LLMs and Diffusion Models in UI/UX: Advancing Human-Computer Interaction and Design

Layla Sun[1], Mengmeng Qin[2] Benji Peng[*, 3],
[1]Simon Fraser University, Canada
[2]City University of Hong Kong, PRC
[3]AppCubic, USA
*Corresponding Email: benji@appcubic.com

Index Terms—**Large Language Models, Human Computer Interaction, Generative Artificial Intelligence, UI/UX, Human Machine Interaction**

Abstract—**The rapid advancements in Generative AI, particularly Large Language Models (LLMs) and Diffusion Models, are transforming UI/UX design and human-computer interaction (HCI). This article explores recent applications of these technologies to augment and automate key stages of the design process, from ideation and prototyping to code generation. By utilizing their strengths in natural language understanding and content generation, LLMs serve as tools for ideation, design enhancement, and as integral components of user interfaces—enabling conversational systems, adaptive UIs, and task automation. Diffusion models, in contrast, focus on generating visual content and assisting in UI prototyping, creating new possibilities for design workflows. Despite these advances, challenges remain, such as maintaining output quality, integrating AI into existing workflows, and addressing ethical issues like data bias and transparency. This article highlights the need to balance human and AI contributions, foster effective human-AI collaboration, and establish robust evaluation criteria. Future research should explore multimodal LLMs, improve transparency and explainability, and democratize access to the design process. The integration of Generative AI into UI/UX design holds significant potential to advance HCI but requires careful consideration of its limitations and societal impacts.**

I. Introduction

A. The Rise of Generative AI

Over the past few years, Generative Artificial Intelligence (GenAI) has gained significant traction across a wide range of fields, propelled by advancements in deep learning architectures and the availability of large datasets [1], [2]. GenAI, represented by models such as ChatGPT [3], [4], Gemini [5], [6], Claude [7], Imagen [8], DALL-E [9], Stable Diffusion [10], and Flux, has demonstrated extraordinary capabilities in producing high-quality content across diverse modalities—text, images, audio, code, and even 3D models [11], [12]. These models, trained on extensive datasets, push the boundaries of creativity and realism in AI-generated content [13].

The accessibility of these models, via user-friendly interfaces and APIs, has democratized AI, empowering individuals and organizations, even with limited technical expertise, to leverage its capabilities [14]. Notable examples of such democratization efforts include tools like Chat-Rec, which facilitates the development of conversational recommendation systems [15], and AutoGen, which supports the creation of LLM applications through multi-agent dialogues [16].

GenAI has seen rapid adoption across multiple industries, including healthcare [17]–[21], architecture [22], and manufacturing [23]. By streamlining workflows and enabling innovation, GenAI is transforming traditional practices across these sectors [24]–[26].

B. The Intersection of AI and UI/UX

The convergence of GenAI and UI/UX design presents transformative opportunities for human-computer interaction [27], [28]. Leveraging their capabilities in understanding and generating text and visuals, AI models can enhance creativity, automate routine tasks, and enable more personalized and engaging user experiences [29], [30]. Imagine a scenario where users describe their desired interface in natural language, and an AI system translates those descriptions into functional prototypes or even complete code [16], [31]. Such advancements could democratize design by empowering non-experts to create their own custom interfaces while freeing designers to focus on higher-level challenges [32].

AI also offers the capability to analyze user interactions and preferences, allowing dynamic adaptation of the interface to provide a more intuitive and personalized experience [33]. For instance, the I-Design system [34] utilizes LLMs to generate personalized 3D interior designs based on user input, showcasing AI's ability to align design solutions with user preferences. Similarly, the study ”Enabling Conversational Interaction with Mobile UI using Large Language Models” [35] explores how LLMs can facilitate conversational interaction with mobile UIs, illustrating the potential for AI to fundamentally reshape user-device interactions [36].

C. Scope and Objectives

This article focuses on the recent advancements and applications of LLMs and Diffusion Models in UI/UX design and human-computer interaction (HCI) [28] [27]. It inspects how they are used to enhance the design processes at different statges, from ideation and prototyping to code generation [29], [30]. It examines the human-AI collaboration in design, analyzing how these technologies are shifting the role of the designer and shaping the future of user experiences [31], [32]. It critically assess the limitations of current applications, including challenges related to accuracy, bias, and the need for robust evaluation frameworks [33], [36]. It also outlines the promising future where everyone benefits from GenAI in UI/UX.

More specifically, this article attempts address the following key research questions:

- How are LLMs and Diffusion Models used to augment and automate different stages of the UI/UX design process?
- What are the benefits and limitations of using such technology during UI/UX design, in terms of creativity, efficiency, usability, and user experience?
- How does GenAI reshape the collaboration between human designers and automated systems?
- What are the ethical concerns when we use GenAI, and how can we ensure responsible development and application of these technologies?
- What is the potential outlook of GenAI in UI/UX design and HCI?

To provide an overview of how LLMs and Diffusion Models are transforming UI/UX design and HCI, we have developed a taxonomy illustrating the key areas of application and research (see **Figure 1**). This taxonomy categorizes the major themes explored in this article, including the use of LLMs as design tools, their integration as components within user interfaces, the application of Diffusion Models in design workflows, and human-AI collaboration.

II. LLMs as Tools for UI/UX Design

LLMs are changing UI/UX design, offering substantial enhancements to various stages of the design process. Their ability to generate natural language and reason through problems makes them valuable tools for designers aiming to boost creativity and efficiency.

A. Idea Generation and Exploration

LLMs are highly effective in supporting ideation, brainstorming, and exploring design concepts. By using their extensive knowledge base and generative capabilities, LLMs help designers overcome conventional thinking and explore a broader range of design possibilities [60]. This is particularly useful during the early stages of the design process, where designers seek inspiration, generate initial concepts, and explore different directions.

1) Overcoming Design Fixation: Design fixation, where designers become attached to a particular approach, limits creativity. LLMs can help break this fixation by generating alternative concepts from initial ideas or constraints [37]. Designers can explore these AI-generated suggestions to challenge their assumptions, consider new directions, and potentially develop more user-centered designs. In 3D design, tools like 3DALL-E can provide reference images to inspire fresh perspectives and prevent fixation [37].

