

A team of three: the role of generative AI in the development of design automation systems for complex products

Alejandro Pradas Gomez ,^{1,✉}, Maximilian Kretzschmar ,², Kristin Paetzold-Byhain ,² and Ola Isaksson ,¹

¹ Chalmers University of Technology, Sweden, ² Technische Universität Dresden, Germany

✉ alejandro.pradas@chalmers.se

ABSTRACT: Given the rise of Generative AI and Large Language Models (LLMs), there is a high interest in their use also in engineering design domain. Current research approaches lack to leverage LLM's new orchestration capabilities and use the LLMs in ways that expose their inherent weaknesses. We present a conceptual model to visualize the contribution of LLMs to design tasks and distribute ownership in the design activities: the triangle of design responsibility. A literature review on the design engineering field presents its current uses in this community. The understanding of the model is validated with industry via survey. We identify future research directions in the field of complex product design. We hope that this model helps design automation developers, researchers and industry practitioners to position and assign responsibility effectively in their design automation implementation.

KEYWORDS: design automation, language models, artificial intelligence, design engineering, computational design methods

1. Introduction

The design automation goal is to offload design tasks to a machine to support the designer or engineer. As Cederfeldt and Elgh (2005) define it, design automation refers to “*Engineering support by implementation of information and knowledge in solutions, tools, or systems, that are pre-planned for reuse and support the progress of the design process*”. But despite its focus on reusability, once automation is implemented in a machine-readable format, its flexibility often comes at the cost of usability. For instance, rigid parameterization can limit adaptability (Rad et al. 2022).

Developing a design automation support system presents a dilemma for developers. On one hand, creating a system that can adapt to any design scenario demands a level of effort and expertise that not all designers may have the time or ability to achieve (Heikkinen 2021). Moreover, such systems assume that users possess sufficient knowledge of the design system to configure the automation. On the other hand, a turn-key design automation system, while easy to use, often lacks the adaptability required to address the specific needs of a product under development, leading to a “black box” effect (Verhagen et al. 2012). This poses a trade-off between user-friendliness and the flexibility of automation.

Within the broader scope of Generative AI (GenAI), Large Language Models (LLMs) have appeared recently as a technology that offers potential in the trade-off between design flexibility and automation. The way these models can support the designer in design automation is by providing an assistance layer between the engineer and conventional methods. LLMs are not only able to independently generate outputs according to specific tasks demands (Generation), but also possess the capability to assist in connecting and integrating conventional engineering methods through interpreting inputs and outputs in accordance with given guidelines (Orchestration)

This duality represents an extension of previous technological possibilities. It also highlights their evolving but underexplored role within the technological agency utilized by engineers in the modern design process. As Leonardi (2011) explains, the imbrication of human and technological agency leads to the formation of new routines in organizational and industrial contexts; in our case, this interplay specifically applies to engineering design, necessitating an analysis to ensure LLM-based systems afford rather than constrain the goals of the engineer. We believe this analysis is of particular importance when considering the integration of non-deterministic models as active participants and therefore partial owners in the design process. The model proposed in this paper is a response to this analysis need.

We define complex products as components that are usually part of complex systems that are difficult to describe, understand, predict, manage, design or change (Sellgren 2005), that additionally require the use of physical simulation models for the validation of the concepts before manufacturing. Their design typically relies on models such as Finite Element Modelling (FEM) and Computational Fluid Dynamics (CFD) to predict performance in relation to neighbouring system components. Examples of such components are jet engine structural parts or truck chassis members which integrate with other subsystems and must perform reliably under dynamic load cases and varying configurations. We envision that LLM-based systems could provide support via Design Automation to orchestrate these activities, providing time savings and improved quality throughout repeated analysis loops.

Other authors have explored the potential of GenAI in Engineering Design (ED) (Chiarello et al. 2024). Their evaluation of the ED literature classified the potential of GenAI (LLMs in particular) to support designers with the generation, evaluation and description of engineering activities. They also highlighted the need for further research in LLMs supporting the ED activities of evaluating and describing. This paper intends to fill the research gap on the support that LLMs can provide to designers on design automation task during the development and evaluation of complex systems, such as structural components of vehicles.

The research reported in this paper aims to answer the following Research Questions (RQs):

RQ1. How can designers - and design automation developers - structure and visualize the contribution of Generative AI, specifically LLMs, within design automation systems?

RQ2. How have LLMs been integrated into design processes within the engineering design community.

RQ3. What are the research gaps in the application of LLMs for engineering design support?

2. Method

To answer RQ1, the experience from developing LLM-based automation applications (Pradas Gómez, Krus, et al. 2024; Pradas Gómez, Panarotto, et al. 2024) has been used to propose a conceptual model, where the contribution of LLM can be easily identified. Additional feedback has been received from academic and industrial practitioners. Moreover, a formal survey has been created to validate the conceptual understanding of the model in industry. Furthermore, the literature survey used to answer RQ2 has been used to further validate the model coherence.

