machines

MDPI

Article

Enhancing Bottleneck Analysis in Ship Manufacturing with Knowledge Graphs and Large Language Models

Yanjun Ma [1], Tao Wu [2,*], Bin Zhou [3], Xiaoyang Liang [4], Jiwang Du [1] and Jinsong Bao [2]

[1] Hudong Shipbuilding, Shanghai 200129, China; kfmyj@163.com (Y.M.); adusdu@163.com (J.D.)
[2] College of Mechanical Engineering, Donghua University, Shanghai 201620, China; bao@dhu.edu.cn
[3] School of Artificial Intelligence, Anhui Polytechnic University, Wuhu 241000, China; bin@ahpu.edu.cn
[4] Shanghai Dongxin Software Engineering Co., Ltd., Shanghai 200090, China; 19821289080@163.com
* Correspondence: researcher_wutao@163.com

Abstract: Ship manufacturing is a critical backbone industry in China, where the nation leads on a global scale in terms of vessel completions and order volumes. However, the high volume of orders often imposes substantial processing loads, increases the risk of equipment failures, and exacerbates production bottlenecks. Despite the accumulation of significant amounts of data in this field, analyzing bottlenecks remains a persistent challenge, primarily due to the presence of heterogeneous, multi-source data and the lack of effective data integration mechanisms. The traditional approaches are largely limited to bottleneck detection, offering minimal capabilities in terms of deep analysis, traceability, and interpretability, which are crucial for comprehensive bottleneck resolution. Meanwhile, extensive knowledge remains underutilized, leading to analytical results that are overly reliant on expert experience and lacking in interpretability. To address these challenges, this research proposes a graph-retrieval-based bottleneck mining method for ship manufacturing, employing large language models and a knowledge graph. The approach integrates a data-driven "turning point" mechanism for dynamic bottleneck detection and the manufacturing process knowledge graph, consisting of process subgraphs and 5M1E (Man, Machine, Material, Method, Measurement, Environment) specification subgraphs. Furthermore, a question-answering chain is introduced to enhance the interaction between the LLMs and the knowledge graph, improving the retrieval and reasoning capabilities. Using practical production data from a Shanghai ship thin plate production line, our method demonstrates a superior performance compared to that of four existing models, validating its effectiveness in throughput bottleneck analysis. This approach provides a scalable and efficient solution for analyzing complex bottleneck issues in industrial production, contributing to enhanced manufacturing efficiency and digital transformation.

Keywords: ship manufacturing process; bottleneck mining; knowledge graph; large language models

check for updates

Academic Editor: Panagiotis Kyratsis

Received: 22 January 2025
Revised: 3 March 2025
Accepted: 6 March 2025
Published: 10 March 2025

Citation: Ma, Y.; Wu, T.; Zhou, B.; Liang, X.; Du, J.; Bao, J. Enhancing Bottleneck Analysis in Ship Manufacturing with Knowledge Graphs and Large Language Models. *Machines* **2025**, *13*, 224. https://doi.org/10.3390/machines13030224

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the advent of the Industrial 4.0 era, China's shipbuilding industry has achieved rapid development, with both the number of ships built and orders received ranking first globally. Although the technology and quality of ship products have significantly improved, there is still a huge gap compared to the shipbuilding level of other advanced countries, most notably in the workshop operation and maintenance management and production efficiency during the ship production and construction process [1]. In China, "digital shipbuilding" has been developed for more than ten years and has established a

good informational infrastructure, resulting in the creation of extensive data on production and processing, models, plans, and reports [2]. These records and results generated from the production process encompass rich practical experience, processing rules, and implementation methods. However, due to their characteristics of multi-source heterogeneity, temporality, and complex semantic relationships, the existing workshop operation and maintenance systems are unable to effectively process and organize these production process data, leading to a significant waste of valuable manufacturing experience and knowledge [3]. In particular, the bottleneck in the manufacturing process's throughput is likely to trigger equipment malfunctions, thereby significantly impacting the efficiency of shipbuilding. Therefore, how to increase the throughput of the production line to complete orders in a shorter time is a major task facing the ship industry.

A ship production line is composed of numerous workstations or nodes. In an ideal situation, all nodes operate normally without bottlenecks, thereby maximizing the throughput. However, in actual production, due to factors such as equipment failure, personnel changes, material shortages, environmental conditions, and order variations, the working status of the production line often experiences short-term fluctuations, leading to bottlenecks. A bottleneck is the node that has the greatest impact on the overall throughput of the production line [4]. Therefore, studying bottleneck mining (detection, analysis, and improvement) holds significant practical value. In brief, the challenge of throughput bottleneck mining for the ship industry mainly involves the effectiveness and interpretability of the bottleneck analysis.

Research on bottleneck detection can be categorized into three types: analytical methods, simulation methods, and data-driven methods. Analytical methods are based on the statistical distribution of machine performance indicators, using Bernoulli [5] or Markov models [6] to describe the basic unit of "workstation-buffer-workstation" in the production line. However, limitations of these methods lie in their reliance on numerous assumptions, which make it difficult to obtain accurate results for complex production lines. Additionally, they are only applicable to steady-state bottlenecks and cannot detect dynamic bottlenecks. Simulation methods, such as General Motors' "C-MORE" model, establish discrete event simulation models that can detect bottlenecks in complex production lines. However, these models require lengthy development times, pose real-time calculation challenges, and involve many approximations and assumptions in model construction. Data-driven methods, exemplified by Li et al.'s "turning point" method, use metrics such as the blocking and starvation times at the nodes to locate bottlenecks, enable dynamic bottleneck detection in linear production lines, and rank bottlenecks by severity using a bottleneck index. Lai [7] improved the "turning point" method, extending its applicability to complex production lines with parallel or reworked loop structures. However, the "turning point" method's shortcoming is that it can only detect bottlenecks in complex production lines but cannot determine the specific causes of the bottlenecks. As a result, it cannot provide effective guidance for the subsequent bottleneck improvement phase [8]. Therefore, bottleneck analysis remains a key issue that needs to be addressed in future research. However, traditional learning methods deal with relatively simple data and problems, with high time costs in model construction, limited to solving single-class problems and having low transferability. In industrial scenarios, the level of informatization is high, with a large amount of data and knowledge stored in information management systems which reflect the real situation of the production process and have not been fully exploited and utilized due to the multi-source and heterogeneous nature of these data. The advantage of knowledge graphs lies in their high retrieval and integration capabilities, enabling the integrated storage and retrieval of multi-source heterogeneous data for data analysis; hence, this technology is widely applied in industrial research for data integration management. Ultimately,

the results of the analysis aim to assist engineers in production optimization. Moreover, the outcomes from the traditional methods are mere specific values, and understanding them requires expert knowledge, with these traditional methods resembling black-box operations that lack corresponding explainability. The advantage of large language models (LLMs) lies in their high capabilities for insight, reasoning, and the induction of production data, and the solutions they provide are relatively interpretable. This technology, combined with knowledge graphs, offers potential solutions to the above problems [9].

Large language models (LLMs) have demonstrated exceptional performance in natural language processing tasks, but they suffer from the issue of hallucination [10]. Retrieval-augmented generation (RAG) [11] mitigates this problem by segmenting and embedding documents into a vector database, enabling LLMs to retrieve relevant text and enhance their responses by using it as context [12]. However, RAG assumes that knowledge can be accurately represented by independent text units, overlooking the relationships between multiple text units. A knowledge graph is a semantic network that reveals relationships between entities [13]. In 2024, Jin et al. [14] proposed the "Graph Chain of Thought", demonstrating the role of knowledge graphs in bridging gaps in the RAG framework and aiding LLM reasoning. Currently, there is no systematic research on the application of LLMs and knowledge graphs to bottleneck detection and analysis.

To solve the above issues, this paper proposes a dynamic bottleneck mining method based on graph retrieval for bottleneck detection, analysis, and improvement in ship manufacturing processes (SPKG). Compared to the traditional bottleneck analysis methods, SPKG introduces manufacturing process knowledge graphs to assist LLMs in analyzing the causes of bottlenecks, achieving the goal of bottleneck improvement by increasing the effectiveness and interpretability of the results of the analysis.

The main contributions of this paper are as follows:

1. A bottleneck mining method for ship manufacturing is proposed that can simultaneously solve the two types of problems of low accuracy in identifying dynamic bottlenecks and insufficient validity in analyzing them, which is named SPKG. The method integrates a data-driven "turning point" approach to identifying dynamic bottlenecks and introduces a production process knowledge graph to assist the large language models (LLMs) in analyzing the causes of bottlenecks, offering guiding suggestions to production decision-makers for improving bottlenecks.
2. A knowledge graph modeling method for ship production processes is developed which reveals the semantic relationships between nodes, providing structured and semantic context information for the LLMs, compensating for the inadequacies in the LLMs' reasoning capabilities, and enabling a more accurate analysis of bottleneck causes.
3. A Question-Query Chain method for SPKG and a Question-Query Space based on large language models have been designed, which, through the use of prompts and query examples, endow LLMs with the capability to query the production process knowledge graph, enabling it to generate accurate Cypher query statements based on decision-makers' questions, thereby obtaining 5M1E information on bottleneck nodes for an effective and interpretable in-depth analysis.

The remaining structure of this paper is as follows: Section 2 introduces the current research status for bottleneck mining methods in ship manufacturing and the progress of large language models. Section 3 provides a detailed description of the overall framework of SPKG presented in this paper and introduces the details and mechanisms of each part. Section 4 conducts a case analysis with real data from a shipbuilding company in Shanghai and experimentally compares the effectiveness of the method proposed in this paper.

Section 5 comprehensively summarizes the purpose, contribution, and future development directions of the method in this paper, providing a reference for subsequent research.

