This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/ICAC65379.2025.11196619, 2025 30th International Conference on Automation and Computing (ICAC)

Market and Production Management of Air Bearings using Ontology-based Knowledge Graph and NLP

Zhiyuan Jia
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
chihyuan02@163.com

Ning Gou
Department of Research and Development
Infina Precision
Jiaxing, China
gou.n@infina-tech.com

Kai Cheng
Mechanical and Aerospace Engineering
Brunel University
London, UK
cheng.kai@brunel.ac.uk

Xiaohui Zhu
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
0000-0003-1024-5442

Yong Yue*
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
Suzhou, China
0000-0001-7695-4538

***Abstract*—This research presents an ontology-based framework designed to enhance market and production management for specialized Small and Medium-sized Enterprises (SMEs) in the precision manufacturing sector, using air bearings as a case study. The framework was validated through real-world applications in a precision manufacturing SME and features a knowledge graph that was used for staff education and improved operational efficiency. Our approach integrates expert-driven ontology modeling with a modular web crawler and advanced Natural Language Processing (NLP) techniques, specifically the BERT-based ELECTRA model, for entity recognition and relationship extraction. This enables the framework to identify potential customers and competitors from unstructured online data, improving strategic decision-making. Key outcomes include enhanced supply chain transparency, more accurate market positioning, and better identification of long-tail demand patterns. The study demonstrates the practical value of combining knowledge graphs with NLP in industrial applications and suggests future extensions such as integration with large language models and the development of a domain-specific search engine.**

***Index Terms*—Air Bearing; Knowledge Graph; Market Intelligence; Ontology; Industrial Management**

I. INTRODUCTION

Air bearings play a crucial role in high-precision manufacturing, particularly in semiconductor production, metrology, and optical processing [1]. By eliminating mechanical contact through a thin film of compressed air, they enable frictionless motion essential for nanometer-level accuracy [2]. This technology reduces mechanical wear, enhances positioning precision, and supports the demands of next-generation fabrication systems.

The development of air bearing technology marked a shift from conventional contact-based support to non-contact solutions. Early theoretical foundations, such as Reynolds' equation, helped establish the behavior of thin gas films and remain central to performance optimization [2]. Progress accelerated during the 1950s with the emergence of systematic studies and international collaboration.

Air bearings are broadly classified into aerostatic, aerodynamic, hybrid, and squeeze film types [2]. Each type offers unique advantages: aerostatic bearings provide high stability through external pressurization, while aerodynamic types excel in high-speed environments. Hybrid and squeeze film types are used for more specialized scenarios. As semiconductor equipment demands greater precision and cleanliness, air bearings are increasingly favored for their durability, zero-contamination characteristics, and ability to meet sub-micron motion requirements.

However, the development and adoption of air bearing technology remain concentrated among a few global players, leaving specialized Small and Medium-sized Enterprises (SMEs) with limited access to structured market intelligence and application-driven strategies. The lack of comprehensive knowledge frameworks restricts these SMEs' ability to position themselves competitively and respond to dynamic market trends.

Despite technological potential, SMEs face common and persistent challenges: fragmented market information scattered across unstructured data sources such as technical reports, online forums, and company disclosures [3], [4]; difficulties in standardizing product lines due to the dominance of highly customized and experimental orders [5]; and intense competition from larger, well-established international players who benefit from economies of scale and brand recognition [6].

To address these issues, this research proposes an integrated framework for intelligent market analysis and production management based on ontology and Natural Language Processing (NLP). Our approach combines expert-driven ontology modeling, modular data acquisition mechanisms, and ELECTRA-based entity recognition [7] to extract, organize, and visualize unstructured industrial data into a coherent and evolving strategic knowledge graph.

Validated through real-world applications in a precision manufacturing SME, the proposed framework not only improved internal staff training and customer targeting but also enhanced technological alignment by revealing key relationships among technical fields, subfields, companies, and products. This structured visibility empowers SMEs to identify underserved market segments, optimize their positioning, and support more agile decision-making processes.

