Human Interaction and Emerging Technologies (IHIET 2025), Vol. 197, 2025, 47–56
https://doi.org/10.54941/ahfe1006698

Designing With 'Intelligence'? Exploring the Limits of AI-Based UX Tools

Saskia Huemer and Benedikt Salzbrunn

University of Applied Sciences Technikum Wien, Austria

ABSTRACT

Artificial intelligence (AI) is increasingly shaping UX practice, with numerous AI-based tools promising to streamline the UX process and improve efficiency, saving time, money, and resources for users. However, how effectively do these tools support UX professionals in UX design and research? This project explores the current limitations and possibilities of using AI-based tools within an integrated, end-to-end human-centered design process. To find suitable tools for this project, 105 available AI-based tools for UX were identified and reviewed. The tools were evaluated according to key criteria: AI functionality, support for UX methods relevant to the case study, GDPR compliance, availability of a free trial, and positive user reviews. Based on these criteria, 23 tools were selected for further assessment, focusing on their AI capabilities and the transparency of AI-generated results. From this assessment, two tools were chosen for use in the case study. The case study utilized a draft of a university website as a test case. It included an AI-based analysis of university websites, followed by AI-moderated interviews with students. Based on the research findings, an interactive prototype was developed and subsequently tested in an AI-moderated usability study with students. The case study revealed gaps between the different phases of the process, rather than an integrated end-to-end workflow. Bridging these gaps required human intervention at several points, particularly when incorporating research results into the prototype. Furthermore, both the methods applied, and the quality of the results produced by the AI-based tools were found to be inferior to the work of experienced UX professionals. This underscores the current limitations of AI-based tools in fully supporting the human-centered design process.

Keywords: Artificial intelligence, User experience, Usability, AI-moderated interviews, AI-generated prototypes, AI-moderated usability tests

INTRODUCTION

Artificial intelligence (AI) has progressed rapidly in recent years, particularly since the release of ChatGPT in 2022 (Dale, 2024; Reuters Media, 2023). AI is now increasingly shaping professional environments (Pereira *et al.*, 2023; Soulami *et al.*, 2024; Statista, 2025) and is exerting a significant influence on UX practice (Stige *et al.*, 2024; Takaffoli *et al.*, 2024; Zhang *et al.*, *2024*). Providers of AI-based tools claim that these tools enhance the efficiency of UX work and offer potential savings in both time and costs (Topp, 2024; Bowen, 2024; Ahmed, 2024).

Despite the advertising promises of these tools and significant AI advancements, current AI-based UX tools still fail to deliver the anticipated comprehensive support in daily workflows (Sponheim and Brown, 2024). Specifically, domain-specific UX tools, such as those for UX design, remain rarely - if ever - used in practice, unlike general-purpose AI tools like ChatGPT. A key concern lies in the quality of AI-generated outputs and their adequacy for professional requirements (Sponheim *et al.*, 2025). Additionally, deploying AI tools necessitates rethinking existing workflows. Established processes, such as the Human-Centered Design (HCD) framework, may require adaptation to enable meaningful integration of AI tools and ensure they generate tangible value (Budiu, 2025).

This paper investigates the current status and limitations of AI in the human-centered design process by addressing three research questions:

- How valid are the outcomes produced by AI-based tools in UX research and design processes?
- What level and forms of human intervention are necessary to ensure effective use of AI-based tools in UX workflows?
- How do AI-based tools enhance or constrain specific phases of the Human-Centered Design process?

To address these questions, we conducted a comprehensive, criteria-based tool selection process followed by a case study leveraging AI-based tools.

RESEARCH AND SELECTION OF AI-BASED TOOLS

Tool research was conducted on IEEE, Scispace, and via Google. On IEEE and Scispace, scientific publications addressing AI-based tools and their applications were identified. Google was used to find websites of tool providers that, according to their own statements, use or are based on AI technologies, as well as overview lists of AI-based UX tools. The Google search primarily followed the Exhaust and Parallel search strategies (Smith, 2012, p. 14).

