REVIEWING THE ROLE OF ARTIFICIAL INTELLIGENCE IN MOBILE APP CREATION

Doan Ngoc Phuong*, Nguyen Thu Phuong, Nguyen Thi Dung
TNU - University of Information and Communication Technology

ARTICLE INFO

Received: 31/5/2025
Revised: 30/6/2025
Published: 30/6/2025

KEYWORDS

AI-powered mobile-app creation
Bibliometric analysis
Systematic review
Generative AI
PRISMA

ABSTRACT

Recent breakthroughs in generative artificial intelligence have turned this technology into the primary engine of mobile app development. Through a combined bibliometric analysis and a PRISMA 2020 review of 56 studies (from January 2020 to May 2025), we surveyed publication years, publisher distribution and keyword co-occurrence networks to redraw the field's development. From there, we identified ten critical gaps, among which the most notable are the lack of artificial intelligence driven user interface testing frameworks and the scarcity of high quality multisensory UI/UX datasets that record visual, tactile and gaze data simultaneously. To address these challenges, two research trends have emerged: prompt engineering, which fine tunes input instructions for large language models, and few shot learning, which enables models to generalize from very few examples. Building on these findings, we propose six future research directions: (1) develop diverse multisensory UI/UX datasets; (2) integrate end to end artificial intelligence testing pipelines; (3) optimize prompts for interface generation; (4) design modular cross platform code architectures; (5) embed explainable artificial intelligence mechanisms in generated interfaces; and (6) deploy large scale quality assurance as a service via automated cloud-based pipelines. These steps lay out a clear roadmap for both academia and practice to advance artificial intelligence driven mobile app creation.

NGHIÊN CỨU TỔNG QUAN VỀ TẠO ỨNG DỤNG DI ĐỘNG BẰNG TRÍ TUỆ NHÂN TẠO

Đoàn Ngọc Phương*, Nguyễn Thu Phương, Nguyễn Thị Dung
Trường Đại học Công nghệ thông tin và Truyền thông – Đại học Thái Nguyên

THÔNG TIN BÀI BÁO

Ngày nhận bài: 31/5/2025
Ngày hoàn thiện: 30/6/2025
Ngày đăng: 30/6/2025

TỪ KHÓA

Tạo ứng dụng di động bằng AI
Phân tích thư mục học
Tổng quan hệ thống
AI tạo sinh
PRISMA

TÓM TẮT

Những bước tiến đột phá trong trí tuệ nhân tạo sinh tạo đã đưa công nghệ này trở thành động cơ chính cho phát triển ứng dụng di động. Thông qua phân tích thư mục học kết hợp tổng quan PRISMA 2020 trên 56 công trình (từ tháng 01 năm 2020 đến tháng 5 năm 2025), chúng tôi đã khảo sát năm xuất bản, phân bố nhà xuất bản và mạng lưới đồng xuất hiện từ khóa để vẽ lại bức tranh phát triển của lĩnh vực. Từ đó, chúng tôi nhận diện mười khoảng trống then chốt, trong đó đáng chú ý nhất là thiếu khung kiểm thử giao diện do trí tuệ nhân tạo điều khiển và sự khan hiếm bộ dữ liệu UI/UX đa giác quan chất lượng cao ghi đồng thời tín hiệu thị giác, xúc giác và ánh nhìn. Để giải quyết các thách thức trong giai đoạn này, hai xu hướng nghiên cứu chủ đạo là tinh chỉnh chỉ dẫn cho mô hình ngôn ngữ lớn và phương pháp học với số lượng mẫu hạn chế. Trên cơ sở những kết quả đó, chúng tôi đề xuất sáu hướng nghiên cứu tương lai: (1) phát triển bộ dữ liệu UI/UX đa giác quan đa dạng; (2) tích hợp quy trình kiểm thử trí tuệ nhân tạo đầu-cuối; (3) tối ưu chỉ dẫn cho sinh giao diện; (4) thiết kế kiến trúc mã module đa nền tảng; (5) tích hợp cơ chế trí tuệ nhân tạo có khả năng giải thích; và (6) triển khai quy mô lớn qua đường ống tự động trên đám mây. Những bước này vạch ra lộ trình rõ ràng cho cả học thuật và thực tiễn thúc đẩy ứng dụng trí tuệ nhân tạo trong phát triển di động.

DOI: https://doi.org/10.34238/tnu-jst.12946

* Corresponding author. *Email: dnphuong@ictu.edu.vn*

1. Introduction

The explosion of generative artificial intelligence (AI) models—most notably ChatGPT and other code-centric Large Language Models (LLMs)—has ushered in a new era in software development, where AI not only assists but also directly participates in interface design and code generation [1], [2]. In the mobile domain, this trend—referred to here as AI-powered mobile-app creation [3]—promises to automate many stages of the traditional toolchain and shorten the journey from idea to product.

During 2020–2021, most studies remained at the proof-of-concept stage, using natural-language understanding to generate small code snippets or simple interfaces [4]. From mid-2021 onward, the maturing code- and user interface (UI)-generation capabilities of multimodal LLMs—combined with the need for automation on resource-constrained devices—have accelerated the rise of no- and low-code platforms such as FlutterFlow and Microsoft Power Platform, illustrating the shift from manual operations to prompt-based, semantics-driven design [5], [6].