This study demonstrates how the integration of ontology and NLP techniques can help SMEs overcome structural information barriers, enhance competitive intelligence, and better adapt to fragmented and rapidly evolving industrial markets. It is particularly effective in addressing long-tail demand patterns, where niche opportunities constitute a significant but often overlooked portion of the high-precision manufacturing landscape [8].

II. Related Work

Air bearing technology, originating from nineteenth-century air lubrication experiments, has undergone continuous evolution to become a foundational component of high-precision motion systems [1]. By eliminating mechanical contact through the maintenance of a stable air film, air bearings enable extremely low friction, high stiffness, and nanometer-level positioning accuracy, making them indispensable in advanced semiconductor equipment, metrology platforms, and optical manufacturing systems [1], [2].

Despite achieving technological maturity in certain sectors, air bearing systems continue to face technical and market challenges. Dynamic behavior modeling under varying loads, thermal management at high operating speeds, and the high cost of integration remain major obstacles to broader adoption across industrial applications [2].

Parallel to the evolution of air bearing technologies, knowledge graphs have emerged as powerful tools for structuring complex industrial knowledge. A knowledge graph organizes entities, attributes, and relationships in a flexible, semantic manner, enabling richer information representation compared to traditional relational databases [12]. Originally developed for web search optimization, knowledge graphs are increasingly applied in industrial domains for tasks such as supply chain visualization, competitive landscape mapping, and technology trend analysis [11], [20].

Recent advances in NLP, particularly the introduction of pre-trained models like ELECTRA [7], have significantly enhanced the efficiency and accuracy of information extraction from unstructured text sources. This has made the automatic construction of domain-specific knowledge graphs more feasible and less reliant on extensive manual curation, especially when integrated with traditional linguistic resources like part-of-speech tagging and dependency parsing [10], [14].

SMEs in high-tech manufacturing sectors, including air-bearing platform providers, face distinct disadvantages in gathering and utilizing market intelligence. Information fragmentation, data sparsity, and the overwhelming presence of larger competitors make traditional intelligence-gathering methods inefficient and costly [3], [5], [6]. Knowledge graph-based methods offer a promising alternative, enabling SMEs to structure fragmented information dynamically, discover hidden opportunities, and simulate market scenarios based on graph analytics [8], [9].

Moreover, in technically dense and rapidly evolving industries, human expertise remains essential for validating and refining automatically extracted knowledge. Human-in-the-loop systems, which integrate expert review into automated processes, have proven effective in enhancing semantic accuracy and maintaining the relevance of evolving knowledge bases [9], [13].

Building upon these foundations, this study presents a practical framework that combines ontology modeling, knowledge graph construction, and NLP-based information extraction. Designed specifically for precision manufacturing SMEs, it aims to bridge the gap between fragmented industrial knowledge and actionable strategic insights, particularly in the complex and opportunity-rich domain of air-bearing applications.

III. Methodology

This section outlines the design and implementation of a modular framework supporting real-time market intelligence gathering and production decision-making for SMEs in the air bearing industry. The framework integrates three core components: ontology-driven knowledge graph construction, automated information acquisition, and an NLP-based entity recognition and integration module [9].

A. Knowledge Graph Construction

The knowledge graph module builds a structured representation of the semiconductor equipment industry using Neo4j, a graph database optimized for complex relationship management. We first defined a domain-specific ontology, capturing concepts such as companies, technical domains, subfields, products, and their interconnections [10].

A top-level classification system comprising 13 technical fields (e.g., Lithography, Etching, CVD) and corresponding subfields (e.g., EUV, Dry Etching) was established. Each field was mapped to key market players like ASML, AMAT, Lam Research, Tokyo Electron, and KLA based on their product specialization and market dominance. Field strengths were encoded with three levels: leading (3), strong participation (2), and involvement (1).

