Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 00 (2025) 000–000

www.elsevier.com/locate/procedia

CIRP Conference on Manufacturing Systems (CIRP CMS)

Automated production line configuration technology through the integration of LLM and mathematical optimization techniques

Kaichiro Nishi[a,*], Daisuke Tsutsumi[a], Takahiro Nakano[a], Youichi Nonaka[a], Gianfranco Pedone[b], András Kovács[b], Krisztián Balázs Kis[b], Ágoston Dalotti[b], József Váncza[b]

[a]*Hitachi, Ltd. Research and Development Group, 292, Yoshida-cho, Totsuka-ku, Yokohama-shi, Kanagawa-ken, 244-0817, Japan*

[b]*Research Laboratory on Engineering & Management Intelligence, Institute for Computer Science and Control (SZTAKI), Hungarian Research Network (HUN-REN), Kende u. 13-17, Budapest, 1111, Hungary*

* Corresponding author. Tel.: +81-80-2098-2561. *E-mail address:* kaichiro.nishi.kt@hitachi.com

Abstract

In manufacturing sites, rapid reconfiguration of production lines is essential to flexibly respond to product variations and demand fluctuations, and to maximize productivity. Line configuration decisions are made in two stages: "production strategy decision" and "line configuration design". However, the formulation of production strategies requires information dispersed across multiple sites and systems, as well as expert knowledge, making automation a long-standing challenge. This study focuses on the natural language processing capabilities of large language models (LLMs) and proposes a method that generates production strategies from user requirements and converts them into parameters for mathematical optimization models used in line configuration design. A key issue arises when knowledge about the correspondence between production strategies and parameters is insufficient, leading to reduced conversion accuracy. To address this, we developed a method that divides LLM processing into three stages "production strategy decision", "parameter item decision", and "parameter value decision". At each stage, the method references hierarchically structured manufacturing knowledge. Furthermore, we integrated causal relationships between ISO 22400 compliant line design parameters (e.g., CT, operating time) and cost KPIs (e.g., labor costs, depreciation) into the knowledge model, enabling prediction of cost fluctuations resulting from parameter changes. This allows LLMs to generate highly feasible strategies that take cost into account. The effectiveness of the proposed method was demonstrated through an industrial case study targeting a battery production line. Compared to the conventional method of manually configuring parameters for optimization engines, the proposed approach reduced the effort required by engineers to design production line configurations by 87.3%.

Keywords: Production Line Configuration; Large Language Models; Mathematical Optimization Model

1. Introduction

In manufacturing sites, it is essential to rapidly reconfigure production lines in order to flexibly respond to various production fluctuations such as changes in product types and demand volume while maximizing productivity. The process of determining line configurations consists of two stages: first, *production strategy decision*, which defines the response policy to fluctuations; and second, *line configuration design*, which involves dividing processes and allocating equipment based on the selected strategy. In the production strategy decision stage, the response policy is defined based on past experience while referencing production information distributed across multiple sites and systems. For example, if product demand increases beyond expectations, strategies such as shortening cycle time, increasing overtime hours, or adding outsourced processes are considered, and the optimal strategy is determined. This task requires expert knowledge based on past experience, and only a limited number of engineers are capable of handling it. In the *line configuration design* stage, detailed tasks are divided into multiple processes, and equipment is allocated to each process. The combinations of process divisions and equipment assignments are vast, making it a time-consuming problem even for experienced engineers.

To address these issues, we consider eliminating dependence on expert knowledge through automation. For the production strategy decision task, the use of large language models (LLMs) capable of natural language processing and external knowledge referencing is a promising approach. For instance, in the task of considering robot motion learning strategies, there is a case study utilizing LLMs [1], suggesting that production strategies can be similarly examined. Line configuration design can be formulated as a combinatorial optimization problem involving process divisions and equipment selection. In our previous research [2], we developed tools that automatically compute configurations using mathematical optimization techniques. By setting appropriate parameters (e.g., CT, operating time) in the mathematical optimization model embedded in the tool, line configurations can be automatically generated.

Furthermore, it is necessary to automatically convert production strategies into parameters required by the *mathematical optimization model*. Identifying parameters related to production strategies and setting appropriate values requires tool-specific expertise, and even skilled engineers have found the task to be time-consuming. When automating this process using LLMs, insufficient knowledge about the correspondence between production strategies and parameters can lead to reduced conversion accuracy. Fig. 1 shows an actual example of a conversion failure. For instance, when the LLM determines a production strategy such as reduce cycle time, it is necessary to select and modify the parameters related to that strategy. However, since there are multiple time-related parameters, the LLM cannot determine which parameter to modify or how to modify it, resulting in incorrect conversion.

