Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10237

Leveraging generative AI tools for design method support: insights, challenges, and best practices

Olga Sankowski✉, Pascal Inselmann and Dieter Krause

TUHH - Hamburg University of Technology, Germany

✉ o.sankowski@tuhh.de

ABSTRACT: Publicly available generative AI tools, such as ChatGPT, Midjourney, and DALL-E 3, have the potential to transform product development by accelerating tasks and improving design ideation. Through case studies of scenario management and persona storyboarding, this research explores the strengths and limitations of generative AI (GenAI) tools. The results highlight GenAI's ability to accelerate routine tasks, improve ideation, and support iterative design, but also reveal limitations in contextual understanding and output quality. Key findings show that effective GenAI integration depends on precise prompt design, iterative interaction and critical validation. Despite their potential, GenAI tools cannot replace human expertise for nuanced design tasks. The study provides actionable insights and best practices for leveraging GenAI tools, paving the way for enhanced human-AI collaboration.

KEYWORDS: artificial intelligence, design methods, human-AI collaboration, case study

1. Introduction

Due to a number of trends and drivers in the areas of market, technology, society, as well as processes and organisation, the complexity in which product development has to generate successful and innovative products keeps increasing (Bender & Gericke, 2021). Various design methods exist to support product development processes. They help identify and pursue development goals, systematically evaluate alternatives, ensure transparency in decision-making, and support collaboration among designers, technical disciplines, and stakeholders (Bender & Gericke, 2021). However, achieving these objectives remains challenging due to growing complexity and growing amount of data being available for product design. Development processes therefore often remain time-consuming, resource-intensive, and influenced by subjective decisions, despite existing methodological support. New potential for supporting design methods is emerging in the form of easy-to-use and publicly accessible Artificial Intelligence (AI) tools, such as ChatGPT, Midjourney or DALL-E 3.

To explain AI, the term Machine Learning (ML) is often used, though it is only one of several approaches. ML enables machines to improve performance through trial-and-error learning from data and experience (Jurafsky & Martin, 2024). In the context of text processing, Natural Language Processing (NLP) refers to teaching programs to understand and process human language, enabling communication between humans and machines (Jurafsky & Martin, 2024). A subset of NLP, Large Language Models (LLMs), like ChatGPT, are trained on massive text datasets to facilitate human interaction (Jurafsky & Martin, 2024). For image generation, Generative Adversarial Networks (GANs) and diffusion models are widely used. Diffusion models, as seen in systems like DALL-E 3, train by converting images into noise and reversing the process to generate new visuals (O'Connor, 2022).

The use of such text- or image-generating AI offers considerable potential in a number of areas. It could be used to boost efficiency, allowing concepts to be generated in seconds or simulations and optimisations to be carried out more quickly with the help of generative AI (GenAI) tools. GenAI tools could be used to create unconventional ideas or to identify as-yet unknown patterns and concepts, thus boosting the innovative

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

strength of the development process. Large amounts of data, e.g. from customer feedback or market analyses, could be analysed more quickly and in a more targeted manner or used as a basis for elaborate simulations, thus promoting a data-driven development process. Finally, AI tools could also be used to improve collaboration between designers or between designers and customers by identifying contradictions in different data sets at an early stage or by quickly visualising concepts for different target groups.

Despite the obvious potential of GenAI tools, these systems also have limitations and risks in terms of applicability and usefulness. Limitations arise from insufficient training data, especially of LLM based chatbots. Furthermore, the data output is often inconsistent and difficult to understand (Bolaños et al., 2024). This manifests itself, as an effect known as 'hallucinations', in which the AI tools tend to generate random and incorrect data (Bolaños et al., 2024). While some design tasks, such as generating a large number of alternative designs, can be best performed by generative AI tools, there are still design tasks that need to be performed by human designers, such as producing a large variety of designs in early phases, selecting designs for iteration, translating abstracted concepts into design principles or achieving a deep understanding, insights and aesthetic sensibility of the design (Ranscombe et al., 2024; Terenzi et al., 2024). Brisco et al. (2023) emphasise that the tools for design idea development that are open access in 2023 should not be used in its current maturity to replace the abilities of a design team to generate concepts. Gmeiner et al. (2023) report that users in this context had problems with 'weird' design outcomes with aesthetic flaws and did not know how to express their goals and ideas to the generative AI tool as accurately to avoid this problem.