To answer RQ2, a literature review was carried out to retrieve documented examples of the application of LLMs within the engineering design community. After the identification of relevant papers, these were analysed to position each contribution within the triangle model, assessing the responsibility attributed to each member. RQ3 was answered from the combination of the literature review and the insights of the survey in industry.

2.1. Survey methodology

The method was presented to industrial participants in two companies: GKN Aerospace (4 different sites and countries) and MAN Truck & Bus SE. For both companies the same format was employed to ensure consistency and facilitate comparative analysis between the two companies. The model was presented for 15 minutes in an online meeting, followed by a digital survey and further 15 minute discussion on the model and the feasibility of GenAI impacting engineering practices.

The survey was designed as a quick and anonymous feedback (2 minutes average) aimed at maximizing responses. Questions 1 to 5 are available in the results section, [Figure 2](#). An additional question asked the participants to provide any comments on the proposed model or the application of GenAI to their engineering tasks, complementing the post-survey group discussion.

2.2. Literature review methodology

The literature review used the PRISMA method (Page et al. [2021](#)) to present and describe the methodology and structure the content of the results and discussion sections. To maintain a focused and contextually relevant scope, we specifically targeted journals and conference papers endorsed by the Design Society (DS) plus an additional journal - JCSIE - due to the relevance of the topic. The search string, repositories used, and selection criteria are available in [Figure 1](#). RQ2 guided the creation of a targeted search string for data collection.

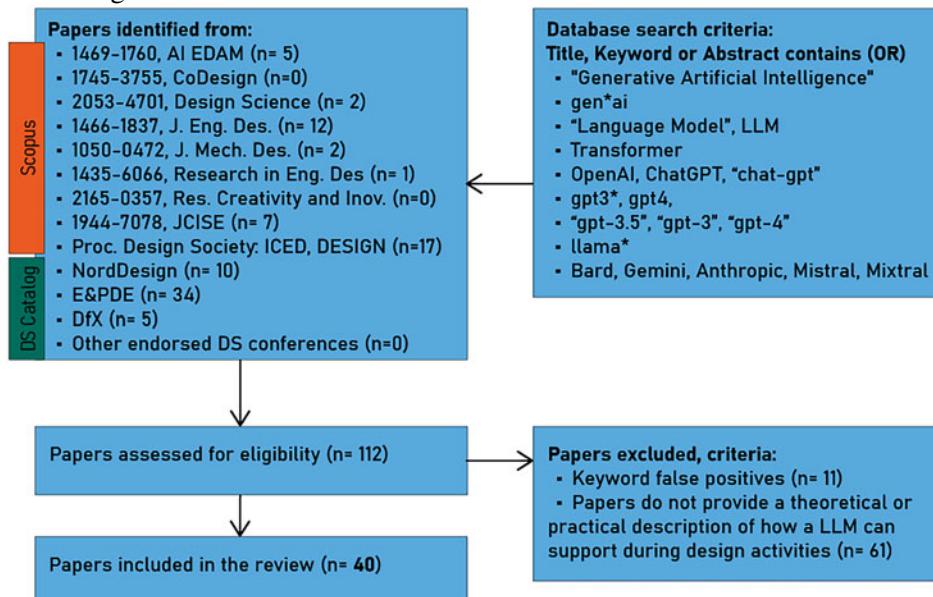


Figure 1. Adapted PRISMA flowchart showing how the 40 papers assessed were selected

Following retrieval of the 112 papers identified, a screening process was performed in which titles and abstracts were analysed for relevance to the research question. The inclusion criteria focused on papers with potential for design automation use. The two main authors acted as the reviewers and evaluated independently a set of 10 papers to test the paper inclusion criteria and scoring logic to position each proposed LLM support on the triangle model. The updated scoring criteria, as presented in [Figure 2](#), was used to evaluate 10 more papers (sampled 5 from conferences and 5 from journals) to confirm the criteria and reduce inter-reviewer bias. The remaining papers (92) were divided evenly and ranked by one reviewer only.

The evaluation criteria focused on three questions, whose purpose was to:

- Question 1: Quantify the split of design responsibility between the designer and the computational support (consisting of LLMs, traditional tools or both)
- Question 2: Quantify the freedom (agency) of the LLM in selecting the steps to achieve the design automation support, vs the static material agency of traditional tools and
- Question 3: Quantify the use of the LLM in their generative vs orchestration capacity.

Question 1: What is the percentage of designer ownership in the design task?	Score
- Designer is not involved at all. Task is fully automated.	0 %
- Designer has limited involvement after an initial task setup. There is no human oversight.	25 %
- Designer and computational support share responsibilities. Designers set goals and make key decisions, with LLMs/tools handling routine aspects.	50 %
- Designer is central to the design process. Human oversight and frequent intervention in decisions are essential; human guides and adapts all outputs.	75 %
- Designer does not use any computational support.	100 %

(Continued)

Continued.