Cypher queries were used to construct the graph, creating nodes and edges representing hierarchical and competitive relationships. The resulting graph not only served as a centralized knowledge repository but also enabled visual exploration, helping SMEs analyze supply chain structures, identify strategic gaps, and support internal training and market planning [11], [12].

B. Information Acquisition

To enrich and update the knowledge graph dynamically, a hybrid information acquisition system was developed, combining automated crawlers with Robotic Process Automation (RPA) tools.

Custom Python-based crawlers were built with a modular architecture, supporting dynamic content parsing, authentication, API communications, and structured data output. They were responsible for scraping structured and semi-structured data from public sources such as company websites, patent databases, and technical forums.

Meanwhile, Octoparse, a commercial RPA tool, was integrated to allow non-technical staff to participate in data acquisition. Its visual interface enabled quick deployment of web scraping workflows without coding, facilitating broader organizational collaboration [13].

To ensure data reliability, a human-in-the-loop verification stage was embedded after data collection. Human reviewers cross-checked information for accuracy, relevance, and uniqueness, filtering out outdated, redundant, or irrelevant records. This dual-process design combined the efficiency of automation with the semantic judgment of domain experts.

C. Entity Recognition and Knowledge Integration

After data collection, unstructured text materials were processed through an advanced NLP pipeline. We implemented a Chinese-language NLP system based on the ELECTRA model [7] and the Language Technology Platform (LTP) [14], capable of performing segmentation, part-of-speech tagging, dependency parsing, and Named Entity Recognition (NER).

The pipeline targeted five main entity types: organizations, persons, locations, technical keywords, and inter-entity relationships. Recognized entities were classified and either mapped to existing nodes in the knowledge graph or used to generate new nodes and edges, allowing the graph to organically expand over time.

The system architecture followed a Django-based Model-View-Controller (MVC) pattern, ensuring modular separation of data processing and interface management. ECharts was used for dynamic, interactive visualization of the knowledge graph, empowering domain experts and business teams to browse, annotate, and analyze complex relationship networks with minimal technical barriers.

By integrating real-time entity extraction with graph updates, the framework enabled SMEs to continuously align internal knowledge with external market signals, responding to new opportunities and competitive threats effectively [15].

D. Summary

The framework's three modules—ontology-based knowledge graph construction, modular information acquisition, and NLP-driven integration—operate independently yet cohesively. Together, they create a scalable, extensible platform for SMEs to bridge fragmented market intelligence, enhance strategic visibility, and support data-informed decision-making across dynamic industrial landscapes.

IV. Results

This section presents the key outcomes of our framework implementation, including the construction of a domain-specific knowledge graph and the insights derived from market and product data analysis related to air-bearing platforms.

A. Knowledge Graph Overview

We constructed a domain-specific knowledge graph to model the structure of the semiconductor equipment industry. This graph was designed to connect companies, product lines, technical fields, and subfields to support strategic insights and market visualization.

The knowledge graph is composed of 79 nodes and 136 relationships, categorized into four major node types: *Company*, *Field*, *Subfield*, and *Product*. It was implemented and visualized using Neo4j. Figure 1 shows a full snapshot of the final graph layout.

Fig. 1. Knowledge graph visualization of the semiconductor equipment supply chain. Each color represents a distinct entity type. Edges indicate structured relationships such as `LEADER_IN`, `BELONGS_TO`, and `HAS_PRODUCT`.

To populate the graph, we collected product and domain data from five dominant vendors: ASML, Applied Materials (AMAT), Tokyo Electron (TEL), Lam Research (LRCX), and KLA. These companies collectively account for over 76.9% of the global semiconductor equipment market [16]. Each company's participation in 13 major fields was quantified according to its level of specialization. Table I summarizes the classification scheme and the assigned field strength scores.

The encoding rule is as follows: 3 represents a leading position (`LEADER_IN`), 2 indicates strong involvement (`STRONG_IN`), and 1 indicates general participation (`INVOLVED_IN`). Based on this mapping, ASML holds a dominant role in Lithography, while KLA leads Metrology. AMAT maintains comprehensive participation across almost all fields except Lithography. TEL and LRCX specialize in process-heavy fields such as Etching, PVD, and Thermal Processing.

