing the user experience in interactive systems. Finally, RQ4 addresses how professionals are utilizing LLMs to support the creation and optimization of their design and development processes. The research questions aim to address the domains of [67] taxonomy in a cross-cutting manner.

3.2 Search

The databases selected for conducting the searches were the ACM Digital Library (Full-Text) and IEEE Xplore. The ACM library was chosen as it is the primary repository hosting the main conferences and journals in the field of HCI, followed by IEEE, which also publishes relevant journals. Both databases support structured searches using boolean operators, provide access to full-text articles, and allow the export of formatted results compatible with tools for conducting reviews and mappings.

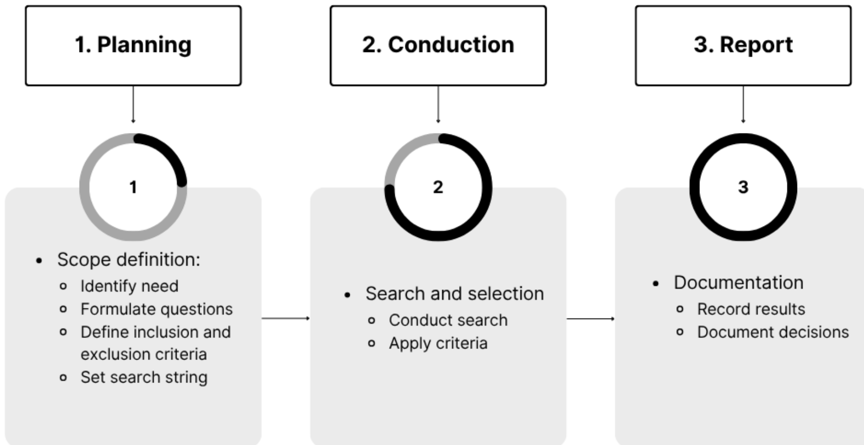

Figure 1: Methodological procedures

The search string was constructed based on the research questions and study objectives, and was structured as follows: ("LLM" OR "generative AI" OR "ChatGPT" OR "IA generativa" OR "Large Language Model" OR "Large Language Models" OR "Modelos de linguagem larga") AND ("accessibility" OR "acessibilidade" OR "usabilidade" OR "usability") AND ("user experience" OR "UX" OR "experiência do usuário" OR "jornada do usuário" OR "user journey").

The search was conducted in both digital libraries on June 23, 2025. In the ACM Digital Library, the platform's internal filter was applied to include only research articles, thereby excluding publications such as posters, demonstrations, short papers, tutorials, extended abstracts, surveys, technical reports, and works in progress.

For this study, only automated search methods were employed in the repositories, using the same search string for both databases. The application of the search string yielded a total of 1,107 articles, 1002 on ACM and 105 on IEEE.

3.3 Study selection and quality assessment

In the ACM Digital Library, a built-in platform filter was applied to exclude works-in-progress, posters, and demonstrations. Following this, inclusion and exclusion criteria were applied to refine the results obtained through the execution of the search string. Articles were assessed based on a review of the title and abstract, in accordance with the predefined inclusion and exclusion criteria. The inclusion criteria (IC) were as follows:

- **IC1:** Papers published in peer-reviewed venues, including workshops, conferences, symposiums, and journals;
- **IC2:** Papers that contain the search string and are related with objective of the mapping;
- **IC3:** Papers written in English and Portuguese.

The exclusion criteria (EC) used to discard studies were:

- **EC1:** Duplicated paper;
- **EC2:** Papers in the form of tutorials, editorials, technical report or surveys;
- **EC3:** Papers not available in full text;
- **EC4:** Papers presented as posters, demonstrations and work in progress;
- **EC5:** Papers that lack an abstract;
- **EC6:** Secondary or tertiary papers (e.g. systematic literature review).

After the initial selection of articles based on the inclusion and exclusion criteria through title and abstract screening, 963 papers were rejected (2 duplicated). An in-depth reading by the first and second author, with dedicated discussion sessions focusing on key points relevant to the mapping study of the 144 resulting papers was conducted to assess quality and relevance to the mapping objectives. Three questions were answered to assess the quality of the studies:

- Does it present a consolidated methodology within activities related to usability, accessibility, and user experience?
- Does it clearly present the motivation or demand for using LLMs within the processes?
- Does it indicate participatory or co-creation practices among UX designers?

The graphic on Figure 2 shows the numbers of the assessment. After applying the inclusion and exclusion criteria, a total of 144

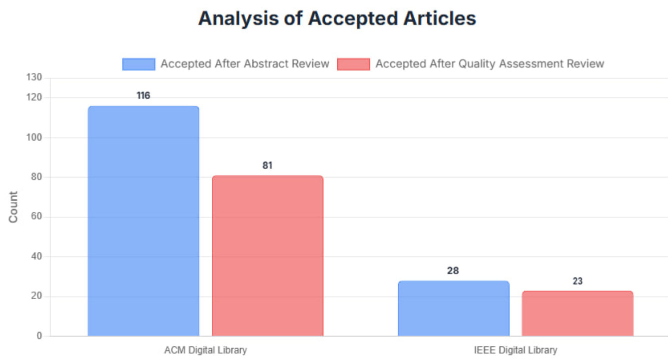

Figure 2: Quality assessment of the accepted papers after the exclusion criteria.

articles were initially accepted—116 from the ACM Digital Library and 28 from IEEE Xplore. These articles underwent a quality assessment through full-text reading, which resulted in the exclusion of 35 ACM articles and 5 IEEE articles. In the end, 104 articles were selected for inclusion in the mapping study.

Figure 3 illustrates all the study selection process, detailing each step involved. This visual representation aims to systematize and increase the transparency of the workflow, contributing to the organization, reproducibility, and traceability of the study.

3.4 Data extraction and Analysis

Data extraction was conducted by analyzing the accepted papers to identify their primary objectives and align them with the corresponding research questions and domain. The literature review conducted by Pang *et al. et al.* [67] proposed a taxonomy for application domain in which HCI researchers have applied LLMs. The proposed taxonomy encompasses *Application Domains, LLM Roles, Limitations and Risks.* However, for the purposes of this mapping study, only the *Application Domain* was considered. We used this taxonomy to categorizes the accepted papers into domains. The extraction process was performed by the third author and reviewed by the first author. In cases of disagreement regarding the association, the second author conducted a secondary review and determined the final classification.

The results of the data extraction will be presented in Section 4, aiming to address the research questions.

4 RESULTS

4.1 Studies Characterization

The articles selected for this review were published between 2021 and June 2025. Figure 4 illustrates the annual distribution of these publications. In the year 2021, no publications were identified that met the inclusion criteria established for this mapping study. The highest concentration of articles is observed in 2024 and 2025, a trend that aligns with the escalating research interest in LLMs in recent years. Furthermore, it is anticipated that the number of publications for the current year, 2025, will exceed the total from 2024 by far.

An analysis of the publication sources shows that the ACM Digital Library holds the majority of the publications (80 out of 104). This is to be expected, as most of the primary conferences and journals in the field of Human-Computer Interaction (HCI) are hosted on this platform. The main publication venues for the accepted articles are listed in Figure 5, where the Conference on Human Factors in Computing Systems (CHI) stands out as the primary venue for the works included in the mapping. The following sections presents the results addressing the research questions and the domain-based categorization.

Furthermore, Figure 5 presents the conferences with the highest recurrence of publications, with CHI standing out as one of the most prominent in the field.

4.2 RQ1 - Where have LLM-based techniques been applied to support usability and accessibility?

The analysis pertaining to RQ1 indicates a significant body of research focused on developing chatbots with enhanced natural language capabilities to improve usability for diverse user groups. For instance, studies describe chatbots designed to provide information retrieval and emotional support for caregivers [26], and to assist patients and their caregivers by facilitating reflection and communication regarding their care [40]. The latter is achieved by improving conversational flow, accurately representing words and meanings, and conveying an appropriate emotional tone. Chatbots are also used for mental health support, where they aim to foster effective and empathetic communication by responding with empathy, cultural sensitivity, and efficacy to enhance system usability [63]. Other identified applications include a news retrieval tool for patients [78] and an e-health service app [19].

To support usability and accessibility, the literature reveals several applications of generative AI (GenAI) [1, 32, 46, 56]. ChatGPT is used to create personas for student learning by simulating users for interview practice [6], and to test various privacy configurations and understand their implications [15]. In the case of interface creation, GenAI is employed to generate prototypes directly from user stories [47] and to create malleable interfaces that users can modify and extend using natural language prompts [13]. For developers, LLMs was used to improve code accessibility and streamline the UI creation process while maintaining usability [60]. Finally, tools have also been developed to address code compliance, such as automatically checking for accessibility violations in both open-source code and code generated by ChatGPT [2].