2) Supporting Divergent and Convergent Thinking: The design process requires both divergent thinking (generating many ideas) and convergent thinking (refining and selecting ideas). LLMs support both. In the divergent phase, they can expand on existing ideas, generate variations, and explore different directions [38]. For instance, an LLM could help create a list of features for a mobile app or suggest visual styles for a website. In collaborative brainstorming, LLMs provide diverse perspectives, enriching the creative pool [61]. During the convergent phase, LLMs can help evaluate and prioritize ideas based on criteria like usability or aesthetics, such as ranking design concepts based on user engagement potential or analyzing feedback to find common themes [62].

B. Prototyping and Design Refinement

Prototyping is crucial for testing and iterating on designs before moving into development. LLMs enhance this stage by automating mockup generation, offering feedback, and simulating user interactions.

1) Automating Mockup Generation: LLMs can generate UI mockups from textual descriptions or user inputs [51]. Tools like PromptInfuser, a Figma plugin, allow designers to link UI elements to LLM prompts, enabling the creation of semi-functional mockups [39]. This reduces the time and effort needed for prototyping, allowing rapid testing of different concepts with users.

2) LLMs as Design Critics: LLMs can also serve as virtual design critics by providing feedback based on established design principles [40]. They can analyze visual elements to identify usability issues, suggest improvements, and provide justifications for these suggestions. This can be helpful for novice designers or those tackling unfamiliar challenges. However, ensuring the reliability and relevance of feedback is still an ongoing research area [58].

C. Automating Code Generation

LLMs have shown proficiency in code generation, potentially accelerating software development, including UI creation. Leveraging their understanding of programming languages and design patterns, LLMs can generate code snippets, complete functions, and even build entire UI components.

1) LLMs for Front-End Development: LLMs are increasingly used to generate code for user interfaces, including HTML, CSS, and JavaScript. Tools like Design2Code uses multimodal LLMs to convert visual designs directly into code [41]. This allows designers to concentrate on the visual aspects while letting AI handle technical implementations.

2) LLMs for Back-End Development: LLMs can also generate server-side logic and database interactions. Although this area is still developing, LLMs might automate the creation of APIs, database schemas, and other back-end components, simplifying complex software systems. However, further research is required to fully assess the capabilities and limitations of LLMs for back-end development [42].

3) Challenges in Code Quality and Accuracy: Despite their impressive abilities, LLMs can generate code that contains errors, inconsistencies, or security vulnerabilities. Ensuring adherence to coding standards is another significant challenge [43], [63]. Careful code review and comprehensive test cases

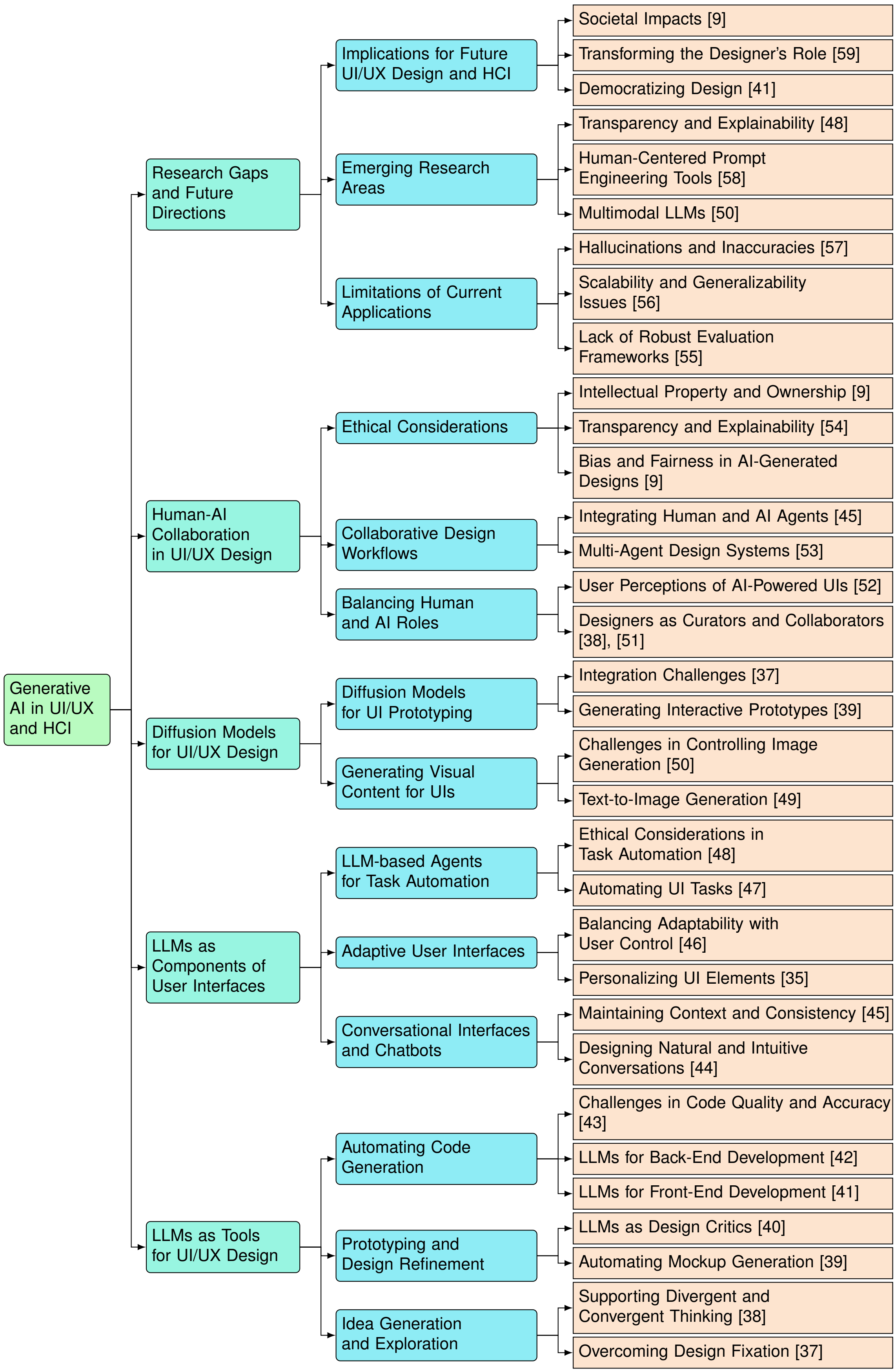

Fig. 1. Taxonomy of LLMs and Diffusion Models in UI/UX Design and HCI

are essential before deploying LLM-generated code in production environments. Active involvement of human developers in code review and testing is crucial to maintain quality and safety.

