The Potential of Generative AI for Systematic Engineering Innovation

Pavel Livotov, Mas'udah

Offenburg University of Applied Sciences, Badstr. 24, 77652 Offenburg, Germany
pavel.livotov@hs-offenburg.de
masudah@hs-offenburg.de

Abstract. Generative AI offers a new path for engineering innovation by automating idea generation and evaluation. This study explores its effectiveness in addressing complex and inventive engineering challenges. Using automated multi-directional and systematic prompt generation, the paper investigates the ability of AI chatbots to autonomously generate and evaluate innovative solution ideas and concepts. Experiments with various LLMs revealed their potential to accelerate the innovation process but also highlighted limitations in generating feasible, ready-to-use solution concepts. To address these challenges, the paper proposes mixed AI innovation teams, where different generative chatbots can complement and monitor each other. This collaborative approach can improve the quality and feasibility of AI-generated solutions. Case studies demonstrate the practical application of these findings and strategies for effective human-AI collaboration in the innovation process. While generative AI holds significant promise, future research should focus on refining AI models and developing frameworks for effective human-AI interaction to ensure the practical feasibility of AI-generated engineering design solutions for inventive problems.

Keywords: Generative AI; Problem-Solving; Engineering Innovation, Inventive Design.

1 Introduction – Background and Related Work

In recent years, generative artificial intelligence (AI) has gained significant attention in engineering due to its ability to autonomously generate content, solve complex problems, and assist in decision-making. In engineering design, it shows potential for enabling rapid prototyping, optimizing designs, and streamlining iterative processes. However, fully realizing its potential requires exploring new methods for creatively and autonomously solving engineering problems beyond traditional paradigms.

The impact of generative AI on engineering innovation has been extensively documented, with tools like Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and Large Language Models (LLMs) demonstrating their efficiency in generating innovative solutions based on training data. Brad [1] explores how inventive principles can enhance activation functions in AI models, increasing their creative capacity. Similarly, Ayaou and Cavallucci [2] propose a framework integrating AI with TRIZ principles to formalize knowledge and link disparate sources for innovative problem-solving. Human-AI collaboration remains crucial in addressing complex engineering challenges. Memmert and Bittner [3] highlight the opportunities of hybrid teams, while Qiu and Jin [4] emphasize the integration of AI with human expertise in enhancing design support systems. Müller, Roth, and Kreimeyer [5] outline barriers to AI-product development integration, such as the lack of standardized processes and documented best practices. Zhu et al. [6] show the success of GPT models in early-stage design concept generation, while Gomez et al. [7] and Ege et al. [8] examine the benefits and limitations of LLMs in complex system design and ideation. Generative AI tools, such as Open AI ChatGPT, Google Gemini, Anthropic Claude or others, can generate ideas and provide guidance, but human intervention is often required for practical implementation. Excessive or insufficient human involvement poses risks of bias or unfeasible designs. The need for behavioral science integration into AI systems is stressed by Van Rooy and Vaes [9], while Boussioux et al. [10] explore scalable human-AI collaboration for sustainable business innovation.

Xu et al. [11] compare ChatGPT's performance with human evaluators in engineering design tasks, highlighting the need for alignment in judgment confidence to improve decision-making. Chiarello et al. [12] discuss the theoretical and practical benefits of LLMs in automating design tasks, increasing efficiency, and balancing computational and human-centric design. Ranscombe et al. [13] evaluate image generative AI for design inspiration boards, noting differences in quantity, variety, and accuracy compared to traditional methods. Studies [14] and [15] emphasize AI's success in generating a variety of ideas during brainstorming and solving technical problems in process engineering.

However, significant challenges remain in creating detailed, practical solution concepts, especially in fields like mechanical engineering. AI-generated designs often lack the technical precision necessary for implementation, requiring clear instructions and technical drawings. Current text-to-image tools typically produce unsatisfactory results for engineering purposes, underscoring the need for more advanced AI capabilities to bridge the gap between concept generation and practical design implementation.

This paper advocates an integrative approach to automated multidirectional prompt generation, drawing from methodologies such as design theory [16], theory of inventive problem solving TRIZ [17], biomimetics, process intensification [18], and other approaches to systematic innovation. The goal is to improve generative AI chatbots' effectiveness in structured collaborative ideation and problem-solving for engineering design. It also explores AI strategies for developing comprehensive solutions by integrating multiple ideas and evaluating their practical applicability. Through controlled experiments, this study identifies patterns in prompting strategies that enhance the creative potential of AI chatbots, whether operating autonomously or in collaborative groups. However, current research limitations may affect the generalizability of the findings, as challenges remain in objectively assessing engineering creativity.