However, in addition to the benefits of speed and convenience, applying AI to the entire application development process also poses requirements for ensuring output quality, source code security, and long-term maintainability [7], [2]. Current code generation systems often lack built-in testing mechanisms and are difficult to scale when faced with complex user scenarios. At the same time, algorithmic transparency, which involves explaining how AI makes design decisions, is limited, affecting trust and compliance with software standards [8], [9].

Some individual studies have proposed solutions, such as using LLM to generate layouts from sketches [10], generating test cases automatically from user logs [11], [12], or combining sketches and prompts to edit interfaces [2]. However, these studies only address discrete aspects, lacking a systematic overview to fully outline the challenges, potentials, and trends in this field [12].

To fill this gap, we conducted a review study combining a quantitative survey of publication data with a rigorous PRISMA screening of scientific articles published in the last five years, with the following main objectives:

• Research question (RQ) 1: analyze publication trends, publisher distribution, and keyword networks during 2020–2025;

• RQ2: code and cluster key challenges in the field;

• RQ3: propose future research directions based on identified gaps.

Finally, by combining the results from the two methods above, we hope to establish a foundation for further research, guide the development of tools to support higher automation, and suggest a comprehensive evaluation framework for future AI solutions to create mobile applications [13].

2. Methods

This study combines two main methods: bibliometric analysis [14] and systematic review [15]. Through bibliometric analysis, we draw an overview of the publication trends, journal distribution, keyword networks, and the development of research topics related to "AI-powered mobile-app creation" in the period from January 2020 to May 2025. At the same time, a systematic review following the PRISMA 2020 framework [15] helps to screen, select, and deeply evaluate 56 criterion papers, thereby extracting key challenges and future research directions in a transparent and reproducible manner

2.1. Data sources and search strategy

The literature for the study was collected from reputable academic databases including Google Scholar, IEEE Xplore, ACM Digital Library, and SpringerLink [5]. To ensure comprehensive coverage, we used a combination of primary keywords ("artificial intelligence" OR "generative AI" OR "large language model" OR "LLM") combined with phrases related to mobile applications ("mobile app" OR "mobile application" OR "UI generation" OR "app creation"). The entire search process was limited to the period from January 2020 to May 2025, considering only peer-reviewed English-language articles, including journals and international conference proceedings [7].

2.2. Inclusion and exclusion criteria

Figure 1 presents the PRISMA 2020 workflow for our study selection process. The literature selection process followed the PRISMA 2020 framework to ensure transparency and reproducibility [15]. Initially, 1012 records were collected from the databases; after removing 210 duplicates and 148 automatically ineligible records, 654 titles and abstracts were reviewed. Next, 295 records unrelated to "AI-powered mobile-app creation" were excluded, leaving 359 full texts retrieved, of which 30 were not retrieved. Of the remaining 329 full texts, 165 were not within the scope of AI-powered mobile-app creation and 108 were of low quality or inaccessible, leaving 56 studies that met the criteria for in-depth analysis [15].

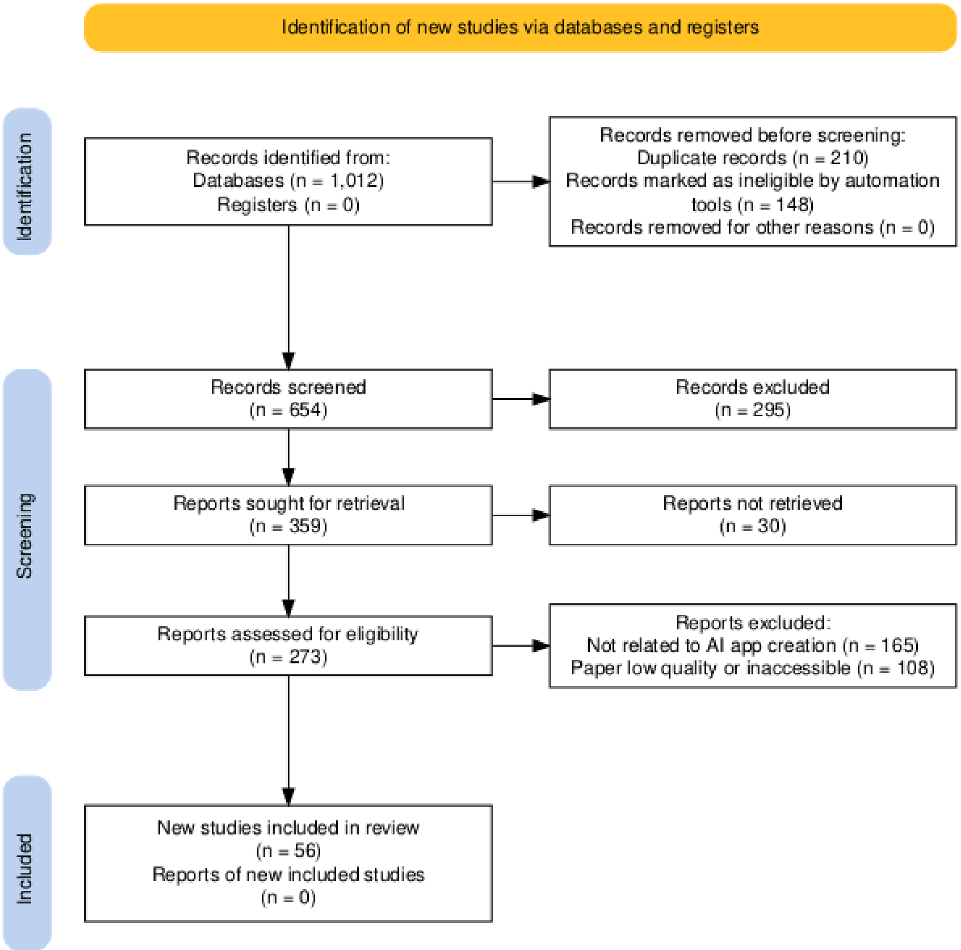

Figure 1. *PRISMA workflow illustrating the identification and selection of articles included in the study*