Question 2: Within the computational support, how much does the LLM orchestrate the task activities or sub-steps?	Score
- LLM not used in orchestrating the workflow. All logic is hard coded into the tool. (e.g. Workflow manager, Python script, KBE, Excel) or directed by the designer.	0 %
- A hardcoded workflow exists, in which the LLM can skip steps based on live results.	25 %
- A basic workflow or series of sub-workflows exist, from which the LLM has the freedom to decide which one to execute next based on the previous step outputs and the overall objective (e.g. graph routing, multi-agent workflow).	50 %
- LLM decides that steps should be taken supported by external guidelines in their input prompt or creates a new workflow given a list of possible steps or creates new steps.	75 %
- LLM decides what steps should be taken independently (e.g. full agentic behaviour).	100 %
Question 3: Within the computational support, how much is the LLM involved in performing the design task?	Score
- LLM not involved on at all. The activity is performed by traditional tools.	0 %
- LLM passes inputs to tools, based on the previous step output or designer instruction.	25 %
- LLM chooses the tools that should be used and prepares their inputs and read outputs.	50 %
- LLM combines the use of tools and self-generated answers.	75 %
- LLM performs the operation without the support of external tools. The next task sees the direct output of the LLM.	%

Example of paper score

Question 1 score (Q1): 20 %
 Question 2 score (Q2): 75 %
 Question 3 score (Q3): 100 %

- Contribution/coordinates of each player:
- 1 - % designer = Q1 = 20%
 - 2 - % LLM = $(100 - Q1) \times Q2 = 60\%$
 - 3 - % tool = $1 - \% \text{ designer} - \% \text{ LLM} = 20\%$

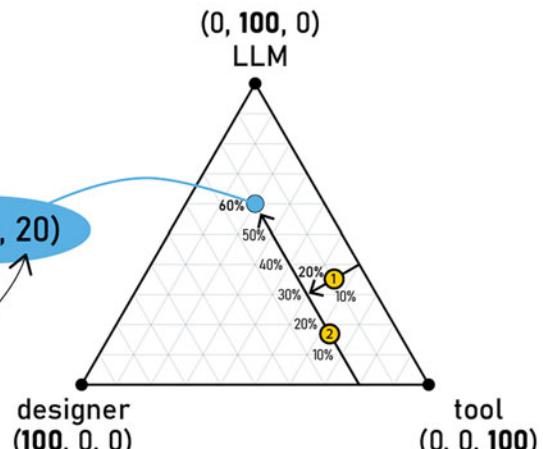


Figure 2. The table above presents the list of questions along with suggested score examples; Below, a calculation example shows how to position a score within the triangular coordinates

3. Results

This section proposes a model to quantify and visualize the contribution of LLMs in engineering design tasks, followed by the results of the industrial feedback and literature mapping.

3.1. The triangle of design responsibility model definition

The triangle of design responsibility addresses challenges in balancing human and technological agency in achieving design automation goals, particularly with the integration of LLMs. Drawing on Leonardi's affordance model (Leonardi 2011), it conceptualizes how the interplay between human designers, traditional design tools and LLMs influence the engineering design process. Responsibility hereby refers to the accountability and control over decision-making and task execution outcome ownership, which is shared among the designer, traditional tools and LLMs.

In essence the design automation goal dictates the required actions to be taken which in turn rely on the interplay between designer and technological agency. Together these agencies create an infrastructure in which the technological agency can either afford or constrain the potential to achieve the desired goal. The technological agency in this context consists of methods, whereby digital engineering tools such as

CAD-software, simulation-software and coded solutions are embodiments of these methods. The use of LLMs in the design process introduces a new and dynamic form of technological agency. Unlike traditional methods which are limited to the static execution of pre-defined tasks, LLMs possess the ability to orchestrate. They can act as collaborative entities, dynamically interpreting inputs, generating outputs and potentially reshaping workflows by not only executing specific tasks but also interacting with other methods. By distinguishing LLMs as a third player in the triangle model, it aims to more accurately capture the evolving and complex interplay between designers, traditional methods and LLMs. In the context of this ICED conference theme - “*Design is a team sport*” - the design owners will be referred from now on as the *players* in the design activity. [Figure 3](#) introduces the model and positions the different types of LLM support frameworks.