Therefore, this paper proposes a method for accurately converting information between different formats: from production strategies expressed in natural language information into mathematical optimization model that require structured information.

2. Related Works

Reducing human burdens both physically and intellectually is a critical challenge in future smart factories [3]. LLMs have the potential to address this challenge, and their application in manufacturing operations is rapidly advancing [4]. In the manufacturing domain, the application of LLMs has been advancing particularly in the areas of product design [5], robotics technology [1], and production system research [6]. These fields have seen significant adoption of systems and tools, and much of the data handled can be converted into text format, which contributes to their high compatibility with LLMs.

Within these, the task of determining line configurations, which falls under the production system domain, has also seen progress in LLM application research. Dimitropoulos et al. [7] proposed a method utilizing LLMs for appropriate task allocation in human-collaborative operations. By instructing the LLM to prioritize human tasks based on user input, task assignments that consider workload are automated. Li et al. [8] proposed a method that enables user interaction in line configuration generation using image generation models, by leveraging LLMs. Specifically, the LLM converts user instruction text into the format required by the image generation model.

A common feature of previous studies utilizing LLMs is that users are expected to formulate specific strategies. In the case of production strategies for line configuration decisions, these strategies are typically developed by integrating information distributed across multiple sites and systems, along with the expertise of skilled engineers. This results in a time-consuming decision-making process. One approach to address this issue is the use of modeling technologies such as SysML [9]. However, defining production strategies for the vast number of production variation patterns is not realistic. Therefore, this study proposes a method in which the LLM consistently performs the entire process from strategy formulation, which corresponds to specific user instructions, to parameter conversion and execution of the mathematical optimization model embedded in the production line configuration generation tool. By adopting this approach, it becomes possible to reduce the burden on line engineers while enabling rapid line configuration decisions, thereby improving responsiveness to production fluctuations.

The contributions of this study are as follows:

- Proposal of a method that enables rapid line configuration decisions by fully automating the process from production strategy to parameter determination for the mathematical optimization model embedded in the line configuration tool.
- Proposal of a knowledge management method that ensures accurate conversion between different information formats, such as natural language (e.g., production strategies) and structured information from mathematical optimization model.
- Evaluation and validation of the proposed method's effectiveness.

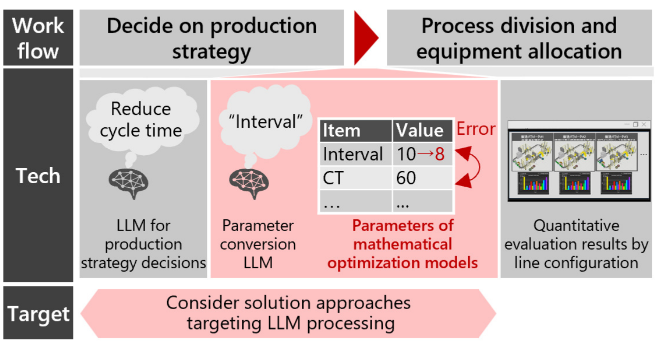

Fig. 1. Workflow for line configuration determination and problem of conventional technologies.

3. Methods

3.1. Overview of proposed method

Fig. 2 shows an overview of the proposed approach for consistently automating the process from production strategy to parameter determination for the mathematical optimization model embedded in the line configuration tool. Based on user input,

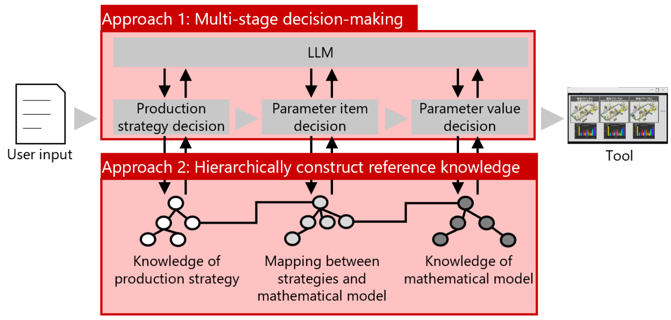

Fig. 2. Workflow of the proposed method.