2 Methodology

2.1 Multidirectional Prompting

There are various approaches to formulating prompts for generative AI chatbots. This paper introduces Multidirectional Prompting (MDP), which applies elementary solution principles to generate innovative solutions. These solutions are defined by the novel, practical, and feasible combination of one or more ideas, specifically tailored to the problem or objectives. MDP explores multiple directions by addressing sub-problems and applying inventive stimuli, allowing AI to generate holistic solutions. This approach enhances the AI's generative capacity by combining solution ideas suited to the specific problem. In MDP, the multiple directions, sub-problems, and inventive stimuli can be selected either by the user or autonomously by the AI chatbots. Typical MDP techniques, including random, systematic, collaborative, and multi-problem prompting, are presented in Table 1.

Table 1. Techniques of Multidirectional Prompting (MDP)

MDP technique	Brief description
1. Random	The AI chatbot simultaneously applies multiple solution principles (SPs) and/or predefined engineering domains without a specific order, generating solution ideas and concepts in a single step.
2. Systematic	The AI chatbot applies a set number of solution principles (SPs) sequentially, generating ideas for each SP. It then combines complementary ideas to create comprehensive solution concepts.
3. Collaborative	Several AI chatbots (e.g., ChatGPT, Google Gemini) independently generate ideas - either randomly or systematically - and then exchange them to develop combined solution concepts.
4. Multiproblem	For complex problems, AI chatbots address prioritized sub-problems either one by one or simultaneously, generating ideas for each sub-problem and combining them into comprehensive solution concepts.

2.2 Automated Formulation of Elementary Creative Stimuli

This paper employs an automated method for generating creative stimuli for product and process design across various engineering domains [15]. Validated in both industrial and educational settings, the method has proven its efficacy in generating innovative solutions and improving the design process. The knowledge base is built on 160 elementary inventive principles [18], enhanced by selected TRIZ tools, including the 40 inventive principles, trends of technical evolution, and standard solutions, along with methodologies like biomimetics, process intensification, and others. Automated idea generation operates at multiple levels, such as: a) improving or transforming system components, b) enhancing useful actions, c) eliminating harmful effects, and d) resolving engineering contradictions. A proposed application for automated prompt generation uses 200 predefined inventive principles and allows user customization for solution search across engineering domains, as illustrated in Figure 1. The prompt composition process follows four key steps, as detailed in Table 2.

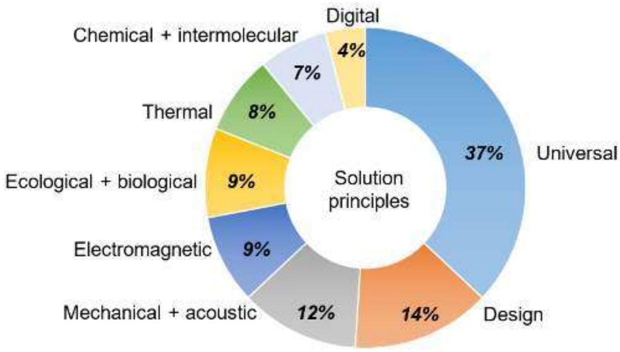

Fig. 1. Variety of the elementary solution principles for multidirectional prompting [15]

Table 2. Main steps of automated prompt generation

Prompting step	Brief description
1. Interactive problem definition	AI chatbot and users collaboratively define the problem, ensuring all relevant information and constraints are considered while avoiding biases. The chatbot refines the problem statement, highlighting key data like the ideal outcome, system components, and undesirable effects. Analyses, such as root cause identification, can be conducted as well.
2. Pre-selection of inventive stimuli and engineering domains	Users / AI chatbots select elementary inventive principles and engineering domains based on the initial problem. There's no limit to the number of principles, and new ones can be added at any stage. Users can also customize their choice based on their expertise or preferences.
3. Step-by-step idea generation	After the problem definition is confirmed, the chatbot generates ideas for each selected solution principle. Users can request more ideas or pause the process to move on to concept creation, ensuring the ideation remains flexible and aligned with user preferences.
4. Solution concept creation	The chatbot combines generated ideas into innovative solution concepts (e.g., five or more), leveraging its broad knowledge base. Users can guide the AI with specific strategies, such as focusing on feasibility, novelty, or a particular technology or core ideas, to ensure the solutions meet the problem's inventive objectives and specific requirements.