2.3. Data analysis

Data analysis was performed on two main streams: quantitative and qualitative. For the quantitative analysis (bibliometric analysis), we used VOSviewer software to build a keyword co-occurrence matrix, draw a topic network map, and calculate the publication index by year, author, journal, and place of publication [14]. This result allows us to identify research trends, connections between main topics, and key groups of authors in the community.

In parallel, a systematic review was conducted through an open coding process in Excel [15]. Each article was labeled according to research challenges and future directions, and then grouped into 10 key challenges and 6 potential research directions. To ensure reliability, two independent reviewers performed the coding, with a Cohen's κ coefficient of agreement of 0.86, and then discussed the final results. Synthesizing these two streams of analysis in parallel not only provides macro data on the publication landscape but also provides insights into the content and research gaps in the field of AI-powered mobile app creation.

3. Results and Discussion

3.1. Publication trends and number of articles on AI-powered mobile-app creation by year

Figure 2 shows three major phases of AI research in mobile app creation: early experimentation (2020–2021), transition (2021–2022), and LLM explosion (2023–mid-2025) [7].

The first phase (2020 – early 2021) was characterized by a limited number of publications, reflecting the fact that AI was mainly applied to single tasks that did not yet interfere with the overall structure of the software development chain [4]. The transition phase (mid-2021 – 2022) marked the popularization of mobile-oriented AutoML pipelines along with lightweight inference technologies such as TensorFlow Lite or Core ML. However, the growth rate temporarily slowed down as the research community began to shift its focus from optimizing embedded models to the more complex goal of creating complete applications through AI.

A clear turning point came in 2023 with the widespread availability of large language models (LLMs) that can generate code directly from natural language descriptions [1], [16]. This marks a shift from a fragmented approach to a comprehensive AI-driven application creation process, and opens up new research directions such as adaptive model compression, federated learning, and privacy protection [17].

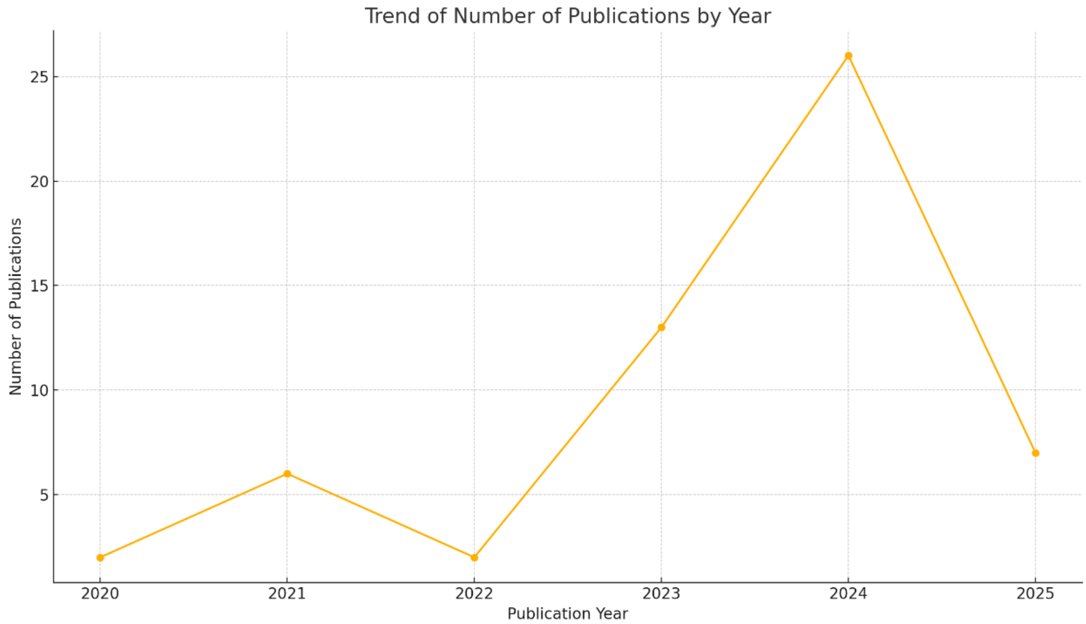

Figure 2. *Number of publications (2020–2025) in the field of AI-Powered Mobile-App Creation*

The period 2024 - first half of 2025 recorded strong progress when three factors: a rich open source data repository, an AI-integrated CI/CD platform, and support from prestigious academic forums converged, promoting research beyond pure code generation capabilities to automated testing, improved user experience, and a complete DevOps toolchain [7], [18]. At the same time, the trend of standardizing LLM-compatible APIs on Android and iOS, along with the emergence of frameworks that optimize inference performance on the device, affirm that AI has become a core factor throughout the life cycle of developing new-generation mobile applications.