As of February 2025, the research yielded a preliminary list of 105 AI-based tools for UX research, UX design, and usability testing. This list was filtered using the following criteria:

- Does the tool explicitly use artificial intelligence?
- Does it support at least one major UX method (e.g. user interview, prototype, usability test)?
- Is it compliant with GDPR, with data storage and processing limited to Europe?
- Is a free trial or pre-purchase testing option available?
- Are there credible positive reviews of the tool?

Tools that failed to meet the requirements or lacked sufficient documentation were excluded, resulting in a shortlist of 23 tools.

The shortlisted tools underwent further analysis focusing on two dimensions:

- AI-driven functionality: The extent to which core features rely on AI.
- Transparency: Whether AI-generated outputs are reproducible and traceable to source data (e.g., via citations or audit trails).

Each tool was rated on a scale from 5 (very good) to 1 (insufficient). A grade of 5 indicated maximal AI functionality and transparency, while a grade of 1 signified no relevant AI functionality or transparency. Tools for which information was missing on the website - for example, regarding the transparency of results - were treated as if they had received a grade of 1.

Table 1 presents an excerpt from the second-stage evaluation. This table summarizes the information according to the defined selection criteria, and only tools that met all initial requirements were included for further analysis.

Table 1: Excerpt from the second-stage evaluation of AI-based tools.

Tool Name	Category	Features	AI-Features	Transparency of Results
Userology	Research, Evaluation	AI-moderated usability tests/ interviews, AI Insights	5	5
Wondering	Research, Evaluation	AI-Interview, AI-analysis, AI-moderated usability tests	5	4
Uizard	Design	AI-generated designs, interactive prototypes	5	n/a
Xelper	Research	AI-Interviews, AI-analysis, reports	4	5
Tellet	Research	AI-Interviews, AI-analysis	4	1*
UX Pilot	Design, Research	AI-generated designs (High-Fidelity, Wireframes)	4	n/a
Kraftful	Research	AI-Interviews	3	1*
UXPin	Design	AI-generated designs, interactive prototypes	3	n/a

Wondering and Userology offer a similar range of AI features. While Userology appears slightly more transparent in reporting their results, Wondering provides a free trial (whereas Userology offers only a demo) and has received more positive reviews. As a result, Wondering (2025) was selected as the top AI-driven UX research tool, and Uizard (2025) as the leading UX design tool for the case study.

CASE STUDY: HUMAN CENTERED-DESIGN PROCESS WITH AI-BASED TOOLS

In our case study, the two AI-based tools were used to recreate the website of the University of Applied Sciences Technikum Wien. The objective was to examine how the human-centered design process can be implemented using AI-based tools, with minimal human intervention.

Analysis of University Websites (ChatGPT)

At the outset of the case study, the existing website of the University of Applied Sciences Technikum Wien and two additional Austrian universities of applied sciences were analysed using ChatGPT-4o. The prompts for ChatGPT were formulated according to the CARE principle (Moran, 2024).

ChatGPT was tasked with analysing the information architecture of the university website. Its output was then manually compared to the actual site structure. The manual review revealed that ChatGPT primarily summarized website content rather than providing a true analysis of navigation and information architecture.

Similarly, the three university websites were analysed together to identify common navigation patterns and available information. ChatGPT's findings were again manually validated against the websites. While the analyses generally matched the actual content, minor inaccuracies were observed. Notably, despite explicit instructions to reference only the specified university websites, ChatGPT occasionally incorporated information from external sources, such as the TU Wien website, potentially affecting the results.

UX-Research With AI-Based Tool (Wondering)

Following the analysis of the existing webpages, the AI study builder in Wondering was used to generate an interview guide for the AI-moderated interviews. A prompt was provided to create three research questions, upon which Wondering's AI generated the interview guide. In addition to the prepared questions, the AI dynamically generated follow-up questions during the interviews based on participants' responses.

AI-moderated interviews were then conducted with high school students nearing graduation, representing a primary user group for a university of applied sciences website. A total of 41 participants completed the AI-moderated interviews. Participants could respond either by typing or by sending a voice message, which was transcribed for inclusion in the interview transcript and subsequent analysis.

Interviews without usable responses were manually excluded, resulting in 20 high-quality interviews for the final analysis. Afterwards, Wondering's AI analysed the transcripts. For each interview section, Wondering's AI generated four to five findings, each supported by citations from the interview transcripts.