2. Related Work

This section will describe the research on the limitations of bottleneck analysis in ship manufacturing and the potential and challenges of LLMs in shipbuilding bottleneck analysis.

2.1. The Limitations of Bottleneck Analysis in Ship Manufacturing

Traditional bottleneck detection methods in ship manufacturing suffer from several limitations that hinder their effectiveness. Firstly, the multi-source, heterogeneous nature of the data, which are voluminous and diverse in their format, structure, and semantics, presents significant challenges for data integration and processing. This complexity often leads to inaccuracies in bottleneck detection and low utilization of available data. For instance, the study by [15] demonstrated that the conventional statistical methods failed to accurately identify bottlenecks in a complex shipbuilding environment due to their inability to handle multi-source data effectively. Secondly, the intricate production processes in ship manufacturing, involving numerous stages and departments, further complicate the detection process. The dynamic nature of the production environment, with its ever-changing conditions, often evades real-time capture by conventional methods. As reported by [16], analytical models based on steady-state assumptions are ill suited to detecting dynamic bottlenecks in shipbuilding. Thirdly, the heavy reliance on expert intuition in the traditional approaches becomes a bottleneck when expert resources are limited or outdated. This was highlighted in a case study by [17], where the overreliance on expert knowledge led to missed or misdiagnosed bottlenecks in ship production. Lastly, the opacity of these methods, functioning akin to a "black box", impedes the interpretation of bottleneck causes, which is essential for continuous process improvement. The review by [18] underscored the need for more transparent and interpretable bottleneck analysis methods in industrial settings.

In summary, these limitations underscore the necessity for novel strategies in bottleneck detection within ship manufacturing, which can address the challenges of data heterogeneity, process complexity, dynamic environments, and the need for interpretability.

2.2. The Potential and Challenges of LLMs in Shipbuilding Bottleneck Analysis

Large language models (LLMs) offer a promising alternative to the traditional bottleneck analysis methods in shipbuilding. The potential of LLMs lies in their ability to process and analyze large volumes of text data, extracting insights and identifying patterns that may be missed by conventional methods. A key advantage is their capability for natural language understanding, which can be leveraged to interpret and analyze production logs and reports, potentially revealing bottlenecks that are not easily detectable through a numerical data analysis alone. However, the application of LLMs in shipbuilding bottleneck analysis is not without its challenges. One of the primary concerns is the issue of hallucination, where LLMs may generate plausible but inaccurate information. To address this, retrieval-augmented generation techniques, as proposed by [19], can be used to enhance the reliability of LLM-generated insights by grounding the model's responses in relevant domain knowledge. Another challenge is the integration of LLMs with domain-specific knowledge, particularly in the form of knowledge graphs. The construction of a manufacturing process knowledge graph that accurately represents the shipbuilding domain is a complex task that requires careful modeling of the semantic relationships between entities. Ref. [20] demonstrated the value of causal knowledge graphs in industrial structure analysis, suggesting a similar approach could be applied to shipbuilding. To quantify the

added value of LLMs in bottleneck analysis, Key Performance Indicators (KPIs) such as a reduction in the bottleneck detection time, an increase in throughput, and an improvement in the production efficiency can be measured. A literature review by [21] indicated that the integration of knowledge graphs with data-driven methods could lead to a 20% increase in the production efficiency by effectively identifying and addressing bottlenecks.

In conclusion, while LLMs present a significant opportunity to enhance bottleneck analyses in ship manufacturing, their application must be approached with an understanding of the challenges involved. The development of a structured knowledge graph and the implementation of retrieval-augmented generation are critical steps toward realizing the potential of LLMs in this domain.

2.3. Research Gaps

Through a comprehensive analysis of the existing literature, this paper identifies several key limitations in the field of shipbuilding bottleneck analysis which hinder further developments in this field:

1. The inadequate utilization of multi-source heterogeneous data: The ship manufacturing process generates a large volume of multi-source heterogeneous data, such as sensor logs, operational reports, and expert notes. However, the existing methods face difficulties in integrating these data sources, which have different formats, structures, and semantics. The lack of a unified data representation framework restricts the depth and breadth of a comprehensive bottleneck analysis and also hinders the extraction of valuable insights from the accumulated data.
2. Excessive reliance on expert knowledge results in subjective analysis outcomes: The prevailing bottleneck analytical methods are largely dependent on the assessments and expertise of domain specialists. This dependence on experts not only introduces subjectivity and increases the labor costs but also limits the scalability of the methods in a large-scale production environment. Manual intervention in the data interpretation and bottleneck diagnosis may lead to inconsistencies and inefficiencies in the analysis results.
3. The limited interpretability and traceability of bottleneck analyses: Traditional bottleneck detection techniques often act as black-box models, lacking transparency and interpretability in their outputs. Despite the capability of discrete event simulation and data-driven approaches to identify bottlenecks, such methods generally do not offer a profound analysis of the underlying causes, nor do they effectively trace back to the origins of problems. The dependence on statistical hypotheses or pre-established rules restricts the application of these methods within intricate and dynamic manufacturing contexts.

The existence of these research gaps underscores the critical necessity for a more potent and dependable approach to bottleneck mining within the shipbuilding process, especially with respect to the analysis of dynamic bottlenecks. This paper aims to fill these research gaps by proposing a graph-retrieval-based bottleneck mining method for ship manufacturing employing large language models and a knowledge graph to enhance the effectiveness and interpretability of bottleneck analyses in the ship manufacturing process, thereby providing guidance suggestions for production decision-makers to improve bottlenecks.

3. Methodology

This section delineates the SPKG method, which is tailored to the bottleneck mining domain, encompassing the detection, analysis, and improvement of bottlenecks. The primary objective of SPKG is to address the critical challenge of throughput bottleneck analysis

in the shipbuilding industry, focusing on enhancing the efficacy and interpretability of the bottleneck analysis after identifying the bottleneck.

The overall framework diagram proposed in this section is illustrated in Figure 1 and consists of three components. Specifically, it includes schema modeling of the manufacturing process knowledge graph, the dynamic bottleneck detection method, and the LLM-based root cause analysis of the bottlenecks. Initially, event logs of orders from the enterprise information system, along with documents and spreadsheets containing 5M1E information about the manufacturing process, are used as the data sources for the graph. The manufacturing process knowledge graph is divided into two subgraphs: a process flow subgraph and a 5M1E specification subgraph. After embedding and indexing, it forms the domain knowledge base for SPKG. Then, the bottleneck detection algorithm locates bottlenecks by monitoring the blocking and starvation times of the nodes in the graph. Once a bottleneck is detected, production decision-makers can inquire about the 5M1E information of the bottleneck node through the LLMs. Ultimately, the LLMs retrieve relevant information from the domain knowledge base using Cypher queries, similarity searches, full-text searches, and other retrieval methods. The retrieved content is used as context for the LLMs to generate more precise responses, assisting decision-makers in identifying the type and cause of a bottleneck. The decision-makers then implement the appropriate improvement measures to eliminate the bottleneck. In summary, the SPKG method has an advantage in its effectiveness and interpretability in addressing throughput bottlenecks in the dynamic manufacturing process. By leveraging knowledge graphs and large language models (LLMs), it outperforms the traditional methods, thereby significantly enhancing the stability and productivity in the domain of intelligent bottleneck mining for industrial processes.

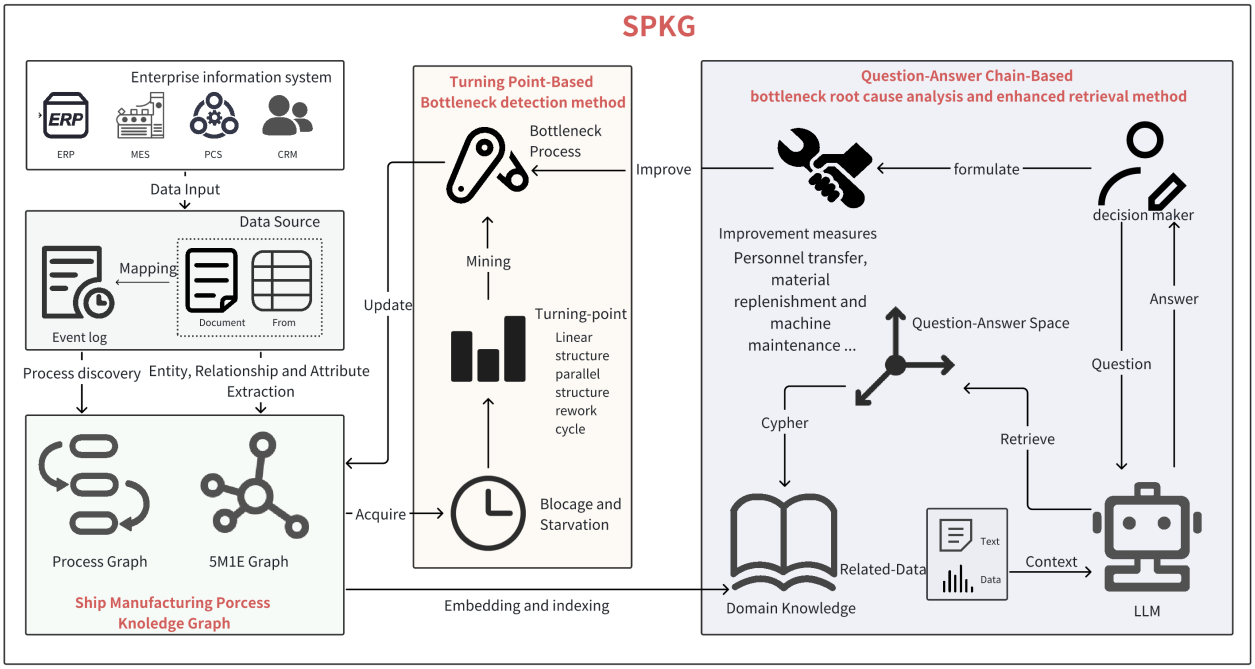


Figure 1. Framework of SPKG.