As an accessibility resource, a range of tools leveraging natural language have been developed for various user groups. For users who are blind or have low vision, applications include a chatbot that creates stories from photos to serve as a form of memory recreation [98], and an app for object detection and scene description [61]. Other tools for this group aim to enable more natural conversational interaction with screen readers [45], and errors in screen reader [105], provide textual cues for navigating UI components and their states [102], and generate tailored video and scenery descriptions [52, 65]. To assist individuals with speech impairments, one study developed a tool that aids communication by interjecting timely, humorous comments into chat conversations [93]. For

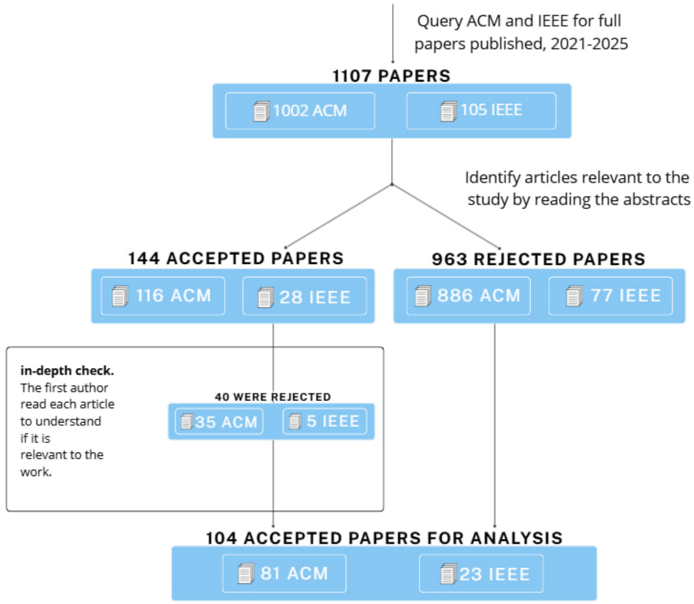

Figure 3: Selection of papers according the inclusion and exclusion criteria

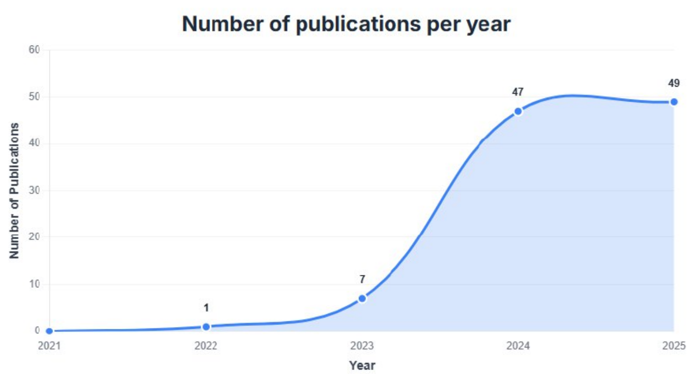

Figure 4: Number of publications per year

assistive technology directed to older adults, an augmented reality (AR) system has been proposed that provides information about surrounding items via image capture, answering related questions in adapted natural language [84].

4.3 RQ2: How have researchers used LLMs for user journey analyses?

To address RQ2, studies were identified that leverage LLM-based tools to support user journey mapping [57] and generate patient personas (e.g., for Parkinson's disease), providing pathways for designing adaptive systems [74]. These include behavior analysis and summarization [101], intention analysis based on retrospective log data [79], and user interaction simulations [11]. In [4], researchers employed LLMs for implicit and real-time user journey analysis through the concept of "step-awareness". LLMs were also used to construct narrative journeys (storyline) and visualize branching narratives, with tools that facilitate the analysis of these journeys [58].

4.4 RQ3: How have LLM-based tools contributed to user experience?

The studies related to RQ3 encompass improvements in communication and more intuitive interactions with AI, optimization of user experience in specific tools, enhancement of workflow and user engagement, personalized support, learning, and Human–AI collaboration. Within these themes, the CrossTalk tool [96] employs the BERT technique (a precursor to LLMs) to recognize user intentions during video-conferences and execute related searches in real time. Also in the context of videoconferencing, a context-sensitive transcription interface has been developed, adapting to user engagement levels throughout meetings. This solution aims to enhance the user experience, particularly in scenarios where meetings are frequently interrupted [49]. For music recommendation, a Conversational Recommendation Systems [100].

In the context of smart cities, to improve user experience LLMs have been employed to process data collected from citizen complaint systems, generating personalized and relevant feedback for users [42]. They have also supported human–home-agent interactions by enabling accessibility-based planning in simulations of real environments [66]. In [30], a vehicle agent based on LLMs was developed to perform multi-agent selection and vehicle control via voice interaction, improving system usability by allowing more natural interaction without the need to manage multiple agents individually.

User experience enhancement has been investigated through the use of chained prompts, response collection, and efficient system feedback integration [95]. Improvements have also targeted blind

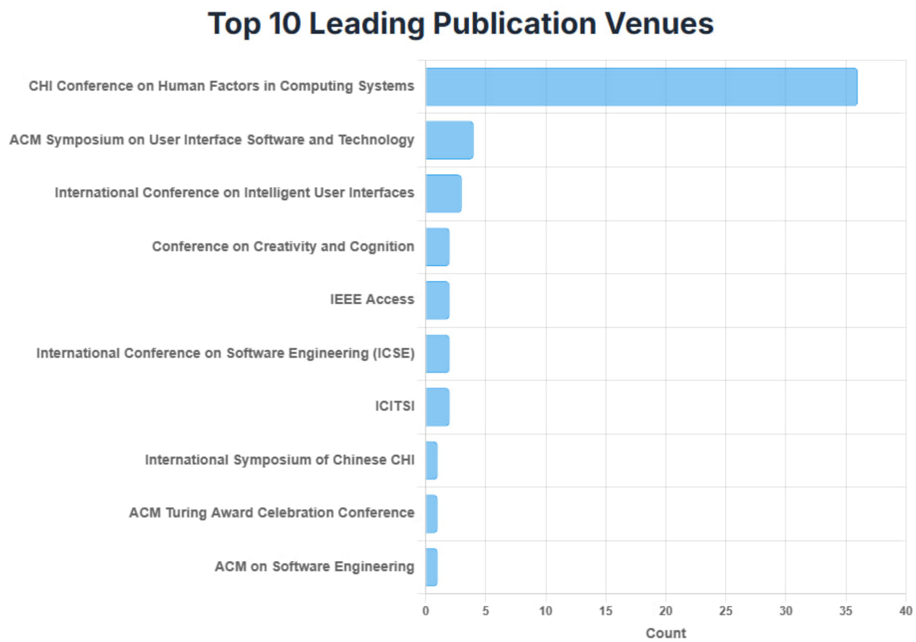

Figure 5: Distribution of accepted papers by publication venues (Top 10).

and low-vision users in text-to-image conversion tasks, enhancing their ability to edit generated items [51]. A personalized advising system for incoming students was developed to simplify access to institutional policies and integrate multiple functionalities [88]. Creation-support tools were identified in [83] to improve writing experiences in complex tasks, and to assist in music composition by transforming images, text, and audio inputs into musical chords [43]. Context-sensitive response generation for users in metaverse environments was also explored [34].

In [73], user–device integration was enhanced to improve conversational interaction between the user and a haptic device, enabling more effective control over dialogues for information access. Audio description generation for videos was investigated in [69], supporting blind users through contextual narration of broadcasted content and addressing the limitations of conventional captioning. User experience was further improved by the framework proposed by Behravan *et al.* [8], which facilitates the generation of 3D objects for augmented reality (AR) environments through natural language and voice-based commands, while [20] create a system to recognizes yoga poses to help professionals in personalized fitness assistance.

Usability-driven tools have contributed to enhancing user experience, such as a system that leverages LLMs to enable natural language interaction for performing a variety of operations within the operating system [38, 62]. For programmers, LLMs have been employed to improve code comprehension by increasing precision and speed, reducing cognitive load, and providing a distraction-free environment with an optimized interface and accessible explanations [99]. In search to create more personalized and accessible educational experiences, a study utilized a participatory design approach to generate stories that merge children's personal interests with literary classics, with the goal of fostering better engagement with reading [14]. Similarly, Chen *et al.* [17] developed an educational tool that delivers an engaging and interactive experience, resulting in personalized learning, faster reasoning, and enhanced comprehension. Addressing the needs of specific user groups, Firsanova [28] created an app for individuals on the autism spectrum. The app aims to create a more empowering information-seeking experience by providing information without the need for a human mediator.

4.5 RQ4: How do UX practitioners use IA in their activities?

To address RQ4, several recent studies were identified that explore the use of LLMs to support design processes and enhance user experience. One example is an interactive system that leverages LLMs to foster collaboration in design spaces, enabling professionals to explore multiple solutions and stimulate creative work [85]. Another proposal introduces a visual-centric tool tailored to UX designers [82]. LLMs have also been employed to encourage creativity and productivity in design processes, particularly among students [89], as well as to track the evolution of design ideas and guide designers in more effective solution exploration [81]. In the prototyping context, an LLM-based tool was developed to identify recurring issues through prompts, helping to prevent incidents and offering early insights into potential failures [92]. LLMs were also applied to automate heuristic evaluations of user interfaces [25, 36].

Other studies focus on inclusion and accessibility, such as the use of LLMs to enhance alternative communication tools for individuals with speech impairments by inserting humorous remarks at appropriate moments in conversations [93]. Additional proposals assist designers in creating graphical user interfaces from basic models, providing customization suggestions via natural language commands [24], and promoting collaboration between designers and clients through annotated feedback in 3D project development [16]. LLMs have been applied to the design of journaling systems for individuals with Parkinson's disease, aiding in persona generation and offering insights for user-centered solutions [74]. Some applications have simulated interactions with ChatGPT 4.0, using the Design Guide to evaluate user experiences [11].

LLMs have also been used by UX practitioners to evaluate user intent across various contexts [33, 59, 80, 101], as well as in systems like PrISM-Q&A, which perform implicit and real-time analysis of user journeys [4]. Another important advancement involves the high-granularity decomposition of tasks using annotated timelines, revision tracking, and contextualized feedback [21]. Finally, noteworthy initiatives integrate human–AI collaboration to improve design quality and user experience through joint exploration of the solution space [18], alongside the development of context-sensitive transcription interfaces adapted to participants' engagement levels during meetings [50].