3.3 Experimental Approach

This section outlines two series of experiments investigating AI-based automated ideation and solution concept generation in engineering design. The experiments involved student projects from a course on AI-aided inventive design at the Offenburg University of Applied Sciences, Germany. Five groups of graduate students and one group of undergraduate students participated, along with a control group of the study's authors. The undergraduates were in their 5th or 6th semester, while the graduate students were pursuing a Master's in Mechanical Engineering and Robotics. All students received training in systematic new product development and TRIZ methodology. Working in groups of 2-3, they used generative AI chatbots with automated multidirectional prompt generation to tackle design challenges.

The experiments were conducted in two series. In the first educational series, all groups were assigned the same problem, the "Twist-off Screw Cap" and same initial prompt composition to use with the AI chatbots. This problem concerns jars or bottles with twist-off caps, which are difficult to open due to the high torque in the thread and the vacuum inside. The goal is to find solutions that make opening easier without additional tools or causing issues for manufacturers or consumers.

In the second series of experiments, six student groups already experienced in generative AI applied their skills to a design problem of their choice. They defined the problem, selected appropriate solution principles for multidirectional prompting, and used AI chatbots of their choice. Participants were free to choose their concept generation strategies, including selecting promising combinations of principles, identifying strong ideas, and setting evaluation criteria. For example, focusing on developing concepts around a core idea appropriate for targeted improvements. Table 3 presents the details of these experiments, including the number of selected solution principles, generated ideas, evaluated concepts, and applied AI chatbots.

Table 3. Experimental scope in the second series of experiments using different AI chatbots.

<table>
<tr><th rowspan="2">Gr.</th><th rowspan="2">Design problem name</th><th colspan="3">Number of</th><th rowspan="2">AI tools applied</th><th rowspan="2">Concept creation strategy</th></tr>
<tr><th>solution principles</th><th>generated ideas</th><th>solution concepts</th></tr>
<tr><td>1</td><td>Smoke detector</td><td>10</td><td>50</td><td>10</td><td rowspan="2">ChatGPT
Gemini</td><td rowspan="6">a) autonomous proposal by AI
b) based on 10 strongest AI ideas selected by AI
c) based on strongest AI ideas selected by the users</td></tr>
<tr><td>2</td><td>Hot drink cup</td><td>10</td><td>45</td><td>12</td></tr>
<tr><td>3</td><td>Quick release chuck</td><td>14</td><td>50</td><td>30</td><td rowspan="4">ChatGPT
Gemini
Claude</td></tr>
<tr><td>4</td><td>Shape adaptive gripper</td><td>10</td><td>196</td><td>18</td></tr>
<tr><td>5.</td><td>Barbeque grill</td><td>22</td><td>150</td><td>33</td></tr>
<tr><td>6.</td><td>Cable winder</td><td>10</td><td>150</td><td>30</td></tr>
</table>

3 Results and Discussion

3.1 AI-driven Ideation and Solution Concept Creation

Generative AI, utilizing multi-directional prompting with elementary solution principles, is capable of formulating up to 100–200 distinct ideas, substantially enhancing both productivity and the diversity of generated ideas, surpassing traditional approaches. By systematically exploring a wide range of potential solutions, this technique significantly strengthens the innovation process, fostering the generation of more varied and inventive concepts. In experiments ChatGPT and Claude performed best with multi-directional prompting, while Gemini, though requiring more interaction, tends to provide more objective evaluations.

In the phase of concept creation different generative AI tools perform variably across tasks, presenting an opportunity for mixed AI teams. Tools like ChatGPT, Google Gemini, and Anthropic Claude complement each other, with each offering unique insights, even though the solutions are often similar. However, AI often introduces hidden biases, so monitoring and adjustments are needed for workable solutions. Human judgment remains crucial, especially in addressing subtle aspects of innovation.

Preliminary results indicate that the most robust AI-driven strategy for solution concept creation involves identifying the most promising core ideas and developing multiple concepts based on them. This approach prioritizes inventive goals, such as usefulness or value (as key metrics for goal achievement), over novelty and feasibility during the initial stages of concept development.