Prototyping With AI-Based Tool (Uizard)

To incorporate the findings from the AI-moderated interviews into the prototype, the interview transcripts were provided to ChatGPT to generate a prompt for Uizard. Since Uizard currently does not support uploading documents with research findings or additional context, and the prompt is limited to 300 characters, this approach represented the only AI-based method for integrating the interview results. The prompt was formulated according to the previously described CARE principle (Moran, 2024). Based on this prompt, Uizard's AI generated the interactive prototype.

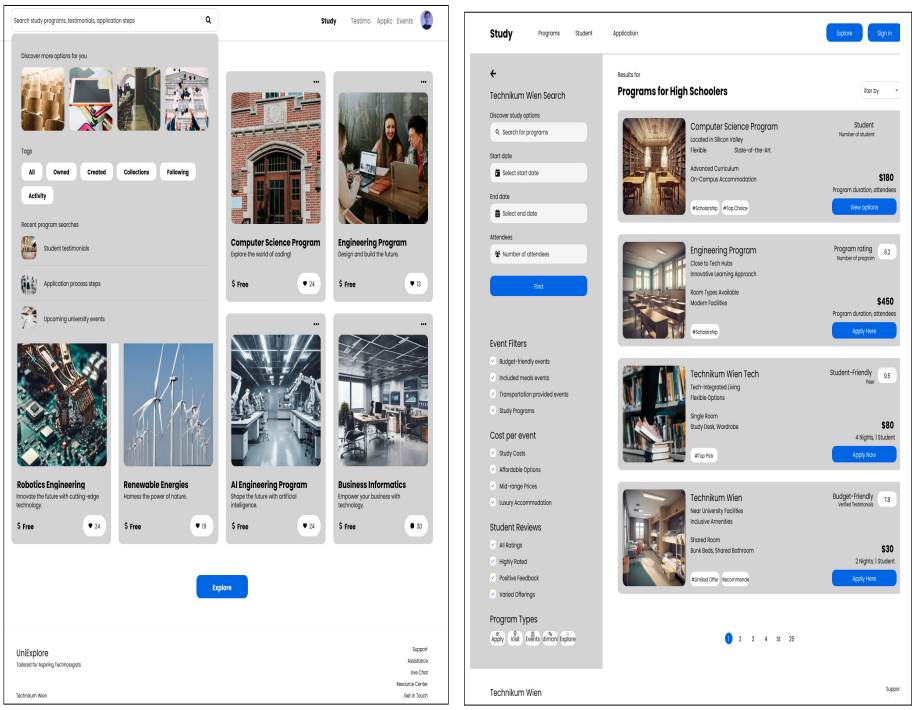

Figure 1: Uizard's first AI-prototype of university website.

To make the prototype testable with the high school students, who had already participated in the AI-moderated interviews, several minor adjustments were made. The prototype texts were adapted to reflect the University of Applied Sciences Technikum Wien by updating the institution's name and aligning the study program names with those offered by the university. These modifications were also performed using Uizard AI, with a separate prompt for each screen.

At the end of the prototyping process, Uizard's Design Review feature was used to evaluate selected screens. The resulting feedback was provided to the Uizard "Autodesigner" to adjust the screens accordingly. Due to technical limitations - specifically, the Design Review feature's restriction to screens with a limited number of elements - not all screens could be reviewed in this manner.

Finally, the prototype was exported from Uizard to Figma, as only Figma prototypes can be tested in Wondering using the Prototype Test feature. Since direct export to Figma was not supported, the Uizard screens were exported as PNG files, converted to SVG files and then imported into Figma.

Usability Test With AI-Based Tool (Wondering)

For the usability test, Wondering AI was again used to create the guide. Based on a prompt, three research questions were generated, which formed the basis for the usability test guide. The guide included participant questions and three prototype tasks. These tasks asked users to find information about the

application process at the University of Applied Sciences, the different study programs, and student opinions from UAS Technikum Wien in the Uizard prototype. Each task was followed by a related follow-up question.