To summarize the accepted papers, the Table 1 grouped them associating the references by research questions.

Table 1: Distribution of the studies by research questions.

Research questions	Paper references
RQ1	[26] [40] [63] [78] [19] [6] [15] [47] [13] [60] [2] [98] [61] [45] [102] [52] [93] [84] [46] [56] [54] [1] [105] [65] [32]
RQ2	[57] [74] [101] [79] [11] [4] [58]
RQ3	[96] [49] [42] [66] [30] [95] [51] [88] [83] [43] [34] [73] [69] [8] [62] [38] [99] [14] [17] [28] [100]
RQ4	[85] [82] [89] [81] [92][25] [36] [93] [24] [16] [74] [11] [101] [59] [80] [33] [4] [21] [18] [50]

4.6 Domains

To improve the categorization of the articles in this mapping study, we also listed them based on application domains. Many of the articles presented in the following sections have already appeared in previous ones, however, this serves as an additional way to search for research sources. The domains used in this mapping study were based on the systematic review conducted by Pang et al. [67]. While the original review identified 10 domains, we selected those most closely aligned with the objectives of our mapping: Augmenting Capabilities, Communication & Writing, Well-being & Health, Accessibility & Aging, Education, Design and User Experience. In some cases, studies were classified under more than one domain.

4.6.1 Augmenting Capabilities. Productivity is a key factor in HCI activities, and researchers have increasingly invested in studies aimed at enhancing the work of designers and UX professionals. A conversational framework that enhances the usability of smart energy system simulations allowing user to describe their simulation scenarios in plain language and GPT seamlessly translates these descriptions into Python scripts, used as inputs to the simulation environment [22]. LayoutCopilot [53] simplifies human-tool interaction by converting natural language instructions into executable script commands, and it interprets high-level design intents into actionable suggestions, significantly streamlining the layout design process.

The work present in [12] investigate a system to search for GUI layout, and Riche *et al.* [76] present a creative AI-assisted design work that refines the intentions' design in creating prompts, the tool leverage LLM's to suggest, vary, and refine new instruments to co-create. Do *et al.* [24] build a system to automate the generation of new options from a initial user interface.

As plugin for Figma, Feng *et al.* [27] create a widget that operationalizes designerly adaptation, Kretzer *et al.* [47] build a plugin to generate prototypes directly from user stories, and [25] build one for automatic heuristic evaluation. For programmers, Zhou *et al.* [106] developed an automated approach that synergizes computer vision (CV), MLLMs, and iterative compiler-driven optimization to generate and refine declarative UI code from designs, and segment of content from podcast [68], while LLMs have been employed to improve code comprehension by increasing precision and speed, reducing cognitive load , and providing a distraction-free environment with an optimized interface and accessible explanations [99].

The chatGPT was used to generate automatic problem suggestions in usability test [48], and improving the quality and reliability of usability evaluations, supporting the development of assessment plans [104]. Beyzaei *et al.* [9] create a technique to efficiently transfer usage-based UI tests across mobile apps, and Turbeville *et al.* [90] present an approach for generating insights from multimodal UX testing data.

4.6.2 Communication & Writing. This domain focuses on improvements in communication and writing tasks, where LLMs have been employed to enhance interaction between users and applications. Conversational agents were used to provide information and emotional support for caregivers [26], and to assist patients and their caregivers by facilitating reflection and communication regarding their care [40]. Emotional support and empathetic communication were explored in [63], while tools for accessing news by diverse user groups were presented in [28, 78]. Researchers developed an application that generates stories from photos for blind users to communicate through memories [98], and another to assist users with speech impairments in expressing humor during chat interactions [93]. Communication in videoconferencing contexts was improved in [96] and [49], and in smart cities to support dialogue between citizens and local governments [42]. Additionally, communication between designers and clients was facilitated through annotation feedback [16].

LLMs have also been applied in writing tools, with a focus on maintaining the natural flow of composition even in complex tasks, assisting users in developing diverse content and employing rhetorical strategies. This approach fosters a more fluid, creative, and productive experience for writers and digital communicators [83].

4.6.3 Accessibility & Aging. LLMs have opened new possibilities for making interactive systems more accessible, inclusive, and adaptable to the needs of users with disabilities. Several initiatives have explored the potential of these technologies to improve usability, user experience (UX), and the integration of inclusive design principles. One such application introduces a text-to-image approach powered by LLMs, which enhances control and editing capabilities of visual content for blind or low-vision users, providing greater autonomy and personalization in digital interface interaction [51]. Additionally, LLMs have been employed in the development of assistive applications capable of detecting objects and describing scenes using natural language, directly benefiting blind and low-vision users [61]. This approach significantly contributes to improving usability and user experience in everyday tasks. For users who are blind or have low vision, applications include a chatbot that creates stories from photos [98], a tool to enable more natural conversational interaction with screen readers [45], provide textual cues for navigating UI components and their states [102], and generate tailored video and scenery descriptions [52].

In development environments, efforts are underway to make programming assistants more accessible to blind developers, increasing their autonomy in software creation and reducing technical barriers [29]. In the context of designing generative AI-based systems, researchers have proposed the creation of clear and actionable design principles focused on accessibility and user experience. These principles include: designing for generative variability, co-creation, mental model construction, and appropriate trust. They aim to support the safe and effective use of AI across different stages of the design process, from ideation to final evaluation [94]. LLMs was used to improve code accessibility and streamline the UI creation process while maintaining usability [60] LLMs are also being used to correct accessibility violations in AI-generated code, such as that produced by ChatGPT, and in open-source projects. This application directly impacts the improvement of accessibility in systems developed by and for people with disabilities, making the code more aligned with inclusive standards [2].

Finally, an innovative initiative uses LLMs to enhance alternative communication tools for individuals with speech disabilities by inserting well-timed humorous comments into conversations, fostering more natural, expressive, and inclusive interactions [93].

4.6.4 Education. The use of LLMs in education has shown great potential to transform pedagogical practices, benefiting both students and instructors. Several interactive systems have been proposed to enhance the learning experience by promoting greater personalization, efficiency, and engagement. Among these proposals is NoTeeline, a system developed to support the creation of personalized real-time notes. Its goal is to reduce students' cognitive load, improve the quality of note-taking, and facilitate later review and organization [37]. Another example is an online programming education system that provides personalized assistance to students through helpful task-completion tips. This type of LLM-powered support enables learners to progress with greater confidence and less frustration, particularly in self-directed learning contexts [10]. A system for incoming students was developed to simplify access to institutional policies and integrate multiple functionalities [88]. To create more personalized and accessible educational experiences, a study utilized a participatory design approach to generate stories that merge children's personal interests with literary classics [14], while Chen *et al.* [17] developed an educational tool that delivers an engaging and interactive experience.

LLMs have also been used to support educators, such as in the case of ReadingQuizMaker, a tool designed to assist instructors in generating high-quality reading comprehension questions. The tool adapts to teachers' natural workflows by automating tasks such as template selection, input configuration, and result editing, thereby streamlining the creation of assessment materials [55]. Additionally, LLMs have been employed in initiatives that combine education and gamification, such as systems for automatically generating themed quizzes to engage museum visitors with educational content [75].

Finally, studies have highlighted the use of LLMs as tools for supporting educational design, helping students engage in creative and productive processes throughout their learning journeys [89], and a study that evaluates the usability of ChatGPT and students' experiences with it on the Moodle platform [23].

4.6.5 Design and User Experience. The Design domain encompasses various types of design work. In the context of our mapping, it specifically includes design efforts aimed at improving the user experience (UX).

The use of LLMs has significantly transformed the field of UX, fostering more adaptive and efficient interfaces. These technologies are applied to reduce technical barriers, enhance personalization, and improve interaction between users and systems. UX enhancement has been investigated through the use of chained prompts, response collection, and efficient system feedback integration [95], integrate human–AI collaboration to improve design quality and user experience through joint exploration of the solution space [18], NLDesign is a user interface (UI) design tool that combines natural language dialogue with generative AI. Its goal is to lower technical barriers for non-expert users by offering intuitive control, personalization support, and simplified generation and editing of interface elements without requiring coding skills. It also promotes component reuse and improves the overall adaptability of the user experience [103]. Creation-support tools were identified in [83] to improve writing experiences in complex tasks.

SimUser is a system that simulates interactions of diverse user profiles with mobile applications, enabling the automated generation of usability feedback. This allows developers to identify and resolve UX issues based on realistic simulated behaviors [97]. VirtuWander explores the potential of LLMs to enhance multimodal interaction in virtual tour guide environments, enriching user experiences with more natural, interactive, and context-aware responses during navigation [91].

In [34], a LLM-based pipeline generates context-aware answers to users when onboarding metaverse, improving the experience and make responses more useful and immersive than those generated without contexts, while in [3] a plug-in for ChatGPT was designed to serve as an intuitive, user-friendly interface between workers and job-related databases, improving the usability and accuracy of queries. For content creation, the system proposed by Jadhav *et* *al.* [39] uses LLM and image generation techniques to create relevant visuals and voiceovers, enhancing content creation efficiency without sacrificing quality.