3.1. The Design and Generation of the Ship Manufacturing Process Knowledge Graph Schema

Compared to traditional data-driven methods, a considerable amount of process knowledge has remained underutilized, which has implications for the accuracy and depth of bottleneck analyses. And knowledge graphs have an advantage in storing distilled knowledge [22]. Therefore, this section constructs a manufacturing process knowledge

graph from top to bottom, thereby providing a solid evidentiary basis for the LLM-based analysis outcomes. This construction encompasses the process flow subgraph and the 5M1E specification subgraph. These two components form the foundation for bottleneck detection and analysis. The schema design involves defining the nodes and relationships, as shown in Tables 1 and 2. The "turning point" method detects bottlenecks by using the blocking and starvation times of the nodes. When a node experiences operational issues due to a downstream node failure or other factors, it is considered blocked; when a node is affected by issues with an upstream node, it is considered starved. A bottleneck causes upstream nodes to become blocked and downstream nodes to become starved, meaning the bottleneck is the "turning point" where widespread blocking turns into widespread starvation.

To detect dynamic bottlenecks, the knowledge graph must include contextual information about ship manufacturing, such as the sequence of steps and the blocking and starvation times for each step. One definition of a bottleneck is the node with the lowest independent productivity [23], so a bottleneck analysis should begin by examining the factors affecting the productivity of each step. These factors are summarized into six categories: Man, Machine, Material, Method, Environment, and Measurement (5M1E). Bottleneck analysis requires examining the 5M1E factors of the bottlenecked step to determine the cause of the bottleneck; thus, the knowledge graph needs to incorporate 5M1E information.

Table 1. Definition of nodes in the manufacturing process knowledge graph.

Node Type	Node Attributes	Indexed Attributes	Index Type
Order	Order Number, Product Name, Product Number, Customer, Color, Cutting Quantity, Delivery Date	Product Name, Customer	Full-Text Index
Workstation	Workstation Name, Workstation Description	Workstation Name	Full-Text Index
Step	Step Name, Start Time, End Time, Step Type	Step Name	Full-Text Index
Buffer	Buffer ID, Maximum Capacity	Buffer ID	Full-Text Index
Blockage and Starvation	Blockage Time, Starvation Time	Blockage Time	Full-Text Index
Man	Number of People, Processing Method, Working Hours, Step Name	Processing Method	Full-Text Index
Machine	Machine ID, Equipment Group, Model, Technical Parameters, Equipment Description	Equipment Description	Vector Index
Material	Material Name, Material Quantity	Material Name	Full-Text Index
Method	Method Name, Technical Parameters, Process Specification	Process Specification	Vector Index
Environment	Parameter Requirements, Standards	Standards	Vector Index
Measurement	Measurement Equipment, Measurement Method, Measurement Standards	Measurement Standards	Vector Index
Issue	Issue Type, Issue Description	Issue Description	Vector Index
Solution	Solution Type, Solution Description	Solution Description	Vector Index

Table 2. Definition of edges in manufacturing process knowledge graph.

Relationship Type	Head Node	Tail Node	Relationship Type	Head Node	Tail Node
Contains	Order	Step	Next	Workstation	Workstation
Next_Step	Step/Buffer	Step	Failure Phenomenon	Workstation	Issue
Next_Buffer	Step	Buffer	Pro_Issu_Analysis	Solution	Step
Blockage & Starvation_Time	Step	Blockage & Starvation	Opr_Issue	Issue	Operation
Works_At	Man	Step	Analysis	Solution	Workstation
Has_Machine	Step	Machine	Solution	Issue	Solution
Has_Material	Step	Material	Has_Environment	Step	Environment
Has_Method	Step	Method	Has_Measurement	Step	Measurement

Starting with customer orders, the manufacturing process is traced, recording the contextual information between steps, the blocking and starvation times of each step, and the location of the buffer zones [24]. The process is then expanded to include 5M1E nodes. By organizing the nodes and relationships shown in Tables 1 and 2, an initial version of the manufacturing process knowledge graph is created. To facilitate SPKG's use of full-text search and similarity search, full-text and vector indices should be created for the indexed attributes in Table 1.

3.2. A Turning-Point-Based Bottleneck Detection Method in the Ship Manufacturing Process

In the ship production line, buffer zones exist between nodes with varying production rates. These buffers temporarily store the in-process products processed by upstream nodes. If over a certain period of time the buffer's stock is consistently neither zero nor at the maximum capacity, it is considered an effective buffer, as it effectively decouples the manufacturing process [25]. When detecting bottlenecks in the entire production line, if there are n effective buffers, the production line can be decoupled into n + 1 independent sections, and bottlenecks can be located within these n + 1 independent sections.

The mathematical definition of a bottleneck in a linear structure is as follows:

Definition 1. *Over a given period, for a production line with n workstations and buffers with a limited capacity, a node is considered a bottleneck if it satisfies any one of Conditions (1)–(3).*

$$\begin{gathered}(t_{B,i} - t_{S,i}) > 0, i \in [1, 2, \ldots, j-1], j \neq 1, j \neq n \\ (t_{B,i} - t_{S,i}) < 0, i \in [j+1, j+2, \ldots, n], j \neq 1, j \neq n \\ t_{B,j} - t_{S,j} < t_{B,j-1} - t_{S,j-1}, j \neq 1, j \neq n \\ t_{B,j} + t_{S,j} < t_{B,j+1} + t_{S,j+1}, j \neq 1, j \neq n\end{gathered} \tag{1}$$

$$\text{if } j = 1: \quad (t_{B,1} - t_{S,1}) > 0 \text{ and } (t_{B,2} - t_{S,2}) < 0 \text{ and } t_{B,1} + t_{S,1} < t_{B,2} + t_{S,2} \tag{2}$$

$$\text{if } j = n: \quad (t_{B,n-1} - t_{S,n-1}) > 0 \text{ and} (t_{B,n} - t_{S,n}) < 0 \text{ and} t_{B,n} + t_{S,n} < t_{B,n-1} + t_{S,n-1} \tag{3}$$

where $t_{B,j}$ and $t_{S,j}$ represent the blocking and starvation times of node j, respectively.

When there are multiple bottlenecks on a production line, the bottlenecks need to be ranked based on a bottleneck severity index. The principle is that bottlenecks cause upstream nodes to be blocked and downstream nodes to starve. Additionally, the sum of the bottleneck's blocking and starvation times is smaller than that of the adjacent nodes. Thus, the heuristic bottleneck indices are defined as shown in Equations (4)–(6):

$$I_1 = \frac{t_{S,2}}{t_{B,2} + t_{S,1}} \tag{4}$$

$$I_i = \frac{t_{B,i-1} + t_{S,i+1}}{t_{B,i} + t_{S,i}}, i = 2, 3, \ldots, n-1 \quad (5)$$

$$I_n = \frac{t_{B,n-1}}{t_{B,n} + t_{S,n}} \quad (6)$$

where I_1 and I_N represent the bottleneck indices for the first node and the n-th node, respectively.

The "turning point" methodology for detecting bottlenecks in a complex production line can be systematically described through the following steps: (1) Decoupling the manufacturing process: Segment the production process into several independent sections by leveraging effective buffers to isolate each section. (2) Simplifying the structural complexity: Within each independent section, treat parallel structures and rework loops as pseudo-nodes. Conduct linear bottleneck detection on the restructured system. (3) Bottleneck identification: If a bottleneck is detected and it is not a pseudo-node, this indicates that the bottleneck for the corresponding independent section has been identified. Otherwise, if the detected bottleneck is a pseudo-node, proceed to the subsequent step. (4) Detailed bottleneck analysis: For any identified pseudo-node, perform further bottleneck detection at a finer level of granularity to uncover the underlying constraints. (5) Global bottleneck determination: Compute the bottleneck indices for all detected bottlenecks. The "turning point" associated with the highest bottleneck index is then recognized as the global bottleneck of the production line.

Upon completing step 4, n local bottlenecks are identified (with $n-1$ representing the number of effective buffers). Upon completing step 5, the unique global bottleneck is detected.

3.3. The Question–Answer Chain of the LLM-Based Bottleneck Root Cause Analysis and an Augmented Retrieval Method

After detecting a bottleneck, the next step is to analyze its root cause. The significant challenge for root cause analysis tasks in manufacturing bottleneck mining is effectiveness and interpretability. Large language models (LLMs) have high interpretability attributed to their technical features, including the display of reasoning processes and evidence chains in multi-turn dialogues, which allows engineers to trace and understand the derivation of conclusions [26,27]. Clear reasoning pathways allow engineers to review and evaluate the credibility of the evidence and reasoning steps referenced in the conversations. A vast knowledge base ensures that the model can draw on a wide range of domain knowledge to support its reasoning and use an augmented retrieval method, as well as employing strategies such as thought chains to simulate human thinking processes, thereby enhancing the transparency and interpretability of the entire reasoning process.