III. LLMs as Components of User Interfaces

Beyond serving as design tools, Large Language Models (LLMs) are increasingly being integrated as core components of user interfaces, enabling innovative approaches to human-computer interaction. By utilizing their capabilities in natural language understanding and generation, LLMs can transform static interfaces into more dynamic, responsive systems that enable human-like interactions, adapt to user preferences, and automate complex tasks. This section explores how LLMs are employed to create more intuitive, efficient, and personalized user experiences, while also addressing the challenges and ethical considerations that arise from their use.

A. Conversational Interfaces and Chatbots

The advancement of LLMs has significantly fueled the development of conversational interfaces and chatbots, enabling users to interact with systems through natural language [35]. These interfaces use LLMs to understand user intent, generate coherent responses, and handle multi-turn conversations, providing a more natural user experience. For example, the Chat-Rec system [15] uses LLMs to create an interactive conversational recommender system, enhancing user engagement with recommendations. This shift toward conversational interaction can make technology more accessible to a broader audience, including those less familiar with traditional graphical interfaces. However, designing effective conversational systems presents challenges, particularly in maintaining context and ensuring consistency across interactions.

Designing natural and intuitive conversational interfaces requires careful consideration of principles that support effective communication between humans and AI. Key factors include understanding user intent, providing clear responses, maintaining a consistent persona, and handling unexpected inputs gracefully [44]. The study by Dang et al. [60] identifies four design goals for interfaces that support prompting: reducing ambiguity, offering feedback on prompt effectiveness, enabling prompt variations, and supporting iterative refinement. Balancing structure with flexibility is another challenge. Too much structure can make interactions rigid, while too little can hinder task completion [64]. Error handling is also critical, requiring clear user feedback, suggestions for correction, and easy ways for users to backtrack or restart the conversation.

Maintaining context and consistency across multiple conversational turns remains a significant challenge for LLM-powered chatbots. LLMs can struggle to retain information from earlier interactions, leading to inconsistent or irrelevant responses. This issue is discussed by Zamfirescu-Pereira et al. [45], who highlight the difficulty of making LLMs respond with "I don't know" when faced with questions beyond their knowledge. Failing to do so often results in inappropriate responses, leading to user frustration. Effective context management requires memory mechanisms, goal tracking, and integrating dialogue history into the LLM's input. Maintaining a consistent persona throughout an interaction is also vital, even as topics change.

B. Adaptive User Interfaces

LLMs are also deployed to create adaptive user interfaces that adjust to user preferences, behaviors, and contexts. By analyzing interactions and learning from feedback, LLMs can personalize the UI, making it more intuitive and efficient for individual users. This adaptability is particularly beneficial for users with diverse needs, as the interface can be tailored to their specific requirements. However, it is important to balance adaptability with user control, ensuring that users feel empowered rather than overwhelmed by the system's capabilities.

LLMs can personalize various UI elements, including layout, color schemes, content recommendations, and micro-interactions. For instance, an LLM could learn a user's preferred color palette and adjust the UI accordingly, or analyze browsing history to provide personalized content suggestions. Wang et al. [35] explored the use of LLMs to personalize mobile UI experiences, generating questions, summaries, and answers based on screen context. While such personalization can create more engaging experiences, it is essential to respect user privacy and avoid unintended consequences like filter bubbles.

Balancing adaptability with user control is a key challenge in designing adaptive UIs. Users must be aware of how the system adapts to their behavior and should have control over these adaptations. Transparency is crucial for building trust, allowing users to understand and manage system behavior. Moreover, unintended consequences, such as limiting exposure to diverse content through personalization, need to be mitigated. Research by Jiang et al. [46] on the ability of LLMs to simulate personality traits raises concerns about the influence these models could have on user behavior, emphasizing the importance of ethical considerations in adaptive systems.

C. LLM-based Agents for Task Automation

LLM-powered agents automate tasks and workflows within user interfaces, reducing cognitive load and making complex tasks more manageable. These agents serve as intelligent assistants, guiding users through processes, offering personalized recommendations, and anticipating whatbuser needs. This can be particularly useful for new users or those struggling with complex tasks, but it also raises ethical questions regarding transparency, AI agency, and unintended consequences.

LLMs can automate a range of repetitive or complex UI tasks, such as form filling, data entry, and navigation. For instance, an LLM-powered agent could automatically complete a user's shipping address based on prior orders or assist in navigating a multi-step process like setting up an account or making a purchase [47]. The ChatDev system [65] illustrates the use of AI agents collaborating to complete software development tasks, demonstrating the potential of LLM-powered agents in automating complex workflows.

The increasing use of LLM-based agents necessitates addressing ethical considerations, including transparency and avoiding bias. As these agents become more capable of making decisions that impact users, their actions must be explainable and aligned with human values. Rastogi et al. [48] explore using LLMs to audit other LLMs for biases, underlining the importance of human oversight in ethical AI design. Developing mechanisms to mitigate bias and ensure that AI agents act ethically is essential for their responsible use.

IV. Diffusion Models for UI/UX Design

Diffusion models have become a powerful tool for generating high-quality images from textual descriptions, opening new avenues for UI/UX design. This section examines the application of diffusion models in creating visual content for user interfaces and facilitating prototyping, while addressing the challenges of controlling image generation and integrating these models into existing design workflows.

A. Generating Visual Content for UIs

Diffusion models can produce a wide array of visual elements for user interfaces, including icons, illustrations, and design mockups. This capability enhances creativity and streamlines design workflows by reducing the manual effort required to create these assets.