The knowledge graph structure reflects these observations. Subfields are connected to their parent fields using `BELONGS_TO`, and products are associated with subfields.

Companies are linked to fields through role-encoded edges. This configuration enables flexible traversal from a macro (field-level) to micro (product-specific) view, supporting a wide range of use cases from customer analysis to strategic planning.

TABLE I
FIELD AND SUBFIELD DISTRIBUTION OF MAJOR SEMICONDUCTOR EQUIPMENT COMPANIES

Field / Company	ASML	AMAT	LRCX	TEL	KLA
Lithography	3	0	0	0	0
Etching	0	3	3	2	0
CVD	0	3	2	2	0
Metrology	1	1	0	0	3
Cleaning	0	2	3	2	0
PVD	0	3	1	1	0
Track	0	0	0	3	0
ALD	0	2	2	1	0
CMP	0	3	0	1	0
Thermal Processing	0	2	0	3	0
Ion Implantation	0	3	0	0	0
ECP	0	3	1	0	0
Dry Strip	0	2	3	1	0

B. Market Insights and Air-Bearing Product Landscape

With the help of the constructed knowledge graph, internal teams gained clarity on the company's position within the broader semiconductor equipment supply chain. This insight allowed technical and business personnel to identify key fields where domestic substitution was feasible, as well as recognize the types of data that were missing. Guided by this structure, the company initiated a targeted data acquisition effort to populate gaps relevant to air-bearing product lines.

The product data presented in this section was collected through a combination of automated web scraping using Python scripts and collaborative efforts from internal sales teams. Non-technical staff utilized RPA tools to assist in collecting product names, model sizes, and brand information from public websites. To improve processing efficiency, we applied NLP-based entity recognition (using ELECTRA and LTP) to parse product tables, filter irrelevant text, and extract valid company-product pairs. The resulting data enabled the construction of a usable and actionable competitor database.

Based on the collected and processed data, we present a structured comparison of product availability across major suppliers in the air-bearing turntable market. The tables below summarize the product offerings in three key air-bearing turntable series: **Measurement**, **Compact**, and **High-Speed**.

These insights reveal underserved market segments and inform future product planning. Most current orders are experimental or customized, suggesting air bearings have not yet achieved widespread adoption. This presents a strategic opportunity to build a case library, provide service-oriented offerings, and support niche customization markets such as research institutes and specialized manufacturing tools [8], [17].

TABLE II
MEASUREMENT SERIES PRODUCT DISTRIBUTION (DIAMETER IN MM)

Company	100	110	150	200	215	250	340
JS1	O		O	O		O	O
WX1	O		O				
SH1	O						
YB		O	O		O		
BJ1		O			O		
BJ1+		O			O		
SY		O			O		
XW			O	O			
HZ			O				
US1			O				
EU1			O				

TABLE III
COMPACT SERIES PRODUCT DISTRIBUTION (DIAMETER IN MM)

Company	100	120	150	200	300
JS1	O		O	O	O
WX1	O		O	O	
SH1	O	O			
XW1			O	O	
SZ1		O			
EU1	O		O		

TABLE IV
HIGH-SPEED SERIES PRODUCT DISTRIBUTION (DIAMETER IN MM)

Company	100
JS1	O
HZ	O
EU1	O

V. MARKET AND TECHNOLOGY ANALYSIS

Understanding the industrial context of air-bearing platforms requires positioning them within the broader digital economy. As shown in Figure 2, the technology value chain extends from mass-market applications (e.g., 5G, AI) to core enablers such as semiconductor manufacturing.