Therefore, on the basis of a knowledge graph, this section aims to use LLMs and augmented retrieval methods to enhance the efficiency and interpretability of analyzing a bottleneck's root cause after its detection as described in Section 3.2. As described in Section 2.1, the graph structure offers the advantages of a high retrieval efficiency and integration of multi-source data, providing a data foundation for the subsequent bottleneck analysis. Meanwhile, since the constructed 5M1E subgraph is in graph structure and the data are stored in a graph database, currently, only Cypher queries can be used to query the graph database, and there is a lack of direct methods for the interaction between large models and graph databases. In addition, Cypher queries intuitively express nodes and relationships, which is crucial for querying and analyzing the complex relationships in knowledge graphs. Therefore, this paper opts to use Cypher to enhance the interaction capability between the LLMs and the graph database. Although the LLMs have some ability to generate Cypher, they lack knowledge of the manufacturing process knowledge graph, making it difficult for them to directly generate accurate Cypher queries. This

paper constructs an SPKG Q&A chain to provide the LLMs with the ability to query the graph. As shown in Figure 2, the prompt includes the schema, the query examples, and the decision-maker's questions. The schema is represented in a symbolic manner similar to Cypher language, including node attributes, relationship attributes, and relationships. The node attributes are represented as follows: NodeTypeAttributeName: DataType, AttributeName: DataType… The relationship attributes are represented as follows: RelationshipTypeAttributeName: DataType, AttributeName: DataType… The relationships are represented as (: NodeType)-[:RelationshipType]->(:NodeType). Following these three rules, all nodes and relationships are converted into strings and passed into the prompt as the schema. The prompt also contains several examples to help the LLMs learn how to generate accurate Cypher queries based on the questions. Each query example includes a question and the corresponding Cypher. When the decision-maker's question is similar to a query example, the LLMs can more easily generate an accurate Cypher. For instance, given the aforementioned prompt, when the decision-maker asks, "What materials are used in cutting?", since the prompt contains "What machines are used in cutting?", the LLMs only need to slightly modify the Cypher in the query example. Furthermore, the Cypher in the example also organizes the retrieved content into a format that the LLMs can easily understand, such as concatenating the retrieved text and data into a semantic string, e.g., RETURN m.model + " is used in cutting".

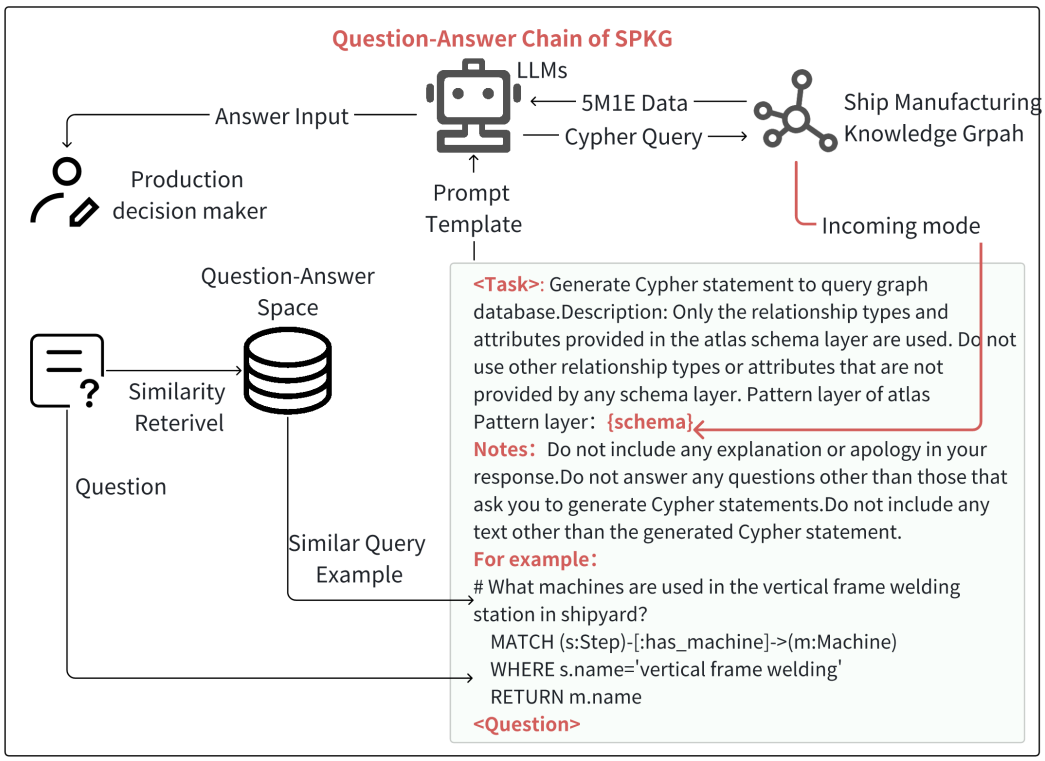


Figure 2. Construction of SPKG Q-A chain in ship manufacturing.

Theoretically, increasing the number of query examples in the prompt could enable the LLMs to address various questions [28,29]. However, including all possible questions in the prompt would inevitably exceed the token limit of the LLMs. Therefore, this paper introduces the concept of the Question-Query Space, which is implemented through the following process: (1) Predefine query examples based on the graph schema; (2) embed and store the questions from all query examples in a vector database; (3) when a user poses a question, perform similarity searches to find similar query examples from the Question-Query Space; and (4) incorporate the similar query examples into the prompt. Once the

SPKG Q&A chain is established, decision-makers can analyze the root causes of bottlenecks through multiple rounds of inquiries. A typical example of multi-round questioning is illustrated in Figure 3. The decision-maker conducts three rounds of inquiries, aiming to obtain information regarding orders, machines, and processes. The LLMs generate corresponding Cypher queries to retrieve the 5M1E information on the bottleneck process as the context for the responses. The Cypher queries involve pattern matching, similarity searches, full-text searches, and multi-hop queries. After the LLMs responds, the decision-maker assesses whether the 5M1E information for the current bottleneck process meets the standards and requirements, systematically identifying causes before implementing corresponding improvement measures, such as personnel adjustments and machine maintenance.

After the original bottleneck has been resolved, it is essential to update the graph to maintain consistency and prepare for the next bottleneck detection and analysis. For instance, the bottlenecks of the cutting machine are indeed influenced by a combination of factors, including, but not limited to, the cutting temperature, machine maintenance history, operator skill level, and material quality. In the case of resolving a bottleneck by changing the cutting temperature, it is crucial to update the knowledge graph to reflect this change. Specifically, the attribute values of the method node related to the cutting temperature should be modified to ensure that the graph accurately represents the current state of the production process. This is essential for maintaining the consistency and accuracy of the knowledge graph, which in turn is vital for the next bottleneck detection and analysis. In summary, while all attributes may not be necessary to determine the bottlenecks at any given time, the knowledge graph is designed to be comprehensive and adaptable. It captures a wide range of attributes to provide a holistic view of the production process, allowing for a nuanced analysis and informed decision-making. The updates to the graph are selective and based on the specific changes in the production environment and the outcomes of the bottleneck resolution measures. Therefore, if the improvement measure involves changing the cutting temperature, the attribute values of the method node should be modified. Additionally, after completing a production shift or workday, the blocking and starvation times of each process step in the graph should be updated based on data from the enterprise information system.

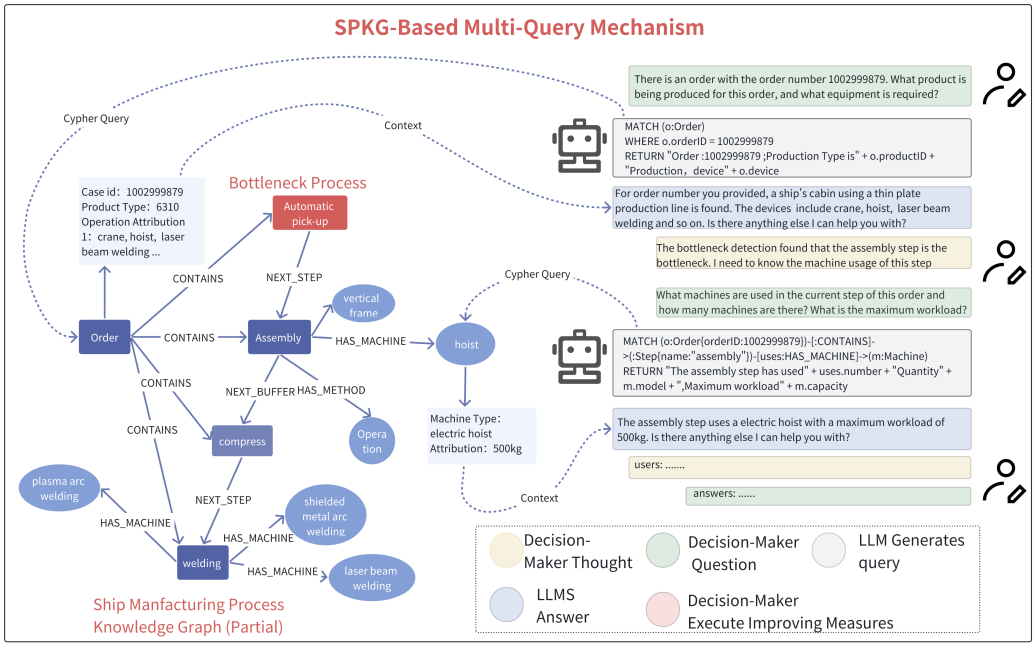


Figure 3. A typical example of multi-query attribution in ship manufacturing.

Figure 3 illustrates the use of the SPKG method to identify and analyze a bottleneck in a vertical frame welding workstation. This study employed the SPKG method to identify and analyze bottleneck issues within the production process for order number 1002999879 at a vertical frame welding workstation. By constructing a manufacturing process knowledge graph incorporating the 5M1E elements and leveraging large language models (LLMs) for an in-depth analysis, this study identified the assembly step as a critical bottleneck. Based on graph queries, the study revealed a mismatch between the maximum 500 kg workload of the electric hoist and the production demands, which guided targeted improvement measures. This case study validates the effectiveness of the SPKG method in identifying and analyzing bottlenecks in dynamic manufacturing environments, enhancing the production stability and efficiency, and providing scientific decision support for bottleneck management in engineering practice.