Thus, the integration of GenAI tools into the product development processes is not a guaranteed success; it cannot replace the designer or expert. Its limitations and potential for this application are not yet fully understood. For the successful use of GenAI tools it is necessary that designers learn to work with such tools, that they acquire the necessary skills, and that they learn about the applications and limitations of GenAI tools in order to understand which tasks they can and cannot perform in the development process (Chaudhry, 2024). In fact, designers need structured approaches or strategies to achieve successful collaboration with GenAI tools. These could help designers to reflect on incorrect or undesirable outcomes, as well as on their level of trust in the tools (Gmeiner et al., 2023). Chiarello et al. (2024) have analysed which design engineering tasks are named how often in the literature and have drawn conclusions regarding the potential use cases for LLMs. According to this, there is further potential to support data analysis and data transformation. Berni et al. (2024) presented a framework of AI-support. They have sorted subtasks of the AI tools in relation to the overarching task of design stimulation for idea and concept generation. They distinguish between tasks that AI can take over, cannot (yet) take over, should take over, and should not do. Accordingly, AI tools should assist the designer and accelerate the process by e.g. varying specific stimuli.

This paper takes a similar approach and examines the extent to which GenAI tools can support and accelerate methodical product development. This is done based on two case studies, in which different product development methods shall be supported by GenAI tools. These example applications are used to derive generally applicable recommendations for the embedding of these tools in other product development methods. The following research questions shall be investigated:

> *RQ 1: To what extent can generative AI tools support structured methods in early-stage product design? What tasks can they perform?*
> *RQ 2: How can design teams effectively integrate generative AI tools into collaborative workflows? What aspects need to be considered?*

To answer these questions, the state of research will first be examined to determine to what extent GenAI tools have already been used to support which task or which design method. The focus here is on the early design phases, as these are the fuzziest and most affected by effective decision making supported by design methods and GenAI tools. In Section 3, the methodological approach of the paper and the design methods used are presented. This is followed by the presentation of the results and a discussion. Finally, the fundamental conclusions for the use of publicly accessible GenAI tools for product development are derived and an outlook on future research directions is given.

2. State of the art

The product development process begins with defining product requirements for the new product. This can be achieved by extracting requirements from the needs expressed by customers and users or by

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

analysing competitor products. For the first task Han et al. (2023) developed a framework and used a pretrained language model to enable automated and large-scale implicit aspect opinion extraction from reviews. This reduces the time and effort required for data preparation and minimises the need for manual extraction of needs and requirements. Chien and Yao (2020) went a step further, assigning the AI chatbot the role of the real user to be simulated through the dialogue. Processing and maintenance of requirements can be further supported by automatic detection of contradiction requirements using LLM (Gärtner & Göhlich, 2024).

The Define phase is followed by the Ideate phase. Considering the context and the given constraints from the requirements, the aim here is to find as many different and innovative product ideas as possible. Kim and Maher (2023) found a positive effect on the number of new ideas when AI was used to support inspiration. Ranscombe et al. (2024), on the other hand, found that this effect could at least not be explained by the use of generative AI tools alone. When it comes to the images generated by the AI, they are of high quality and can be produced in large numbers, but the designs presented have a low degree of diversity. The participants in a study by Kaljun and Kaljun (2024) created fewer product designs, but these were better elaborated. The authors suspect that the AI helped them to carry out more iterations in the design process. Hwang and Won (2021) argue that the participants in their study achieved better results because they were less concerned about being judged by the AI tools than by other humans. Overall, several authors argue that humans and AI should work together to achieve better results in ideation and the overall design process (see e.g. (Ege et al., 2024; Wieland et al., 2022)).

Based on these studies, the question arises as to how humans and AI should collaborate in the design process and what effects this interaction has on humans. Gmeiner et al. (2023) found that designers working with GenAI tools were unclear about what contextual information they needed to provide to the AI tool, and wished the tools had more contextual knowledge about the solution being sought, similar to a human. Chang and Kuo (2024) examined a tool that transforms text into images and came to the conclusion that engaging with the AI tool, which required the precise description of the desired representation in written form, had a positive influence on design cognition. With regard to the tasks that AI tools can perform in collaboration with designers, Song et al. (2024) created a framework that enables classification based on three dimensions: the initiation spectrum, i.e. whether the human or the AI acts as a prompter, the intelligence scope distinguishes whether the AI's knowledge base is specialised or general, and finally, the cognitive model distinguishes whether the AI is performing an analysis-oriented or synthesis-oriented task. Chong et al. (2022) examined the influence of designers' self-confidence and AI performance in various combinations on the quality of the results in the context of human-AI collaboration. From this, it can be inferred that less competent designers are unable to identify gaps and flaws in the AI-generated result. Thus, a certain level of competence on the side of the designer is still needed to achieve effective and efficient collaboration.

Apart from general tasks in the design process that can be supported or taken over by AI, there are investigations into the extent to which AI can support specific design or process methods. Hassani et al. (2024) were able to reduce manual effort and time in the analysis phase of a failure mode and effects analysis by using LLM. Spreafico and Sutrisno (2023) designed a collection of targeted questions to be asked to a chatbot (ChatGPT) that could improve the automatic identification of failures in social failure mode and effect analysis. Koh (2024) tested the performance of LLMs in the automatic generation of Design Structure Matrices (DSM) and was able to reproduce over 70% of the DSMs.