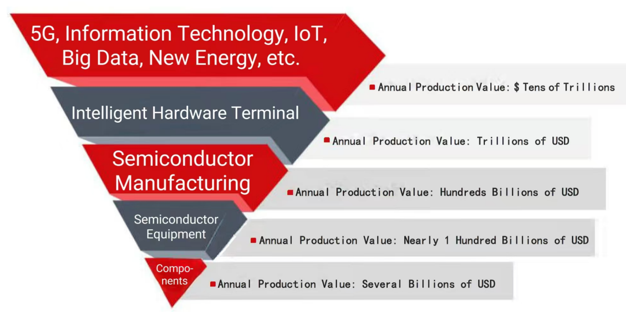

Fig. 2. Global Technology Industry Funnel from Application to Core Components

Semiconductor manufacturing, a \$1 trillion segment, relies on precision equipment valued at nearly \$100 billion annually [16]. Air-bearing motion systems, though representing a smaller market, are critical for enabling ultra-stable, vibration-free motion, increasingly demanded by Moore's Law scaling, AI acceleration, and advanced packaging [18]. Positioned at the interface of physical mechanics and computational needs, air-bearing platforms align naturally with semiconductor equipment as the most strategic entry point.

A. Downstream Market Overview

Air-bearing systems are applicable in semiconductor production, display manufacturing, ultra-precision machining, and automation. Among these, semiconductors represent the largest and most policy-supported downstream opportunity [16].

As shown in Figure 3, the semiconductor equipment market is dominated by fields such as Etching (25.0%) and Lithography (19.7%), with five major companies—ASML, AMAT, TEL, LRCX, and KLA—controlling over 76.9% of the market.

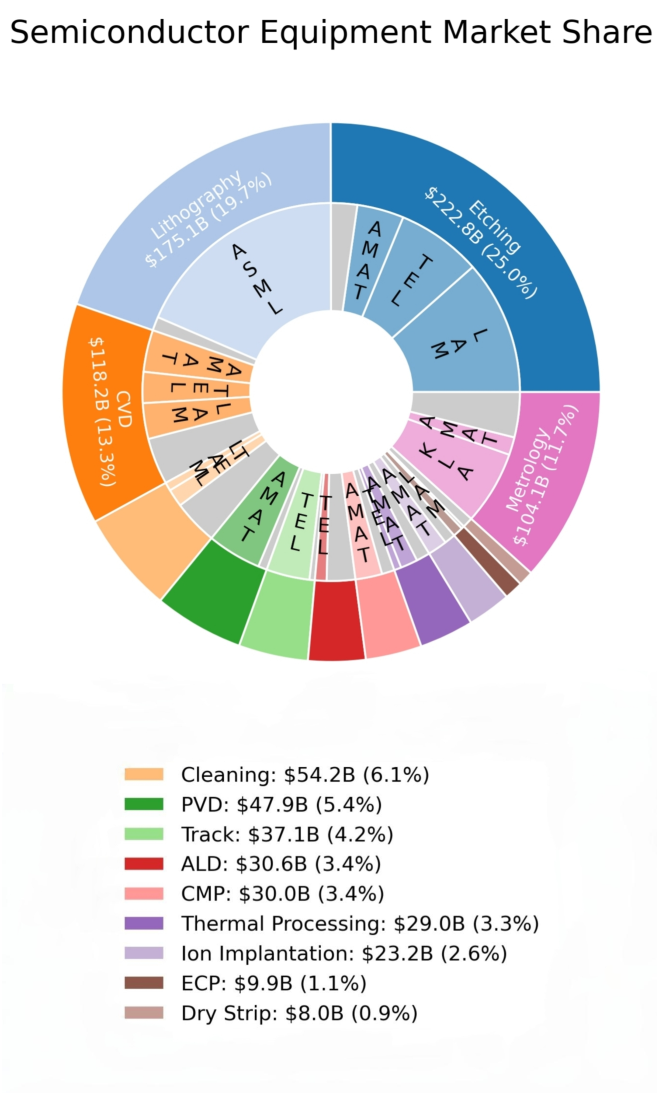

Fig. 3. Semiconductor Equipment Market Share by Field and Vendor

Table V highlights flagship equipment from these vendors, illustrating the need for ultra-high precision motion stages.

