Automated Report Generation and Knowledge Management System for Photovoltaic Power Stations using Knowledge Graphs

1st Aili Pang[1,a], 2nd Shihao Wang[1,b*], 3rd Muhuan Wu[2,c], 4th Junrui Zhang[3,d]

[1]State Grid Shanghai Economic Research InstituteState Grid Corporation of China Co., Ltd, Shanghai, China
[2]Department of Engineering MathmaticsUniversity of Bristol, Bristol, UK
[3]Moham Digital Technology Co., Ltd, Shanghai, China

[a]Pang_aily@163.com, [b]Wangshh13@foxmail.com, [c]dn22438@bristol.ac.uk,[d]zhangjunrui986@163.com,

Abstract. In the context of swift expansion in the photo- voltaic (PV) industry, developing new PV power stations hascreated substantial volumes of heterogeneous, multi-source data across design, construction, and operational stages. Presently,the potential of this data remains largely untapped due to the inadequacy of existing document and knowledge management systems to integrate and manage it effectively, leavingit fragmented and underutilized. Recognizing the absence ofa comprehensive tool for the full-spectrum management and integration of data from project information to specific electrical designs in PV power station projects, this paper introduces a systematic solution designed to confront the intricate challenges of managing such diverse data. The system adeptly processesunstructured data through the adoption of automated knowledgegraphs for data organization and analysis, facilitating the generation of professional engineering reports and signifying a notableadvancement in knowledge management. The application of advanced ontological techniques for transforming unstructured data into a structured Resource Description Framework (RDF) format provides a sturdy foundation that enhances data integrityand accessibility while considerably reducing the manual labor associated with data processing. Additionally, the capability of this system to preserve complex data relationships within a graphdatabase denotes a significant enhancement over conventional tabular data storage methods, improving project management efficiency. This progress enables effortless data visualization and querying, allowing for smooth navigation through the knowledge graph to locate specific information and understand its interconnections. Automatically creating detailed, precise projectengineering reports from structured data and insights extracted from the knowledge graphs significantly uplifts communication and documentation standards within PV project management. This advancement offers satisfactory results for a broad spectrumof stakeholders needing rapid, comprehensive project insightsto inform their decision-making processes. It is a pivotal step towards optimizing the use and management of data in renewableenergy projects.

© The Author(s) 2024
T. Yao et al. (eds.), *Proceedings of the 2024 3rd International Conference on Engineering Management and Information Science (EMIS 2024)*, Advances in Computer Science Research 111,
https://doi.org/10.2991/978-94-6463-447-1_39

Keywords: Data Fusion, Knowledge Graph, Knowledge Management, Automated Report Generation

1 Introduction

The rapid evolution of the PV industry necessitates innovative approaches to PV power station project knowledge management and analysis. While supporting standard data storage and retrieval, traditional information and document management systems are inadequate in addressing the multi- source, heterogeneous PV project data [1]. Previous works forknowledge management in the power industry often focused on a pure knowledge modelling perspective [2]–[4]. A scalableframework that consists of real-time data updating, multi- source data integration and project report generation is yetto be developed. This paper introduces a novel knowledge management system framework with a report generation function that leverages the power of knowledge graphs and graph databases to enhance PV station project management.

The system addresses these challenges by automating project information collection, organization, and analysis. Using advanced ontological methods [5], multi-source, heterogenous data is converted into a structured Resource Description Framework (RDF) format [6], facilitating the construction of comprehensive knowledge graphs. These graphs are then stored in a graph database, which, unlike traditional tabular databases, excels in querying and preserving complex relation-ships between data points [7]. This approach not only enhancesdata integrity and accessibility but also significantly reduces the time and effort involved in data processing.

Moreover, the ability of this system to visualize data and provide query capabilities for all collected information marks a significant advancement in making complex data more understandable and actionable. Users can easily navigate the knowledge graph to find specific information and explore its connections to other relevant data. This feature is particularly beneficial for stakeholders who require quick access to com- prehensive project insights to make informed decisions.

Another standout feature of this system is its ability togenerate engineering project reports automatically. Detailed, accurate, and visually appealing reports are produced by leveraging the structured data and insights derived from the knowledge graphs. This automation not only streamlines the reporting process but also enhances the quality and professionalism of the documents, contributing to more effective communication and documentation practices within the PV project management domain.