The AI-moderated usability tests involved the same participants (n = 30) as the AI-moderated interviews. During the tasks, participants' screens were recorded to capture their click paths. Additionally, participants were asked to think aloud, and their verbalizations were recorded.

Some usability tests were excluded from analysis due to early termination or lack of task engagement. Ultimately, 18 responses were included in the analysis. Wondering AI generated findings for each usability test section. For the prototype tasks, Wondering also measured three metrics: average task duration, average missed clicks, and task success rate.

RESULTS

The results section is divided into three parts: first, an overview of how AI-based tools can be integrated into the human-centered design process; second, a summary of the practical implementation and required human intervention; and third, an evaluation of the AI tools' effectiveness and quality, based on expert reviews and participant feedback.

AI-Supported Human-Centered Design Process

The human-centered design (HCD) process consists of four main steps: understanding and analysing the context of use, specifying user requirements, producing design solutions, and evaluating the designs (International Organization for Standardization, 2025, p. 21).

In the case study, AI-based tools were applied at each stage. ChatGPT supported the initial analysis of the UAS Technikum Wien website and comparable systems, efficiently processing and summarizing relevant information. User requirements were then explored through AI-moderated interviews with high school students using Wondering. For the design phase, Uizard was used to generate an interactive prototype, with key research findings summarized and transferred using ChatGPT due to limited integration options. Finally, the prototype was evaluated through AI-moderated usability testing with Wondering.

Overall, the structure of the HCD process remained largely unchanged when using AI-based tools. The main challenge was the manual transfer of results between process steps, highlighting the current lack of seamless integration between different AI tools.

Necessary Human Intervention in the AI-Supported Human-Centered Design Process

The case study revealed that, due to a lack of integration between the AI-based tools, human intervention was required at several key stages of the human-centered design process (see Figure 2). At the outset, a prompt had to be manually formulated for the evaluation of the existing website using generative AI. For the AI-moderated interviews, not only was a prompt needed to generate the interview guide, but it was also necessary to

manually review and exclude interviews that were incomplete or unsuitable for analysis.

After the interviews, handing over the research findings to the prototyping tool required a workaround: since direct transfer was not possible, generative AI (ChatGPT) was used to condense the requirements into a prompt for the AI-based prototyping tool. The creation of the AI-generated prototype also involved several manual steps, including exporting the prototype to Figma, importing screens as SVG files, and reconnecting all screens to prepare for usability testing. Integrating the prototype into the usability test and defining success screens for each task also required manual intervention.

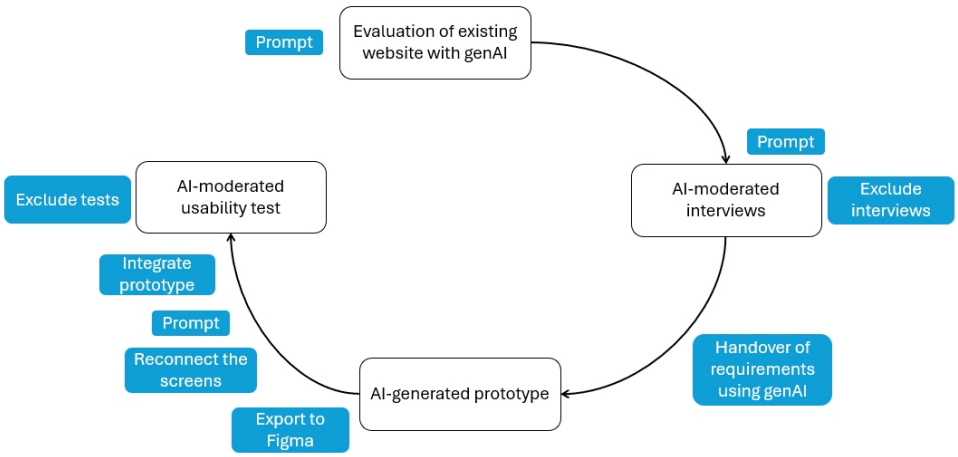

Figure 2: Necessary human intervention in the AI-supported human-centered design process.

Finally, following the AI-moderated usability tests, all test sessions had to be manually reviewed, and low-quality or incomplete tests excluded before the AI could analyse the results. These repeated points of human intervention highlight current limitations in the seamless application of AI-based tools throughout the human-centered design process.