Further studies examined publicly accessible GenAI tools in their suitability to support the product development process. Generative AI tools can be used to inspire ideas or visualise impressions so easily that people without formal design training can visualise product ideas, as seen in a collaborative ideation course at Hamburg Open Online University (Hoou, 2024)). Al-sa'di and Miller (2023) were also able to use an AI-Chatbot (ChatGPT) to support designers in the define and ideation stages by providing suggestions and feedback. Ege et al. (2024) had ChatGPT create a design for a spring-loaded launcher from scratch as part of a hackathon and then compared its performance with human developers. Chuma and Oliveira (2023) used ChatGPT as a decision-making tool in business and found that it could not fully replace a decision-making expert, but it could support them. Terenzi et al. (2024) also emphasise that generative AI tools such as Midjourney can certainly be used to visualise design ideas, but that this requires an iterative process of prompt revision and optimisation, for which the designer needs to be trained in the skills of prompt design. An interesting application of AI tools can be found in (Edwards et al., 2024). To avoid the difficulty for the user of having to put their design ideas into precise words,

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

they had hand sketches created and fed them into GPT-4V(ision). GPT-4V then created prompts from this, which could be used for image generation via DALL-E 3.
According to the state of research, there is particular, though not exclusive, potential for integrating GenAI tools in the early stages of development, the Define and Ideate phases. Thus, GenAI should support the analysis of implicit data and the derivation of product requirements in the Define phase, thus simplifying and accelerating the process for the designer. In the Ideate phase, the GenAI tools should support an iterative design process and, in line with the recommendation of Berni et al. (2024), support the designer by quickly varying stimuli in the idea generation. This interaction between humans and AI should promote design cognition (Chang & Kuo, 2024). According to the framework of (Song et al., 2024), the following two categories are addressed in this context of human-AI collaboration: I. category (Analysis + general knowledge base + Human as Prompter), in which AI acts as a search engine or consultant, and V. category (Synthesis + general knowledge base + Human as Prompter), in which AI acts as an ideator or generator.
To further explore and answer the research questions, Section 3 presents two case studies with two different design methods that are relevant at different points in the early design phase. The extent to which different GenAI tools can support these is explored in Section 4. A comparison with the classical approach without AI support or, alternatively, a comparison of the performance of different tools shall be used to determine where their application has limitations or generates errors and whether general conclusions for prompt design or for working with GenAI can be drawn from this.

3. Methodology

Two design methods are chosen to represent the first two phases of the product development process, Define and Ideate. Their procedures and challenges are outlined before detailing the study. The studies were conducted in the end of 2023 and the beginning of 2024. The versions used were ChatGPT 4, DALL-E 3, Midjourney version 5.2 for Leonardo.Ai and Canva, no version numbers are known.

3.1. Scenario management in the define phase

The aim of scenario management is to extrapolate a range of possible future scenarios and to define actions or draw conclusions for one's own product. A time horizon of 5 to 20 years is considered. We used the scenario management approach according to (Gausemeier et al., 1998). This is a 5-step procedure, consisting of scenario preparation (1), scenario field analysis (2), scenario prognostic (3), scenario development (4) and scenario transfer (5).
The challenge in implementing scenario management is to identify influencing factors and extrapolate them into consistent scenarios, as these are based on a large amount of information. In some cases, expert knowledge can be used, but as the influencing factors come from different areas such as society, politics, technology and trends, extensive research on these topics is often unavoidable. The use of text-generating AI should support this and speed up the research and analysis of data. In the best case, the tools can directly generate plausible scenarios or at least support individual steps in the process.
To check and compare the quality of the results, the scenario management is carried out in three iterations. The first iteration is carried out without GenAI support. In the second iteration, ChatGPT is asked to provide scenarios for the same use case, analogous to the results of the first iteration without AI support. In the third and final iteration, the scenario management is performed step by step with the help of ChatGPT. The analysis will compare the results and quality differences between the three iterations. We expect the comparison of the second and third iterations to provide insights into which steps are particularly suitable for GenAI support and what might explain this effect.