Introducing this innovative system marks a significant leap forward in PV station project management. By addressing the critical challenges of data processing, visualization, and reportgeneration, a comprehensive solution is offered that not only improves efficiency and accuracy but also paves the way for more strategic and informed project management practices.

The rest of the paper is structured as follows: Section II details existing research and provides a comprehensive literature review, identifying gaps in the current understanding.Section III describes the system design, delving into the architectural and functional aspects of the developed system. Section IV presents an evaluation of the system, outlining the methodologies used for testing and discussing the outcomesto assess the effectiveness of this system. Finally, the paper concludes in Section V with a discussion on the implications of this research and explores potential avenues for future work.

2 Literature Review

In this section, a critical examination of existing literature related to knowledge graphs and their application within photovoltaic power systems is conducted. Key limitations in current methodologies are identified, and advancements indata fusion techniques are explored, paving the way for the introduction of contributions that address these identified gaps.

2.1 Related Work

Knowledge graph for knowledge management

In the field of power engineering, the application of knowledge graph technology has garnered widespread attention. As a potent tool for data integration and analysis, knowledge graphs effectively manage complex project information, facilitating the discovery and utilization of relationships between data [8], [9]. Knowledge graphs have demonstrated proficiency across various applications, notably semantic search [10], recommender systems [11] and E-commerce [12]. The construction of knowledge graphs is based on robust ontologies and semantic networks, enabling structured representation of complex relationships between concepts and entities. For instance, Zhou investigated how ontologies could be utilized to build a knowledge graph in the PV domain to enhance data utilization in project management [1]. Constructing an ontology involves not only a deep understanding of domain knowledge but also accurate descriptions of relationships between data, which is crucial for improving the quality of project decision-making. In the practical application of knowledge graphs, data fusion and analysis are key steps. By integrating data from diverse sources into a unified knowledge graph, project managers can gain more intuitive access to project status and performance indicators. A case was represented where knowledge graphs were used for data fusion, significantly improving data utilization efficiency and project management outcomes [13]. Despite the considerable potential of knowledge graphs in PV power engineering project management, some challenges persist in their practical application. Issues related to data quality and consistency, the costs of updating and maintaining knowledge graphs, and the need for advanced data analysis skills are major obstacles that need to be overcome. Future research should focus on improving the automation level of data integration, optimizing knowledge graph updating mechanisms, and developing more user-friendly query and analysis tools. In conclusion, knowledge graphs offer a new perspective and tool for PV power engineering project management, effectively integrating and analyzing complex data within projects. With further research and technological innovation, knowledge graphs are expected to play a more significant role in enhancing project management efficiency and decision- making quality.

Data fusion and analysis

Integrating multi-source, heterogeneous data has emerged as a pivotal area of research, offering profound insights into integrating diverse data types forenhanced understanding and analysis. Notably, using knowledge graphs in heterogeneous data fusion presents a significantadvantage by providing a structured framework to represent entities and relationships across different data modalities [14].Data fusion involves integrating various data types, such as text, images, audio, and video, to understand information comprehensively. By combining different modalities, the goal is to leverage the richness and complementary nature of each data type to enhance the accuracy and completeness of data analysis [15]. As for the role of knowledge graphs in multimodal data fusion, they are a powerful tool that structuresinformation into entities and relationships. These graphs rep- resent complex interconnections between different data types, enabling a holistic view of information integration. Knowledgegraphs provide a structured representation of entities and relationships, allowing for a more organized and interpretable view of multimodal data. Researchers can conduct sophisticated data inference processes across multiple modalities by leveraging knowledge graphs, enabling more profound insightsand predictions. It has efficient Data Integration, streamlining the process of integrating diverse data types by offering a unified framework for storing and processing information fromvarious sources [16]. Integrating multi-source, heterogeneous data with knowledge graphs provides a promising avenue for advancing data analysis and understanding.