TABLE V
KEY PRODUCTS OF MAJOR SEMICONDUCTOR EQUIPMENT VENDORS

Company	Product	Field	Subfield
ASML	NXE:3400B	Lithography	EUV
AMAT	Centris Sym3	Etching	Dry Etch
LRCX	Kiyo Series	Etching	Plasma Etch
TEL	TELINDY Plus	Thermal Processing	RTP
KLA	CIRCL Series	Metrology	Optical

While adjacent markets such as OLED displays and automation exist, they are typically more fragmented and less scalable compared to the semiconductor industry.

B. Supply Chain Positioning

The precision equipment value chain comprises OEMs, system integrators, subsystem suppliers, unit suppliers, and element providers. SMEs can strengthen their role as Tier 2 or Tier 3 suppliers by focusing on modular air-bearing units, avoiding direct competition with full-system integrators while contributing to the high-end supply chain [19].

C. Technology and Policy Drivers

Two major trends further favor air-bearing adoption:

- **AI Computing Boom:** Accelerated demand for GPUs and sub-5nm chips drives investment in EUV lithography and metrology equipment, both requiring high-stability stages;
- **Domestic Substitution:** National policies in China promoting semiconductor independence stimulate demand for local motion control subsystems [17].

D. Strategic Summary

In conclusion, semiconductors offer the most scalable and strategically aligned growth path for air-bearing technologies. By refining their specialization and leveraging structured knowledge graph systems, SMEs can enhance their visibility and resilience within this complex, high-value ecosystem.

VI. CONCLUSION

This study presents an integrated framework for market and production management of air-bearing systems, combining ontology-based knowledge graphs and NLP techniques. It contributes both a practical tool for SMEs in precision manufacturing and a methodological reference for data-driven decision-making in fragmented industrial environments.

A. Key Findings

The research identified significant opportunities for air-bearing applications, particularly within the semiconductor equipment industry. Through the construction of a domain-specific knowledge graph, SMEs were able to visualize complex industry structures, improve internal training, and enhance strategic targeting. Additionally, the analysis revealed that most air-bearing orders remain highly customized, highlighting a ”long-tail” market structure where service-oriented business models and specialized solutions can yield a competitive advantage [8].

B. Theoretical and Practical Contributions

From a theoretical perspective, this work demonstrates how ontology-based knowledge graphs can organize fragmented industrial information into coherent structures, supporting semantic reasoning and strategic insight. Practically, the framework offers SMEs a way to build internal knowledge capabilities, bridge market intelligence gaps, and improve responsiveness to shifting industry demands. The approach is also adaptable to adjacent fields where precision motion and supply chain complexity are critical factors.

C. Limitations and Future Work

While the framework effectively integrates knowledge graph construction, modular information acquisition, and NLP-based entity recognition, limitations remain in multilingual data support, dynamic update automation, and semantic search capabilities. Future enhancements will focus on:

- Integrating Retrieval-Augmented Generation (RAG) techniques, such as LightRAG [21], to evolve into a domain-specific knowledge search system;
- Expanding entity recognition to cover more technical domains, enabling cross-industry strategic mapping;
- Incorporating global data sources to enhance international market intelligence tracking.

D. Final Remarks

This research highlights the strategic value of combining structured knowledge representation with practical industrial data. In a rapidly evolving global manufacturing landscape, frameworks like the one developed here provide SMEs with the necessary tools to achieve greater agility, strategic foresight, and sustainable growth.

Acknowledgment

This work was partially supported by the Suzhou Municipal Key Laboratory for Intelligent Virtual Engineering (SZS2022004).

References

[1] F. Wardle, *Ultra-precision bearings*. Elsevier, 2015.

[2] F. Al-Bender, *Air bearings: theory, design and applications*. John Wiley & Sons, 2021.