4. The Case Study

This study focuses on a ship factory in Shanghai, applying the SPKG bottleneck analysis method to the factory's daily production scheduling. Choosing this as a typical example is due to the fact that the shipbuilding industry is a bellwether of China's advanced manufacturing capabilities, where the optimization of production processes carries substantial weight in maintaining the country's leading position in global vessel production [30]. The complexity inherent in ship manufacturing, with its myriads of interconnected processes, makes it an ideal test bed for innovative analytical methods such as SPKG. Specially, this shipyard workshop has a complete production line and a high level of equipment intelligence, including raw material yards, preprocessing lines, cutting equipment, sheet metal flow lines, pre-assembly lines, section turning frames, and other main equipment. Among these, this paper takes thin plate production line as the research object, and the main production process includes steel plate pre-cutting, milling and plate assembly, cutting, vertical frame welding, transverse frame assembly, transverse frame welding, repair, polishing, etc. And other major process and flow data (event logs) [31], fault documents, and operation and maintenance tables are used, among other multi-data. Therefore, this research focuses on the analysis of production process bottlenecks in the context of ship production and assembly. Selecting the production data on this workshop as a case study is not only a typical demonstration case but also the lessons learned have strong portability, thereby providing new directions for other shipyards and even other manufacturing sectors in promoting the integration of production bottleneck analysis and large language models.

4.1. The Data Description and Experimental Setup

The experimental period spans 80 working days, during which bottleneck detection and analysis are conducted every two working days. After implementing bottleneck improvement measures, the manufacturing process knowledge graph is updated. Two working days later, the blocking and starvation times of each process step are updated before conducting another bottleneck detection. In total, 40 rounds of bottleneck detection and analysis are performed. In terms of the workshop in the thin plate production line that is the focus of the data collection, a partial layout of this production line is depicted in Figure 4 below:

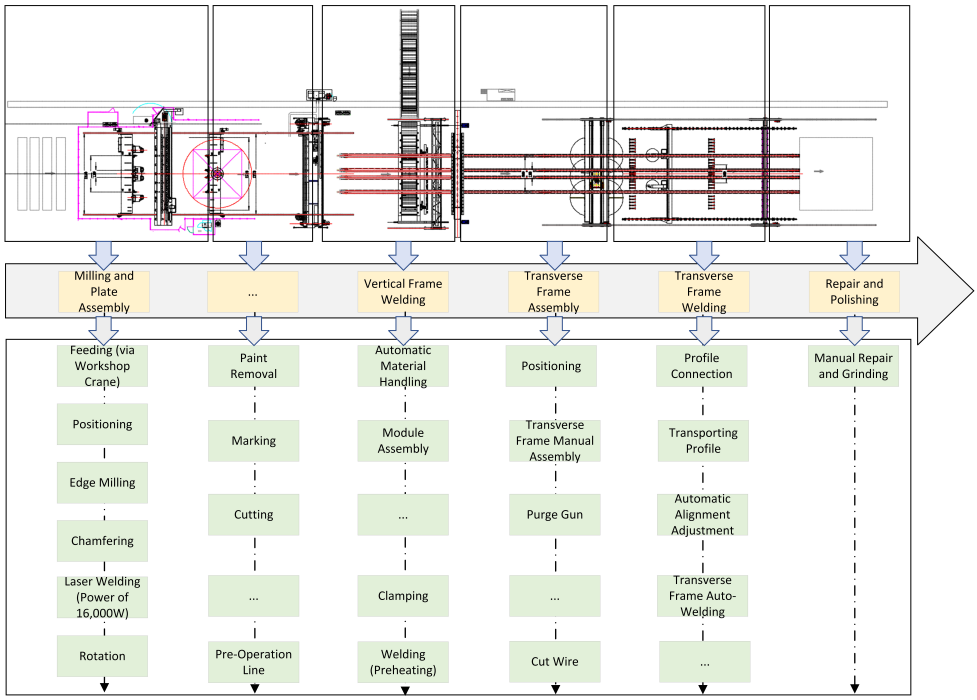


Figure 4. Partial layout for the thin plate production line in ship manufacturing.

As shown in Figure 4, the entire production line consists of 6 workstations. Among these, the milling and plate assembly workstation is composed of 14 processes, including feeding (via workshop crane), positioning, and chamfering, among others. The cutting workstation consists of 12 processes, including paint removal, marking, and pre-operation, among others.

4.1.1. Data Description

The data used in this paper fall into two categories: the production process event log data contained within the production management system and ship operation and maintenance documents. The production process event logs are records of production trajectories used for the study of ship production bottleneck identification. The operation and maintenance documents contain the historical knowledge essential for ship bottleneck analysis, which supports the bottleneck analysis results of the large language models. This paper stores both types of data in the form of triplets and unifies them in the graph database Neo4j [32]; a sample of a partial ship knowledge graph is shown in Figure 5.

Figure 5 illustrates the structured organization of the data facilitating the bottleneck analysis within the specified vertical frame welding workstation through the knowledge graph method delineated in Section 3.1. In this figure, blue nodes represent workstations, and green nodes denote processes, while the remaining nodes signify the pertinent knowledge required for a comprehensive analysis. Figure 5 presents the example of a specific workstation, namely the vertical frame welding station, with the knowledge graph constructed in accordance with the methodology delineated in Section 3.1. This figure elucidates the structured organization of data that facilitates the subsequent bottleneck analysis. In the depicted diagram, blue nodes correspond to workstations, and green nodes signify processes, while the other nodes represent the pertinent knowledge essential for the integrated analysis.

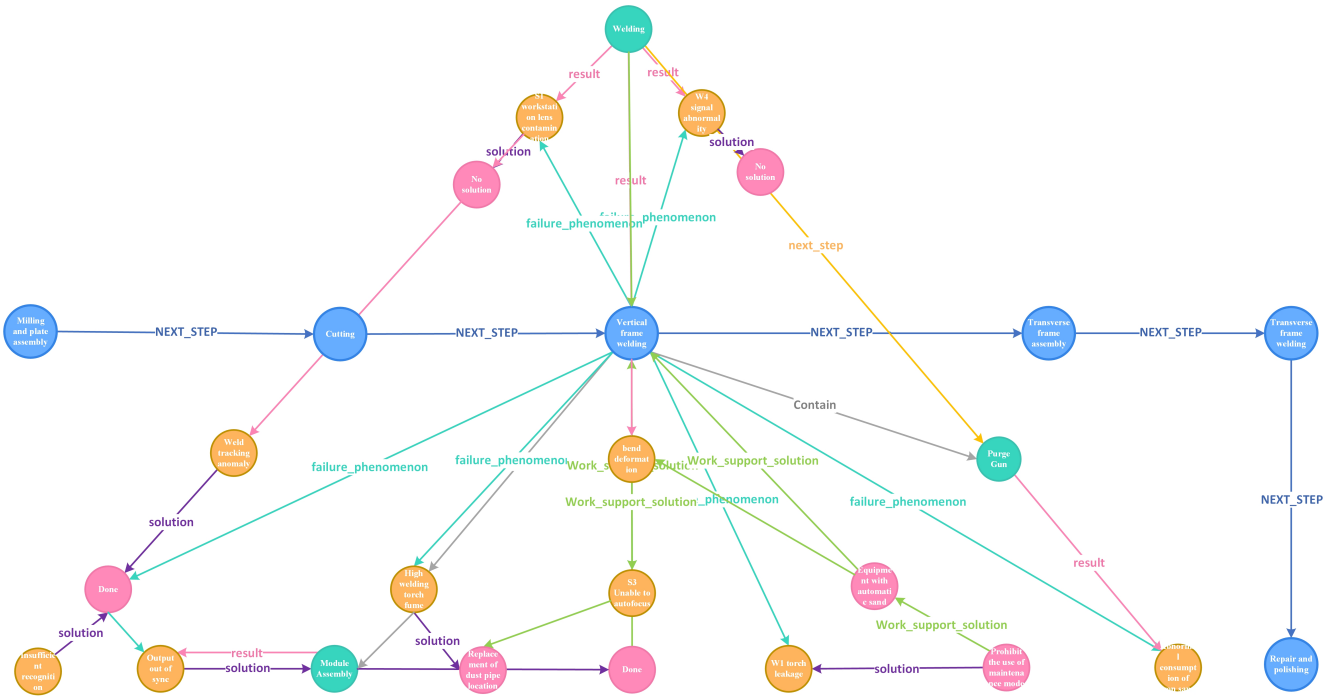


Figure 5. The construction of process knowledge graph in ship manufacturing.

The partial production process for ships collected in this paper mainly includes 6 one-level workstations and 52 two-level production process. Specifically, part of the main process is described in the following table, including edge milling, CNC plasma cutting, transverse frame manual assembly, profile connection, and transverse frame auto-welding, manual repair and grinding, positioning, etc. The core equipment and the number of processing personnel involved are listed in Table 3. And shipbuilding thin plate production records are described in Table 4. In particular, IMG refers to the fully automatic welding machine for plate assembly; ESAB is a welding and cutting machine [33].

Table 3. Main ship manufacturing process.