Text-to-image diffusion models offer designers a unique method to visualize ideas directly from textual descriptions. This approach helps explore diverse design concepts and overcome design fixation by generating unexpected visual outputs. In architectural design, researchers have demonstrated the use of diffusion models to produce massing models and facade designs from textual prompts, incorporating daylight-driven strategies to improve visual renderings [49]. Similarly, diffusion-based AI art platforms like Midjourney, DALL-E 2, and Stable Diffusion have been applied to interior and exterior design workflows [22]. In fashion design, these models are used to generate new clothing designs and enable virtual try-on experiences, allowing users to visualize different styles [66].

Beyond inspiration, diffusion models can generate specific UI elements, such as buttons, icons, and layout grids, significantly reducing design time. Researchers have explored using diffusion models to generate mobile GUI designs from textual descriptions, demonstrating their potential for faster and more cost-effective prototyping [67]. This approach harnesses the ability of diffusion models to generate diverse images from text prompts, offering an alternative to traditional manual design processes.

Despite these possibilities, controlling the output of diffusion models to align with specific design intent remains challenging. The inherent stochastic nature of these models can result in unpredictable outcomes, complicating efforts to ensure visual coherence across multiple generated images. In fashion design, for example, diffusion models may generate novel yet inconsistent outputs that do not fully reflect user preferences or domain-specific needs [38]. Translating abstract design ideas into precise textual prompts is often difficult, especially for complex designs. Users also have limited control over aspects like pose and structural features, which presents further challenges in guiding the model's output [50].

B. Diffusion Models for UI Prototyping

Diffusion models can be used to create interactive prototypes and early design concepts, which can then be tested and evaluated by users. By combining diffusion models with web development frameworks, designers can generate semi-functional mockups that include interactive UI elements. These prototypes allow for valuable feedback on the usability and appeal of designs before substantial development resources are committed.

Researchers have developed tools like the Figma plugin PromptInfuser, which enables designers to create interactive mockups by linking UI elements to prompts [68]. This integration of UI design and AI-driven generation allows for rapid iteration on both visual and interactive components, resulting in more effective user-centered prototypes.

Diffusion models also support the exploration of design variations based on the same textual input, helping designers consider a wider range of aesthetic and functional possibilities. For instance, researchers have applied these models to generate different layout options for interior design, incorporating diverse functional strategies and user preferences [69]. Similar techniques have been used for webpage layout generation, producing diverse layouts that adhere to user guidelines [70].

However, the integration of diffusion models into existing UI/UX design tools presents several challenges. Many current tools are optimized for manual design processes, which can make the seamless incorporation of generative models difficult. The computational demands of diffusion models often necessitate specialized hardware or cloud-based services, which may not be accessible to all designers. Additionally, the lack of standardized methods for representing and manipulating outputs within existing workflows creates further barriers to integration. For example, researchers faced technical issues like software crashes and limitations in handling complex designs when trying to integrate text-to-image AI into 3D design workflows [37].

To address these challenges, researchers are developing techniques to improve the controllability and consistency of diffusion models in UI/UX design. One approach involves fine-tuning models on domain-specific datasets to generate more relevant and coherent outputs. Another strategy focuses on advancing prompt engineering techniques to guide the model's output more precisely.

The use of diffusion models in UI/UX design also raises ethical concerns. The potential for these models to generate biased or inappropriate content necessitates careful curation and filtering of outputs. Additionally, the increasing use of AI-generated designs may impact the role of human designers, raising important questions about authorship and originality in creative work.

As diffusion models continue to evolve, their integration into UI/UX design workflows is likely to become more seamless. Future research directions include developing more intuitive interfaces for interacting with these models, enhancing their ability to generate consistent design elements across multiple outputs, and creating hybrid approaches that combine the strengths of AI-generated designs with human creativity and expertise.

V. Human-AI Collaboration in UI/UX Design

The advancements in LLMs and diffusion models have introduced new opportunities for human-AI collaboration in UI/UX design, creating a shift in how design and development are approached. This collaboration necessitates a re-evaluation of the designer's role and careful consideration of ethical implications.

A. Balancing Human and AI Roles

The emergence of GenAI tools requires a new perspective on the designer's role. Rather than replacing designers, AI empowers them to act as curators, collaborators, and strategists, utilizing AI capabilities to enhance their work. This shift involves understanding both the strengths and limitations of these technologies.

LLMs can serve as creative partners by supporting ideation, providing inspiration, and automating repetitive tasks. Studies have shown that LLMs help designers overcome design fixation by generating diverse concepts [51]. Artists view LLMs as tools for automating tasks, expanding ideas, and facilitating communication [38]. In software development, the AI-augmented double diamond framework has been proposed to structure LLM use for tasks like text summarization and idea generation [59].

LLMs can also assist in evaluating design ideas by offering feedback based on predefined criteria, making assessments more objective and enabling better comparisons [71]. For instance, prompt patterns have been used to improve code quality, refactoring, and design evaluations [64]. Maintaining a balance between AI assistance and human creativity is crucial to ensure that LLMs augment, rather than replace, the designer's role.

User perceptions of AI-powered UI/UX designs are also key to successful adoption. Trust, transparency, and control are critical factors. Studies show that users are more likely to trust AI systems if they understand how they work and can control their behavior [52]. Providing transparency in AI-generated design decisions and allowing users control over AI suggestions can foster trust and user empowerment. Research indicates that users feel a stronger sense of ownership when selecting from AI suggestions rather than having text continuously generated [72]. Therefore, interfaces should support user control and agency in interactions with GenAI.

B. Collaborative Design Workflows

Incorporating LLMs and diffusion models into collaborative design workflows requires new approaches to communication and task allocation between human designers and AI agents.

Multi-agent design systems, such as the DesignGPT framework, simulate different design roles with AI agents to facilitate collaboration [53]. These agents interact in natural language, enabling better communication and knowledge sharing. For example, ChatDev, a virtual chat-powered software development company, uses AI agents to simulate different roles within a development workflow [65]. Similarly, combining interactive evolution with LLMs has been used to co-create game designs [73].

However, designing effective multi-agent systems requires careful consideration of agent capabilities, communication protocols, and task allocation strategies to ensure seamless collaboration. Research is needed to determine how to best balance AI capabilities with human expertise, ensuring a smooth and productive collaboration.