Graph database

Neo4j, a leading graph database, has emerged as a cornerstone technology for developing knowledge graph systems in various domains, including power engineering. Its unique ability to model intricate relationships between data entities in a natural and intuitive manner makesit particularly well-suited for applications where complex connections between data points are paramount [17]. Within the domain of PV power systems, Neo4j offers significant advantages. It enables the integration of diverse data sources, including real-time sensor data, historical performance metrics, and environmental factors, into a unified and actionable data landscape. The graph-based nature of Neo4j facilitates efficient querying and analysis, allowing for complex datainference processes and pattern recognition tasks that are often challenging with traditional relational databases [18]. Moreover, Neo4j provides superior performance in traversing relationships and conducting analyses, thereby enhancing the scalability and responsiveness of knowledge graph systemsin managing large-scale PV project data. Its adoption within the PV power system domain underscores its potential to revolutionize knowledge management and decision-making processes in the field.

2.2 Research Gap

The integration of knowledge graphs into PV power systemmanagement presents promising opportunities, yet critical gapsremain in current research and practice. Specifically, there is alack of robust mechanisms for automating real-time data up- dates within knowledge graph systems tailored for PV power systems. Additionally, scalability and performance issues hinder the effective management of large-scale PV project data. Furthermore, the absence of efficient tools for generating real-time project reports from knowledge graph data and the limitedexploration of predictive analytics for hazard identificationposes significant challenges. Addressing these gaps is essentialfor advancing PV power system management practices and unlocking the full potential of knowledge graphs in the field.

3 System Design

3.1 System Overview

The system design uses knowledge graph and graph modeling techniques to enhance PV project knowledge management. Prior to this project, a domain-specific ontological model was meticulously developed for the PV engineering domain, serving as the foundation for structuring the knowledge graph. This model defines a comprehensive set of entities, relationships, and attributes, capturing the multifaceted nature of PV projects [5]. Fig 1 illustrates the architecture of this system, depicting the integration of advanced data extraction techniques to analyze technical documents related to PV projects. Structured and unstructured data sources are parsed to identify relevant entities and relationships, facilitating the automated population of the knowledge graph. This automation minimizes manual effort and reduces the risk of errors. The resulting knowledge graph acts as a dynamic repository of interconnected data and serves as the basis for visualization and reporting tools. Users can explore the graph via an intuitive interface, gaining insights, identifying issues, and uncovering optimization opportunities. Moreover, the system utilizes the structured nature of the knowledge graph to automatically generate professional engineering reports, synthesizing complex data into tailored analyses for stakeholders.

In essence, the system architecture capitalizes on the strengths of Neo4j in handling complex data relationships, supported by a meticulously crafted ontological model specific to PV engineering. By strategically organizing data into a comprehensive knowledge graph, the system offers unique capabilities in data visualization and automated reporting, thereby advancing PV project management practices and setting new standards for knowledge graph applications in engineering domains.

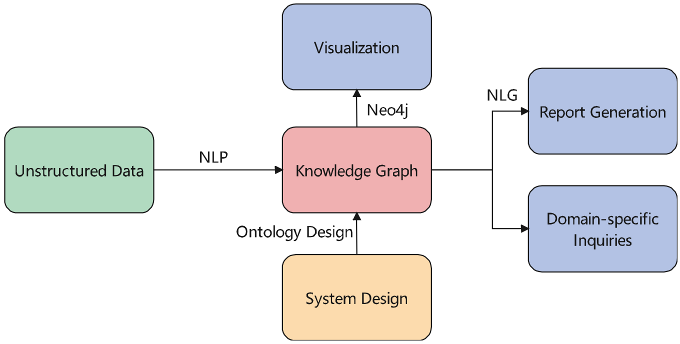

Fig. 1. System overview

3.2 Automated Knowledge Graph Construction

The overall technical framework of the knowledge graph section of this project is shown in Fig 2. For the data source selection, constructing a comprehensive knowledge graph for PV systems begins with the meticulous selection of data sources. Given the complexity and diversity of PV systems, data is gathered from various sources, including technical manuals, system performance reports, and real-time sensor data. These sources encompass structured data from databases, unstructured data from documents and reports, and semi- structured data from logs and system alerts. The aim is to cover all the knowledge pertinent to PV systems, from installation and maintenance to performance evaluation and troubleshooting. The data source will be organized according to the pre-designed ontology architecture. Ontology is the foundational schema of knowledge graph that understands and categorizes the PV system data. The ontology design process involves defining a set of concepts and categories in the PV domain, such as components (busbar, power station), operations (installation, maintenance), and performance metrics (efficiency, output). These categories are interconnected through relationships, enabling the knowledge graph to model complex interactions within the system. This structure facilitates data integration and retrieval and supports reasoning about the state of this system and potential issues. The process of organizing data through ontology involves data integration and data cleaning. Data integration involves harmonizing information from diverse sources into a unified model defined by ontology. This process includes transforming data into a standardized format, often RDF, and linking related entities based on common properties or identifiers. Data cleaning is integral to this stage, where inconsistencies, duplicates, and errors are identified and corrected. Techniques such as entity resolution and deduplication are employed to ensure the accuracy and reliability of the knowledge graph.