Production Workstation	Process	Operation Type
Milling and Plate Assembly	Positioning Edge Milling	Semi-Automatic Semi-Automatic
Cutting	CNC Plasma Cutting Flame Slat Cutter CNC Plasma Cutting	Program Regulation Program Regulation Program Regulation
Transverse Frame Assembly	Transverse Frame Manual Assembly Purge Gun Cut Wire	Semi-Automatic Program Regulation Program Regulation
Vertical Frame Welding	Module Assembly Material Handling	Program Regulation
Transverse Frame Welding	Profile Connection Transverse Frame Auto-Welding	Semi-Automatic Program Regulation
Repair and Polishing	Manual Repair And Grinding	Semi-Automatic

Table 4. Shipbuilding thin plate production records.

Target Block	Index	Panel Name	Panel Length	Panel Width (Seam Direction)	Number of Seams
Block 1	1	FG01C_10F	Next	Workstation	2
	2	CG22C_C	Failure Phenomenon	Workstation	3
	3	CG22C_68FA	Pro_Issu_Analyasis	Solution	2
	4	BD21C_1X	Opr_Issue	Issue	6
Block 2	1	AS12P_C	Analysis	Solution	3
	2	ES34P_U	Solution	Issue	4
	3	AS42P_B	Has_Environment	Step	1
	4	ES61P_H	4288	8500	2
	5	ED52C_J	3393	7697	1
	6	CD31C_61F	3267	10,018	3
	7	ES31P_186FD	2593	6050	1
	8	CD12C_7L	3544	7200	2
	9	BD12C_M	2788	7950	1
	10	CG51C_J	3776	10,650	2

For order 1002999879, the partial production data event log trajectories are shown in Table 5, mainly including the process, activity, start timestamp, end timestamp, material, and operation, all of which need to be synchronized with the process knowledge graph for the bottleneck analysis.

In addition, the operational knowledge required for the bottleneck analysis in this paper is primarily obtained from equipment failure report forms. From such data, specific causes and solutions can be identified, which support the results generated by the large models [34]. Moreover, the analysis plans summarized from these reports are more interpretable, making them easier for engineers to understand and analyze.

Table 5. Partial event log structure for the process.

Case ID	Production Process	Activity	Start Timestamp	End Timestamp	Attribution	…
1002999879	Steel Pretreatment	438101HCNR01.gen	2019-03-06 10:52	2019-03-06 11:03	8.71	…
1002999879	Chamfering	1438101HCNR02.gen	2019-03-06 11:10	2019-03-06 11:55	43.265	…
1002999879	Rotation	1438101HCNR03.gen	2019-03-06 13:30	2019-03-06 14:03	29.503	…
1002999879	Vertical Frame Welding	1438101HCNR04.gen	2019-03-06 14:15	2019-03-06 14:48	29.503	…
…	…	…	…	…	…	…
1002999879	Transverse Frame Assembly	1438101HCNR06.gen	2019-03-06 16:10	2019-03-06 17:00	42.061	…
…	…	…	…	…	…	…

4.1.2. Experimental Model Selection and Parameter Settings

To evaluate the effectiveness of SPKG in bottleneck analysis, this study establishes two control groups: the baseline LLMs and traditional RAG. An ablation study control group without the Question-Query Space is included to assess its role in helping SPKG generate more accurate queries. To address data privacy and cybersecurity concerns, our study employs an offline deployment of LLMs with anonymized and synthetic data, alongside rigorous security protocols to protect intellectual property and prevent breaches.

For the model hyper-parameter selection, the baseline LLMs are set with a temperature of 0. The embedding model used is BGE-M3, and the langchain framework is employed to implement the functionalities of retrieval-augmented generation (RAG). Milvus is utilized as a vector database [26], while Neo4j serves as the graph database [35]. The Python environment is managed using Anaconda, and the graph query language Cypher is executed within the Neo4j database server. The Question-Query Space is populated with a total of 900 query examples, and the experiments are conducted using Python version 3.10.

1. **Baseline LLMs:** This study employs GPT-3.5-Turbo, ChatGLM-4-9b-Chat [36], and Qwen2-72B-Instruct as the baseline models. Decision-makers ask the LLMs questions directly to analyze the root causes of bottlenecks without providing an external knowledge base containing process information, relying only on appropriate prompts.
2. **Traditional RAG:** The 5M1E information and data regarding the enterprise's process are split and embedded in pure text format, with the resulting high-dimensional vectors stored in a vector database as an external knowledge base. When answering questions, the LLMs first retrieve relevant content from this external knowledge base to serve as context.
3. **SPKG:** This approach provides the LLMs not only with the most similar text to the question as context but also with paths containing rich semantic information, namely the nodes and relationships within the manufacturing process knowledge graph. As the length and number of paths increase, the information provided becomes increasingly comprehensive.
4. **SPKG (without the Question-Query Space):** This variant excludes the Question-Query Space and provides only a few typical query examples within the prompt.

4.1.3. Evaluation Criteria

To facilitate the assessment of the effectiveness of the aforementioned methods in analyzing bottlenecks, this study sets the maximum number of inquiry rounds to 8. If decision-makers can identify the root cause of the bottleneck within 8 rounds of Q&A based on the LLMs' responses, they are considered effective. Moreover, the evaluation metrics in this section are inspired by the literature [37,38].

4.2. Results and Discussion

This subsection delineates the specific processes and activities within the ship manufacturing workflow that were identified as bottlenecks and examines their implications for operational decision-making. The SPKG method was applied to the thin plate production line, which encompasses a series of complex processes from steel plate pre-cutting to final assembly. The analysis revealed several critical bottlenecks:

- **The milling and plate assembly workstation:** A bottleneck was identified in the edge milling process, where the machine's operational efficiency was suboptimal, resulting in a delay in the subsequent assembly process;
- **The cutting workstation:** The CNC plasma cutting machine was identified as a bottleneck due to its high maintenance requirements and frequent downtime, which disrupted the flow of materials to subsequent workstations;
- **Transverse frame assembly:** The manual assembly of the transverse frame was found to be time-consuming and prone to errors, creating a bottleneck that impacted the overall production schedule.

The identification of these bottlenecks had a direct impact on operational decision-making:

- **Resource allocation:** Decision-makers reassigned resources to increase the maintenance frequency of the CNC plasma cutting machine, ensuring smoother operation and reducing downtime;
- **Process optimization:** The edge milling process was optimized through adjustments to the machine settings and operator training, which enhanced the efficiency of the milling and plate assembly workstation;
- **Automation:** The manual assembly process was partially automated, which reduced the error rate and shortened the assembly time for the transverse frame.

Table 6 presents the performance of the four models in the bottleneck analysis. The results indicate that SPKG outperforms both the baseline LLMs and traditional RAG in this context. The baseline LLMs demonstrate poor performance in the bottleneck analysis due to their lack of ship domain knowledge necessary to assist decision-makers in answering questions. SPKG excels in the bottleneck analysis because it provides the LLMs with not only context that is similar to the question but also structured and semantic context. As a result, its performance is superior to that of traditional RAG. However, without the aid of the Question-Query Space, SPKG still finds it challenging to respond flexibly to the diverse questions posed by production decision-makers, relying only on a few query examples provided in the prompt. Although SPKG shows the best performance among the models tested, the absolute success rate in the bottleneck analysis remains relatively low.

Table 6. Success rate in bottleneck analysis of 4 models.

Baseline LLMs	Bottleneck Analysis Success Rate (%)			
	Baseline LLMs	RAG	SPKG	SPKG (Without Question-Query Space)
GPT-3.5-Turbo	2.5	17.5	52	25
ChatGLM-4-9b-Chat	0	22.5	62.5	30
Qwen2-72B-Instruct	5	22.5	65	32.5

When investigating cases of failure in the SPKG bottleneck analysis, it was found that the LLMs often failed to successfully generate graph query statements. This issue is attributed not only to the limited reasoning capabilities of the LLMs but also to the lack of corresponding query examples in the Question-Query Space. After replacing the query examples in Question-Query Space, further experiments were conducted to assess the success rate of the LLMs in generating query statements when the questions and query examples involved different types of nodes, as shown in Figure 6. The results indicate that the success rate for the LLMs in generating query statements is highest when the query examples include the types of nodes relevant to the questions. Conversely, when the node types differ, the absolute success rate remains low. To improve the success rate of the SPKG bottleneck analysis, the Question-Query Space should encompass query examples that involve a variety of node types.

The Q&A chain was constructed to evaluate the root cause analysis performance of the LLMs after identifying bottlenecks, which enhanced the interactive ability between the LLMs and the knowledge graph. The essence of the Q&A chain was to translate the text inquiries into Cypher instructions to help the LLMs understand and retrieve the knowledge graph constructed in Section 3.1, thereby obtaining efficiency in the analysis results. For example, the following sequence of inquiries was used to analyze the bottleneck at the cutting workstation: (1) What is the current downtime frequency of the CNC plasma cutting machine? (2) What are the common causes of downtime for this machine? (3) How does the maintenance schedule affect the machine's operational efficiency? (4) What measures can be taken to reduce downtime and improve efficiency? The operating information of the Q&A chain is detailed in Section 3.3.

Each inquiry was designed to extract pertinent information from the knowledge graph, providing decision-makers with actionable insights to address the identified bottlenecks.

Adjusting the number of query examples in the Question-Query Space reveals the performance of SPKG in the bottleneck analysis, as illustrated in Figure 7. With an increase in the query examples, the success rate of the bottleneck analysis for all three LLMs rises significantly, but the rate of the increase gradually levels off. Once the number of query examples approaches a certain critical value, the success rate of the bottleneck analysis

no longer changes appreciably, and the curve tends to flatten. This phenomenon occurs because a knowledge graph, with a defined number of node and relationship types, as well as attribute types, requires only a limited number of query examples to thoroughly describe its pattern layer. For the manufacturing process knowledge graph defined in this paper, this critical value is approximately 600 examples. After reaching this threshold, the success rate of the bottleneck analysis is constrained by the reasoning and code generation capabilities inherent to the LLMs themselves.