Integrating human and AI agents in collaborative workflows presents unique challenges. Achieving the final outcomes can be unpredictable, highlighting the need for robust communication tools to facilitate clear exchanges between humans and AI [45]. Communication breakdowns may arise due to differing styles and reasoning capabilities, and misaligned goals can cause conflicts. Trust issues are also a concern if users find the AI unreliable or opaque. Developing shared mental models, establishing clear communication protocols, and designing transparent interfaces can help address these challenges.

C. Ethics

The use of GenAI in UI/UX design raises ethical concerns that must be addressed to ensure responsible and equitable outcomes.

GenAI models are trained on large datasets that may contain biases, which can be reflected in their outputs. This can lead to UI/UX designs that perpetuate societal inequalities [9]. For example, a text-to-image model trained on a dataset biased toward Western fashion aesthetics may produce culturally insensitive designs. Addressing data bias requires careful curation of training data and techniques to mitigate bias in model outputs.

Transparency and explainability are essential for users to understand how AI-generated designs are produced. When users understand the rationale behind AI decisions, they are more likely to trust and accept the results. This is especially important in sensitive fields like healthcare [54]. Techniques such as saliency maps and counterfactual explanations can provide insights into AI reasoning, although developing effective, understandable explanations remains a challenge.

The issue of intellectual property rights and ownership of AI-generated designs is complex. As AI becomes more capable of creating novel content, questions about ownership arise: does the copyright belong to the AI developer, the user who provided prompts, or the AI itself? This uncertainty requires new legal and ethical frameworks to clarify ownership and ensure fair attribution [9].

VI. Research Gaps and Future Directions

While advancements in LLMs and diffusion models have shown significant promise for UI/UX design, several research gaps and challenges remain. These limitations hinder the full realization of their potential and raise critical questions about the future of human-AI collaboration in design.

A. Limitations of Current Applications

Despite rapid progress in GenAI, current applications in UI/UX design face several limitations that need further research and development. These challenges stem from the inherent nature of these models, their training data, and the design of current interfaces.

One major challenge is the lack of robust evaluation frameworks for AI-generated UI/UX designs. Traditional usability testing methods are often insufficient for capturing the unique characteristics of AI-generated content [55]. Evaluating the creativity and novelty of AI-generated designs requires new metrics that go beyond traditional usability measures [64]. There is also a need for standardized benchmarks and datasets to enable fair comparisons between different AI models, similar to the RTLLM benchmark proposed by [42]. Without such frameworks, it is difficult to assess the true impact of GenAI in design, as seen in studies like [74].

Another key limitation is the scalability and generalizability of current LLM and diffusion model applications. While these models perform well on specific tasks and datasets, such as generating code for targeted development tasks [3], their performance often degrades with more complex or diverse design problems [56]. This limitation is partly due to the reliance on large amounts of training data, which may not be available across all design domains. Additionally, the high computational costs of training and deploying these models can be prohibitive for smaller organizations. Addressing these challenges requires more efficient training methods and exploring techniques for transferring knowledge across design domains, as seen in cross-domain recommendation research by [15].

Another significant concern is the tendency of LLMs and diffusion models to produce inaccurate or nonsensical outputs, often referred to as ”hallucinations” [57], [75]. These hallucinations arise from biases in the training data, ambiguous prompts, or limitations in the model's understanding. In the context of UI/UX design, such inaccuracies can lead to unusable or harmful interfaces [76]. Addressing this issue requires new techniques for detecting and mitigating hallucinations, such as incorporating human feedback loops, improving data quality, and developing more robust evaluation methods that can identify inaccuracies [50].

B. Emerging Research Areas

The usage of GenAI in UI/UX design is rapidly evolving, with several promising research areas emerging that could significantly shape the future of human-computer interaction.

Multimodal LLMs, which can process and generate content across multiple modalities such as text, image, and code, open up new possibilities in UI/UX design [50]. These models could enable the seamless integration of different design elements, allowing designers to craft holistic user experiences. As an example, a multimodal LLM could generate a website layout, relevant images, and the underlying code, similar to the capabilities explored in [16]. However, developing such models requires overcoming challenges in data alignment, model training, and ensuring consistency across modalities.

Another emerging area is human-centered prompt engineering tools. Crafting effective prompts can be challenging for non-experts, requiring an understanding of the model's capabilities and limitations [58]. Developing intuitive prompt engineering tools is crucial for democratizing the use of GenAI in design. Such tools could provide interactive guidance, suggest relevant prompts, and visualize the design space, helping users understand the impact of different prompts on generated outputs. Research in this area should focus on making these tools accessible and user-friendly, as suggested by design guidelines for leveraging LTGMs [38].

Transparency and explainability are also crucial for building trust in AI systems and empowering users to control the design process [48]. Current GenAI models are often black boxes, making it difficult for designers to understand their decision-making processes. Developing transparent AI systems involves creating methods for visualizing AI decisions, providing explanations for generated designs, and enabling user interaction to refine AI understanding. Techniques from Explainable AI (XAI), like those suggested for skill estimation in code generation [43], could be adapted to enhance transparency in UI/UX design.

C. Implications for the Future of UI/UX Design and HCI

GenAI in UI/UX design has significant implications for the future of the field, potentially democratizing design, transforming the role of the designer, and affecting broader societal aspects.

GenAI tools could democratize design by making it accessible to non-experts, allowing them to create and customize UI/UX designs [41]. This could lead to a more diverse design landscape, with users having more control over their digital experiences. However, ensuring that these tools are accessible to users with varying levels of technical expertise is essential. Intuitive interfaces, clear instructions, and addressing potential biases in AI models are crucial to achieving this goal. Studies like [35] illustrate the potential of LLMs to enable conversational interaction, further empowering non-expert users.

The rise of GenAI is also transforming the designer's role, requiring new skills and collaborative workflows [59]. Designers must learn to effectively interact with AI systems, craft prompts, and evaluate AI-generated outputs. This shift involves embracing AI as a creative partner, as illustrated by ChatGPT's ability to perform design-related tasks [44].

Lastly, the societal impacts of AI-generated UIs extend beyond the design community, raising ethical, economic, and creative concerns [9]. Biases in AI models could lead to discriminatory designs, and automating design tasks might displace human designers. Highly personalized user experiences also bring privacy and data security concerns. Addressing these issues requires a multi-disciplinary approach involving ethicists, policymakers, and the public to ensure responsible AI use. Future research should focus on ethical guidelines for AI-powered design and mechanisms for accountability and transparency in AI deployment.