3.2. Core publishers and journals driving research on AI-powered mobile-app creation

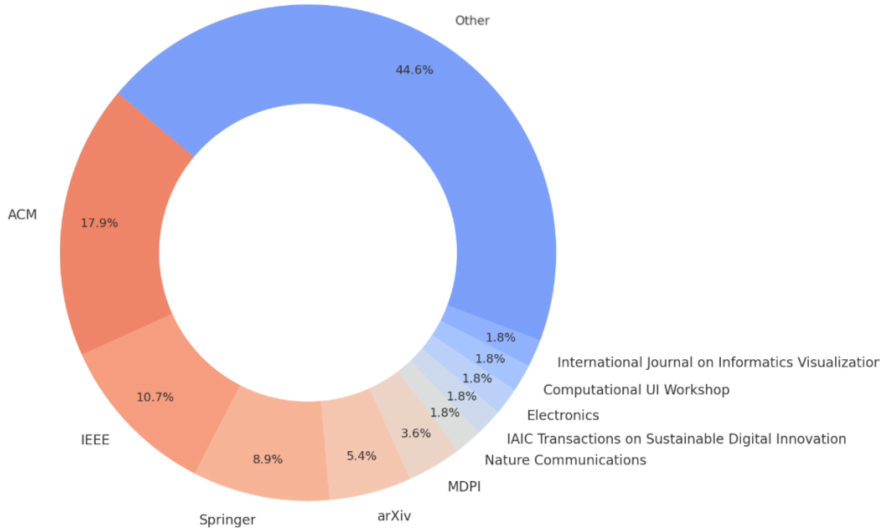

Figure 3. *Distribution of selected articles by publisher*

Figure 3 showing the proportion of articles published by the place of publication of research works in the field of AI-powered mobile-apps shows a rather typical two-tier structure [14], and Table 1 provides the exact counts of publications per venue. At the core, major publishing channels such as ACM, IEEE, and Springer act as the main academic anchors, attracting the majority of research with high technical content or proposing new fundamental methods [7]. The fact that works on creating AI-powered mobile apps are concentrated here shows that this field is gradually taking shape as a research branch with depth and entering a period of academic stability [7].

Table 1. *Three leading publishers and their publication shares*

Publisher	Number of articles	Percentage
ACM	10	17.9
IEEE	6	10.7
Springer	5	8.9
Others	35	62.5

However, Figure 3 also reflects a significant fragmentation at the outer layer, where a range of smaller journals and conferences contribute to the expansion of the space for discussion and experimentation [14]. The proportion of articles in the "Other" category accounts for a large proportion of the total number of publications, which is not simply a reflection of fragmentation, but rather shows that the field is still in a dynamic development phase, where new topics, interdisciplinary approaches, and open access models have the opportunity to thrive. From this, it can be seen that the publication ecosystem in the field operates according to a "stable center – flexible periphery" model: large channels shape research norms, while open repositories and small conferences enable rapid innovation and experimentation.

3.3. Keyword co-occurrence network analysis in AI-powered mobile-app creation

To clarify the knowledge structure and relationships between topics, we construct a keyword co-occurrence network (Figure 4) based on 230 standardized labels [14]. At the center, the AI node plays a core role, connecting to multidisciplinary topic clusters such as natural language processing, android application security, digital health, and agricultural technology [7], demonstrating the influence beyond pure software development. The large language models cluster is closely associated with GPT, transformer, interface, code generation, and generative AI, indicating the explosion of LLM applications in automatic interface and source code generation [10], [19]. At the same time, the adaptive interface and prototyping concepts reveal the trend of building flexible applications that adapt to user behavior in real time. In addition, the mobile application development node is associated with explainable AI and emerging technologies, reflecting the parallel concerns between performance and transparency of AI-generated systems [8].

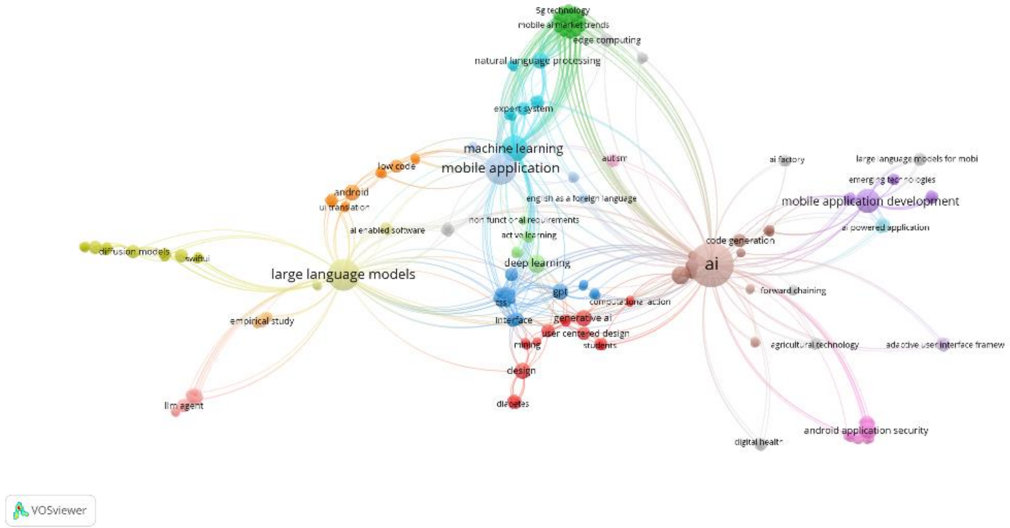

Figure 4. *Keyword co-occurrence network in research on AI and mobile-app development*

To confirm these points, we further identified the top ten keywords by total link strength (Figure 5). Terms such as AI, large language models, mobile application indicate research focuses and at the same time expose notable gaps [14].