4.6.6 Well-being & Health. The healthcare domain has significantly benefited from the use of LLMs, with applications focused on emotional support, accessibility, personalized care, and enhancing user experience in digital health systems. One such application is CareJournal, an app developed to support caregivers through sensitive and contextualized conversations. The tool has an ability to accurately represent word meanings and maintain an appropriate emotional tone, fostering a more empathetic and humanized experience in daily care routines [40], in the same direction a chatbot designed to provide information retrieval and emotional support for caregivers [26], and for patients [78]. Another advancement is the creation of a user-centered teleneurology platform that employs a GPT-based chatbot for triage and consultations. This system is distinguished by its ability to provide clear and reliable responses, relevant guidance to patients, and a design aimed at reducing anxiety during interactions [72]. LLMs are also used to generate personas with Parkinson's disease, offering valuable insights for the design of health-focused journaling systems. This enables the development of solutions that are better aligned with the specific needs of this population [74]. The implementation of medical chatbots has also been explored as a way to deliver personalized responses to health inquiries, supporting user experience with context-specific and relevant information [41, 86]. Furthermore, LLMs have been applied to improve the experience of elderly users in e-health applications by simplifying interfaces and adopting chat-based interaction models, thereby reducing the cognitive effort required to use these tools [19].

The Table 2 summarize the papers discussed by domain application.

5 DISCUSSIONS

This mapping study aimed to investigate how LLMs are being applied in HCI activities, with a particular focus on usability, accessibility, user experience, and professional design practices. The analysis was guided by 4 Research Questions (RQs) and Domain applications defined by a taxonomy. The analysis reveals a significant and expanding body of research, primarily addressing how these advanced AI capabilities are enhancing interactions across diverse user groups and professional practices.

5.1 Approaches to Improving Usability and Accessibility

In answering RQ1, a prominent theme was the development of sophisticated chatbots, leveraging natural language capabilities to improve usability for various audiences. This includes providing critical support for vulnerable populations such as caregivers, patients, and also people with disabilities. The emphasis on improving conversational flow, accurate semantic representation, and appropriate emotional tone highlights a move towards more empathetic and effective human-AI communication. Furthermore, the application of chatbots extends to mental health support, where the focus is on fostering empathetic and culturally sensitive interactions to enhance system usability. The results enabled us to answer the research questions and revealed, among other findings, a shift in the terminology related to chatbots. Several studies have begun referring to them as "agents" indicating a possible evolution in the understanding of these tools.

5.2 Methods for Analyzing the User Journey

Addressing RQ2, the mapping study identified a diverse set of approaches in which LLMs are leveraged to support user journey analysis. These approaches reflect a shift from static, manual mapping techniques to dynamic, data-driven methods that enable deeper insights into user behavior, intention, and experience. Simulated user personas, like an user with certain disease, and interaction modeling tools highlight the potential of LLMs to support user journey analysis. These studies perform journey mapping and allow developers to anticipate diverse user behaviors, summarize interaction patterns, and identify usability or accessibility issues early in the design process. These techniques support more empathetic and context-aware design decisions.

5.3 Enhancing User Experience in Interactive Systems

LLMs have been widely adopted to improve user experience (RQ3) by enabling adaptive, conversational, and context-aware interfaces. These contributions span diverse domains, from videoconferencing and smart cities to education, accessibility, and creative support, highlighting the versatility of LLMs and related technologies in shaping Human–AI collaboration. Tools are used to recognize user intentions during video conferences, exemplify the shift toward real-time, context-aware AI support. Similarly, adaptive transcription interfaces respond to user engagement levels, mitigating the impact of interruptions and enhancing the fluidity of communication. These innovations reflect a growing emphasis on seamless integration of AI into everyday workflows. In smart city contexts, LLMs have been employed to process citizen complaints and generate personalized feedback fostering more responsive governance. Human–home-agent interactions have also benefited from accessibility-based planning in simulated environments promoting inclusive design. The development of vehicle agents capable of voice-based multi-agent control further illustrates how natural language interfaces can simplify complex systems and improve usability.

Several studies target accessibility improvements, such as audio description generation for blind users and enhanced text-to-image editing for low-vision individuals. These tools reduce barriers to digital content and empower users to engage more fully with technology. The use of natural language to generate 3D objects in AR environments and yoga pose recognition for personalized fitness assistance further demonstrate the potential of AI to support diverse user needs. Educational applications of LLMs include personalized advising systems, participatory storytelling tools that blend children's interests with literary classics, and interactive learning environments that enhance reasoning and comprehension. These systems prioritize engagement and adaptability, aligning educational content with individual learner profiles. Notably, the development of an app for individuals on the autism spectrum underscores the

Table 2: Distribution of the studies by domain

Domains	Paper references
Augmenting Capabilities	[22] [53] [12] [76] [24] [27] [47] [25] [106] [68] [99] [48] [104] [9] [90]
Communication & Writing	[26] [40] [63] [78] [28] [98] [93] [96] [49] [42] [16] [83]
Accessibility & Aging	[51] [61] [98] [45] [102] [52] [29] [94] [60] [2] [93]
Education	[37] [10] [88] [14] [17] [55] [75] [89] [23]
Design & User Experience	[95] [18] [103] [83] [97] [91] [34] [3] [39]
Well-being & Health	[40] [26] [78] [72] [74] [86] [41] [19]

importance of empowering users through autonomous information access.

System-level improvements include natural language interfaces for operating system tasks, chained prompts for efficient feedback integration, and context-sensitive responses in metaverse environments. These innovations reflect a broader trend toward conversational and adaptive systems that prioritize user comfort and control.

5.4 Supporting UX Professionals in Design and Development

RQ4 explores how LLMs are being used to assist professionals in UX and design-related activities. The reviewed studies reveal a growing ecosystem of tools and frameworks that integrate LLMs into creative workflows, prototyping, evaluation, and collaboration. LLM-powered systems have been developed to foster creativity and support exploration of multiple design solutions. These tools enable designers to engage in iterative ideation, guided by AI-generated suggestions that expand the solution space. Visual-centric platforms tailored to UX designers further demonstrate how LLMs can adapt to domain-specific needs, offering intuitive interfaces and actionable insights. LLMs have also been employed to track the evolution of design ideas and guide professionals toward more effective solutions. By analyzing design trajectories and providing contextual feedback, these systems help designers refine their concepts and avoid common pitfalls. In prototyping contexts, prompt-based tools identify recurring issues and offer early warnings about potential failures, contributing to more resilient and user-centered designs.

The automation of heuristic evaluations through LLMs represents a significant advancement in UX assessment. These tools streamline the evaluation process, reduce manual effort, and ensure consistency in interface analysis. Additionally, systems perform implicit and real-time analysis of user journeys, enabling practitioners to evaluate user intent and experience dynamically across various contexts. LLMs have also been used to generate personas supporting the design of journaling systems that align with specific user needs. These applications reflect a broader commitment to accessibility and empathetic design.

LLMs facilitate collaboration between designers and clients through annotated feedback project environments, and assist in generating graphical user interfaces from basic models using natural language commands. These tools reduce technical barriers and promote co-creation, making design processes more inclusive and efficient. Advanced systems have introduced high-granularity task decomposition using annotated timelines, revision tracking, and contextualized feedback. These features support detailed analysis of design workflows and help practitioners maintain clarity and coherence throughout iterative development. Context-sensitive transcription interfaces further enhance collaboration by adapting to participant engagement levels during meetings.

5.5 Domain Application

The graphic in Figure 6 presents the distribution of articles across application domains. It is evident that the Augmenting Capabilities domain concentrates a bigger number of related studies. This indicates a research effort aimed at enhancing productivity within HCI activities.

The increasing incorporation of LLMs into design and usability evaluation tools marks a significant advancement in how professionals interact with computational systems. The solutions demonstrate that LLMs automate technical tasks and expand users' creative and strategic capabilities. Tools highlight the potential of LLMs to interpret design intentions and convert them into executable actions, thereby bridging the gap between conceptualization and implementation. This ability to transform natural language commands into actionable suggestions enhances the fluidity of the creative process and contributes to the democratization of design. Furthermore, systems that search, refine, and vary graphical user interface (GUI) layouts based on prompts illustrate how LLMs can act as co-creators, fostering experimentation and innovation.

The integration of LLMs into platforms like Figma, through plugins that generate prototypes, perform heuristic evaluations, and adapt interfaces based on user stories—reinforces the trend toward increasingly responsive and user-centered tools. For developers, approaches that combine computer vision, multimodal LLMs, and iterative compiler-driven optimization offer promising pathways for the automated generation of declarative UI code, while facilitating the comprehension of complex systems through accessible explanations and optimized interfaces. For usability, the use of LLMs to automatically suggest problems during testing and support the development of evaluation plans points to greater reliability and efficiency in validation processes. Techniques that enable the transfer of usage-based UI tests across mobile applications, and the extraction of information from multimodal UX testing data enhance the analytical capabilities of UX professionals, promoting a more integrated and context-aware approach.