REFERENCES

[1] Y. Zhang, C. Liu, W. Ding, *et al.*, "A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt," *arXiv preprint arXiv:2303.04226*, 2023.

[2] X. Wang, Y. Xu, Y. Zhu, *et al.*, "Generative ai for synthetic data across multiple medical modalities: A systematic review of recent developments and challenges," *arXiv preprint arXiv:2407.00116*, 2024.

[3] Y. Shen, K. Song, X. Tan, D. Li, W. Lu, and Y. Zhuang, "HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging F ace," *Neural Information Processing Systems*, 2023.

[4] T. Wu, S. He, J. Liu, S. Sun, K. Liu, Q.-L. Han, and Y. Tang, "A brief overview of chatgpt: The history, status quo and potential future development," *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 5, pp. 1122–1136, 2023.

[5] G. Team, R. Anil, S. Borgeaud, Y. Wu, J.-B. Alayrac, J. Yu, R. Soricut, J. Schalkwyk, A. M. Dai, A. Hauth, *et al.*, "Gemini: a family of highly capable multimodal models," *arXiv preprint arXiv:2312.11805*, 2023.

[6] M. Reid, N. Savinov, D. Teplyashin, D. Lepikhin, T. Lillicrap, J.-b. Alayrac, R. Soricut, A. Lazaridou, O. Firat, J. Schrittwieser, *et al.*, "Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context," *arXiv preprint arXiv:2403.05530*, 2024.

[7] A. Anthropic, "The claude 3 model family: Opus, sonnet, haiku," *Claude-3 Model Card*, vol. 1, 2024.

[8] C. Saharia, W. Chan, S. Saxena, L. Li, J. Whang, E. L. Denton, K. Ghasemipour, R. Gontijo Lopes, B. Karagol Ayan, T. Salimans, *et al.*, "Photorealistic text-to-image diffusion models with deep language understanding," *Advances in neural information processing systems*, vol. 35, pp. 36479–36494, 2022.

[9] Z. Epstein, A. Hertzmann, M. Akten, H. Farid, J. Fjeld, M. R. Frank, M. Groh, L. Herman, N. Leach, R. Mahari, O. Pentland, H. Schroeder, and A. Smith, "Art and the science of generative AI," *Science*, vol. 380, pp. 1110–1111, June 2023.

[10] R. Gozalo-Brizuela and E. C. Garrido-Merchan, "ChatGPT is not all you need. A State of the Art Review of large Genera tive AI models," *arXiv.org*, 2023.

[11] H. Liu, Y. Li, Y. Wang, *et al.*, "A survey on generative diffusion models," *arXiv preprint arXiv:2209.02646*, 2023.

[12] X. Li, F. Tao, W. Ye, A. Nassehi, and J. W. Sutherland, "Generative manufacturing systems using diffusion models and ChatGPT," *arXiv.org*, 2024.

[13] Y. Wang, H. Liu, Y. Li, *et al.*, "The road ahead: Emerging trends, unresolved issues, and concluding remarks in generative ai—a comprehensive review," 2024.

[14] X. Zhu, P. Ghamisi, B. Runkle, *et al.*, "Generative ai for geoscience: Foundations, trends, and future challenges," *arXiv preprint arXiv:2402.03349*, 2024.

[15] Y. Gao, T. Sheng, Y. Xiang, Y. Xiong, H. Wang, and J. Zhang, "Chat-REC: Towards Interactive and Explainable LLMs-Augmented Recommend er System," *arXiv.org*, 2023.

[16] Q. Wu, G. Bansal, J. Zhang, Y. Wu, B. Li, E. Zhu, L. Jiang, X. Zhang, S. Zhang, J. Liu, A. H. Awadallah, R. W. White, D. Burger, and C. Wang, "AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversati on," *arXiv.org*, 2023.

[17] B. Meskó, "The Impact of Multimodal Large Language Models on Health Care's Future," *Journal of Medical Internet Research*, vol. 25, p. e52865, Nov. 2023.

[18] Y. Liu, Z. Zhang, Y. Zhang, *et al.*, "Generative ai in healthcare: A survey on large language models," *arXiv preprint arXiv:2310.12724*, 2024.

[19] B. Peng, Z. Bi, P. Feng, Q. Niu, J. Liu, and K. Chen, "Emerging techniques in vision-based human posture detection: Machine learning methods and applications," *Authorea Preprints*, 2024.

[20] Q. Niu, K. Chen, M. Li, P. Feng, Z. Bi, J. Liu, and B. Peng, "From text to multimodality: Exploring the evolution and impact of large language models in medical practice," *arXiv preprint arXiv:2410.01812*, 2024.

[21] Q. Niu, J. Liu, Z. Bi, P. Feng, B. Peng, and K. Chen, "Large language models and cognitive science: A comprehensive review of similarities, differences, and challenges," *arXiv preprint arXiv:2409.02387*, 2024.

[22] J. Ploennigs and M. Berger, "AI art in architecture," *AI in Civil Engineering*, vol. 2, Aug. 2023.

[23] L. Makatura, M. Foshey, B. Wang, F. Hähnlein, P. Ma, B. Deng, M. Tjandrasuwita, A. Spielberg, C. Owens, P. Y. Chen, A. Zhao, A. Zhu, W. Norton, E. Gu, J. Jacob, Y. Li, A. Schulz, and W. Matusik, "How Can Large Language Models Help Humans in Design And Manufacturing? Part 2: Synthesizing an End-To-End LLM-Enabled Design and Manufacturi ng Workflow," *Harvard Data Science Review*, May 2024.

[24] Y. Zhang, C. Liu, and W. Ding, "Emerging trends and impacts: A bibliographic analysis of generative ai in marketing," 2024.

[25] M. Li, K. Chen, Z. Bi, M. Liu, B. Peng, Q. Niu, J. Liu, J. Wang, S. Zhang, X. Pan, *et al.*, "Surveying the mllm landscape: A meta-review of current surveys," *arXiv preprint arXiv:2409.18991*, 2024.