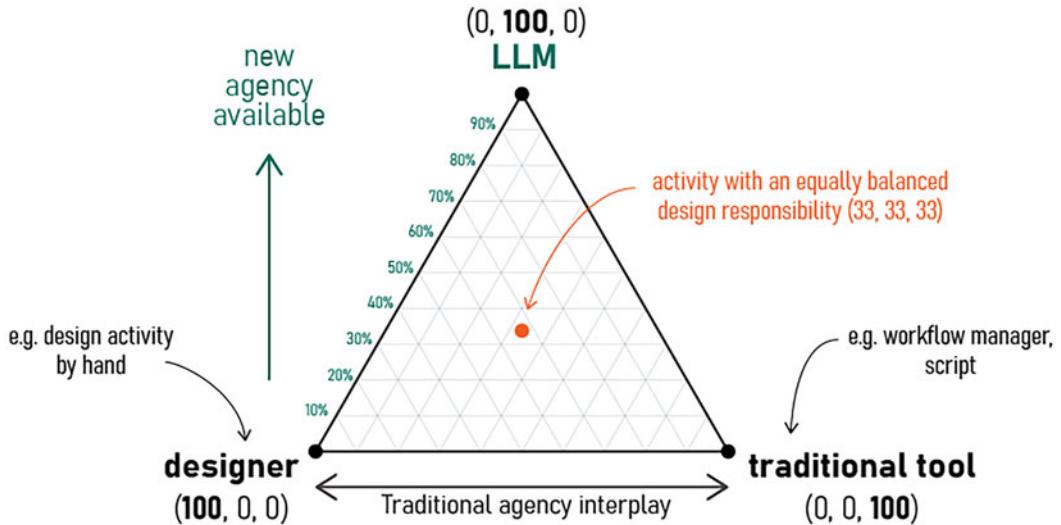


Figure 3. Conceptual model of the triangle of design responsibility, with an example in orange of an execution of a design activity where the responsibility is equally shared by the three players

The vertices of the ternary diagram symbolize the three players. The designer embodies human agency, responsible for setting objectives, able to make decisions and guiding the overall design process. Tools represent traditional methods, able to execute predefined, hardcoded tasks with narrowly defined parameters. The language model emerges as the novel player capable of orchestrating workflows and collaborating with both designers and tools to enhance automation. The diagram illustrates the distribution of responsibility, with ownership percentages totalling 100% and varying based on the design task and product. For example, a chat interaction between a designer and a LLM-based system - with knowledge of design practices - suggesting next activities could be positioned between the designer and the LLM around (50, 50, 0). Likewise, a LLM system embedded in a workflow with the purpose of routing inputs to the appropriate traditional tool would be positioned in the surrounding of (0, 50, 50) if they have a balanced contribution.

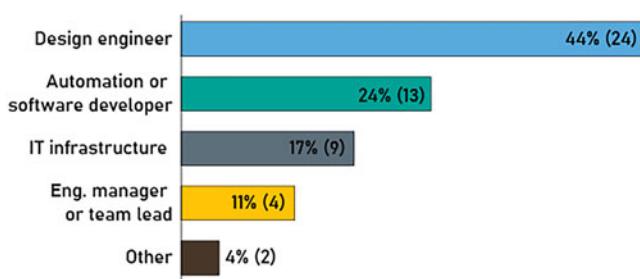
3.2. Survey

The results from the 54 respondents (34 from GKN, 20 from MAN) from 5 independent offices are presented in [Figure 4](#). In addition to the questions presented, participants were encouraged to make use of a 6th free-text field to share further insights, which were used to supplement and contextualize the survey findings.

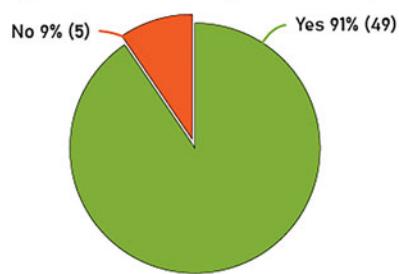
3.3. Literature review

The scores of the 40 papers assessed against the 3 questions and criteria described in [Figure 2](#) returns 3 different dimensions ranging from 0 to 100%. Q2 and Q3 dimension scores were combined for visualization purposes acting as coordinates. Papers were then grouped based on those coordinates and represented as blue bubbles in the chart, with the diameter representing the quantity of papers at that coordinate. In addition, each proposition of LLM design support was mapped against one or more development phases of the engineering design process (Pahl and Beitz [2007](#)).

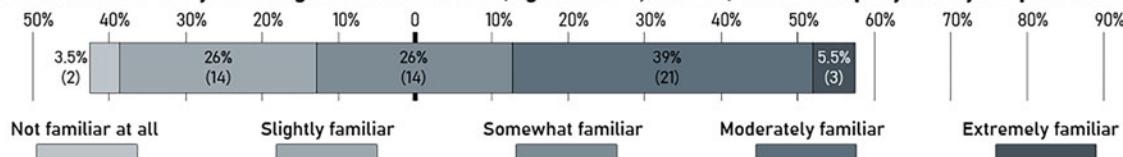
Q1: Survey participant role distribution (n = 54)



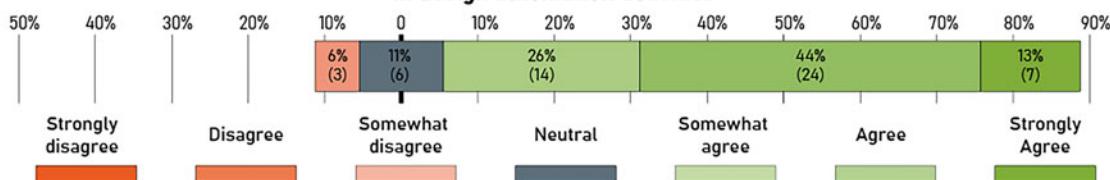
Q2: Did you understand the concept of the triangle model as it was presented to you?