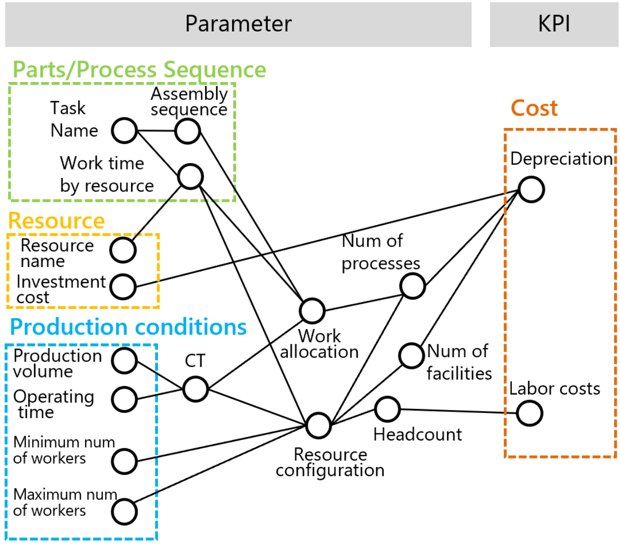

Fig. 3. Knowledge of production strategy.

a single LLM dynamically switches the referenced knowledge and sequentially determines both the production strategy and the parameters for the optimization model, ultimately generating the line configuration automatically using the tool. There are two main approaches.

The first approach involves structuring the process into multiple stages. LLMs exhibit higher inference performance when instructions are clarified in detail [10]. Additionally, when skilled engineers determine parameters using the line configuration tool, their thought process typically involves first identifying which parameters are related to the production strategy, followed by determining the parameter values. Therefore, the previously unified parameter conversion process is divided into two distinct steps: parameter item decision and parameter value decision. While this separation enhances the inference performance of the LLM, it introduces a challenge in maintaining consistency across the entire process.

This issue is addressed by the second approach: constructing hierarchical reference knowledge. By not only building and referencing knowledge appropriate to each LLM process, but also linking the knowledge across stages, the method ensures overall consistency, thereby enabling both improved accuracy and comprehensive automation. The element that plays the role of linking knowledge is *mapping between strategies and mathematical model*. Specifically, by clarifying the relationships between production strategies and the parameters of mathematical optimization model, inconsistencies in information across processes are prevented.

Details of each LLM's processing will be described in the following sub sections.

3.2. Production strategy decision

Fig. 3 shows a part of the knowledge of production strategy. This knowledge integrates the causal relationships between line design parameters (e.g., CT, operating time) and cost KPIs (e.g., labor cost, depreciation cost), both defined in accordance with ISO 22400, while also considering assembly specific constraints such as assembly sequence. Based on these integrated causal relationships, the LLM can predict cost fluctuations resulting from parameter changes and generate highly feasible production strategies that take cost into account. To allow the LLM to refer to this knowledge, the causal relationships described in text are expressed in JSON format, enabling the LLM to retrieve highly relevant information based on similarity to the user's input text.

For example, when the user inputs a sentence such as "Product demand has increased, and production volume will rise by 20%", the LLM begins by exploring nodes related to production volume parameter. At this point, the CT parameter is referenced first. The CT parameter is connected via edges to the depreciation and labor costs parameters through several intermediate parameters, and the edges indicate that changes in CT will affect both depreciation and labor costs. Based on this relational information, the LLM can predict that shortening the CT will increase the number of processes, leading to higher depreciation costs, and also require more workers, resulting in increased labor costs. With this understanding, the LLM can output a production strategy proposal in text format, such as considering CT reduction, together with the reasoning behind that decision. Outputting the rationale together with the result is important for enabling users to understand the basis of the production strategy determined by the LLM, and it enhances the transparency of the decision-making process. When multiple production strategies can be considered, the LLM preferentially outputs the candidate that results in the minimum cost while satisfying the given constraints.

3.3. Parameter item decision

Fig. 4 shows a part of the mapping between strategies and mathematical model. This knowledge describes the correspondence between production strategies and parameters of the mathematical optimization model, enabling the determination of which parameters should be reflected based on the selected strategy.