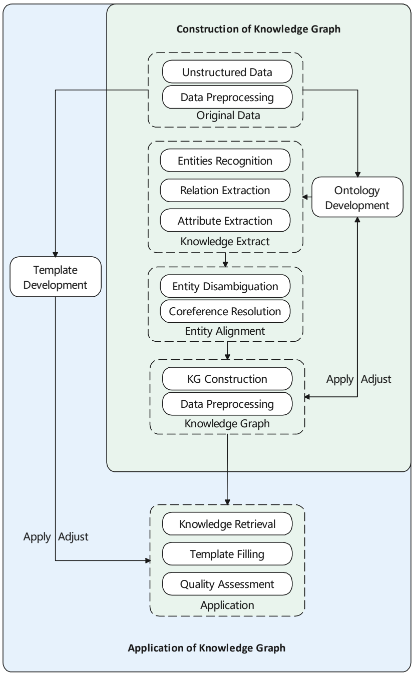

Fig. 2. Knowledge Graph Construction

The knowledge graph is designed to meet the complex data needs of PV systems by providing a holistic and interconnected view of all relevant information. This enables:

Efficient Querying. Complex queries about the components of this system, performance, and interrelations can be answered quickly and accurately.

Advanced Analytics. The graph structure facilitates the analysis of relationships and patterns within the data,aiding in predictive maintenance, optimization, and fault detection.

Enhanced Decision Making. The knowledge graph sup-ports informed decision-making regarding system design,operation, and maintenance by integrating and making sense of vast data.

In summary, constructing a knowledge graph for PV system project management involves a systematic approach to data source selection, ontology design, data integration, and

cleaning. This foundation enables the knowledge graph toeffectively meet the complex data needs of PV systems, driving efficiency, reliability, and performance improvements.

3.3 User Query and Report Generation

A pivotal feature of the system is its capability to automatically generate detailed engineering reports by leveraginga sophisticated template-based approach. This functionalityharnesses the structured data organized within the knowledge graph, ensuring that relevant information is dynamically inserted into predefined report templates. This process employs advanced natural language generation (NLG) techniques [19] to contextualize and articulate the data in a coherent and professionally formatted document.

The system design incorporates an intuitive front-end user interaction interface to visualize the complex relationships of this PV system as represented in the graph database. This interface, built upon user-centred design principles, enables users to navigate the intricate web of data relationships througha graphical representation, mirroring the structure of the underlying Neo4j graph database.

These system functionalities, the automatic engineering report generation and the front-end user interaction interface, stand as core components of the proposed knowledge graph- based system for PV system project management. They exemplify how the system innovatively approaches knowledge management, visualization, and reporting, embodying potential for significant enhancements in efficiency, accuracy, and user engagement in the management of PV projects.

4 Implementation

4.1 Tech Stack

In implementing the proposed system, selecting the technical stack is pivotal to achieving the data collection of this system, processing, visualization, and storage objectives. Core technologies such as HTML, CSS, JavaScript, and React were deployed for front-end development, while Neo4j was utilized as the primary database technology for the back-end. This combination facilitated the creation of an interactive and user- friendly interface, ensuring robust knowledge management and analysis capabilities. Fig 3 illustrates how the front-end and back-end are structured.

The front-end of this system is developed utilizing HTML, CSS, and JavaScript, with a significant focus on React. React, a widely-used JavaScript library for building user interfaces enabled the dynamic and responsive design of the web ap- plication. The decision to employ React was influenced byits component-based architecture, which permits the modular development of UI components. These components can be easily reused and managed, leading to a maintainable code- base and a streamlined development process. Through React,a template engine rendering mechanism was implemented,enabling the dynamic presentation of data. Such a mechanism proves especially effective in intuitively rendering complex data structures, thereby enhancing user interaction and data interpretability.