Q3: How familiar are you with generative AI tools (e.g. ChatGPT, CoPilot) at the company or in your personal life?



Q4: The triangle model is a useful model to position and visualize the contribution of Generative AI in design automation activities



Q5: How useful do you think GenAI frameworks that combine human, generative AI and your usual tools could be in supporting part of your daily engineering tasks?

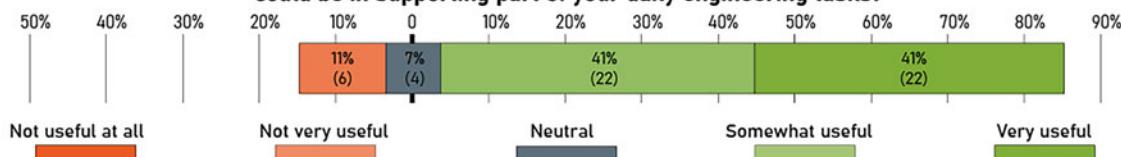


Figure 4. Survey results

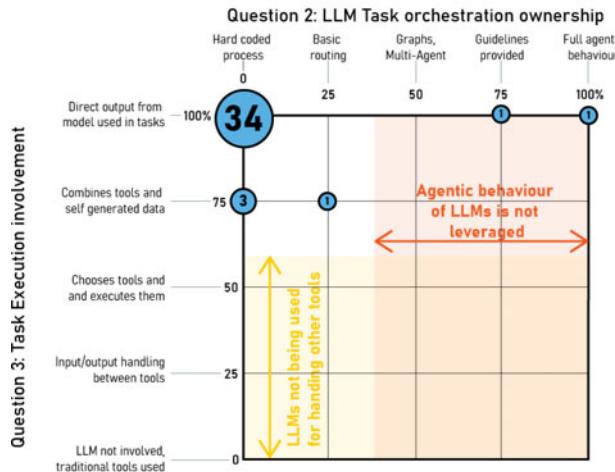
4. Discussion

4.1. Survey results interpretation

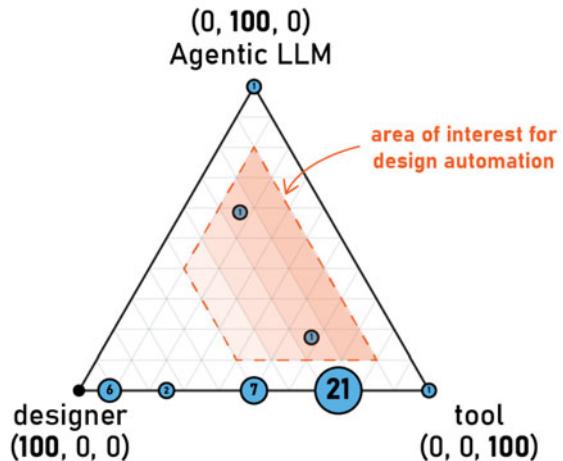
The Q1 results show a varied participant role distribution, covering all stakeholders involved in design automation activities, with those most directly affected by the integration of GenAI into engineering tasks making up the largest groups of respondents. The survey results showed that most participants were to some extent familiar with GenAI language models, with only 3.5% reporting no familiarity at all. While most have some kind of experience with GenAI language models, there is a considerable variation in familiarity which may also represent a challenge in adopting GenAI technology.

A 91% understanding rate is considered sufficient by the authors to conclude that the model is simple to understand in industry. Participants largely agreed on the usefulness of the triangle model, with 83% expressing a positive sentiment. This suggests that the model provides a practical and effective way to frame and leverage the interaction between designers, traditional methods as well as language models for solving engineering problems in their agentic capabilities. One Respondent (ID#36) noted, “*This is exactly what we need to implement the GenAI ideas which can increase the efficiency of users [...]!*” underscoring the model’s potential to guide the integration of GenAI into engineering workflows.

(a) Mapping literature scores of Q2 vs Q3



(b) Mapping scores of Q1 vs Q2 on presented model



(c) In which phases of the product development process does the literature propose to use LLMs?