The domain of Communication & Writing reveals the way users interact with digital systems, particularly by fostering more natural, empathetic, and expressive exchanges. Conversational agents have emerged as key tools in this domain. These systems go beyond information delivery, offering empathetic responses that contribute

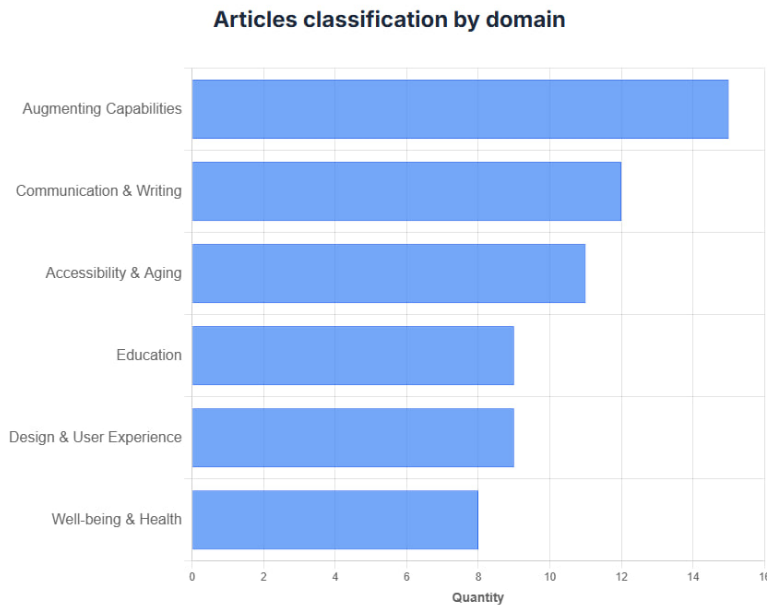

Figure 6: Number of Publications by domain

to a more humanized experience of care. Emotional support mechanisms have also been explored in contexts, highlighting the potential of LLMs to simulate affective communication. Efforts to make communication more inclusive are evident in applications designed for blind users, such as systems that generate narrative stories from photos. Similarly, tools that assist users with speech impairments in expressing humor during chat interactions demonstrate how LLMs can enrich social engagement in digital conversations. Solutions have been applied to improve communication in videoconferencing environments and facilitate dialogue between citizens and local governments in smart city initiatives. Beyond conversational interfaces, LLMs have been integrated into writing tools that support users in maintaining the natural flow of composition, even in cognitively demanding tasks. By assisting with rhetorical strategies and content diversification, these systems promote a more fluid, creative, and productive writing experience.

Across domains, LLM adoption reflects a balance between automation and augmentation. While automation accelerates evaluation and prototyping, augmentation enhances human creativity and decision-making. The challenge ahead lies in aligning these capabilities with ethical, participatory, and inclusive design principles, especially in sensitive domains like healthcare and education.

5.6 Limitations of the Mapping Study

While this mapping study provides a broad overview of how Large Language Models (LLMs) are being applied across various domains of HCI, several limitations must be acknowledged.

5.6.1 Scope of Literature Sources. The study focused exclusively on articles indexed in the ACM Digital Library and IEEE Xplore between 2021 and 2025. Although these databases are highly reputable and central to the HCI field, relevant works published in other venues (e.g., Springer, or domain-specific journals) may have been excluded. This is a firm limitation, and to mitigate the constraint of selecting only two databases, the ACM Digital Library was chosen as one of them which publishes the main conferences and journals in HCI field. To identify duplicate articles across both databases, we used the Parsifal [1] software to systematize the process of identification and summarizing all retrieved publications.

5.6.2 Domain Categorization Constraint. The categorization of articles into application domains was based on the taxonomy proposed by Pang et al. [67] published in 2025, which, while robust, may not fully capture the evolving nature of LLM applications. Some studies addressed multiple domains simultaneously, and assigning them to a single category may oversimplify their contributions. In such cases, cross-domain overlaps were noted, but the granularity of classification remains a challenge. To enrich the mapping the research questions were also addressed, listing the papers related to each question.

5.6.3 Interpretation of Research Contributions. The initial phase of article inclusion/exclusion was based on titles and abstracts, and only after the quality assessment was made by an in-depth reading. While this approach enables large-scale analysis, it may overlook nuanced insights found only in full-text content. In some cases, there are variations in terminology, particularly those related to LLMs, and certain studies may have been excluded due to abstracts that did not reflect the targeted keywords. To mitigate this issue, multiple synonyms were incorporated into the search string to account for variations in key terms.

5.6.4 Lack of Empirical Validation. This study does not include empirical validation of the effectiveness or impact of the LLM-based

[1] https://parsif.al/

systems reviewed. The mapping is descriptive rather than evaluative, and future work should consider systematic comparisons, user studies, or meta-analyses to assess the practical outcomes of these technologies in real-world HCI contexts.

6 CONCLUSIONS

This study systematically mapped 104 peer-reviewed articles covering the past five years (2021–2025) on the use of LLMs to support usability, accessibility, and UX activities in HCI. The study selected relevant papers based on data extracted from the ACM Digital Library and IEEE Xplore, according the research objectives following the application of inclusion and exclusion criteria..

The results provided answers to the research questions and revealed, among other findings, a observed shift in the use of "agent" as a trend. The use of LLMs in user journey analysis introduces a movement in how designers understand and model user experiences. These tools enable real-time, personalized, and predictive insights that go beyond surface-level metrics. As LLMs continue to evolve, their integration into UX workflows may lead to more responsive, inclusive, and emotionally intelligent systems. The breadth of applications explored illustrates the transformative potential of LLMs in enhancing UX across domains. By enabling more natural, personalized, and accessible interactions, these tools redefine the boundaries of Human–AI collaboration.

The integration of LLMs into UX and design workflows signals toward AI-augmented creativity and evaluation. These tools automate routine tasks enabling professionals to focus on strategic and empathetic aspects of design. As LLMs become more sophisticated, their role in facilitating real-time collaboration, inclusive design, and adaptive prototyping will continue to expand.

It was also observed the growing use of LLMs in the healthcare domain, where these models have been employed to provide more personalized support to patients. However, gaps were identified across the reviewed studies. For instance, a limited number of investigations addressed the needs of individuals with disabilities. Most research efforts focused on accessibility for blind and low-vision users, with few studies aiming to reduce barriers for other disabilities.

From a practical perspective, the mapping reveals three key takeaways: (i) integration into workflows, LLMs are most impactful when embedded into the daily processes of design and evaluation, reducing friction between ideation, prototyping, and testing; (ii) accessibility gaps, despite promising results, most research focuses on blind and low-vision users, with limited exploration of solutions for other disabilities or intersectional needs; and (iii) design augmentation, LLMs are emerging as creative partners, complementing human expertise rather than replacing it, particularly in exploratory and collaborative design contexts.

Future research should address the following directions:

- Conduct longitudinal and empirical studies on the sustained adoption of LLM-based tools in professional HCI practice.
- Expand accessibility research to include a wider range of disabilities and contexts.
- Investigate ethical and participatory approaches for journey modeling and adaptive systems, ensuring equitable and inclusive outcomes.

ACKNOWLEDGMENTS

We thank Professor Ismayle de Sousa Santos for his inspiring guidance and for sharing valuable feedback on conducting a systematic mapping study. This research is academically and financially supported by the Programa de Pós-Graduação em Ciência da Computação at UECE, as well as by the Instituto Atlântico (IA). In this article, Microsoft Copilot was used as a tool for translating the authors' native language into English and refining grammatical structures throughout the text.

REFERENCES

[1] Wajdi Aljedaani, Marcelo Medeiros Eler, and PD Parthasarathy. 2025. Enhancing accessibility in software engineering projects with large language models (llms). In *Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 1*. 25–31.

[2] Wajdi Aljedaani, Abdulrahman Habib, Ahmed Aljohani, Marcelo Eler, and Yunhe Feng. 2024. Does ChatGPT Generate Accessible Code? Investigating Accessibility Challenges in LLM-Generated Source Code. In *Proceedings of the 21st International Web for All Conference* (Singapore, Singapore) *(W4A '24)*. Association for Computing Machinery, New York, NY, USA, 165–176. https://doi.org/10.1145/3677846.3677854

[3] Rubén Alonso, Danilo Dessí, Antonello Meloni, Marco Murgia, Diego Reforgiato Recupero, and Giuseppe Scarpi. 2024. A Seamless ChatGPT Knowledge Plug-in for the Labour Market. *IEEE Access* 12 (2024), 155821–155837. https://doi.org/10.1109/ACCESS.2024.3485111

[4] Riku Arakawa, Jill Fain Lehman, and Mayank Goel. 2024. Prism-q&a: Step-aware voice assistant on a smartwatch enabled by multimodal procedure tracking and large language models. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 4 (2024), 1–26.

[5] Aram Bahrini, Mohammadsadra Khamoshifar, Hossein Abbasimehr, Robert J. Riggs, Maryam Esmaeili, Rastin Mastali Majdabadkohne, and Morteza Pasehvar. 2023. ChatGPT: Applications, Opportunities, and Threats. In *2023 SIEDS*. 274–279. https://doi.org/10.1109/SIEDS58326.2023.10137850

[6] Jose Barambones, Cristian Moral, Angélica de Antonio, Ricardo Imbert, Loïc Martínez-Normand, and Elena Villalba-Mora. 2024. ChatGPT for Learning HCI Techniques: A Case Study on Interviews for Personas. *IEEE Transactions on Learning Technologies* 17 (2024), 1460–1475. https://doi.org/10.1109/TLT.2024.3386095

[7] Simone Barbosa and Bruno Silva. 2010. *Interação humano-computador*. Elsevier Brasil.

[8] Majid Behravan and Denis Gračanin. 2025. From Voices to Worlds: Developing an AI-Powered Framework for 3D Object Generation in Augmented Reality. In *2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 150–155.

[9] Benyamin Beyzaei, Saghar Talebipour, Ghazal Rafiei, Nenad Medvidović, and Sam Malek. 2025. Automated Test Transfer across Android Apps using Large Language Models. *Proc. ACM Softw. Eng.* 2, ISSTA, Article ISSTA098 (June 2025), 24 pages.

[10] Anastasiia Birillo, Elizaveta Artser, Anna Potriasaeva, Ilya Vlasov, Katsiaryna Dzialets, Yaroslav Golubev, Igor Gerasimov, Hieke Keuning, and Timofey Bryksin. 2024. One step at a time: Combining llms and static analysis to generate next-step hints for programming tasks. In *Proceedings of the 24th Koli Calling International Conference on Computing Education Research*. 1–12.