Selected	Keyword	Occurrences	Total link strength
☑	ai	24	33
☑	machine learning	8	22
☑	large language models	12	21
☑	mobile application	12	19
☑	interface	2	9
☑	user interface	2	9
☑	android application security	3	8
☑	code vulnerability	3	8
☑	deep learning	4	8
☑	gpt	3	8

Figure 5. *Top 10 keywords by total link strength*

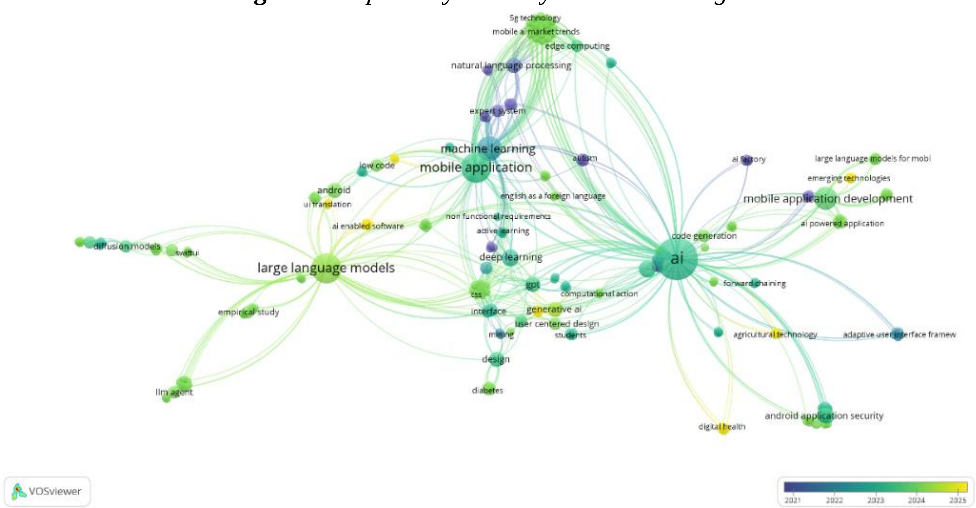

Figure 6. *Year-layered keyword co-occurrence network*

Finally, to observe the emergence of topics over time, we overlay the keyword network with time (Figure 6). The results show that from 2024, prompt engineering, automated UI testing, and few-shot learning emerged, confirming the shift from theoretical research to the implementation of automated interface testing tools and prompt optimization models for UI/UX [2].

In summary, from basic network analysis, tracking trends over time, and assessing link strength, we have drawn a comprehensive picture of the relationships between key topics, and at the same time pointed out potential research gaps for further work [14].

3.4. Concept evolution analysis in AI-powered mobile-app creation

Following the analysis of knowledge structure from keyword networks, we examine the evolution of concepts over time through the academic vocabulary evolution chart.

Figure 7 shows the evolution of the academic vocabulary in articles related to the topic of "AI-powered application development" over time from 2020 to 2025, classified by word class: noun (blue), verb (orange), and adjective (green). The layout of Figure 7 suggests a semantic movement that fairly accurately reflects the development of the field in terms of both content breadth and technical depth [14].

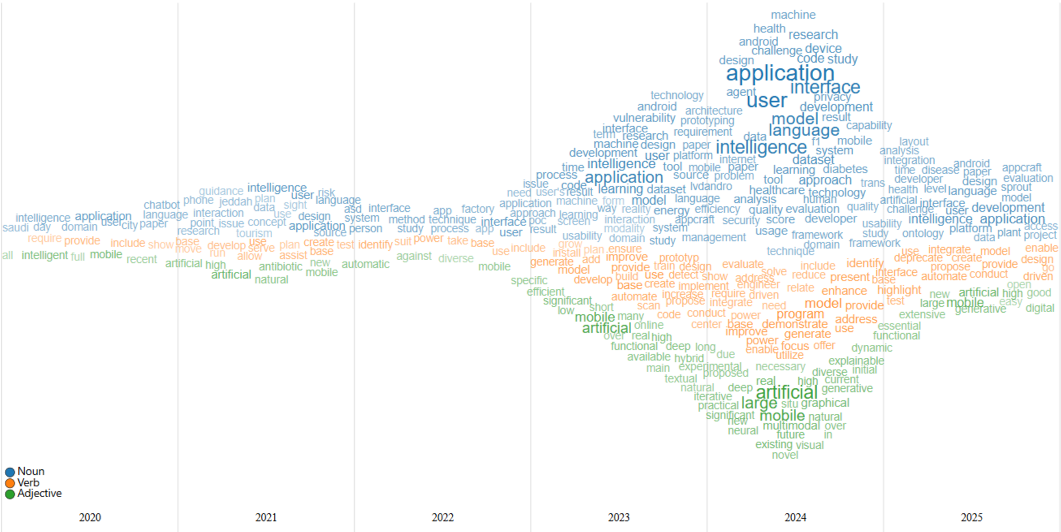

Figure 7. *Evolution of research topics in AI-Powered Mobile-App Creation (2020–2025)*