[3] S. Durst and I. R. Edvardsson, "Knowledge management in SMEs: a literature review," *Journal of Knowledge Management*, vol. 16, no. 6, pp. 879–903, 2012.

[4] N. Ayinuer, A. Ruxianguli, and A. Yasen, "Design and research of unstructured data knowledge graph tool based on Neo4j graph database," in *Proc. 11th Int. Conf. Communications, Circuits and Systems (ICCCAS)*, 2022.

[5] M. Levy, P. Powell, and P. Yetton, "The dynamics of SME information systems," *Small Business Economics*, vol. 19, pp. 341–354, 2002.

[6] Y. Liao and J. Barnes, "Knowledge acquisition and product innovation flexibility in SMEs," *Business Process Management Journal*, vol. 21, no. 6, pp. 1257–1278, 2015.

[7] K. Clark, "ELECTRA: Pre-training text encoders as discriminators rather than generators," *arXiv preprint arXiv:2003.10555*, 2020.

[8] H. Weeds, "Superstars and the long tail: The impact of technology on market structure in media industries," *Information Economics and Policy*, vol. 24, no. 1, pp. 60–68, 2012.

[9] B. Zhang, A. Meroño-Peñuela, and E. Simperl, "Towards explainable automatic knowledge graph construction with human-in-the-loop," *HHAI 2023: Augmenting Human Intellect*, IOS Press, 2023.

[10] T. Wu, G. Qi, C. Li, and M. Wang, "A survey of techniques for constructing Chinese knowledge graphs and their applications," *Sustainability*, vol. 10, no. 9, p. 3245, 2018.

[11] A. Hogan, E. Blomqvist, M. Cochez, *et al.*, "Knowledge graphs," *ACM Computing Surveys*, vol. 54, no. 4, pp. 1–37, 2021.

[12] L. Ehrlinger and W. Wöß, "Towards a definition of knowledge graphs," *SEMANTiCS (Posters, Demos, SuCCESS)*, vol. 48, pp. 1–4, 2016.

[13] S. Rahman, F. Choi, H. Kim, D. Zhang, and E. Hruschka, "Knowledge Acquisition and Integration with Expert-in-the-loop," *arXiv preprint arXiv:2402.03291*, 2024.

[14] W. Che, Y. Feng, L. Qin, and T. Liu, "N-LTP: An Open-source Neural Language Technology Platform for Chinese," in *Proc. EMNLP 2021 System Demonstrations*, pp. 42–49, 2021.

[15] H. Xiong, J. Bian, Y. Li, *et al.*, "When search engine services meet large language models: visions and challenges," *IEEE Transactions on Services Computing*, 2024.

[16] S. Ou, Q. Yang, and J. Liu, "The global production pattern of the semiconductor industry: an empirical research based on trade network," *Humanities and Social Sciences Communications*, vol. 11, no. 1, pp. 1–13, 2024.

[17] H. Liu, H. Zhao, and S. Li, "Future social change of manufacturing and service industries: Service-oriented manufacturing under the integration of innovation-flows drive," *Technological Forecasting and Social Change*, vol. 196, p. 122808, 2023.

[18] Y. Chen, "Research on Transformation and Upgrade Paths of China's Manufacturing Industry Based on the Sophisticated Integration of Industrial Chain and Innovation Chain," *Sustainable Development*, vol. 13, pp. 232–240, 2023.

[19] K.-H. Chang, Y.-J. Sun, C.-A. Lai, *et al.*, "Big data analytics energy-saving strategies for air compressors in the semiconductor industry–an empirical study," *International Journal of Production Research*, vol. 60, no. 6, pp. 1782–1794, 2022.

[20] X. Zou, "A survey on application of knowledge graph," *Journal of Physics: Conference Series*, vol. 1487, p. 012016, 2020.

[21] Z. Guo, L. Xia, Y. Yu, T. Ao, and C. Huang, "LightRAG: Simple and Fast Retrieval-Augmented Generation," *arXiv preprint arXiv:2410.05779*, 2024.