3.2 Limitations in AI Evaluation of Ideas and Concepts

During the evaluation phase, the chatbots autonomously assessed their ideas and solution concepts using the following criteria: Feasibility (0 = unviable, 1 = feasible with effort, 2 = easily implementable), Novelty (0 = common, 1 = moderately novel, 2 = highly original), and Usefulness (0 = irrelevant, 1 = moderately useful, 2 = highly useful). In both experimental series, the chatbots consistently overvalued their concepts compared to participant ratings, particularly in usefulness and feasibility. The use of finer rating scales (e.g., 5- or 10-point scales) in the second series did not significantly reduce overestimation with ChatGPT, whereas Claude demonstrated better accuracy. It's also noted that AI evaluations can vary slightly, typically by ±1 point, between repeated evaluations in the same or separate chat sessions. Additionally, individual concept ratings may differ from aggregate ratings of multiple concepts. Finer scales, such as 10-point ratings, provide more nuanced and consistent assessments, helping to mitigate this variability.

Moreover, generative AI models appear to exhibit a moderate "Not Invented Here" effect when evaluating solutions proposed by other AI chatbots or engineers in concept evaluation across all experiments. For example, in a pairwise comparison, both ChatGPT4.0 and Gemini rate the usefulness of their own concepts higher. The authors consider this phenomenon useful, as it promotes a more balanced assessment of ideas and concepts when different AI chatbots operate as virtual teams of specialists, either autonomously or in collaboration with engineers.

3.3 Feedback and Observations from Experiments

Different generative AI tools perform variably across tasks, presenting an opportunity for mixed AI teams. Tools like Open AI ChatGPT, Google Gemini, and Anthropic Claude complement each other, with each offering unique insights despite similar solutions. ChatGPT and Claude perform best with multi-directional prompting, while Gemini, though requiring more interaction, tends to provide more objective evaluations. However, AI often introduces hidden biases, so monitoring and adjustments are needed for workable solutions.

The results of an anonymous survey conducted among the 17 participants at the end of the second series of experiments are particularly interesting. The participants rated their responses using a 10-point scale: 1-2 (very low), 3-4 (low), 5-6 (medium), 7-8 (high), and 9-10 (very high), with the mean values presented in Table 4.

Table 4. Results of an anonymous survey on the performance of generative AI

No.	Survey question: How do you rate the following aspects in application of generative AI …	Mean values (17 participants)
1	contribution of AI to increasing your personal inventive CREATIVITY?	7.1 SD=1.6
2	performance of AI in terms of the ideas USEFULNESS?	6.1 SD=2.1
3	performance of AI in terms of the ideas NOVELTY?	6.7 SD=1.9
4	performance of AI in terms of the ideas FEASIBILITY?	4.7 SD=1.9
5	overall performance of AI in the solution concept development phase?	6.1 SD=2.0
6	level of detail of the solution concepts proposed by the AI, so that designers can quickly implement a solution concept?	4.7 SD=2.0
7	accuracy of the evaluation of solution concepts by AI?	4.2 SD=1.7

AI chatbots received the highest rankings for enhancing participants' personal inventive creativity, but the lowest for evaluation accuracy and the level of detail in solution concepts needed for quick implementation.

Interestingly, participants also reported difficulties in personally evaluating the large number of ideas generated by AI. Subsequent analysis of the students' protocols by supervisors revealed that many novel and useful ideas were not recognized as promising and were excluded from concept creation. This highlights a key challenge in applying AI to the innovation process: engineers or students often expect ready-to-use solutions and struggle to thoroughly process numerous ideas generated by AI. This challenge highlights the need for a systematic exploration of collaborative frameworks and models for AI and human interaction in the inventive design process. Future research should prioritize developing and refining these collaboration models to optimize the integration of AI technologies in human-centered innovation, while also fostering the acceptance of design concepts created autonomously by AI.

4 Concluding Remarks and Outlook

The results of this study reveal key insights for applying generative AI in inventive engineering design. First, multi-directional prompting with elementary solution principles greatly boosts productivity and variety in idea generation, surpassing traditional methods like brainstorming or classical TRIZ. The challenge now shifts to selecting strong ideas and developing effective solution concepts. A key question is finding the optimal balance of human involvement in AI-assisted problem-solving. Second, AI chatbots tend to overestimate the feasibility of their concepts, highlighting the need for better self-evaluation algorithms. Bridging the gap between AI and human evaluation is crucial for real-world application. Third, varying degrees of overestimation between AI models (e.g., ChatGPT versus Gemini) show that model architecture impacts assessment accuracy. Future research should focus on minimizing these biases. Finally, the gap between AI-generated ideas and practical implementation remains a challenge. Advances in AI's ability to produce technically feasible solutions, including text-to-CAD tools will be essential for improving AI's role in engineering design and inventive problem solving.