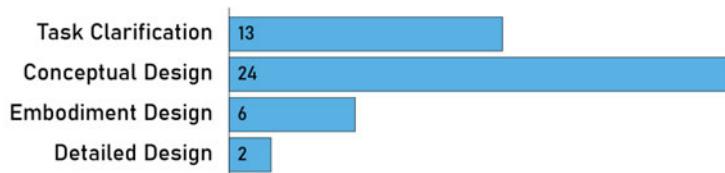


Figure 5. Top-left (a) shows a count of papers positioned against Q2 and Q3; Top-right (b) shows the reviewed papers positioned on the triangle of design responsibility; Bottom (c) maps the activities against the design development phases

However, the same respondent also acknowledged that “[...] there are a lot of challenges to get GENAI implementation 100% as we can't replace sometimes conventional way of working [...]”. This perspective, although stemming from an individual automation/software developer, captures a recurring theme in the survey: despite the positive reception, many participants highlighted the challenges of implementation, particularly in environments with numerous traditional methods and workflows. The GenAI examples that industry practitioners have experienced (e.g. ChatGPT) were not connected at all with the tools that must be used in real product development activities.

As a Design Engineer (ID#38) stated, “[...] it is going to be very difficult to integrate generative AI models into our processes as one complete system because we use so many different design tools at different stages.” This highlights the perceived complexity of utilizing GenAI's orchestration capabilities within complex engineering workflows. Echoing the need for practical solutions, a team lead (ID#48) suggested, “[...] to come up with very specific and simple use cases that could be demonstrated”, highlighting the importance of concrete examples to build confidence in the technology in industry. The literature review reveals a lack of papers investigating let alone demonstrating the potential of integrating GenAI within engineering processes while combining the use of traditional methods.

Regarding the general usefulness of GenAI frameworks that combine human, traditional methods and GenAI LLMs, the survey revealed an overall positive and optimistic view, with 82% expressing a favourable sentiment and only 11% who slightly oppose. However, comments revealed a more nuanced perspective, with recurring concerns about reliability and risks. For example, a Design Engineer (ID#39) remarked, “We need measures to make sure GenAI is more robust and reliable for engineering if it is to be used at all” reflecting a recurring theme of scepticism about reliability. Another participant with a software development background (ID#34) added “This shares my vision of how generative AI can be useful as well, but I'm concerned about the risk of the AI misinterpreting or hallucinating with regards to the instructions given, and don't see how that risk is mitigated.”

These mixed responses highlight two key aspects: strong enthusiasm for GenAI's potential to transform engineering workflows as well as persistent concerns about reliability, integration challenges, and ensuring the technology affords rather than constrains the goals of the engineer. Addressing these issues is crucial for the successful adoption of GenAI as mirrored in previous ED research by Kretzschmar et al. (2024a, 2024b). The triangle model seeks to contribute to this by providing a way to conceptualize and

balance the contributions of human expertise, traditional methods, and GenAI language models in engineering design, to determine their successful interplay in the design process.

4.2. Literature review interpretation

The results in [Figure 5 \(c\)](#) show that only 8/40 papers (20%) cover the later phases of the design (embodiment and detailed). These phases are precisely the most time consuming and most complex to perform and automate for the type of products that this paper focuses on (complex products, physical models). Therefore, we identify a research gap that could be further researched.

The combination of the literature Q2 and Q3 presented in [Figure 5 \(a\)](#) highlights two gaps. First, most studies (38/40) use LLMs or GenAI models as tools, leveraging only their material agency and not using their dynamic capabilities ([Leonardi 2011](#)). An example from other fields where LLMs are explored to behave in human agency is Voyager ([Wang et al. 2023](#)). Second, none of the papers allowed the LLMs the freedom to use the tools needed to execute the design task. Only 4 of them combined the generation of the LLM output and some external, hard-coded rationale or support. This is important, as the results of the survey and discussion with industry show the importance of keeping traditional tools in any future GenAI framework. Other literature outside the ED research area that would fill the bottom right area of [Figure 5 \(a\)](#) are MechAgents ([Ni and Buehler 2024](#)) a framework that leverages a multi-agent structure to solve mechanical problems via FEM coding, or the use of RAG and simulation files to evaluate heat treatments effects on 3D parts ([Sun et al. 2024](#)).

Finally, [Figure 5 \(b\)](#) shows that since LLMs are being used as traditional tools, and prevented from executing their agentic properties, they lie on the bottom line on the design responsibility triangle, with the rest of traditional tools. Only two papers give LLMs, designers and tools a shared responsibility on design activities ([Pradas Gómez, Panarotto, et al. 2024](#); [Li et al. 2024](#)). Only one paper, at the top of the triangle, fully delegated design responsibility to the LLM. This paper best exemplifies this corner of the model, where during a hackathon competition researchers delegated all the design decisions to ChatGPT and acted as physical executors of the LLM's decisions in a workshop environment ([Ege et al. 2024](#)). An example in the “gap zone” of [Figure 5 \(a\)](#) and [\(b\)](#) from outside the ED community combining agentic behaviour and external tool usage is HuggingGPT ([Shen et al. 2023](#)). This implementation uses a LLM to process top level tasks request from the human and answers them using a generic 4 step process: *task planning* where it decides what to do and in what order, *model (tool) selection*, *execution of the task* and finally *response generation*; gathering all results of the workflow and presenting them to the user.