In 2020–2021, the terms that emerged were mostly descriptive and general, with nouns like “application”, “user”, “language”, “intelligence” indicating an initial focus around basic user interface and functional structure. Verbs like “develop”, “use”, “assist” and adjectives like “automatic”, “intelligent” reflect early efforts to integrate small AI components into traditional application systems [1], [4]. From 2022 onwards, the vocabulary density increases significantly, leading to an expansion in topics and depth of expertise. Clusters such as “interface design”, “machine learning”, “dataset”, “platform”, “framework” and “energy efficiency” represent a shift from interface construction to optimization of the entire development chain. In particular, the period 2023–2024 witnessed the dense appearance of technical and systemic terms such as “federated learning”, “prototype”, “evaluation”, “ontology”, “explainable AI”, showing that the field is entering a phase of shaping specialized methodologies and standardizing tools [17].

From 2024 to early 2025, many keywords such as “generative”, “iterative”, “novel”, “practical usage” or “visual layout” appear with suggestive nature and mark the shift from technical exploration to user experience enhancement, paving the way for end-to-end generative systems. Overall, Figure 7 not only reflects the process of increasing technical complexity, but also shows the convergence of emerging research directions such as federated machine learning, explainability, and multimodal design - which are the premise for the sustainable development of the field of AI applications in creating applications on smart devices [2].

3.5. Synthesis of challenges and future research directions

Based on a high-fidelity open-coding process (Cohen's κ = 0.86) [15], we identified ten key challenges in AI-powered mobile-app creation (Table 2). Each challenge and its corresponding future research direction are underpinned by representative sources from the recent literature, ensuring the synthesis is grounded in empirical or conceptual evidence. Table 2 therefore not only lists the issues but also highlights the studies that motivate both the challenges and the proposed avenues for future work. Among them, ignoring usability (29 times) [8], the scarcity of large-scale, noise-filtered UI datasets such as MUD [20], and not fully exploiting the power of AI/ML (26 times) [7] were the two most frequently mentioned issues, indicating the urgent need for a standard usability-guideline framework and an effective AI/ML integration mechanism.

Table 2. *Ten key challenges, gaps, and future research directions*

Challenge (Occurrences)	Core gap	Future research direction
Neglect of Functional Usability [8] (29)	Models focus on aesthetics while omitting usability during interface generation.	Exploit task-oriented knowledge embeddings: inject domain ontologies and usability factors into the UI generation process [8]
Under-utilization of AI/ML Capabilities [7] (26)	Lack of optimal architecture for large models on resource-constrained devices.	Develop lightweight-LUIMs combining hybrid LLM+LVM+LCM architectures with parameter sharing and adapter fusion for edge devices [7]
Limited Code Input & Native Output [21] (22)	Fragmented code generation for HTML/CSS/Flutter; lacking end-to-end "prompt-to-code" compilation to native apps.	Build Compiler-in-the-Loop for Generative UI: create an intermediate AST and compile to target platforms, enabling "write once, deploy anywhere." [21]
Missing Standardized UI/UX Guidelines [9] (22)	Generative models are not fine-tuned on industry standards; usability attributes are not integrated into UI generation.	Study prompt-engineering derived from UI/UX design patterns: map patterns to instruction templates and embed usability attributes in generation [9]
Lack of User-Centred Design Tools (21)	Gap between user-centred design (UCD) and UI generation.	Develop Human-in-the-Loop co-creation tools: support drag-and-drop, natural-language refinement, and version control of UI states [22]
Scalability & Adaptation in Dynamic Environments (17)	No mechanisms for real-time layout restructuring, explainability, or ethical compliance in interfaces.	Build Adaptive-Ethical LUIMs with context-aware personalization, real-time explainability, and guarantees for ethics, security, and privacy [4]
Lack of Personalization & User Interaction [2] (15)	Learning-by-operating and on-device online learning not yet exploited.	Design multimodal dynamic personalization models combining behavioral and affective signals with on-device continual learning [1]
Lack of Testing & Validation Support [11] (14)	No quality-assurance process for automatically generated UI/UX.	Develop QA-as-a-Service for LUIMs: build an automated pipeline to generate and test interfaces, leveraging LLM self-refinement [5]
Scarcity of Multisensory Data & UX Feedback [20] (13)	Absence of large multimodal datasets (eye-tracking, touch heat-map, SUS) for training LUIMs.	Establish a large-scale multimodal UI/UX dataset with interaction videos, eye-tracking, touch heat-maps, and SUS scores for LUIM training [20]
Poor Comprehension of Complex & Implicit Layout Semantics [16] (13)	Inability to understand multi-dimensional dynamic layouts.	Research Semantic Layout Transformer: integrate graph attention and scene graphs to model complex interface structures accurately [23]

Based on the above challenges, six future research axes are proposed in Table 3. Specifically, contextual model personalization and explainable AI are two priority directions to overcome the limitations of transparency and reproducibility [14]. At the same time, extending modularity to no-/low-code frameworks and integrating federated learning are also considered as key opportunities to optimize the performance and security of AI-driven UI pipelines [21].