3.2. Persona method in the ideate phase

Personas are stereotypical archetypes of product users (Long, 2009). They are intended to represent realistic people, but not real users (Adlin & Pruitt, 2010). The results of qualitative user research are usually used to create personas, which are usually presented in the form of a profile. Persona profiles contain information about the stereotypical user, including their age, gender, education, hobbies, place of residence and personal needs. Personas are given specific information, such as real names and photos. This representation aims to continuously incorporate the needs of real users into the development process, for example by playing a kind of role-playing game with the persona, in which the design of the product can be continuously compared with the needs of the users (Long, 2009). They

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

can also be guided through various interactions with the product in the form of a storyboard (Adlin & Pruitt, 2010). A storyboard is a scribbled or comic-like representation that tells a story through a series of images.
Aside from the difficulty of deriving personas from empirical data, which becomes even greater the more accurate they are to be (see Chapman & Milham, 2006), the graphical representation is a challenge. While the persona profile itself only requires a portrait image of a person, which can either be drawn or taken from stock photos, storyboards require a certain amount of drawing skills. This is especially true if the storyboard needs to be created quickly, but still be detailed enough to help designers identify potential usability issues in different situations or come up with new ideas for user-friendly features.
The potential of GenAI is explored in this case by comparing different GenAI tools in terms of their ability to create images that show personas interacting with a product in a usage context. A total of four tools were selected: Leonardo.ai, Canva, DALL-E 3 and Midjourney. The first three offer a limited number of free image generations, while Midjourney requires payment from the first generated image but is widely regarded for its superior image quality. The comparison is based on a standardized prompts across four usage contexts. A wide variety of places, people and products are considered to see if there are any limitations or boundaries to what can be depicted.

4. Results

The results for the design methods and areas of support via GenAI tools are presented individually. The respective use cases for the methods are briefly summarised at the beginning.

4.1. Applying scenario management for an aircraft galley-lavatory-monument

The use case for the scenario management is a product family of aircraft galleys, which can be optionally configured with a variable number of trolleys with food and drinks, various electrical appliances, such as ovens, fridges or beverage makers, as well as with a variable number of different storage compartments. In addition, part of the galley can be replaced by a lavatory with one or two cabins. The aim of the application of scenario management was to decide which trends should be addressed more strongly in the coming product development cycles to be ideally positioned for the future market. When applying scenario management to individual products, the challenge is that the scenarios start at a high level and, as part of the transfer step (5), must be broken down into statements that can be applied to a specific system within the aircraft.
In the first iteration, the scenario management was carried out purely by hand, i.e. based on literature review and team discussions. This resulted in a list of influencing factors at both global and industry level in the scenario field analysis (2), from which 12 relevant key factors, including ‘financial strength of the population’ or ‘sustainability of aviation’, were selected. Dimensions were then used to create various possible projections for these (3). For example, financial strength could increase, but the population's willingness to invest in air travel could still be low, or the opposite could be the case, or both could be low, or both could be high. A total of five possible future scenarios were then derived from the combination of projections (4). E.g. ‘Green Age of Growth & Greener Skies’ describes a future in which flying is only allowed on certain routes while, due to global political pressure, aircraft fly with a high proportion of sustainable aviation fuels. This would increase the cost pressure on short and medium-haul flights and result in further reductions to in-flight services, with the possibility of them being cancelled altogether. The galley would no longer have to be served exclusively by the crew but would increasingly develop into a self-service kitchen similar to a vending machine. To this end, it would have to be equipped with the appropriate self-service units and advertising space (5).
For the second iteration, ChatGPT is used to generate scenarios directly. The prompt entered contains the following components: a short definition of the term scenario, a specification of the content and writing style, a specification of the time horizon and the number of scenarios desired, a reference to the possible sources for influencing factors, such as politics and technologies, as well as the objective of a future-robust galley design and an exemplary mention of areas on which influences are expected, such as business models of airlines. This prompt structure took several iterations and led to four scenarios. In terms of content and style, the formulated scenarios met the objectives. They first describe the global political situation, then address the effects of the aviation industry and then, more specifically, the design of aircraft galleys. The overarching description of the scenarios is initially plausible, but decreases as the

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

scenarios become more specific. Some of the descriptions, such as 'vertical gardens and hydroponic systems are an integral part of the kitchen', 'passengers can personalise their meals in advance using virtual platforms' and 'the galleys are equipped with advanced 3D printers' or 'the galleys are open platforms for culinary creativity' are unclear or simply unrealistic.

The third iteration explored whether the quality of the results could be improved by developing the scenarios step by step. We started with the scenario field analysis (2) and then proceeded step by step to the end. In the scenario field analysis (2), 25 influencing factors were generated at the global level, of which 22 had significant overlaps in content with the influencing factors created in the first iteration. At the industry level, there were still 9 out of 25 that had significant overlaps in content with the influencing factors created in the first iteration. In most cases, the AI-generated influencing factors would serve as a plausible supplement to the influencing factors created by hand. For example, ChatGPT assigned greater importance to nutrition and service topics. From these 50 possible influencing factors, ChatGPT then independently selected 15 relevant key factors in response to the given prompts. A comparison of these with scenarios from the first iteration showed that in six cases the key factor also appeared in the scenarios created by hand and in four cases the key factor was not explicitly named but was also addressed in the scenarios created by hand. In total, five key factors, such as 'mobile ordering and delivery services' or 'passenger feedback and ratings,' had no equivalent in the scenarios created by hand. When creating the scenario projections (3), it is apparent that ChatGPT continued to work with the previously created key factors and dimensions, but it did not do it systematically and consistently, with the result that the projections are often not completely transparent.