[11] Mariana Gomes Borges, Claiton Marques Correa, and Milene Selbach Silveira. 2024. Human+ Human or Human+ AI: A Comparative Analysis of a Narrative Visualization Design Guide. In *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems*. 1–13.

[12] Paul Brie, Nicolas Burny, Arthur Sluÿters, and Jean Vanderdonckt. 2023. Evaluating a Large Language Model on Searching for GUI Layouts. *Proc. ACM Hum.-Comput. Interact.* 7, EICS, Article 178 (June 2023), 37 pages. https://doi.org/10.1145/3593230

[13] Yining Cao, Peiling Jiang, and Haijun Xia. 2025. Generative and Malleable User Interfaces with Generative and Evolving Task-Driven Data Model. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–20.

[14] Inês Carmo, Patrícia Costa, and Pedro Santana. 2024. Boosting children's reading motivation with LLM-generated story crossovers. In *2024 International Conference on Graphics and Interaction (ICGI)*. IEEE, 1–8.

[15] Chaoran Chen, Weijun Li, Wenxin Song, Yanfang Ye, Yaxing Yao, and Toby Jia-Jun Li. 2024. An empathy-based sandbox approach to bridge the privacy gap among attitudes, goals, knowledge, and behaviors. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–28.

[16] Chen Chen, Cuong Nguyen, Thibault Groueix, Vladimir G Kim, and Nadir Weibel. 2024. Memovis: A GenAI-powered tool for creating companion reference images for 3d design feedback. *ACM Transactions on Computer-Human Interaction* 31, 5 (2024), 1–41.

[17] Liuqing Chen, Zhaojun Jiang, Duowei Xia, Zebin Cai, Lingyun Sun, Peter Childs, and Haoyu Zuo. 2024. BIDTrainer: An LLMs-driven Education Tool for Enhancing the Understanding and Reasoning in Bio-inspired Design. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 676, 20 pages. https://doi.org/10.1145/3613904.3642887

[18] Pei Chen, Jiayi Yao, Zhuoyi Cheng, Yichen Cai, Jiayang Li, Weitao You, and Lingyun Sun. 2025. CoExploreDS: Framing and Advancing Collaborative Design Space Exploration Between Human and AI. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 22, 20 pages. https://doi.org/10.1145/3706598.3713869

[19] Sharfuddin Khan Chisty, Anika Tahsin Miami, and Jannatun Noor. 2025. Smart Sheba: Enhancing Elderly User Experience with LLM-Enabled Chatbots and User-Centered Design. In *Proceedings of the 13th International Conference on Information & Communication Technologies and Development* (Nairobi, Kenya) *(ICTD '24)*. Association for Computing Machinery, New York, NY, USA, 69–83. https://doi.org/10.1145/3700794.3700803

[20] Sravan Kumar Chittimalla and Leela Krishna M Potluri. 2025. Enhancing Yoga Pose Recognition Using Large Language Models with Real-World Dataset Integration for Personalized Fitness Assistance. In *2024 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*. 1–6. https://doi.org/10.1109/ITIKD63574.2025.11004828

[21] Nicholas Davis, Jacob Sherson, and Janet Rafner. 2025. The Co-Creative Design Framework for Hybrid Intelligence. In *Proceedings of the 2025 Conference on Creativity and Cognition (C&C '25)*. Association for Computing Machinery, New York, NY, USA, 560–572. https://doi.org/10.1145/3698061.3726934

[22] Gabriel Dengler, Peter Bazan, Reinhard German, Pooia Lalbakhsh, and Ariel Liebmann. 2024. A Conversational Human-Computer Interface for Smart Energy System Simulation Environments. In *Proceedings of the Winter Simulation Conference* (San Antonio, Texas, USA) *(WSC '23)*. IEEE Press, 2978–2989.

[23] Ati Suci Dian Martha, Sri Widowati, Desy Puspa Rahayu, Muhammad Ilham Nursyawal, and Rayhan Rizqial Hariz. 2024. Examining Usability and Student Experience of ChatGPT as a Learning Tool on Moodle in Higher Education. In *2024 International Conference on Information Technology Systems and Innovation (ICITSI)*. 8–13. https://doi.org/10.1109/ICITSI65188.2024.10929431

[24] Hyo Jin Do, Michelle Brachman, Casey Dugan, James M Johnson, Julia Lauer, Priyanshu Rai, and Qian Pan. 2024. Grounding with Structure: Exploring Design Variations of Grounded Human-AI Collaboration in a Natural Language Interface. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–27.

[25] Peitong Duan, Jeremy Warner, Yang Li, and Bjoern Hartmann. 2024. Generating Automatic Feedback on UI Mockups with Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 6, 20 pages. https://doi.org/10.1145/3613904.3642782

[26] Fernanda Espinoza, Darren Cook, Chris R Butler, and Rafael A Calvo. 2023. Supporting dementia caregivers in Peru through chatbots: generative AI vs structured conversations. In *Electronic Workshops in Computing*. BCS Learning & Development, 89–98. https://doi.org/10.14236/ewic/BCSHCI2023.11

[27] K. J. Kevin Feng, Q. Vera Liao, Ziang Xiao, Jennifer Wortman Vaughan, Amy X. Zhang, and David W. McDonald. 2025. Canvil: Designerly Adaptation for LLM-Powered User Experiences. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, Article 932, 22 pages.

[28] Victoria Firsanova. 2023. Towards Building a Mobile App for People on the Spectrum. In *Companion Proceedings of the ACM Web Conference 2023* (Austin, TX, USA) *(WWW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 555–559. https://doi.org/10.1145/3543873.3587533

[29] Claudia Flores-Saviaga, Benjamin V. Hanrahan, Kashif Imteyaz, Steven Clarke, and Saiph Savage*. 2025. The Impact of Generative AI Coding Assistants on Developers Who Are Visually Impaired. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 1164, 17 pages. https://doi.org/10.1145/3706598.3714008

[30] Toru Furusawa. 2025. In-Vehicle Artificial Intelligence Agent System with Local Large Language Models for Agent Selection and Vehicle Control. In *2025 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 259–264.

[31] Muhammet Kürşat Görmez, Murat Yılmaz, and Paul M Clarke. 2024. Large language models for software engineering: A systematic mapping study. In *European Conference on Software Process Improvement*, Vol. 2179. Springer, Cham, 64–79. https://doi.org/10.1007/978-3-031-71139-8_5

[32] Joshua Gorniak, Yoon Kim, Donglai Wei, and Nam Wook Kim. 2024. VizAbility: Enhancing Chart Accessibility with LLM-based Conversational Interaction. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Article 89, 19 pages.

[33] Yi Guo, Danqing Shi, Mingjuan Guo, Yanqiu Wu, Nan Cao, and Qing Chen. 2024. Talk2Data: A Natural Language Interface for Exploratory Visual Analysis via Question Decomposition. *ACM Trans. Interact. Intell. Syst.* 14, 2, Article 8 (April 2024), 24 pages. https://doi.org/10.1145/3643894

[34] Jihyeong Hong, Yokyung Lee, Dae Hyun Kim, DaEun Choi, Yeo-Jin Yoon, Gyucheol Lee, Zucheul Lee, and Juho Kim. 2024. A context-aware onboarding agent for metaverse powered by large language models. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 1857–1874.

[35] Xinyi Hou, Yanjie Zhao, Yue Liu, Zhou Yang, Kailong Wang, Li Li, Xiapu Luo, David Lo, John Grundy, and Haoyu Wang. 2024. Large Language Models for Software Engineering: A Systematic Literature Review. *ACM Trans. Softw. Eng. Methodol.* 33, 8, Article 220 (Dec. 2024), 79 pages. https://doi.org/10.1145/3695988

[36] Forrest Huang, Gang Li, Tao Li, and Yang Li. 2024. Automatic Macro Mining from Interaction Traces at Scale. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1038, 16 pages. https://doi.org/10.1145/3613904.3642074

[37] Faria Huq, Abdus Samee, David Chuan-En Lin, Alice Xiaodi Tang, and Jeffrey P Bigham. 2025. NoTeeline: Supporting Real-Time, Personalized Notetaking with LLM-Enhanced Micronotes. In *Proceedings of the 30th International Conference on Intelligent User Interfaces*. 1064–1081.

[38] Srujana Inturi, Swarna Latha Potluri, Guda Vanitha, and B Indira. 2025. VirtualMe: A Large Language Model-Powered Chatbot for Enhanced Human Computer Interaction. In *2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)*. IEEE, 626–634.

[39] Dhiraj Jadhav, Sakshee Agrawal, Sakshi Jagdale, Pranav Salunkhe, and Raj Salunkhe. 2024. AI-Driven Text-to-Multimedia Content Generation: Enhancing Modern Content Creation. In *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*. 1610–1615. https://doi.org/10.1109/I-SMAC61858.2024.10714771

[40] Jazette Johnson, Hira Jamshed, Rachael Zuppke, Amanda Leggett, Emily Mower Provost, and Robin N. Brewer. 2025. "I felt listened to": Evaluating an AI-Powered Reflection Tool for Care Partners. *ACM Trans. Access. Comput.* (April 2025). Just Accepted.

[41] Aayush Kapoor and Sujala D. Shetty. 2024. Enhancing Healthcare Information Accessibility Through a Generative Medical Chatbot. In *2024 ICETCS*. 1–4.

[42] Kai Karren, Michael Schmitz, and Stefan Schaffer. 2024. Improving conversational user interfaces for citizen complaint management through enhanced contextual feedback. In *Proceedings of the 6th ACM Conference on Conversational User Interfaces*. 1–11.