Table 3. *Six future research axes for AI-Powered Mobile-App Creation*

Research axis	Detailed description
QA-as-a-Service	AI-based automated UI testing and evaluation services providing an end-to-end pipeline from layout generation through test execution to self-healing.
Semantic Layout Transformer	Models that convert semantic specifications into interface layouts using attention over scene-graph representations to generate complex yet coherent layouts.
Multisensory UI/UX Dataset	A multimodal repository (interaction videos, eye-tracking, touch heat-maps, SUS ratings) for training and benchmarking next-generation UI/UX models.
Multi-modal Dynamic Personalization	Interface personalization driven by behavioral signals and real-time context (e.g., emotion, ambient light), coupled with on-device online learning.
Compiler-in-the-Loop	A "prompt-to-code" pipeline via an intermediate AST that translates visual structure and semantics into native Android, iOS, or Flutter source code.
Human-in-the-Loop Co-creation	Collaborative tools between designers and AI enabling drag-and-drop, natural-language editing, and instant feedback to optimize UX.

In summary, challenges in automated UI testing, contextual personalization, modularity in code generation, scarcity of quality UI/UX data, and lack of integration of guidelines into prompt engineering are limiting the potential of AI-powered mobile-app creation. Limited device resources and lack of human-in-the-loop tools also slow down the transition from research to practical application. Building a QA-as-a-Service platform, developing a Semantic Layout Transformer, and establishing a multimodal data repository will be key steps to overcome these bottlenecks and accelerate the evolution of AI-powered mobile app generation.

4. Conclusion

Using bibliometric analysis and a PRISMA guided review, this study answers three research questions. For RQ1, we outline three distinct developmental phases of AI in mobile app creation, tracing the journey from early experimentation to the era of large language models. For RQ2, we identify ten main obstacles, the most critical being the absence of a standard user interface testing framework and the shortage of multisensory UI and UX datasets. For RQ3, we present six promising research paths, such as QA as a Service, a compiler in the loop approach, and prompt optimisation for interface design, each aligned with short term, medium term, and long term milestones. The ordered list of challenges offers a clear priority map for engineers, while the six paths provide a practical roadmap for embedding AI into continuous integration and delivery pipelines, generating cross platform code, and producing explainable interfaces. We also encourage the release of rich multisensory datasets and the creation of benchmarks linked to quality assurance to close the gap between prototypes and production. By connecting academic insight with practical needs, this work aims to help researchers and industry adopt AI more effectively and sustainably in mobile application development.

REFERENCES

[1] D. Chen, X. Zhang, J. Lee, *et al.*, "LLM for mobile: An initial roadmap," *ACM Trans. Softw. Eng. Methodol.*, vol. 34, no. 5, pp. 1–29, May 2025, doi: 10.1145/3708528.

[2] J. Wei, A.-L. Courbis, T. Lambolais, G. Dray, and W. Maalej, "On AI-inspired UI design," *IEEE Softw.*, vol. 42, no. 3, pp. 50–58, 2025, doi: 10.1109/MS.2025.3536838.

[3] D. Dao, J. Y. C. Teo, W. Wang, and H. D. Nguyen, "LLM-Powered Multimodal AI Conversations for Diabetes Prevention," *in Proc. 1st ACM Workshop AI-Powered Q&A Syst. Multimedia (AIQAM '24)*, Phuket, Thailand, 10 Jun. 2024, pp. 1–6, doi: 10.1145/3643479.3662049.