Table 1. Comparison of different approaches of creating scenarios with ChatGPT

2nd iteration without methodological framework	**3rd iteration wi t h methodological framework**
Scenario 1: Sustainability revolution	Scenario 4: Environmentally friendly era and culinary diversity
	Global level
- Global transformation towards sustainability	- International efforts lead to environmentally friendly standards
- Stricter environmental regulations from politics	
- Strong social environmental awareness	
- Shift towards a circular economy	
	Avaiation industry level
- Electric aircraft become standard	- Electric aircraft widely used and subsidized
- Ecological requirements for aircraft drastically tightened	
	Product level: Galley
- Galleys are centres of sustainable food production	- Health awareness, innovative storage technologies and digital payment process esenable a greater variety of healthier and sustainable multicu ltural menu options
- Vertical gardens, hydroponic systems integrated into galleys	- Electric kitchen appliances become standard
- Airlines focus on resource efficiency; no disposable packaging	- Storage technologies enable storage of fresh ingredients
	- Digitalis ation enables precise planningand inventory management

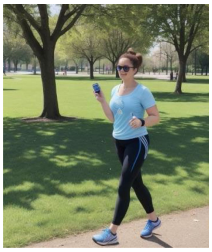
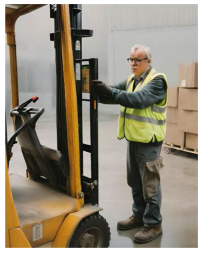
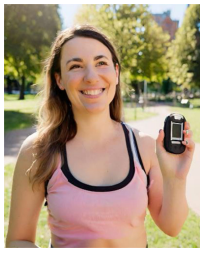
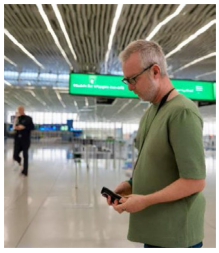
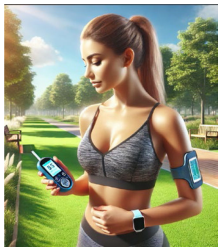
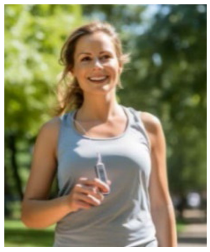
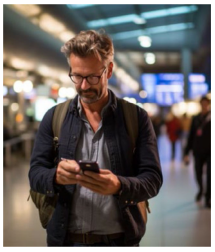
For the last steps, scenario development (4) and scenario transfer (5), the analysis of the scenarios shows that although they follow the requirement to serve different system levels, they focus on a specific key factor instead of describing the possible futures in a multidimensional way. An exemplary comparison of the scenarios created in the third iteration using the step-by-step approach with those created immediately in the second iteration can be seen in Table 1. The plausibility of the statements regarding the product 'galley' has increased significantly. However, there are hardly any precise statements regarding product design here either.

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

4.2. Accelerating visualisation of personas in contexts of use

This study focused on comparing four image-generating AI tools that were used to visualise four use cases. The structure of the prompts for the different use cases was created to be comparable. Attention was paid to a certain variation of people in terms of gender and age, as well as location and product (household appliance, industrial appliance, medical appliance, digital appliance). They should include a brief description of the person to be portrayed, the environment and situation, as well as the product and task. The image style should be specified in the prompt and, if possible, in the style settings, as highly artistic styles are less helpful for product development. When designing the prompt, it is important to keep sentences short and to the point. Long and convoluted descriptions were not always processed well by these tools, or information was seemingly ignored.

Table 2. Comparison of various generative AI tools in terms of their performance in visualising personas in the context of use

AI-Tool	Use case 1	Use case 2	Use case 3	Use case 4
Leonardo.Ai (Absolute Reality v.16)				
Canva (Foto)				
DALL-E 3, accessed via ChatGPT				
Midjourney version 5.2 (Hyperrealistic)				

The images shown in Table 2 were generated without iteration. The prompts for the four use cases are:

1. A real and modern photo of a man at the age of 10. He is in the house. The place is bright, and it is sunny. He is using a vacuum cleaner. The person is holding the tube of the vacuum cleaner in his hand.