For example, when the strategy to shorten the CT is selected, the system refers to the *cycle_time entry* in this knowledge to determine which parameter in the mathematical optimization

```
"cycle_time": {
    "a": "Maintain the existing cycle time restriction.",
    "b": "Set the cycle time limit to {AAA} seconds to match the
        required adjustment in production volume, which
        inversely correlates with cycle time.",
    "c": "Configure the cycle time limit to zero, indicating no
        specific restriction. This setting is valid when optimizing
        for the shortest possible makespan."
},
"init_configuration": {
…
```

Fig. 4. Mapping between strategies and mathematical model.

model should be modified and how. In this example, the *cycle_time* parameter in the mathematical optimization model is defined with three modification options. In case a, the parameter is not modified, which applies when the strategy to shorten CT has not been selected. In case b, the parameter value is determined in inverse proportion to the change in production volume, which applies when the strategy to shorten CT has been selected. In case c, the value is set to zero, allowing the tool to calculate and apply the optimal CT automatically. This applies when the user provides an instruction to optimize without specifying a concrete production volume. Based on this modification information, along with the user input and the selected production strategy, the LLM can infer that the parameter value should be determined according to option b, in inverse proportion to the change in production volume.

The selected parameters to be modified and their corresponding modification details are output in JSON format.

3.4. *Parameter value decision*

Fig. 5 shows a part of the knowledge of mathematical model. This knowledge includes information such as the values, meanings, and formats of parameters used in the mathematical optimization model. Based on this knowledge and the output from the preceding LLM process, parameter values are configured in a format executable by the tool.

In this example, since *cycle_time* needs to be set inversely proportional to a 20% increase in production volume, the value is changed from 60 to 48. This modification is output in JSON format.

Finally, by having the tool read and execute the updated parameter where *cycle_time* is set to 48, the line configuration can be redesigned. In this way, by referencing hierarchically structured knowledge at each LLM processing stage, it becomes possible to accurately convert information even when the format differs between the input text and the required parameter representation. Even if an infeasible parameter value is output due to issues such as hallucinations, the tool returns an error, allowing the user to recognize that the LLM has produced an inappropriate parameter value.

#	Parameter	Value	Definition	...	Format
1	interval	10	Changeover time	...	Int
2	cycle_time	60	Time required to produce one unit	...	Double
...	...	...	...	...	...
200	worker_position	[0,0]	Worker's X and Y coordinate positions	...	Int Array

Fig. 5. Knowledge of mathematical model.

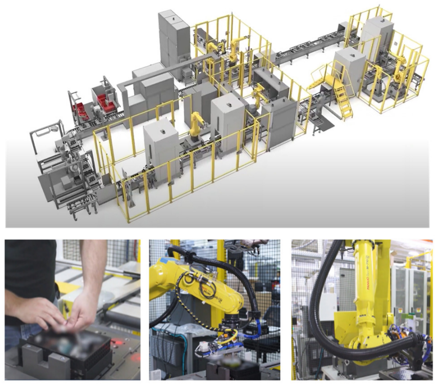

Fig. 6. Outline of battery assembly line.

4. Experimental evaluation

4.1. *Verification Target*

In this study, the proposed method was verified using a battery assembly line as the target. An overview of the target battery assembly line is shown in Fig. 6. This line consists of 51 tasks executed by eight types of main resources. In the initial line configuration, the 51 tasks are allocated across 16 stations to meet a production volume of 480 units per shift (8 hours), which corresponds to a cycle time constraint of 60 seconds.

4.2. *Preliminary Evaluation*

In the preliminary evaluation, a comparative experiment was conducted to verify the effectiveness of the proposed approach by varying the completeness of the mapping between strategies and mathematical model. In Pattern A, only the cycle time parameter was constructed within the mapping between strategies and mathematical model, whereas in Pattern B, all parameters related to the validation UCs were included. In Pattern A, for parameters other than cycle time, the LLM inferred the relevance to production strategies based on the parameter names.

As verification UCs, five cases were defined based on interviews with experts from the business division, as shown in Table 1. The outputs of each LLM were evaluated in terms of syntax and content. GPT-4o was used without fine-tuning.

- **Syntax:** In the processes of the parameter item decision and the parameter value decision, the LLM is instructed to output in JSON format. The evaluation was conducted to determine whether the output could be correctly parsed as a JSON file by the program.

- **Content:** The output was evaluated by comparing it with model answers created by experienced engineers for production line configuration, assessing whether the content was equivalent.

4.3. Preliminary Evaluation Results

Table 2 shows the results of the preliminary evaluation. From the perspective of syntax, correct output was confirmed for all verification use cases under both conditions. This is attributed to the fact that LLMs are trained using large volumes of textual data such as JSON during model development.