References

1. Brad, S.: Enhancing Creativity in Deep Learning Models with SAVE-Inspired Activation Functions. In: Cavallucci, D., Livotov, P., Brad, S. (eds) Towards AI-Aided Invention and Innovation. TFC 2023. IFIP Advances in Information and Communication Technology, vol 682. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-42532-5_12
2. Ayaou, I., Cavallucci, D.: Multi-domain and Heterogeneous Data Driven Innovative Problem Solving: Towards a Unified Representation Framework. In: Cavallucci, D., Livotov, P., Brad, S. (eds) Towards AI-Aided Invention and Innovation. TFC 2023. IFIP Advances in Information and Communication Technology, vol 682. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-42532-5_10
3. Memmert, L., Bittner, E.: Complex Problem Solving through Human-AI Collaboration: Literature Review on Research Contexts. Proceedings of the 55th Hawaii International Conference on System Sciences, pp 378-387 (2022). http://hdl.handle.net/10125/79376
4. Qiu, Y., Jin, Y.: ChatGPT and finetuned BERT: A comparative study for developing intelligent design support systems. Intelligent Systems with Applications, 21, 200308 (2024). https://doi.org/10.1016/j.iswa.2023.200308

5. Müller, B., Roth, D., Kreimeyer, M.: Barriers to the Use of Artificial Intelligence in the Product Development – A Survey of Dimensions Involved. Proceedings of the Design Society,3, 757-766 (2023). doi:10.1017/pds.2023.76
6. Zhu, Q., Luo, J.: Generative transformers for design concept generation. Journal of Computing and Information Science in Engineering, 23(4), Article 041003 (2023). DOI: 10.1115/1.4056220
7. Gomez, A.P., Krus, P., Panarotto, M., Isaksson, O.: Large language models in complex system design. Proceedings of the Design Society, 4, 2197-2206 (2024). doi:10.1017/pds.2024.222
8. Ege, D.N., Øvrebø, H.H., Stubberud, V., Berg, M.F., Steinert, M., Vestad, H.: Benchmarking AI design skills: insights from ChatGPT's participation in a prototyping hackathon. Proceedings of the Design Society, 4, 1999-2008 (2024). doi:10.1017/pds.2024.202
9. Van Rooy, D., Vaes, K.: Harmonizing human-AI synergy: behavioral science in AI-integrated design. Proceedings of the Design Society, 4, 2287-2296 (2024). doi:10.1017/pds.2024.231
10. Boussioux, L., Lane, J.N., Zhang, M., Jacimovic, V. and Lakhani, K.R.: The Crowdless Future? Generative AI and Creative Problem Solving. Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-005, July 01, 2024). http://dx.doi.org/10.2139/ssrn.4533642
11. Xu, W., Kotecha, M.C., McAdams, D.A.: How good is ChatGPT? An exploratory study on ChatGPT's performance in engineering design tasks and subjective decision-making. Proceedings of the Design Society, 4, 2307-2316 (2024). doi:10.1017/pds.2024.233
12. Chiarello, F., Barandoni, S., Majda Škec, M., Fantoni, G.: Generative large language models in engineering design: opportunities and challenges. Proceedings of the Design Society, 4, 1959-1968 (2024). doi:10.1017/pds.2024.198
13. Ranscombe, C., Tan, L., Goudswaard, M., Snider, C.: 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 (2024). doi:10.1017/pds.2024.223
14. Mas'udah, Livotov, P.: Nature's lessons, AI's power: sustainable process design with generative AI. Proceedings of the Design Society, vol. 4, pp. 2129–2138, Cambridge University Press (2024). doi:10.1017/pds.2024.215
15. Livotov, P.: Enhancing Engineering Creativity with Automated Formulation of Elementary Solution Principles. Proceedings of the Design Society, vol. 3, 1645-1654 (2023). doi:10.1017/pds.2023.165
16. Kannengiesser, U., Gero, J.S.: Can Pahl and Beitz' systematic approach be a predictive model of designing? Design Science, vol. 3, e24, Cambridge University Press (2017). doi:10.1017/dsj.2017.24
17. Altshuller, G.S.: Creativity as an exact science: the theory of the solution of inventive problems, Gordon and Breach Science Publishers, New York, ISSN 0275-5807 (1984)
18. Livotov, P., Chandra, S.A.P., Mas'udah, Law, R., Reay, D., et al.: Eco-innovation in process engineering: Contradictions, inventive principles and methods, Thermal Science and Engineering Progress, Volume 9, Pages 52-65 (2019). https://doi.org/10.1016/j.tsep.2018.10.012