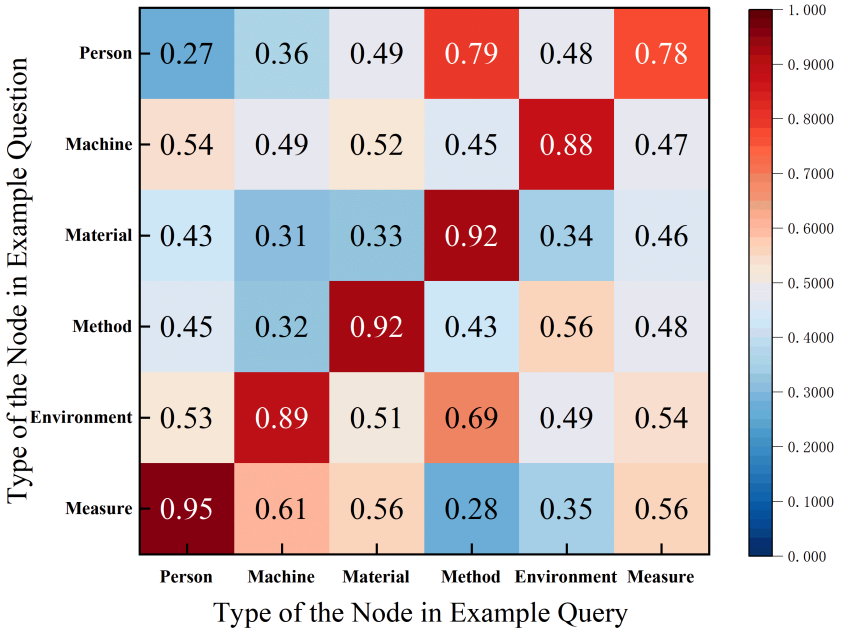


Figure 6. Success rates of LLMs generating Cypher when node types of questions are different from those of query examples.

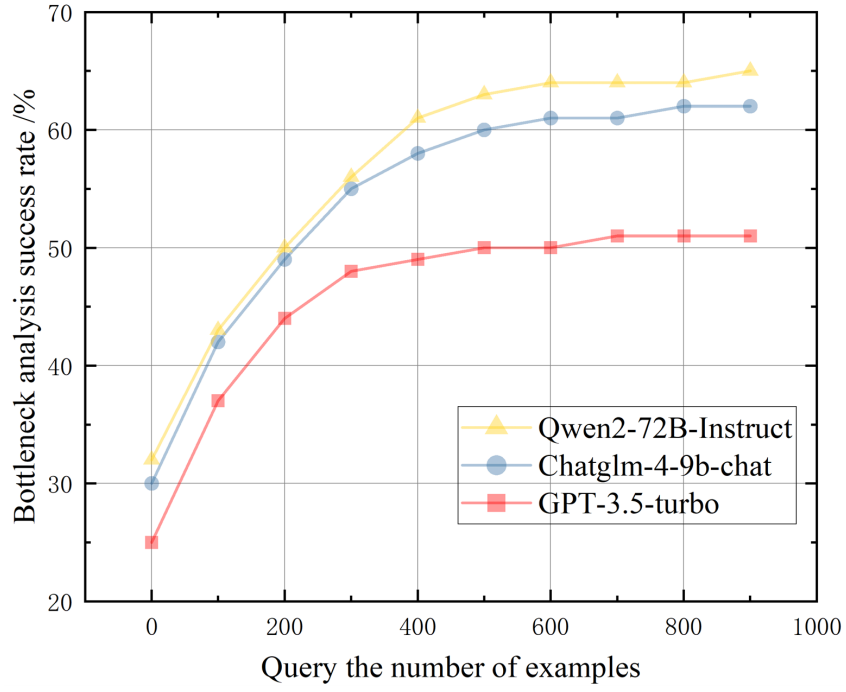


Figure 7. Effect of the number of query examples on success rates of bottleneck analysis.

In brief, the case study validates the efficacy of the SPKG method in identifying and addressing bottlenecks within the ship manufacturing process. By integrating domain-specific

knowledge and leveraging large language models, the SPKG method not only enhances the interpretability of the bottleneck analysis but also informs operational decision-making towards more efficient production practices.

5. Conclusions

The primary aim of SPKG proposed in this study is to solve the challenge of throughput bottleneck identification in the shipbuilding industry, focusing on enhancing the efficiency and interpretability of bottleneck analyses. This method integrates the "turning point" bottleneck detection approach, utilizing contextual information from a manufacturing process knowledge graph to analyze the causes of bottlenecks and provide guidance on improvements to production decision-makers. The experimental results demonstrate the effectiveness of SPKG in identifying and analyzing bottlenecks in ship production processes, offering a significant improvement in the traceability and interpretability of the analysis outcomes. The case studies illustrate how SPKG can facilitate the identification of bottleneck causes and suggest improvements, contributing to a more efficient production process and serving as a reference for digital transformation in the industry.

To solve the above issues, the main research findings of this paper are summarized as follows:

1. SPKG demonstrates superior effectiveness and interpretability in the bottleneck analysis compared to those of the standard LLMs and conventional RAG methods. This indicates that the integration of the production process knowledge graph to aid the LLMs in the analysis of causes of bottlenecks can offset their shortcomings in terms of inferential abilities, leading to the more precise identification of causes of bottlenecks. The experimental results in Section 3.2 show that SPKG as proposed in this paper had success rates of 52%, 62.5%, and 65% in the bottleneck analysis for GPT-3.5-Turbo, ChatGLM-4-9b-Chat, and Qwen2-72B-Instruct, respectively. Compared to the two baseline methods of the LLMs and RAG, the performance was improved by 13 to 20.8 times and 2.7 to 2.9 times, respectively. Specifically, SPKG based on the Qwen2-72B-Instruct model has the highest performance and the strongest analytical capability. Therefore, this paper selected this model for bottleneck mining, which improved the accuracy of the analysis.
2. The construction of the SPKG question–answer chain is essential for the improvement of its accuracy. Without the assistance of the question–answer chain in the ship thin plate production line, the success rate of the LLMs generating Cypher query statements is low, making it difficult to effectively analyze the causes of bottlenecks. The experiment in Section 3.2 shows that the SPKG method proposed in this paper improves the performance by 10% to 108% compared to the removal of the Question-Query Space. Therefore, this paper selected SPKG with the Question-Query Space for bottleneck mining, which could improve the accuracy of bottleneck issue analysis in ship manufacturing.
3. For ship manufacturing, the design of the Question-Query Space is crucial to the precision performance of SPKG. The quantity and quality of query examples will affect the success rate of LLMs generating Cypher query statements and consequently the success rate of the bottleneck analysis. The experiment described in Figure 7 shows that the upper limit of the question–answer queries of SPKG proposed in this paper depends on the reasoning capability constraints of the LLMs themselves and the construction of the schema layer of the process knowledge graph. Among these, SPKG proposed in this paper exhibits the highest performance in its bottleneck analysis success rate when the LLM chosen is the Qwen2-72B-Instruct model; as the number of query examples increases proportionally, all of the models gradually fall into a locally

optimal state, and their performance ceases to improve after the number of queries reaches 600. Therefore, the number of multi-turn inquiries is limited in this paper to ensure the effectiveness of the bottleneck analysis in ship manufacturing.

However, the research in this paper also has its limitations, which mainly depend on the quality of the construction of the process knowledge graph and the design of the prompt engineering that the LLMs can support. The primary aim of this research is to use the ship production field, which is highly dynamic and diverse, as a case study to provide an application reference for bottleneck analyses and digital transformation in other industrial sectors. The analysis of failures in the SPKG bottleneck analysis primarily indicated that the LLMs incorrectly generated query statements, failing to retrieve context relevant to the questions. Future research will focus on three main areas: (1) expanding the corpus of query examples within the Question-Query Space to refine the inferential capabilities of the LLMs; (2) developing structured methods to enhance the pattern layer description of the manufacturing process knowledge graph; and (3) advancing the use of LLMs with enhanced instruction-following and logical reasoning abilities to further refine the analytical capabilities of SPKG. These directions aim to solidify the scientific foundation of the method and broaden its applicability across various industrial sectors.

Author Contributions: Conceptualization: Y.M. and T.W. Methodology: T.W. Software: Y.M. Validation: B.Z. and J.B. Formal analysis: Y.M. Investigation: Y.M. Resources: J.D. and T.W. Data curation: B.Z. and J.B. Writing—original draft preparation: X.L. and T.W. Writing—review and editing: B.Z. and J.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Municipal Natural Science Foundation of Shanghai (21ZR1400800).

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to data privacy and security concerns restrictions in the Shipbuilding company.