On the other hand, focusing on the content, the production strategy decision process was correctly output in all validation use cases, whereas the parameter item decision process achieved only a 40% success rate in Pattern A. Furthermore, the parameter value decision process produced correct outputs for all use cases where the preceding process had succeeded. This is likely because the task of the production strategy decision process primarily involves textual input and output, which aligns with the domain where LLMs excel. Similarly, the parameter value decision process deals with data formats similar to those used in the preceding process, and the task itself is relatively less complex.

In contrast, the parameter item decision process requires structured information as output, whereas the preceding process involves natural language input. Therefore, the presence or absence of explicitly defined correspondence information significantly affects the success rate.

Based on the preliminary evaluation results, it was confirmed that our development approach improves the conversion success rate from 40% to 100%, even when dealing with different information formats such as natural language data like production strategies and structured data from mathematical optimization models.

4.4. Main Evaluation

In the main evaluation, we conducted an integrated assessment covering the entire process from production strategy decision to parameter determination for the line configuration tool [2], and finally to line configuration generation through tool execution. The evaluation UCs were the same five cases shown in Table 1, as used in the preliminary evaluation.

In the evaluation, skilled engineers used only the line configuration tool, while non-experts utilized both the tool and the LLM-based development approach. For skilled engineers, the tasks between production strategy decision and parameter value decision for the line configuration tool were performed manually. Since the tool requires setting as many as 188 input parameters, even skilled engineers are expected to spend a significant amount of time configuring them appropriately. The evaluation metric was the effort required to produce a line configuration of comparable quality to that of a skilled engineer. GPT-4o was used without fine-tuning.

- **Effort:** The time taken from the start of line configuration consideration to the point where the configuration was finalized by modifying parameters in the tool and generating the output. For non-experts, the process was repeated until the configuration was approved by a skilled engineer.

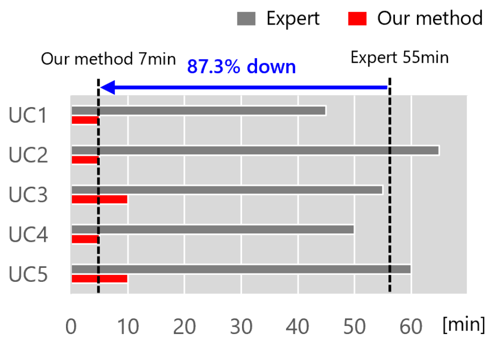

Fig. 7. Comparison of man-hours by expert and proposed method.

4.5. Main Evaluation Results

Fig. 7 presents the results of the main evaluation. This figure compares the effort required between non-experts using the proposed development technology and skilled engineers not using it. While skilled engineers required an average of 55 minutes, non-experts using the proposed method completed the task in an average of 7 minutes, confirming a reduction in effort of 87.3%. Although part of this difference is due to the time skilled engineers spend on parameter value decision, the most significant factor is the number of iterations involved in reworking the task. Skilled engineers tend to repeat the cycle of production strategy decision, parameter value decision, and result verification until a satisfactory line configuration is achieved, which increases the overall effort. In contrast, when non-experts use the proposed development technology, the LLM consistently automates the entire process from production strategy decision to parameter value decision and execution resulting in a substantial reduction in effort. Furthermore, the LLM can refer to standardized models and propose multiple feasible production strategies. As a result, it becomes possible to generate multiple line configurations at once, compare them quantitatively in terms of productivity and investment cost, and quickly determine the optimal configuration.

5. Conclusions and future work

In this study, we proposed a method that enables consistent execution from production strategy to parameter determination for a mathematical optimization model using LLMs, by employing multi-stage processing and hierarchical reference

Table 1. Verification UC

UC	User Input text
1	Event: Forecasts show an increase in the demand of battery packs. Objective: We want to modify our production system configuration to increase the long-term production of battery packs volume by 20%, keeping additional investment costs to a minimum.
2	Event: Forecasts show a decrease in the demand of battery pack products. Objective: We want to reduce the production of battery packs by 30%. However, due to employment regulations, no worker has to be dismissed. We must keep the current number of workers and reorganize the production line.
3	Event: The worker in Cell 1 has called in sick. Objective: Take the necessary measures to ensure that the production levels are maintained unchanged.
4	Event: The production costs have continuously increased in the last 6 months. Objective: We want to lower the production costs by outsourcing part of the production line process. Identify which part of the process should (or could) be outsourced.
5	Event: The robot in the fourth cell has broken down. Objective: We want to shorten the period when production is at a minimum. The delivery time for a robot of the same model as the one that broke is six months. If it is a different model, the delivery time is one week because the manufacturer has stock. However, the work time for a different model is 1.2 times longer.