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

2. A real and modern photo of a man aged 61. He is on company premises. The place is dark, and it is foggy. He is wearing glasses. He is wearing gloves and work shoes. He is using an entry and exit aid for a forklift truck. The person is standing next to the forklift.
3. A real and modern photo of a woman aged 33. She is in the park. The place is bright, and it is sunny. She is using an insulin pump. The person has the insulin pump in her hand before jogging.
4. A real and modern photo of a man aged 47. He is in the airport hall. The lighting is mediocre. He is using WhatsApp/ Messenger. The person is holding the cell phone in his hand and WhatsApp is open.

The quality of the resulting representations varies greatly. For example, the degree of realism achieved by DALL-E 3 seems rather mixed, or at least less than with the other tools. However, this is probably not problematic at all when using the images as a storyboard. More difficult are the obvious presentation errors, which are particularly noticeable in Use Case 1 with DALL-E 3 (cut vacuum cleaner hose) and in Use Case 2 with Leonardo.Ai (forklift within a forklift). However, the incorrect or wrong product interaction, which does not match the instructions in the prompt, is particularly problematic when using the images in the context of a storyboard. This is particularly true for Use Case 2, which in three out of four cases shows an obviously incorrect interaction with a forklift entry/exit aid. Only in the Midjourney is the person trying to get into the forklift from the side.

5. Discussion

This section addresses the research questions and compares study results with previous research. We also present conclusions for working with AI tools and a summary of our limitations.

RQ 1: To what extent can generative AI tools support structured methods in early-stage product design? What tasks can they perform? GenAI tools, such as ChatGPT and DALL-E 3, can assist in several key tasks during early product development phases. For Scenario Management in the Define phase, ChatGPT is effective for rapid literature review and generating influencing factors. However, the plausibility of its outputs declines as tasks demand nuanced or context-specific knowledge, such as when extrapolating these factors into actionable scenarios. The Persona method in the Ideate phase revealed that while AI-generated images provide rapid prototyping and flexibility, their accuracy and realism vary significantly across tools. Midjourney consistently delivered the highest quality outputs, yet limitations in representing complex usage contexts persist among all generative AI tools.

RQ 2: How can design teams effectively integrate generative AI tools into collaborative workflows? What aspects need to be considered? Effective collaboration with publicly accessible GenAI tools hinges on well-designed prompts and iterative refinement. The quality of outputs improves when users provide detailed, structured input and carefully validate results. Designers must adopt a critical stance, particularly in tasks involving data interpretation, where hallucinations or inaccuracies can occur. Tools such as Midjourney offer user-friendly mechanisms for refining outputs, but their reliance on external platforms (e.g., Discord for public image sharing) raises concerns about data security and privacy. This requires strategic considerations for tool selection based on task requirements and organizational policies.

The results of the two studies align with and extend the findings of prior research. For instance, the efficacy of ChatGPT in rapid data synthesis and influencing factor identification mirrors the findings of Han et al. (2023), who highlighted AI's capability to streamline data extraction and analysis in product development. Similarly, the study's observations regarding the iterative refinement of AI-generated outputs resonate with Terenzi et al. (2024), who emphasized the necessity of iterative prompt design for achieving usable results in generative AI-supported design processes. Furthermore, the limitations observed in representing nuanced or underrepresented product contexts echo the challenges identified by Ranscombe et al. (2024) and Gmeiner et al. (2023), who noted gaps in AI's contextual understanding and its dependence on the availability of high-quality training data. The superior performance of Midjourney in creating realistic persona visuals aligns with Brisco et al. (2023), who reported that higher-quality AI tools generally deliver more accurate outputs, albeit at the expense of accessibility and data privacy considerations. These connections underscore the broader applicability of the study's insights and reaffirm the critical role of human expertise in leveraging AI tools effectively.

From these findings four best practices can be derived for prompt design and tool selection:

- Prompt Precision: Effective prompts are specific, concise, and structured. Adding contextual details improves GenAI understanding and output relevance. This is becomes more critical the more specific the context is.

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

- Task Fit: Tasks involving systematic analysis or predictable outputs (e.g., influencing factor identification) are well-suited for GenAI, whereas creative synthesis often requires human oversight.
- Iterative Interaction: Iterations allow for refinement, particularly in visual tasks. Tools with built-in feedback mechanisms, like Midjourney, facilitate this process better than static systems.
- Tool Appropriateness: Midjourney excels in visual realism but comes with privacy trade-offs. ChatGPT is ideal for data collection but requires cautious interpretation.