Conflicts of Interest: Author Yanjun Ma and Jiwang Du were employed by the company Hudong Shipbuilding. Author Xiaoyang Liang was employed by the company Shanghai Dongxin Software Engineering Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Taşdemir, A.; Nohut, S. An Overview of Wire Arc Additive Manufacturing (WAAM) in Shipbuilding Industry. *Ships Offshore Struct.* **2021**, *16*, 797–814. [CrossRef]
2. Wu, Z.; Liang, C. A Review and Prospects of Manufacturing Process Knowledge Acquisition, Representation, and Application. *Machines* **2024**, *12*, 416. [CrossRef]
3. Song, D. A Literature Review, Container Shipping Supply Chain: Planning Problems and Research Opportunities. *Logistics* **2021**, *5*, 41. [CrossRef]
4. Rolli, F.; Parretti, C.; Barbieri, R.; Polidoro, A.; Bindi, B. RDF Knowledge Graphs Designed with Axiomatic Methodology to Enhance Industry 4.0. *Machines* **2025**, *13*, 58. [CrossRef]
5. Dogan, O.; Areta Hiziroglu, O. Empowering Manufacturing Environments with Process Mining-Based Statistical Process Control. *Machines* **2024**, *12*, 411. [CrossRef]
6. Çelik, D.; Meral, M.E.; Waseem, M. Investigation and Analysis of Effective Approaches, Opportunities, Bottlenecks and Future Potential Capabilities for Digitalization of Energy Systems and Sustainable Development Goals. *Electr. Power Syst. Res.* **2022**, *211*, 108251. [CrossRef]
7. Lai, X.; Shui, H.; Ding, D.; Ni, J. Data-Driven Dynamic Bottleneck Detection in Complex Manufacturing Systems. *J. Manuf. Syst.* **2021**, *60*, 662–675. [CrossRef]
8. Tonmoy, S.M.T.I.; Zaman, S.M.M.; Jain, V.; Rani, A.; Rawte, V.; Chadha, A.; Das, A. A Comprehensive Survey of Hallucination Mitigation Techniques in Large Language Models. *arXiv* **2024**, arXiv:2401.01313. [CrossRef]

9. Zhou, B.; Li, X.; Liu, T.; Xu, K.; Liu, W.; Bao, J. CausalKGPT: Industrial Structure Causal Knowledge-Enhanced Large Language Model for Cause Analysis of Quality Problems in Aerospace Product Manufacturing. *Adv. Eng. Inform.* **2024**, *59*, 102333. [CrossRef]
10. Pan, Y.; Zhang, L. A BIM-data Mining Integrated Digital Twin Framework for Advanced Project Management. *Autom. Constr.* **2021**, *124*, 103564. [CrossRef]
11. Ding, H.; Pang, L.; Wei, Z.; Shen, H.; Cheng, X. Retrieve Only When It Needs: Adaptive Retrieval Augmentation for Hallucination Mitigation in Large Language Models. *arXiv* **2024**, arXiv.2402.10612. [CrossRef]
12. Wu, S.; Xiong, Y.; Cui, Y.; Wu, H.; Chen, C.; Yuan, Y.; Huang, L.; Liu, X.; Kuo, T.W.; Guan, N.; et al. Retrieval-Augmented Generation for Natural Language Processing: A Survey. *arXiv* **2024**, arXiv.2407.13193. [CrossRef]
13. Rakin, S.; Shibly, M.A.R.; Hossain, Z.M.; Khan, Z.; Akbar, M.M. Leveraging the Domain Adaptation of Retrieval Augmented Generation Models for Question Answering and Reducing Hallucination. *arXiv* **2024**, arXiv:2410.17783. [CrossRef]
14. Jin, B.; Xie, C.; Zhang, J.; Roy, K.K.; Zhang, Y.; Li, Z.; Li, R.; Tang, X.; Wang, S.; Meng, Y.; et al. Graph Chain-of-Thought: Augmenting Large Language Models by Reasoning on Graphs. *arXiv* **2024**, arXiv:2404.07103. [CrossRef]
15. Lang, H.; Yang, G.; Li, C.; Xu, J. Multisource heterogeneous transfer learning via feature augmentation for ship classification in SAR imagery. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5228814. [CrossRef]
16. Hu, Y.; Miao, X.; Si, Y.; Pan, E.; Zio, E. Prognostics and health management: A review from the perspectives of design, development and decision. *Reliab. Eng. Syst. Saf.* **2022**, *217*, 108063. [CrossRef]
17. Benz, L.; Münch, C.; Hartmann, E. Development of a search and rescue framework for maritime freight shipping in the Arctic. *Transp. Res. Part A Policy Pract.* **2021**, *152*, 54–69. [CrossRef]
18. Zhang, C.; Zhang, X.; Gao, G.; Lang, H.; Liu, G.; Cao, C.; Song, Y.; Guan, Y.; Dai, Y. Development and application of ship detection and classification datasets: A review. *IEEE Geosci. Remote Sens. Mag.* **2024**, *12*, 12–45. [CrossRef]
19. Zhang, Y.; Li, S.; Liu, J.; Yu, P.; Fung, Y.R.; Li, J.; Li, M.; Ji, H. Knowledge overshadowing causes amalgamated hallucination in large language models. *arXiv* **2024**, arXiv:2407.08039.
20. Li, J.; Horiguchi, Y.; Sawaragi, T. Counterfactual inference to predict causal knowledge graph for relational transfer learning by assimilating expert knowledge–Relational feature transfer learning algorithm. *Adv. Eng. Inform.* **2022**, *51*, 101516. [CrossRef]
21. Xu, Z.; Dang, Y. Data-driven causal knowledge graph construction for root cause analysis in quality problem solving. *Int. J. Prod. Res.* **2023**, *61*, 3227–3245. [CrossRef]
22. Hou, B.; Lin, Y.; Li, Y.; Fang, C.; Li, C.; Wang, X. KG-PLPPM: A Knowledge Graph-Based Personal Learning Path Planning Method Used in Online Learning. *Electronics* **2025**, *14*, 255. [CrossRef]
23. Wu, T.; Li, J.; Bao, J.; Liu, Q. ProcessCarbonAgent: A Large Language Models-Empowered Autonomous Agent for Decision-Making in Manufacturing Carbon Emission Management. *J. Manuf. Syst.* **2024**, *76*, 429–442. [CrossRef]
24. Liu, J.; Zheng, D.; Wang, K.; Li, Z.; Liu, S.; Peng, L.; Yang, D. Evolutionary Manufacturing Approaches for Advancing Flexible Perovskite Solar Cells. *Joule* **2024**, *8*, 944–969. [CrossRef]
25. Ugurlu, H.; Cicek, I. Analysis and Assessment of Ship Collision Accidents Using Fault Tree and Multiple Correspondence Analysis. *Ocean. Eng.* **2022**, *245*, 110514. [CrossRef]
26. Chen, M.; Tao, Z.; Tang, W.; Qin, T.; Yang, R.; Zhu, C. Enhancing emergency decision-making with knowledge graphs and large language models. *Int. J. Disaster Risk Reduct.* **2024**, *113*, 104804. [CrossRef]
27. Singh, C.; Inala, J.P.; Galley, M.; Caruana, R.; Gao, J. Rethinking interpretability in the era of large language models. *arXiv* **2024**, arXiv:2402.01761.
28. Li, B.; Jiang, Y.; Gadepally, V.; Tiwari, D. LLM Inference Serving: Survey of Recent Advances and Opportunities. *arXiv* **2024**, arXiv:2407.12391. [CrossRef]
29. Friha, O.; Ferrag, M.A.; Kantarci, B.; Cakmak, B.; Ozgun, A.; Ghoualmi-Zine, N. Llm-Based Edge Intelligence: A Comprehensive Survey on Architectures, Applications, Security and Trustworthiness. *IEEE Open J. Commun. Soc.* **2024**, *5*, 5799–5856. [CrossRef]
30. Fonseca, Í.A.; Gaspar, H.M. Challenges When Creating a Cohesive Digital Twin Ship: A Data Modelling Perspective. *Ship Technol. Res.* **2021**, *68*, 70–83. [CrossRef]
31. Schuster, D.; van Zelst, S.J.; van der Aalst, W.M.P. Utilizing Domain Knowledge in Data-Driven Process Discovery: A Literature Review. *Comput. Ind.* **2022**, *137*, 103612. [CrossRef]
32. Van Landuyt, D.; Wijshoff, V.; Joosen, W. A Study of NoSQL Query Injection in Neo4j. *Comput. Secur.* **2024**, *137*, 103590. [CrossRef]
33. Laher, B.; Buzzi, C.; Brunnhofer, P.; Leitner, M. Effect of Grinding on the Fatigue Strength of Mild and High Strength Steel Fillet Weld Joints. *Results Eng.* **2024**, *22*, 102135. [CrossRef]
34. Wang, D.; Ouyang, D.; Chen, Y.; Xu, S.; Lv, Q.; Zhai, J.; Liu, M.; Ruan, S. Laser Transmission Welding of Glass-Fiber-Reinforced Polymer and Stainless Steel with a Three-Dimensional-Printing Surface Pretreatment. *Opt. Laser Technol.* **2024**, *168*, 109967. [CrossRef]

35. Mansoor, M.H.; Lu, X.; Woksepp, H.; Sattar, A.; Humak, F.; Ali, J.; Li, R.; Bonnedahl, J.; Mohsin, M. Detection and Genomic Characterization of Klebsiella Pneumoniae and Escherichia Coli Harboring Tet (X4) in Black Kites (Milvus Migrans) in Pakistan. *Sci. Rep.* **2024**, *14*, 9054. [CrossRef]
36. GLM, T.; Zeng, A.; Xu, B.; Wang, B.; Zhang, C.; Yin, D.; Zhang, D.; Rojas, D.; Feng, G.; Zhao, H.; et al. ChatGLM: A Family of Large Language Models from GLM-130B to GLM-4 All Tools. *arXiv* **2024**, arXiv:2406.12793. [CrossRef]
37. Guo, T.; Chen, X.; Wang, Y.; Chang, R.; Pei, S.; Chawla, N.V.; Wiest, O.; Zhang, X. Large Language Model Based Multi-Agents: A Survey of Progress and Challenges. *arXiv* **2024**, arXiv:2402.01680. [CrossRef]
38. Mehandru, N.; Miao, B.Y.; Almaraz, E.R.; Sushil, M.; Butte, A.J.; Alaa, A. Evaluating Large Language Models as Agents in the Clinic. *npj Digit. Med.* **2024**, *7*, 84. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.