Table 2. Preliminary verification result

Valuation target	Pattern A		Pattern B	
	Syntax	Content	Syntax	Content
Production strategy decision	-	5/5	-	5/5
Parameter item decision	5/5	2/5	5/5	5/5
Parameter value decision	5/5	2/5	5/5	5/5

knowledge. The proposed method was evaluated in the context of line configuration tasks for a battery manufacturing line. As a result, it was confirmed that non-expert workers were able to generate line configurations of comparable quality to those produced by experts, while reducing the required effort by 87.3%.

This outcome not only extends the capabilities of non-experts, but also contributes to reducing the workload of skilled engineers by decreasing the number of review iterations. Furthermore, by alleviating the burden on engineers, line configurations can be determined more quickly, enabling more agile responses to production fluctuations.

While ISO 22400 provides a foundational KPI framework, real-world decisions such as equipment selection or process consolidation require integration of additional domain knowledge, including lifecycle costs (CAPEX/OPEX), OEE factors, risk metrics associated with operator proficiency, and reliability data (e.g., mean time between failures). Future research will explore expansion of the knowledge model to incorporate these traditionally unstructured factors through integration with enterprise-level data system.

Acknowledgement

This paper is the result of the research funded by the joint project between Hitachi, Ltd. and HUN-REN SZTAKI. Hungarian authors have been partially supported also by the European Union project RRF-2.3.1-21-2022-00004 within the framework of the Artificial Intelligence National Laboratory, and by the National Research, Development and Innovation Office project TKP2021-NKTA-01, on “Research of Cooperative Manufacturing and Logistics Systems to Support a Competitive and Sustainable Economy”.

References

[1] Y. Yin, K. Wan, C. Li, P. Zheng, An LLM-enabled human demonstration-assisted hybrid robot skill synthesis approach for human-robot collaborative assembly, CIRP Annals 74 (1) (2025) 1–5.

[2] D. Tsutsumi, J. Bergmann, P. Dobrovoczki, N. Hayashi, A. Kovács, Z. Szádoczki, Ádám Szaller, Managing fluctuations in production via an optimal portfolio of assembly line configurations, Procedia CIRP 130 (2024) 1372–1377, 57th CIRP Conference on Manufacturing Systems 2024 (CMS 2024).

[3] L. Wang, R. X. Gao, J. Krüger, J. Váncza, Human-centric assembly in smart factories, CIRP Annals 74 (2) (2025) 789–815.

[4] Y. Li, H. Zhao, H. Jiang, Y. Pan, Z. Liu, Z. Wu, P. Shu, J. Tian, T. Yang, S. Xu, Y. Lyu, P. Blenk, J. Pence, J. Rupram, E. Banu, N. Liu, L. Wang, W. Song, X. Zhai, K. Song, D. Zhu, B. Li, X. Wang, T. Liu, Large language models for manufacturing (2024). arXiv:2410.21418.

[5] Z. Jiang, A. Liu, D. Zhang, X. Xu, Y. Dai, Customization and personalization of large language models for engineering design, CIRP Annals 74 (1) (2025) 191–195.

[6] Y. Xia, M. Shenoy, N. Jazdi, M. Weyrich, Towards autonomous system: flexible modular production system enhanced with large language model agents, in: 2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA), 2023, pp. 1–8.

[7] N. Dimitropoulos, M. Kaipis, S. Giartzas, G. Michalos, Generative ai for automated task modelling and task allocation in human robot collaborative applications, CIRP Annals 74 (1) (2025) 7–11.

[8] X. Li, A. Nassehi, S. J. Hu, B. G. Joung, R. X. Gao, A large manufacturing decision model for human-centric decision-making, CIRP Annals 74 (1) (2025) 621–625.

[9] C. Steimer, J. Fischer, J. C. Aurich, Model-based design process for the early phases of manufacturing system planning using sysml, Procedia CIRP 60 (2017) 163–168, complex Systems Engineering and Development Proceedings of the 27th CIRP Design Conference Cranfield University, UK 10th – 12th May 2017.

[10] P. Sahoo, A. K. Singh, S. Saha, V. Jain, S. Mondal, A. Chadha, A systematic survey of prompt engineering in large language models: Techniques and applications (2025). arXiv:2402.07927.