Evaluation of AI-Based UX Tools

In the initial phase of evaluating the UX research tool, participants (n = 41) rated their experience with the AI-moderated interviews. The responses were predominantly critical, with most participants expressing dissatisfaction and only a minority reporting a positive experience. The AI analysis of follow-up questions revealed that participants were particularly frustrated by the frequent repetition of questions, which also reduced their motivation. While a few appreciated being able to respond at their own pace and feeling acknowledged, there was a general desire for greater variety and clearer structure in the interview questions.

The method execution of the UX research tool Wondering, used to design and conduct interviews and usability tests, currently remains unsuitable for practical application. The interview questions are overly generic and

imprecise, with inconsistent terminology usage. Similarly, the usability test tasks lack specificity and clear endpoints, making it difficult to derive meaningful usability insights. Consequently, the results and AI-generated findings are superficial and provide limited novel insights. These conclusions were validated through an expert review involving a senior methodology, psychology, and social sciences specialist with 20+ years of experience.

A similar pattern emerged with the initial Uizard prototype created during the case study. The prototype was functionally inadequate, violating multiple Gestalt principles, containing numerous UI errors (such as poorly chosen colour schemes), and failing to align with the institutional identity of a university of applied sciences. The prototype was also clearly identifiable as AI-generated from an expert perspective, further undermining its practical utility. This assessment was corroborated by a lead UX designer with 10+ years of experience in managing complex design projects.

Subsequent tests with additional prototyping tools, UX Pilot and UXPin, yielded similarly limited results. Despite their advanced features, both tools produced prototypes that suffered from comparable shortcomings in usability, design quality, and integration.

DISCUSSION

The purportedly ready-to-use AI for UX tools currently available are, in practice, only suitable to a very limited extent. Methods are executed imprecisely, and the quality of the results is insufficient - this applies to both the research and design aspects. As demonstrated in the case study, interview guides and usability tasks generated by AI tools such as Wondering were overly generic, lacked methodological rigor, and produced only superficial insights. Similarly, the design prototypes created with tools like Uizard exhibited significant usability and design flaws, including violations of Gestalt principles and poor alignment with institutional requirements.

Another unresolved issue concerns the target audience for such tools. For professionals, these tools are overly restrictive and prescriptive, limiting expert judgment and flexibility. For laypersons, while the tools may appear accessible, the methodological shortcomings identified in this study make them unsuitable for reliable UX work.

How reliable and valid are the outcomes produced by AI-based tools in UX research and design processes? Current AI-based tools deliver results that are often generic, superficial, and lack the methodological rigor required for reliable and valid UX research and design outcomes. **What level and forms of human intervention are necessary to ensure effective use of AI-based tools in UX workflows?** Extensive human intervention is required at multiple stages—including prompt formulation, quality control, data transfer, and expert review—to compensate for the limitations and lack of integration in current AI-based tools. **How do AI-based tools enhance or constrain specific phases of the Human-Centered Design process?** While AI-based tools can accelerate and partially automate tasks in the HCD process, they are constrained by integration gaps, limited methodological depth, and quality issues, thus requiring significant human oversight in every phase.

Limitations

The selection of tools for this study was conducted with the utmost diligence, thorough research, and professional judgment to ensure a representative and comprehensive evaluation. However, it is acknowledged that other AI-based UX tools not included in this study may offer higher quality or better integration capabilities. Furthermore, it is important to emphasize that AI, in principle, can and should serve as a valuable complement for UX professionals. When used by experienced practitioners who understand where AI can enhance efficiency or foster creativity, AI tools can be highly beneficial.

Future Aspects

It should be noted that our research and tool evaluation were conducted in February 2025. Given the rapid pace of AI development, it is reasonable to assume that many of these tools have already improved since then. We look forward with interest to observing how future iterations of AI-based UX tools will evolve and the extent to which they will be able to meet the demands of professional practice.

ACKNOWLEDGMENT

Thank you to Axel from Wondering for providing a license for our study.