[4] M. Hasan, K. S. Mehrab, W. U. Ahmad, and R. Shahriyar, “Text2App: A framework for creating Android apps from text descriptions,” *arXiv preprint arXiv:2104.08301 [cs.SE]*, Apr. 16, 2021.
[5] S. Böhm and S. Graser, “AI-based mobile app prototyping: Status quo, perspectives, and preliminary insights from experimental case studies,” in *Proc. 16th Int. Conf. Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services (CENTRIC 2023)*, Valencia, Spain, Nov. 13–17, 2023, pp. 29–37.
[6] L. Alwakeel, K. Lano, and H. Alfraihi, “Towards integrating machine learning models into mobile apps using AppCraft,” in *Proc. Agile Model-driven Engineering Workshop (Agile MDE) at the Software Technologies: Applications and Foundations (STAF) Federated Conferences*, Leicester, U.K., Jul. 20, 2023, CEUR Workshop Proc., vol. 3620, pp. 7–10.
[7] Y. Li, X. Dang, H. Tian, *et al.*, “An empirical study of AI techniques in mobile applications,” *J. Syst. Softw.*, vol. 219, Art. no. 112233, Jan. 2025, doi: 10.1016/j.jss.2024.112233.
[8] A. Namoun, A. Alrehaili, Z. U. Nisa, H. Almoamari, and A. Tufail, “Predicting the usability of mobile applications using AI tools: The rise of large user-interface models, opportunities, and challenges,” *Procedia Comput. Sci*., vol. 238, pp. 671–682, 2024, doi:10.1016/j.procs.2024.06.076.
[9] M. H. Miraz, M. Ali, and P. S. Excell, “Cross-cultural usability evaluation of AI-based adaptive user interface for mobile applications,” *Acta Sci. Technol.*, vol. 44, no. 1, Art. no. e61112, Jul. 2022, doi: 10.4025/actascitechnol.v44i1.61112.
[10] J. Wei, A.-L. Courbis, T. Lambolais, B. Xu, P.-L. Bernard, and G. Dray, “Boosting GUI prototyping with diffusion models,” in *Proc. 2023 IEEE 31st Int. Requirements Engineering Conf. (RE)*, Hannover, Germany, Sept. 4–8, 2023, pp. 275–280, doi: 10.1109/RE57278.2023.00035.
[11] J. Senanayake, H. Kalutarage, M. O. Al-Kadri, A. Petrovski, and L. Piras, “Android code vulnerabilities early detection using AI-powered ACVED plugin,” in *Proc. 37th Annu. IFIP WG 11.3 Conf. Data and Applications Security and Privacy (DBSec 2023)*, Sophia Antipolis, France, Jul. 19–21, 2023, pp. 339–357, doi: 10.1007/978-3-031-37586-6_20.
[12] L. Alwakeel, K. Lano, and H. Alfraihi, “AppCraft: Model-driven development framework for mobile applications,” *IEEE Access*, vol. 13, pp. 23658–23699, Feb. 2025, doi: 10.1109/ACCESS.2025.3536321.
[13] Y. Gui, Y. Wan, Z. Li, Z. Zhang, D. Chen, H. Zhang, *et al.*, “UICoPilot: Automating UI synthesis via hierarchical code generation from webpage designs,” *in Proc. ACM Web Conf. 2025 (WWW ’25)*, Sydney, NSW, Australia, 28 Apr.–2 May 2025, pp. 1–10, doi: 10.1145/3696410.3714891.
[14] I. H. Sarker, M. M. Hoque, M. K. Uddin, and T. Alsanoosy, “Mobile data science and intelligent apps: Concepts, AI-based modeling and research directions,” *Mobile Netw. Appl.*, vol. 26, no. 1, pp. 285–303, Jan. 2021, doi: 10.1007/s11036-020-01650-z.
[15] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, *et al.*, “The PRISMA 2020 statement: An updated guideline for reporting systematic reviews,” *BMJ*, vol. 372, Art. no. n71, Mar. 2021, doi: 10.1136/bmj.n71.
[16] M. Xing, R. Zhang, H. Xue, Q. Chen, F. Yang, and Z. Xiao, “Understanding the weakness of large language model agents within a complex Android environment,” in *Proc. 30th ACM SIGKDD Conf. Knowl. Discov. Data Min. (KDD ’24)*, Barcelona, Spain, Aug. 25–29, 2024, pp. 6061–6072, doi: 10.1145/3637528.3671650.
[17] J. Senanayake, H. Kalutarage, L. Piras, M. O. Al-Kadri, and A. Petrovski, “Assuring privacy of AI-powered community-driven Android code vulnerability detection,” in *Proc. ESORICS Int. Workshops 2024*, Lecture Notes in Computer Science, vol. 15264, Springer, Bydgoszcz, Poland, Sept. 16–20, 2024, pp. 457–476, doi: 10.1007/978-3-031-82362-6_27.
[18] S. Petridis, M. X. Liu, A. J. Fiannaca, V. Tsai, M. Terry, and C. J. Cai, “In situ AI prototyping: Infusing multimodal prompts into mobile settings with MobileMaker,” in *Proc. 2024 IEEE Symp. Visual Languages and Human-Centric Comput. (VL/HCC)*, Liverpool, U.K., Sept. 2–6, 2024, pp. 121–133, doi: 10.1109/VL/HCC60511.2024.00023.
[19] F. Huang, G. Li, X. Zhou, J. F. Canny, and Y. Li, “Creating user interface mock-ups from high-level text descriptions with deep-learning models,” *arXiv preprint arXiv:2110.07775 [cs.HC]*, Oct. 14, 2021.
[20] S. Feng, S. Ma, H. Wang, D. Kong, and C. Chen, “MUD: Towards a large-scale and noise-filtered UI dataset for modern style UI modeling,” in *Proc. 2024 CHI Conf. Human Factors Comput. Syst. (CHI ’24)*, Honolulu, HI, USA, May 11–16, 2024, doi: 10.1145/3613904.3642350.

[21] S. N. Ardini, S. Sunarya, and K. Latifah, "Development of mobile application through the concept of artificial intelligence to enhance pronunciation skill in EFL," *KnE Social Sciences*, vol. 9, no. 6, pp. 56–66, Mar. 2024, doi: 10.18502/kss.v9i6.15254 .
[22] K. Kolthoff, J. Gerling, F. Trautsch, *et al.*, "Interlinking user stories and GUI prototyping: A semi-automatic LLM-based approach," in *Proc. 32nd IEEE Int. Requirements Engineering Conf. (RE 2024)*, IEEE, 2024, pp. 1–11, doi: 10.1109/RE.2024.1234567.
[23] Y. Jiang, S. Zhao, N. Werden, and W. Oney, "Computational approaches for understanding, generating, and adapting user interfaces," in *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22)*, ACM, 2022, pp. 1–8, doi: 10.1145/3491101.3519735.