[43] Yewon Kim, Sung-Ju Lee, and Chris Donahue. 2025. Amuse: Human-AI Collaborative Songwriting with Multimodal Inspirations. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–28.

[44] Barbara Kitchenham, Stuart Charters, et al. 2007. *Guidelines for performing systematic literature reviews in software engineering*. Technical Report. Technical report, ver. 2.3 ebse technical report. ebse.

[45] Satwik Ram Kodandaram, Utku Uckun, Xiaojun Bi, IV Ramakrishnan, and Vikas Ashok. 2024. Enabling Uniform Computer Interaction Experience for Blind Users through Large Language Models. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) *(ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, Article 73, 14 pages. https://doi.org/10.1145/3663548.3675605

[46] Rand Koualty, Nien-Ying Chou, and Suleiman Alabdallah. 2024. Generative AI Agents, Build a Multilingual ChatGPT-based Customer Service Chatbot. In *2024 2nd International Conference on Foundation and Large Language Models (FLLM)*. 5–10. https://doi.org/10.1109/FLLM63129.2024.10852436

[47] Felix Kretzer, Kristian Kolthoff, Christian Bartelt, Simone Paolo Ponzetto, and Alexander Maedche. 2025. Closing the loop between user stories and GUI prototypes: an LLM-based assistant for cross-functional integration in software development. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–19.

[48] Emily Kuang, Minghao Li, Mingming Fan, and Kristen Shinohara. 2024. Enhancing UX Evaluation Through Collaboration with Conversational AI Assistants: Effects of Proactive Dialogue and Timing. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, Article 3, 16 pages.

[49] Geonsun Lee, Yue Yang, Jennifer Healey, and Dinesh Manocha. 2025. Since U Been Gone: Augmenting Context-Aware Transcriptions for Re-Engaging in Immersive VR Meetings. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–20.

[50] Geonsun Lee, Yue Yang, Jennifer Healey, and Dinesh Manocha. 2025. Since U Been Gone: Augmenting Context-Aware Transcriptions for Re-Engaging in Immersive VR Meetings. In *Proceedings of the 2025 CHI Conference on Human*

Factors in Computing Systems (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 785, 20 pages. https://doi.org/10.1145/3706598.3714078

[51] Seonghee Lee, Maho Kohga, Steve Landau, Sile O'Modhrain, and Hari Subramonyam. 2024. AltCanvas: a tile-based editor for visual content creation with generative AI for blind or visually impaired people. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–22.

[52] Chaoyu Li, Sid Padmanabhuni, Maryam S Cheema, Hasti Seifi, and Pooyan Fazli. 2025. VideoA11y: Method and Dataset for Accessible Video Description. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 1055, 29 pages. https://doi.org/10.1145/3706598.3714096

[53] Bingyang Liu, Haoyi Zhang, Xiaohan Gao, Zichen Kong, Xiyuan Tang, Yibo Lin, Runsheng Wang, and Ru Huang. 2025. LayoutCopilot: An LLM-Powered Multiagent Collaborative Framework for Interactive Analog Layout Design. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 44, 8 (2025), 3126–3139. https://doi.org/10.1109/TCAD.2025.3529805

[54] Michael Xieyang Liu, Savvas Petridis, Vivian Tsai, Alexander J Fiannaca, Alex Olwal, Michael Terry, and Carrie J Cai. 2025. Gensors: Authoring Personalized Visual Sensors with Multimodal Foundation Models and Reasoning. In *Proceedings of the 30th International Conference on Intelligent User Interfaces*. 755–770.

[55] Xinyi Lu, Simin Fan, Jessica Houghton, Lu Wang, and Xu Wang. 2023. ReadingQuizMaker: A Human-NLP Collaborative System that Supports Instructors to Design High-Quality Reading Quiz Questions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 454, 18 pages. https://doi.org/10.1145/3544548.3580957

[56] Matteo Manzoni, Sergio Mascetti, Dragan Ahmetovic, Ryan Crabb, and James M. Coughlan. 2025. MapIO: A Gestural and Conversational Interface for Tactile Maps. *IEEE Access* 13 (2025), 84038–84056. https://doi.org/10.1109/ACCESS.2025.3566286

[57] Yihan Mei, Zhao Wu, Junnan Yu, Wenan Li, and Zhibin Zhou. 2025. GeneyMAP: Exploring the Potential of GenAI to Facilitate Mapping User Journeys for UX Design. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–22.

[58] Aditi Mishra, Frederik Brudy, Qian Zhou, George Fitzmaurice, and Fraser Anderson. 2025. WhatIF: Branched Narrative Fiction Visualization for Authoring Emergent Narratives using Large Language Models. In *Proceedings of the 2025 Conference on Creativity and Cognition*. 590–605.

[59] Aditi Mishra, Frederik Brudy, Qian Zhou, George Fitzmaurice, and Fraser Anderson. 2025. WhatIF: Branched Narrative Fiction Visualization for Authoring Emergent Narratives using Large Language Models. In *Proceedings of the 2025 Conference on Creativity and Cognition (C&C '25)*. ACM, 590–605.

[60] Peya Mowar, Yi-Hao Peng, Jason Wu, Aaron Steinfeld, and Jeffrey P Bigham. 2025. CodeA11y: Making AI Coding Assistants Useful for Accessible Web Development. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 45, 15 pages. https://doi.org/10.1145/3706598.3713335

[61] Geetika Munjal, Austin Dious, Devansh Agarwal, Rahul Sharma, and Gunjan Jarawat. 2025. Blindviz—Intelligent Navigation and Environmental Awareness for the Visually Impaired. In *2025 3rd International Conference on Disruptive Technologies (ICDT)*. IEEE, 492–496.

[62] Kálmán Balázs Neszlényi, Alex Milos, and Attila Kiss. 2024. AssistantGPT: Enhancing User Interaction with LLM Integration. In *2024 IEEE 22nd Jubilee International Symposium on Intelligent Systems and Informatics (SISY)*. IEEE, 000619–000624.

[63] Shi Hui Ng, Lay-Ki Soon, and Tin Tin Su. 2023. Emotion-Aware Chatbot with Cultural Adaptation for Mitigating Work-Related Stress. In *Proceedings of the Asian HCI Symposium 2023* (Online, Indonesia) *(Asian CHI '23)*. ACM, 41–50.

[64] Jakob Nielsen. 1994. *Usability engineering*. Morgan Kaufmann.

[65] Zheng Ning, Brianna L Wimer, Kaiwen Jiang, Keyi Chen, Jerrick Ban, Yapeng Tian, Yuhang Zhao, and Toby Jia-Jun Li. 2024. SPICA: interactive video content exploration through augmented audio descriptions for blind or low-vision viewers. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.

[66] Ziqi Pan, Xiucheng Zhang, Zisu Li, Zhenhui Peng, Mingming Fan, and Xiaojuan Ma. 2025. ACKnowledge: A Computational Framework for Human Compatible Affordance-based Interaction Planning in Real-world Contexts. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–20.

[67] Rock Yuren Pang, Hope Schroeder, Kynnedy Simone Smith, Solon Barocas, Ziang Xiao, Emily Tseng, and Danielle Bragg. 2025. Understanding the LLM-ification of CHI: Unpacking the Impact of LLMs at CHI through a Systematic Literature Review. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 456, 20 pages. https://doi.org/10.1145/3706598.3713726

[68] Jimin Park, Chaerin Lee, Eunbin Cho, and Uran Oh. 2024. Enhancing the Podcast Browsing Experience through Topic Segmentation and Visualization with Generative AI. In *Proceedings of the 2024 ACM International Conference on Interactive Media Experiences* (Stockholm, Sweden) *(IMX '24)*. Association for Computing Machinery, 117–128.

[69] Jaehyeong Park, Juncheol Ye, Seungkook Lee, Hyun W. Ka, and Dongsu Han. 2025. NarrAD: Automatic Generation of Audio Descriptions for Movies with Rich Narrative Context. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. 409–419.

[70] Kai Petersen, Sairam Vakkalanka, and Ludwik Kuzniarz. 2015. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and software technology* 64 (2015), 1–18. https://doi.org/10.1016/j.infsof.2015.03.007

[71] Mauro Pezzè, Silvia Abrahão, Birgit Penzenstadler, Denys Poshyvanyk, Abhik Roychoudhury, and Tao Yue. 2025. A 2030 Roadmap for Software Engineering. *ACM Trans. Softw. Eng. Methodol.* 34, 5, Article 118 (may 2025), 55 pages. https://doi.org/10.1145/3731559

[72] Wasifur Rahman, Abdelrahman Abdelkader, Sangwu Lee, Phillip Yang, Md Saiful Islam, Tariq Adnan, Masum Hasan, Ellen Wagner, Sooyong Park, E. Ray Dorsey, Catherine Schwartz, Karen Jaffe, and Ehsan Hoque. 2024. A User-Centered Framework to Empower People with Parkinson's Disease. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 4, Article 175 (Jan. 2024), 29 pages. https://doi.org/10.1145/3631430

[73] Shwetha Rajaram, Hemant Bhaskar Surale, Codie McConkey, Carine Rognon, Hrim Mehta, Michael Glueck, and Christopher Collins. 2025. Gesture and Audio-Haptic Guidance Techniques to Direct Conversations with Intelligent Voice Interfaces. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–20.

[74] Mashrur Rashik, Shilpa Sweth, Nishtha Agrawal, Saiyyam Kochar, Kara M Smith, Fateme Rajabiyazdi, Vidya Setlur, Narges Mahyar, and Ali Sarvghad. 2025. AI-Enabled Conversational Journaling for Advancing Parkinson's Disease Symptom Tracking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM, Article 1029, 23 pages.