4.3. Limitations

The survey was subject to response biases (acquiescence and courtesy), mitigated by providing the survey to be anonymous and framing it as a constructive feedback exercise. Another limitation is the short and shallow exposure of the participants to the model (15 minute explanation). Though they grasped the concept, more time could have been spent to give higher quality feedback. This activity is planned as part of future research, when participants will interact with the model via their example cases to score and position them themselves in the triangle model. The feedback will be used to improve the definition of the questions and provide score example guidance, increasing model interpretability.

The literature review focused on a limited subset of databases and research areas, meaning the conclusions apply primarily to the Design Society’s scope. This limitation arose from time constraints and the number of reviewers. An initial query search on Scopus, unrestricted by specific journals and including terms like “product development” or “complex products”, yielded thousands of results, many being false positives. Refining the search using filters still left over 300 documents, leading to a narrower scope aligned with the Design Society’s focus and the conference theme. The ArXiv repository was excluded as a primary search database to balance content quality and a manageable paper count, though relevant papers may exist there. Finally, while the scoring criteria for the papers were explicitly defined, we acknowledge that an element of subjectivity remains inherent in the scoring process.

4.4. Answer to research questions

To answer RQ1 (How can designers and design automation developers structure and visualize the contribution of Generative AI, specifically LLMs, within design automation systems?) a model was developed and validated through an industry survey, with relevant ED literature positioned against it. The

literature review for RQ2 (How have LLMs been integrated into design processes within the engineering design community?) also answers RQ3 (What are the research gaps in the application of LLMs for engineering design support?). Observing [Figure 5](#), it shows that LLMs have primarily been used as traditional tools, with varying levels of design delegation, generally without support from additional tools. Their implementations have not leveraged agentic capabilities, show different levels of design delegation between human designers and these tools, and primarily focus on early design phases.

Combined with the industrial feedback, we identify a research gap for the design automation of complex products, such as the structural components of vehicle systems throughout every phase of the design process, but with a higher interest in the later phases of the design, due to the amount of engineering hours available to automate. There is a lack of research on frameworks that combine traditional tools to ensure the accuracy and performance of the tasks with LLMs providing flexibility on the execution of the task and adapting to designer instructions. The designer, as unanimously described by industry practitioners, shall remain in the driver's seat. Finally, industry is eager for a practical implementation of such a framework, ensuring research is not only theoretical but also executable and testable. This will help assess whether the technology meets societal expectations and measure its effectiveness in practice.

5. Conclusion

A model for positioning the support of Generative AI in design automation has been presented. It highlights the different nature of design responsibility ownership of each contributor: Designers, traditional tools and LLMs in their agentic role. The proposed model, together with the ranking criteria, is simple to understand and can be used in industry to clarify the expectation of each player in design automation activities. As a result, each of the three team members will be played using their strengths, resulting in more effective automation systems. This combination can support designers in a novel way and allow them to increase the flexibility of the design automation. Also, they can use the benefits of highly hard-coded tools without the need for the system users to be experts on the system itself. Therefore, we see a research opportunity to leverage LLMs agentic capabilities to automate some of the tasks that are difficult to do with traditional automation techniques, due to complexity and development effort. Future research must also address the challenge of ensuring reliability and reproducibility in LLM-driven automation, as designers and engineers must be able to trust that results are both consistent and dependable, a critical requirement for design automation solutions.