REFERENCES

Ahmed, T. (2024), "The Future of Product Discovery: AI-Driven Workflows and AI-Enhanced Creativity | Interview with Tash Ahmed, User Evaluation", available at: https://www.brilliantexperience.com/blog/the-future-of-product-discovery-ai-driven-workflows-and-ai-enhanced-creativity-interview-with-tash-ahmed-user-evaluation (accessed 27 April 2025).

Bowen, J. (2024), "AI Tools That Merge Research Speed with Human Insight | Interview with Jack Bowen, CoLoop", available at: https://www.brilliantexperience.com/blog/ai-tools-that-merge-research-speed-with-human-insight-interview-with-jack-bowen-coloop (accessed 27 April 2025).

Budiu, R. (2025), "Why I'm Not Worried About My UX Job in the Era of AI", available at: https://www.nngroup.com/articles/ux-job-with-ai/ (accessed 16 May 2025).

Dale, R. (2024), "A year's a long time in generative AI", *Natural Language Engineering*, Vol. 30 No. 1, pp. 201–213.

International Organization for Standardization (2025), *ISO 9241 Ergonomics of human-system interaction Part 210: Human-centred design for interactive systems*.

Moran, K. (2024), "CARE: Structure for Crafting AI Prompts", available at: https://www.nngroup.com/articles/careful-prompts/ (accessed 8 March 2025).

Pereira, V., Hadjielias, E., Christofi, M. and Vrontis, D. (2023), "A systematic literature review on the impact of artificial intelligence on workplace outcomes: A multi-process perspective", *Human Resource Management Review*, Vol. 33 No. 1, p. 100857.

Reuters Media (2023), "ChatGPT sets record for fastest-growing user base - analyst note", available at: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/ (accessed 27 April 2025).

Smith, A. G. (2012), "Internet search tactics", *Online Information Review*, Vol. 36 No. 1, pp. 7–20.

Soulami, M., Benchekroun, S. and Galiulina, A. (2024), "Exploring how AI adoption in the workplace affects employees: A bibliometric and systematic review", *Frontiers in artificial intelligence*, Vol. 7.

Sponheim, C. and Brown, M. (2024), "AI UX-Design Tools Are Not Ready for Primetime: Status Update", available at: https://www.nngroup.com/articles/ai-design-tools-not-ready/ (accessed 28 April 2025).

Sponheim, C., Brown, M. and Dykes, T. (2025), "AI Design Tools Are Marginally Better: Status Update", available at: https://www.nngroup.com/articles/ai-design-tools-update-2/ (accessed 16 May 2025).

Statista (2025), "Künstliche Intelligenz in der Arbeitswelt", available at: https://de.statista.com/themen/13316/ki-in-der-arbeitswelt/ (accessed 27 April 2025).

Stige, Å., Zamani, E. D., Mikalef, P. and Zhu, Y. (2024), "Artificial intelligence (AI) for user experience (UX) design: A systematic literature review and future research agenda", *Information Technology & People*, Vol. 37 No. 6, pp. 2324–2352.

Takaffoli, M., Li, S. and Mäkelä, V. (2024), "Generative AI in User Experience Design and Research: How Do UX Practitioners, Teams, and Companies Use GenAI in Industry?", in Vallgårda, A., Jönsson, L., Fritsch, J., Alaoui, S. F. and Le Dantec, C. A. (Eds.), *Designing Interactive Systems Conference, 01 07 2024 05 07 2024, Copenhagen Denmark*, ACM, New York, NY, USA, pp. 1579–1593.

Topp, K. (2024), "Get Insights in Hours and Virtual Consumer Feedback | Interview with Kathryn Topp, Yabble", Available at: https://www.brilliantexperience.com/blog/get-insights-in-hours-and-virtual-consumer-feedback-interview-with-kathryn-topp-yabble (accessed 27 April 2025).

Zhang, Y., Atiq, A. and Chow, W., "Exploring the Role of AI in UX Research", in Cochrane, Narayan, Bone, Deneen, Saligari, Tregloan and Vanderburg (Eds.), *Navigating the Terrain: Emerging frontiers in learning spaces, pedagogies, and technologies: ASCILITE 2024 Conference Proceedings*, pp. 556–560.