[26] J. Gao, T. Kärkkäinen, and M. Turpeinen, "Promise and challenges of generative ai in healthcare information systems," 2024.

[27] Q. Yang, A. Steinfeld, and J. Zimmerman, "Ai-augmented user interface design: A survey and vision for the future," *ACM Transactions on Computer-Human Interaction*, vol. 30, no. 3, pp. 1–36, 2023.

[28] Z. Xu, Y. Zhang, Y. Li, *et al.*, "Generative ai for ui/ux design: A comprehensive survey," *arXiv preprint arXiv:2310.04454*, 2023.

[29] Y. Liu, X. Wang, and L. Zhang, "Exploring the potential of large language models in ui/ux design," *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2023.

[30] X. Zhang, H. Chen, and Y. Liu, "Ai-powered user interface design: Challenges and opportunities," *IEEE Transactions on Human-Machine Systems*, vol. 54, no. 1, pp. 76–87, 2024.

[31] J. Li, Y. Wang, and Q. Zhang, "Ai-assisted ui design: A comparative study of different generative ai approaches," *International Journal of Human-Computer Studies*, vol. 171, p. 102930, 2024.

[32] S. Park, J. Kim, and S. Lee, "Exploring the impact of ai-generated ui/ux designs on user satisfaction and engagement," *International Journal of Human–Computer Interaction*, vol. 39, no. 14, pp. 2089–2105, 2023.

[33] X. Chen, Y. Liu, and Z. Wang, "Adaptive ui/ux design using reinforcement learning: A case study in mobile applications," *IEEE Access*, vol. 12, pp. 12345–12356, 2024.

[34] A. Çelen, G. Han, K. Schindler, L. Van Gool, I. Armeni, A. Obukhov, and X. Wang, "I-Design: Personalized LLM Interior Designer," *arXiv.org*, 2024.

[35] B. Wang, G. Li, and Y. Li, "Enabling Conversational Interaction with Mobile UI using Large Languag e Models," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing S ystems*, pp. 1–17, ACM, Apr. 2023.

[36] J. Gao, X. Li, and M. Galley, "Conversational ui design: Leveraging large language models for natural interactions," *Foundations and Trends® in Human–Computer Interaction*, vol. 17, no. 1, pp. 1–108, 2024.

[37] V. Liu, J. Vermeulen, G. Fitzmaurice, and J. Matejka, "3DALL-E: Integrating Text-to-Image AI in 3D Design Workflows," in *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, ACM, July 2023.

[38] H.-K. Ko, G. Park, H. Jeon, J. Jo, J. Kim, and J. Seo, "Large-scale Text-to-Image Generation Models for Visual Artists' Creati ve Works," in *Proceedings of the 28th International Conference on Intelligent User I nterfaces*, ACM, Mar. 2023.

[39] S. Petridis, M. Terry, and C. J. Cai, "PromptInfuser: Bringing User Interface Mock-ups to Life with Large Lan guage Models," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Comp uting Systems*, ACM, Apr. 2023.

[40] J. D. Weisz, J. He, M. Muller, G. Hoefer, R. Miles, and W. Geyer, "Design Principles for Generative AI Applications," in *Proceedings of the CHI Conference on Human Factors in Computing System s*, vol. 10, pp. 1–22, ACM, May 2024.

[41] C. Si, Y. Zhang, Z. Yang, R. Liu, and D. Yang, "Design2Code: How Far Are We From Automating Front-End Engineering?," *arXiv.org*, 2024.

[42] Y. Lu, S. Liu, Q. Zhang, and Z. Xie, "RTLLM: An Open-Source Benchmark for Design RTL Generation with Large L anguage Model," in *2024 29th Asia and South Pacific Design Automation Conference (ASP-DAC)*, IEEE, Jan. 2024.

[43] P. Denny, J. Prather, B. A. Becker, J. Finnie-Ansley, A. Hellas, J. Leinonen, A. Luxton-Reilly, B. N. Reeves, E. A. Santos, and S. Sarsa, "Computing Education in the Era of Generative AI," *Communications of the ACM*, vol. 67, pp. 56–67, Jan. 2024.

[44] A. B. Kocaballi, "Conversational AI-Powered Design: ChatGPT as Designer, User, and Produ ct," *arXiv.org*, 2023.

[45] J. Zamfirescu-Pereira, H. Wei, A. Xiao, K. Gu, G. Jung, M. G. Lee, B. Hartmann, and Q. Yang, "Herding AI Cats: Lessons from Designing a Chatbot by Prompting GPT-3," in *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, ACM, July 2023.

[46] H. Jiang, X. Zhang, X. Cao, C. Breazeal, D. Roy, and J. Kabbara, "PersonaLLM: Investigating the Ability of Large Language Models to Expr ess Personality Traits," *arXiv.org*, 2023.

[47] A. Bastola, H. Wang, J. Hembree, P. Yadav, Z. Gong, E. Dixon, A. Razi, and N. McNeese, "LLM-based Smart Reply (LSR): Enhancing Collaborative Performance with ChatGPT-mediated Smart Reply System," *arXiv.org*, 2023.

[48] C. Rastogi, M. Tulio Ribeiro, N. King, H. Nori, and S. Amershi, "Supporting Human-AI Collaboration in Auditing LLMs with LLMs," in *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 913–926, ACM, Aug. 2023.

[49] P. Li and B. Li, "Generating Daylight-driven Architectural Design via Diffusion Models," *arXiv*, 2024.

[50] J. J. Y. Chung and E. Adar, "PromptPaint: Steering Text-to-Image Generation Through Paint Medium-li ke Interactions," in *Proceedings of the 36th Annual ACM Symposium on User Interface Softwar e and Technology*, ACM, Oct. 2023.

[51] M. P. Verheijden and M. Funk, "Collaborative Diffusion: Boosting Designerly Co-Creation with Generati ve AI," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Comp uting Systems*, ACM, Apr. 2023.

[52] A. Al-sa'di and D. Miller, "Exploring the Impact of Artificial Intelligence language model ChatGPT on the User Experience," *International Journal of Technology, Innovation and Management (IJTIM)*, vol. 3, pp. 1–8, May 2023.