However, there are limitations to the validity of this study. The research consists of two case studies which have primarily qualitative implications and hardly allow for quantitative conclusions. The second case study has a certain degree of replicability due to its simple structure, the comparison of four use cases and the short prompts, but the first case study on scenario management in the aviation industry can only be regarded as an exemplary study. It is unclear whether the quality of the scenarios generated step-by-step by ChatGPT would be of comparable quality for other products and industries, such as the cockpit of a car. It should be noted that the results of GenAI tools used in the studies are fundamentally unreproducible due to their mode of operation, wherein the results are regenerated each time by the prompt and are not stored in their entirety.

6. Conclusion

This paper highlights the growing usefulness of publicly available GenAI tools in supporting early stages of product development. GenAI accelerates routine tasks such as data synthesis and visualisation, giving designers more bandwidth for strategic thinking. However, its limitations - particularly in dealing with ambiguity or delivering high-quality, context-specific output - underline the irreplaceable role of human expertise in creative and critical tasks.

Further research should focus on developing robust frameworks for AI-human collaboration in design tasks and extending the capabilities of GenAI tools for underrepresented or novel contexts. The findings encourage designers to embrace GenAI as a complementary tool, while remaining mindful of its limitations. Through iterative learning and adaptation, GenAI can significantly enhance the efficiency and creativity in the design processes.

References

Adlin, T., & Pruitt, J. (2010). The essential persona lifecycle: Your guide to building and using personas (1st edition). Morgan Kaufmann.

Al-sa'di, A., & Miller, D. (2023). Exploring the Impact of Artificial Intelligence language model ChatGPT on the User Experience. *International Journal of Technology, Innovation and Management (IJTIM)*, 3(1), 1–8. https://doi.org/10.54489/ijtim.v3i1.195

Bender, B., & Gericke, K. (2021). Einleitung. In B. Bender & K. Gericke (Eds.), Pahl/Beitz Konstruktionslehre (pp. 1–6). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-57303-7_1

Berni, A., Borgianni, Y., Rotini, F., Gonçalves, M., & Thoring, K. (2024). Stimulating design ideation with artificial intelligence: present and (short-term) future. *Proceedings of the Design Society*, 4, 1939–1948. https://doi.org/10.1017/pds.2024.196

Bolaños, F., Salatino, A., Osborne, F., & Motta, E. (2024). Artificial intelligence for literature reviews: opportunities and challenges. *Artificial Intelligence Review*, 57(10). https://doi.org/10.1007/s10462-024-10902-3

Brisco, R., Hay, L., & Dhami, S. (2023). Exploring the Role of Text-to-Image AI in Concept Generation. *Proceedings of the Design Society*, 3, 1835–1844. https://doi.org/10.1017/pds.2023.184

Chang, H.-Y., & Kuo, J.-Y. (2024). Exploring metacognitive processes in design ideation with text-to-image AI tools. *Proceedings of the Design Society*, 4, 915–924. https://doi.org/10.1017/pds.2024.94

Chapman, C. N., & Milham, R. P. (2006). The Personas' New Clothes: Methodological and Practical Arguments against a Popular Method. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(5), 634–636. https://doi.org/10.1177/154193120605000503

Chaudhry, B. M. (2024). Concerns and Challenges of AI Tools in the UI/UX Design Process: A Cross-Sectional Survey. *In Chi'24: Extended abstracts of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1–6). The Association for Computing Machinery. https://doi.org/10.1145/3613905.3650878

Chiarello, F., Barandoni, S., Majda Škec, M., & Fantoni, G. (2024). Generative large language models in engineering design: opportunities and challenges. *Proceedings of the Design Society*, 4, 1959–1968. https://doi.org/10.1017/pds.2024.198

Chien, Y.-H., & Yao, C.-K. (2020). Development of an AI Userbot for Engineering Design Education Using an Intent and Flow Combined Framework. *Applied Sciences*, 10(22), 7970. https://doi.org/10.3390/app10227970

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press

Chong, L., Kotovsky, K., & Cagan, J. (2022). Are Confident Designers Good Teammates to Artificial Intelligence? A Study of Self-Confidence, Competence, and Collaborative Performance. *Proceedings of the Design Society*, 2, 1531–1540. https://doi.org/10.1017/pds.2022.155

Chuma, E. L., & Oliveira, G. G. de (2023). Generative AI for Business Decision-Making: A Case of ChatGPT. *Management Science and Business Decisions*, 3(1), 5–11. https://doi.org/10.52812/msbd.63

Edwards, K. M., Man, B., & Ahmed, F. (2024). Sketch2Prototype: rapid conceptual design exploration and prototyping with generative AI. *Proceedings of the Design Society*, 4, 1989–1998. https://doi.org/10.1017/pds.2024.201

Ege, D. N., Øvrebø,H. H., Stubberud, V., Berg, M. F., Steinert, M., & Vestad, H. (2024). Benchmarking AI design skills: insights from ChatGPT's participation in a prototyping hackathon. *Proceedings of the Design Society*. 4, 1999–2008. https://doi.org/10.1017/pds.2024.202