References

- Cederfeldt, Mikael, and Fredrik Elgh. 2005. "Design Automation in SMEs—Current State, Potential, Need and Requirements." *In DS 35: Proceedings ICED 05*.
- Chiarello, Filippo, Simone Barandoni, Marija Majda Škec, and Gualtiero Fantoni. 2024. "Generative Large Language Models in Engineering Design: Opportunities and Challenges." *Proceedings of the Design Society 4* (May): 1959–68. doi: [10.1017/pds.2024.198](https://doi.org/10.1017/pds.2024.198).
- Ege, Daniel Nygård, Henrik H. Øvrebø, Vegar Stubberud, Martin Francis Berg, Martin Steinert, and Håvard Vestad. 2024. "Benchmarking AI Design Skills: Insights from ChatGPT's Participation in a Prototyping Hackathon." *Proceedings of the Design Society*.
- Heikkinen, Tim. 2021. "Transparency of Design Automation Systems Using Visual Programming—within the Mechanical Manufacturing Industry." *Proceedings of the Design Society 1*. Cambridge University Press.
- Kretzschmar, Maximilian, Maximilian Peter Dammann, Sebastian Schwoch, Elias Berger, Bernhard Saske, and Kristin Paetzold-Byhain. 2024a. "Key Concepts, Potentials and Obstacles for the Implementation of Large Language Models in Product Development." *In Entwerfen Entwickeln Erleben 2024: Menschen, Technik Und Methoden in Produktentwicklung Und Design*, 381–93. Technische Universität Dresden. doi: [10.25368/2024.EEE.031](https://doi.org/10.25368/2024.EEE.031).
- Kretzschmar, Maximilian, Maximilian Peter Dammann, Sebastian Schwoch, Felix Braun, Bernhard Saske, and Kristin Paetzold-Byhain. 2024b. "Evaluating the Role of Generative AI in Product Development and Design - A Systematic Review." *In Proceedings of NordDesign 2024*, 21–30. The Design Society. doi: [10.35199/NORDDESIGN2024.3](https://doi.org/10.35199/NORDDESIGN2024.3).
- Leonardi, Paul M. 2011. "When Flexible Routines Meet Flexible Technologies: Affordance, Constraint, and the Imbrication of Human and Material Agencies." *MIS Quarterly* 35 (1). Management Information Systems Research Center, University of Minnesota: 147–67. doi: [10.2307/23043493](https://doi.org/10.2307/23043493).

- Li, Mingrui, Zuoxu Wang, Zhijie Yan, Xinxin Liang, and Jihong Liu. 2024. "Promoting Knowledge Recommendation in Innovative Engineering Design: A BERT-GAT-Based Patent Representation Learning Approach." *Journal of Engineering Design*, April ,1–26. doi: [10.1080/09544828.2024.2339713](https://doi.org/10.1080/09544828.2024.2339713).
- Ni, Bo, and Markus J. Buehler. 2024. "MechAgents: Large Language Model Multi-Agent Collaborations Can Solve Mechanics Problems, Generate New Data, and Integrate Knowledge." *Extreme Mechanics Letters* 67 (March): 102131. doi: [10.1016/j.eml.2024.102131](https://doi.org/10.1016/j.eml.2024.102131).
- Page, Matthew J., Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, et al. 2021. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." *BMJ* 372 (March). British Medical Journal Publishing Group: n71. doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71).
- Pahl, G., and W. Beitz. 2007. *Engineering Design: A Systematic Approach*. 3rd ed. London: Springer.
- Pradas Gómez, Alejandro, P. Krus, M. Panarotto, and O. Isaksson. 2024. "Large Language Models in Complex System Design." In , 4:2197–2206. doi: [10.1017/pds.2024.222](https://doi.org/10.1017/pds.2024.222).
- Pradas Gómez, Alejandro, Massimo Panarotto, and Ola Isaksson. 2024. "Evaluation of Different Large Language Model Agent Frameworks for Design Engineering Tasks." *DS 130: Proceedings of NordDesign 2024*, Reykjavik, Iceland, 12th-14th August 2024, 693–702.
- Rad, Mohammad Arjomandi, Kent Salomonsson, Mirza Cenanovic, Henrik Balague, Dag Raudberget, and Roland Stolt. 2022. "Correlation-Based Feature Extraction from Computer-Aided Design, Case Study on Curtain Airbags Design." *Computers in Industry* 138. Elsevier: 103634.
- Sellgren, Ulf. 2005. "A Situated Question-Driven and Model-Based Approach to Design Reasoning." In *DS 35: Proceedings ICED 05*.
- Shen, Yongliang, Kaitao Song, Xu Tan, Dongsheng Li, Weiming Lu, and Yueling Zhuang. 2023. "HuggingGPT: Solving AI Tasks with ChatGPT and Its Friends in Hugging Face." arXiv. doi: [10.48550/ARXIV.2303.17580](https://arxiv.org/abs/2303.17580).
- Sun, Yixiao, Xusheng Li, Chao Liu, Xiaohu Deng, Wenyu Zhang, Jiangang Wang, Zeyu Zhang, Tengyang Wen, Tianyu Song, and Dongying Ju. 2024. "Development of an Intelligent Design and Simulation Aid System for Heat Treatment Processes Based on LLM." *Materials & Design* 248 (December): 113506. doi: [10.1016/j.matdes.2024.113506](https://doi.org/10.1016/j.matdes.2024.113506).
- Verhagen, W.J.C., P. Bermell-Garcia, R.E.C. Van Dijk, and R. Curran. 2012. "A Critical Review of Knowledge-Based Engineering: An Identification of Research Challenges." *Advanced Engineering Informatics* 26 (1): 5–15. doi: [10.1016/j.aei.2011.06.004](https://doi.org/10.1016/j.aei.2011.06.004).
- Wang, Guanzhi, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. 2023. "Voyager: An Open-Ended Embodied Agent with Large Language Models." arXiv. <http://arxiv.org/abs/2305.16291>.