[53] S. Ding, X. Chen, Y. Fang, W. Liu, Y. Qiu, and C. Chai, "DesignGPT: Multi-Agent Collaboration in Design," in *2023 16th International Symposium on Computational Intelligence and De sign (ISCID)*, IEEE, Dec. 2023.

[54] S. Moore, R. Tong, A. Singh, Z. Liu, X. Hu, Y. Lu, J. Liang, C. Cao, H. Khosravi, P. Denny, C. Brooks, and J. Stamper, "Empowering Education with LLMs - The Next-Gen Interface and Content Ge neration," in *Communications in Computer and Information Science*, pp. 32–37, Springer Nature Switzerland, 2023.

[55] C. Kulkarni, S. Druga, M. Chang, A. Fiannaca, C. Cai, and M. Terry, "A Word is Worth a Thousand Pictures: Prompts as AI Design Material," *arXiv.org*, 2023.

[56] Z. Englhardt, R. Li, D. Nissanka, Z. Zhang, G. Narayanswamy, J. Breda, X. Liu, S. Patel, and V. Iyer, "Exploring and Characterizing Large Language Models for Embedded System Development and Debugging," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pp. 1–9, ACM, May 2024.

[57] Z. Lin, "Why and how to embrace AI such as ChatGPT in your academic life," *Royal Society Open Science*, vol. 10, Aug. 2023.

[58] A. Mishra, U. Soni, A. Arunkumar, J. Huang, B. C. Kwon, and C. Bryan, "PromptAid: Prompt Exploration, Perturbation, Testing and Iteration usi ng Visual Analytics for Large Language Models," *arXiv.org*, 2023.

[59] S. G. Bouschery, V. Blazevic, and F. T. Piller, "Augmenting human innovation teams with artificial intelligence: Explor ing transformer-based language models," *Journal of Product Innovation Management*, vol. 40, pp. 139–153, Jan. 2023.

[60] H. Dang, L. Mecke, F. Lehmann, S. Goller, and D. Buschek, "How to Prompt? Opportunities and Challenges of Zero- and Few-Shot Lear ning for Human-AI Interaction in Creative Applications of Generative M odels," *arXiv.org*, 2022.

[61] O. Shaer, A. Cooper, O. Mokryn, A. L. Kun, and H. Ben Shoshan, "AI-Augmented Brainwriting: Investigating the use of LLMs in group idea tion," in *Proceedings of the CHI Conference on Human Factors in Computing System s*, vol. 14, pp. 1–17, ACM, May 2024.

[62] G. Di Fede, D. Rocchesso, S. P. Dow, and S. Andolina, "The Idea Machine: LLM-based Expansion, Rewriting, Combination, and Sug gestion of Ideas," in *Creativity and Cognition*, ACM, June 2022.

[63] B. Peng, Z. Bi, Q. Niu, M. Liu, P. Feng, T. Wang, L. K. Q. Yan, Y. Wen, Y. Zhang, and C. H. Yin, "Jailbreaking and mitigation of vulnerabilities in large language models," *arXiv preprint arXiv:2410.15236*, 2024.

[64] J. White, S. Hays, Q. Fu, J. Spencer-Smith, and D. C. Schmidt, "ChatGPT Prompt Patterns for Improving Code Quality, Refactoring, Requi rements Elicitation, and Software Design," *arXiv.org*, 2023.

[65] C. Qian, W. Liu, H. Liu, N. Chen, Y. Dang, J. Li, C. Yang, W. Chen, Y. Su, X. Cong, J. Xu, D. Li, Z. Liu, and M. Sun, "ChatDev: Communicative Agents for Software Development," *arXiv.org*, 2023.

[66] K. S. I. Mantri and N. Sasikumar, "Interactive Fashion Content Generation Using LLMs and Latent Diffusion Models," *arXiv.org*, 2023.

[67] J. Wei, A.-L. Courbis, T. Lambolais, B. Xu, P. L. Bernard, and G. Dray, "Boosting GUI Prototyping with Diffusion Models," in *2023 IEEE 31st International Requirements Engineering Conference (RE)*, IEEE, Sept. 2023.

[68] S. Petridis, M. Terry, and C. J. Cai, "PromptInfuser: How Tightly Coupling AI and UI Design Impacts Designers ' Workflows," in *Designing Interactive Systems Conference*, vol. 16, pp. 743–756, ACM, July 2024.

[69] Z. He, X. Li, L. Fan, and H. J. Wang, "Revamping Interior Design Workflow Through Generative Artificial Intel ligence," in *Communications in Computer and Information Science*, pp. 607–613, Springer Nature Switzerland, 2023.

[70] C.-Y. Cheng, F. Huang, G. Li, and Y. Li, "PLay: Parametrically Conditioned Layout Generation using Latent Diffus ion," *International Conference on Machine Learning*, 2023.

[71] S. Suh, M. Chen, B. Min, T. J.-J. Li, and H. Xia, "Luminate: Structured Generation and Exploration of Design Space with L arge Language Models for Human-AI Co-Creation," *arXiv*, 2023.

[72] F. Lehmann, N. Markert, H. Dang, and D. Buschek, "Suggestion Lists vs. Continuous Generation: Interaction Design for Wri ting with Generative Models on Mobile Devices Affect Text Length, Word ing and Perceived Authorship," in *Mensch und Computer 2022*, ACM, Sept. 2022.

[73] P. L. Lanzi and D. Loiacono, "ChatGPT and Other Large Language Models as Evolutionary Engines for On line Interactive Collaborative Game Design," in *Proceedings of the Genetic and Evolutionary Computation Conference*, ACM, July 2023.

[74] L. Lian, B. Li, A. Yala, and T. Darrell, "LLM-grounded Diffusion: Enhancing Prompt Understanding of Text-to-Imag e Diffusion Models with Large Language Models," *Trans. Mach. Learn. Res.*, 2023.

[75] B. Peng, K. Chen, M. Li, P. Feng, Z. Bi, J. Liu, and Q. Niu, "Securing large language models: Addressing bias, misinformation, and prompt attacks," *arXiv preprint arXiv:2409.08087*, 2024.

[76] S. Harrer, "Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine," *eBioMedicine*, vol. 90, p. 104512, Apr. 2023.