Gärtner, A. E., & Göhlich, D. (2024). Towards an automatic contradiction detection in requirements engineering. *Proceedings of the Design Society*, 4, 2049–2058. https://doi.org/10.1017/pds.2024.207

Gausemeier, J., Fink, A., & Schlake, O. (1998). Scenario Management. *Technological Forecasting and Social Change*, 59(2), 111–130. https://doi.org/10.1016/S0040-1625(97)00166-2

Gmeiner, F., Yang, H., Yao, L., Holstein, K., & Martelaro, N. (2023). Exploring Challenges and Opportunities to Support Designers in Learning to Co-create with AI-based Manufacturing Design Tools. *In ACM Digital Library, Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–20). Association for Computing Machinery. https://doi.org/10.1145/3544548.3580999

Han, Y., Bruggeman, R., Peper, J., Ciliotta Chehade, E., Marion, T., Ciuccarelli, P., & Moghaddam, M. (2023). Extracting Latent Needs From Online Reviews Through Deep Learning Based Language Model. *Proceedings of the Design Society*, 3, 1855–1864. https://doi.org/10.1017/pds.2023.186

Hassani, I. E., Masrour, T., Kourouma, N., Motte, D., & Tavcar, J. (2024). Integrating large language models for improved failure mode and effects analysis (FMEA): a framework and case study. *Proceedings of the Design Society*, 4, 2019–2028. https://doi.org/10.1017/pds.2024.204

Hoou. (2024). Collaborative Ideation: Gemeinsam Ideen entwickeln: Online - Postkarten. Hamburg Open Online University. https://learn.hoou.de/course/view.php?id=483&section=2

Hwang, A. H.-C., & Won, A. S. (2021). IdeaBot: Investigating Social Facilitation in Human-Machine Team Creativity. In *CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan.

Jurafsky, D., & Martin, J. H. (2024). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models (3rd edition, Online manuscript released August 2024). https://web.stanford.edu/~jurafsky/slp3.

Kaljun, K. K., & Kaljun, J. (2024). Enhancing Creativity in Sustainable Product Design: The Impact of Generative AI Tools at the Conceptual Stage. In *47th MIPRO ICT and Electronics Convention (MIPRO)* (pp. 451–456). IEEE. https://doi.org/10.1109/MIPRO60963.2024.10569541

Kim, J., & Maher, M. L. (2023). The effect of AI-based inspiration on human design ideation. *International Journal of Design Creativity and Innovation*, 11(2), 81–98. https://doi.org/10.1080/21650349.2023.2167124

Koh, E. C. (2024). Auto-DSM: Using a Large Language Model to generate a Design Structure Matrix. *Natural Language Processing Journal*, 9, 100103. https://doi.org/10.1016/j.nlp.2024.100103

Long, F. (2009). Real or Imaginary; The Effectiveness of Using Personas in Product Design. In *Proceedings of the Irish Ergonomics Society Annual Conference*, Dublin, Ireland. https://www.frontend.com/thinking/using-personas-in-product-design/

O'Connor, R. (2022, May 12). Introduction to Diffusion Models for Machine Learning. https://www.assemblyai.com/blog/diffusion-models-for-machine-learning-introduction/

Ranscombe, C., Tan, L., Goudswaard, M., & Snider, C. (2024). Inspiration or indication? Evaluating the qualities of design inspiration boards created using text to image generative AI. *Proceedings of the Design Society*, 4, 2207–2216. https://doi.org/10.1017/pds.2024.223

Song, B., Zhu, Q., & Luo, J. (2024). Human-AI collaboration by design. *Proceedings of the Design Society*, 4, 2247–2256. https://doi.org/10.1017/pds.2024.227

Spreafico, C., & Sutrisno, A. (2023). Artificial Intelligence Assisted Social Failure Mode and Effect Analysis (FMEA) for Sustainable Product Design. *Sustainability*, 15(11), 8678. https://doi.org/10.3390/su15118678

Terenzi, B., Menchetelli, V., Pagnotta, G., & Avallone, L. (2024). Connection between AI and product design - Potentials and critical issues in the text-to-image software-assisted design experience. In *AHFE International, Intelligent Human Systems Integration (IHSI 2024)*. AHFE International. https://doi.org/10.54941/ahfe1004511

Wieland, B., Wit, J. de, & Rooij, A. de (2022). Electronic Brainstorming With a Chatbot Partner: A Good Idea Due to Increased Productivity and Idea Diversity. *Frontiers in Artificial Intelligence*, 5, 880673. https://doi.org/10.3389/frai.2022.880673

https://doi.org/10.1017/pds.2025.10237 Published online by Cambridge University Press