[75] Marcos Ribeiro, Joana Santos, João Lobo, Sílvia Araújo, Luís Magalhães, and Telmo Adão. 2024. VR, AR, gamification and AI towards the next generation of systems supporting cultural heritage: addressing challenges of a museum context. In *Proceedings of the 29th International ACM Conference on 3D Web Technology* (Guimarães, Portugal) *(Web3D '24)*. Association for Computing Machinery, New York, NY, USA, Article 9, 10 pages. https://doi.org/10.1145/3665318.3677172

[76] Nathalie Riche, Anna Offenwanger, Frederic Gmeiner, David Brown, Hugo Romat, Michel Pahud, Nicolai Marquardt, Kori Inkpen, and Ken Hinckley. 2025. AI-Instruments: Embodying Prompts as Instruments to Abstract & Reflect Graphical Interface Commands as General-Purpose Tools. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, Article 1104, 18 pages.

[77] Y. Rogers, H. Sharp, and J. Preece. 2013. *Design de Interação - Além da Interação Humano-Computador* (3th. ed.). Bookman Editora.

[78] Phillip Schneider, Nils Rehtanz, Kristiina Jokinen, and Florian Matthes. 2023. Voice-Based Conversational Agents and Knowledge Graphs for Improving News Search in Assisted Living. In *Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments* (Corfu, Greece) *(PETRA '23)*. Association for Computing Machinery, New York, NY, USA, 645–651. https://doi.org/10.1145/3594806.3596534

[79] Chirag Shah, Ryen White, Reid Andersen, Georg Buscher, Scott Counts, Sarkar Das, Ali Montazer, Sathish Manivannan, Jennifer Neville, Nagu Rangan, et al. 2023. Using large language models to generate, validate, and apply user intent taxonomies. *ACM Transactions on the Web* (2023).

[80] Chirag Shah, Ryen White, Reid Andersen, Georg Buscher, Scott Counts, Sarkar Das, Ali Montazer, Sathish Manivannan, Jennifer Neville, Nagu Rangan, Tara Safavi, Siddharth Suri, Mengting Wan, Leijie Wang, and Longqi Yang. 2025. Using Large Language Models to Generate, Validate, and Apply User Intent Taxonomies. *ACM Trans. Web* (May 2025). https://doi.org/10.1145/3732294 Just Accepted.

[81] Hanshu Shen, Lyukesheng Shen, Wenqi Wu, and Kejun Zhang. 2025. Ideationweb: Tracking the evolution of design ideas in human-ai co-creation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–19.

[82] Xinyu Shi, Yinghou Wang, Ryan Rossi, and Jian Zhao. 2025. Brickify: Enabling Expressive Design Intent Specification through Direct Manipulation on Design Tokens. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–20.

[83] Momin N Siddiqui, Roy D Pea, and Hari Subramonyam. 2025. Script&Shift: A layered interface paradigm for integrating content development and rhetorical strategy with llm writing assistants. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–19.

[84] Tianyu Song, Zhengyi Liu, Ruibin Zhao, and Jie Fu. 2024. ElderEase AR: Enhancing Elderly Daily Living with the Multimodal Large Language Model and

Augmented Reality. In *Proceedings of the 2024 International Conference on Virtual Reality Technology*. 60–67.

[85] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured generation and exploration of design space with large language models for human-ai co-creation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–26.

[86] Xin Sun, Yunjie Liu, Jos A. Bosch, and Zhuying Li. 2025. Interface Matters: Exploring Human Trust in Health Information from Large Language Models via Text, Speech, and Embodiment. *Proc. ACM Hum.-Comput. Interact.* 9, 2, Article CSCW116 (May 2025), 32 pages.

[87] Macy Takaffoli, Sijia Li, and Ville Mäkelä. 2024. Generative AI in User Experience Design and Research: How Do UX Practitioners, Teams, and Companies Use GenAI in Industry?. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 1579–1593. https://doi.org/10.1145/3643834.3660720

[88] Michael Tamascelli, Olivia Bunch, Blake Fowler, Maryam Taeb, and Achraf Cohen. 2025. Academic Advising Chatbot Powered with AI Agent. In *Proceedings of the 2025 ACM Southeast Conference*. 195–202.

[89] Xilin Tang, Jerrod Windham, and Benjamin Bush. 2024. Pre-AI and post-AI design: balancing human Creativity and AI Tools in the Industrial Design Process. In *Proceeding of the 2024 International Conference on Artificial Intelligence and Future Education*. 100–108.

[90] Kelsey Turbeville, Jennarong Muengtaweepongsa, Samuel Stevens, Jason Moss, Amy Pon, Kyra Lee, Charu Mehra, Jenny Gutierrez Villalobos, and Ranjitha Kumar. 2024. LLM-powered Multimodal Insight Summarization for UX Testing. In *Proceedings of the 26th International Conference on Multimodal Interaction (ICMI '24)*. Association for Computing Machinery, 4–11.

[91] Zhan Wang, Lin-Ping Yuan, Liangwei Wang, Bingchuan Jiang, and Wei Zeng. 2024. VirtuWander: Enhancing Multi-modal Interaction for Virtual Tour Guidance through Large Language Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 612, 20 pages. https://doi.org/10.1145/3613904.3642235

[92] Zijie J Wang, Chinmay Kulkarni, Lauren Wilcox, Michael Terry, and Michael Madaio. 2024. Farsight: Fostering responsible ai awareness during ai application prototyping. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–40.

[93] Tobias M Weinberg, Kowe Kadoma, Ricardo E. Gonzalez Penuela, Stephanie Valencia, and Thijs Roumen. 2025. Why So Serious? Exploring Timely Humorous Comments in AAC Through AI-Powered Interfaces. In *Proceedings of the 2025 CHI (CHI '25)*. ACM, Article 759, 19 pages.

[94] Justin D. Weisz, Jessica He, Michael Muller, Gabriela Hoefer, Rachel Miles, and Werner Geyer. 2024. Design Principles for Generative AI Applications. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 378, 22 pages. https://doi.org/10.1145/3613904.3642466

[95] Tongshuang Wu, Michael Terry, and Carrie Jun Cai. 2022. Ai chains: Transparent and controllable human-ai interaction by chaining large language model prompts. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–22.

[96] Haijun Xia, Tony Wang, Aditya Gunturu, Peiling Jiang, William Duan, and Xiaoshuo Yao. 2023. CrossTalk: Intelligent substrates for language-oriented interaction in video-based communication and collaboration. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–16.

[97] Wei Xiang, Hanfei Zhu, Suqi Lou, Xinli Chen, Zhenghua Pan, Yuping Jin, Shi Chen, and Lingyun Sun. 2024. SimUser: Generating Usability Feedback by Simulating Various Users Interacting with Mobile Applications. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 9, 17 pages. https://doi.org/10.1145/3613904.3642481

[98] Shuchang Xu, Chang Chen, Zichen Liu, Xiaofu Jin, Lin-Ping Yuan, Yukang Yan, and Huamin Qu. 2024. Memory Reviver: Supporting Photo-Collection Reminiscence for People with Visual Impairment via a Proactive Chatbot. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–17.

[99] Litao Yan, Alyssa Hwang, Zhiyuan Wu, and Andrew Head. 2024. Ivie: Lightweight Anchored Explanations of Just-Generated Code. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 140, 15 pages. https://doi.org/10.1145/3613904.3642239

[100] Sojeong Yun and Youn-kyung Lim. 2025. User Experience with LLM-powered Conversational Recommendation Systems: A Case of Music Recommendation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 898, 15 pages. https://doi.org/10.1145/3706598.3713347

[101] Guanhua Zhang, Mohamed Adel Naguib Ahmed, Zhiming Hu, and Andreas Bulling. 2025. Summact: Uncovering user intentions through interactive behaviour summarisation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–17.

[102] Mengxi Zhang, Huaxiao Liu, Changhao Du, Tengmei Wang, Han Li, Pei Huang, and Chunyang Chen. 2025. Distinguishing GUI Component States for Blind Users using Large Language Models. *ACM Transactions on Software Engineering and Methodology* (2025).

[103] Tianhao Zhang, Fu Peiguo, Jie Liu, Yihe Zhang, and Xingmei Chen. 2024. NLDesign: A UI Design Tool for Natural Language Interfaces. In *Proceedings of the ACM Turing Award Celebration Conference - China 2024* (Changsha, China) *(ACM-TURC '24)*. Association for Computing Machinery, New York, NY, USA, 153–158. https://doi.org/10.1145/3674399.3674455

[104] Qingxiao Zheng, Minrui Chen, Pranav Sharma, Yiliu Tang, Mehul Oswal, Yiren Liu, and Yun Huang. 2025. EvAlignUX: Advancing UX Evaluation through LLM-Supported Metrics Exploration. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM, Article 1051, 25 pages.

[105] Mingyuan Zhong, Ruolin Chen, Xia Chen, James Fogarty, and Jacob O Wobbrock. 2025. ScreenAudit: Detecting Screen Reader Accessibility Errors in Mobile Apps Using Large Language Models. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–19.

[106] Ting Zhou, Yanjie Zhao, Xinyi Hou, Xiaoyu Sun, Kai Chen, and Haoyu Wang. 2025. DeclarUI: Bridging Design and Development with Automated Declarative UI Code Generation. *Proc. ACM Softw. Eng.* 2, FSE, Article FSE011 (June 2025), 23 pages.