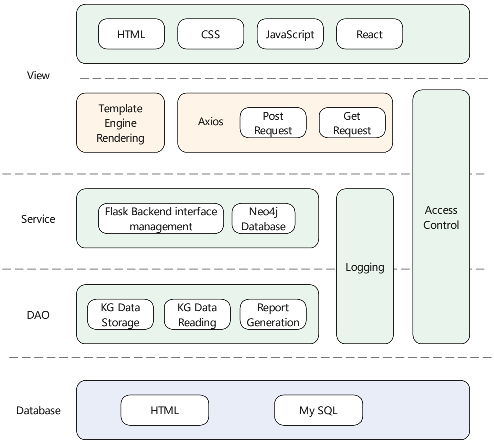

Fig. 3. System Structure

Moreover, integrating advanced JavaScript features withthe state management capabilities of React ensured seamless communication between the front-end and back-end. This is crucial for handling service requests, fetching data from the server, and updating the UI in real-time. The asynchronous nature of JavaScript, coupled with the efficient rendering lifecycle of React, provided a smooth and interactive userexperience, even when managing large datasets.

Neo4j was selected for back-end and knowledge management. The graph-based model of Neo4j aligns perfectly with the requirements of this project for analyzing and visualizing intricate networks within PV station project management. Thechoice of Neo4j was made due to its intuitive representation and querying of relationships, a fundamental need for constructing the knowledge graph.

Accompanying Neo4j, OWL (Web Ontology Language) andRDF Schema (RDFS) were employed to define and structure the ontology of project. These technologies enabled the creation of a flexible and extensible ontology system, accurately representing the concepts and relationships within the domain.OWL and RDFS provided a semantic framework that supports the integration, interpretation, and analysis of heterogeneous and unstructured data.

Integrating Neo4j with the project-specific ontology facilitated efficient reading, cleaning, organizing, storing, and summarizing data, transforming scattered and unstructured information into a structured and coherent knowledge graph. Furthermore, the capability of this system to generate final engineering documents and present data within the graph database was significantly enhanced by the querying capabilities of Neo4j and the semantic richness afforded by OWL and RDFS.

In summary, the implementation of the system relied ona meticulously chosen technical stack comprising HTML, CSS, JavaScript, and React for front-end development, complemented by Neo4j for back-end knowledge management. Using OWL and RDFS further enabled the system to effectively manage complex data relationships and ontologies. This technological foundation addressed PV station project management challenges by offering a comprehensive solution for automated information collection, efficient data integration,and intuitive data visualization.

4.2 Knowledge Graph and User Interface Implementation

The integration of the knowledge graph with the aforementioned technical stack is a pivotal aspect of the design of this system, as shown in Fig 4.

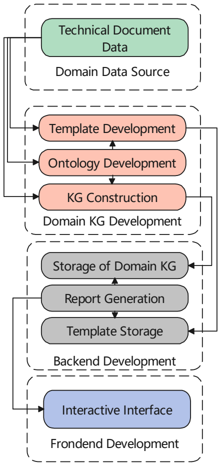

Fig. 4. Knowledge Graph Implementation

The detailed knowledge graph integration implementation with the proposed tech stack is shown in Fig 4. The transformation of unstructured data into a structured format conducive to graph database storage involves the application of NLP embedding techniques. These techniques collect word vectors with similar semantic meanings, aligning with the structure of this ontology. By employing embedding methods, words and phrases from technical documents are analyzed and con- verted into a format that mirrors the conceptual and relational schema defined by the ontology. This process ensures that the data retains its original context and is enriched with semantic meaning, making it suitable for incorporation into the knowledge graph. Once processed, these data are inserted into the graph database and linked according to the relationships and entities defined within the ontology of this system. This methodology allows for efficient data organization, enabling complex relationships between different pieces of information to be represented and queried effectively.

The knowledge graph is a fundamental system element,facilitating various functionalities that enhance utility and efficiency. One of the primary applications of the knowledge graph is in the generation of professional engineering documents. Utilizing templates developed with natural language generation (NLG) techniques [20], the system can produce professional but also clear and concise records. This capabilitystems from the structured and enriched data provided by the knowledge graph, which allows for the automatic generation ofcontent that accurately reflects the domain-specific knowledgeencapsulated within the graph.

Furthermore, the data of knowledge graph is accessible to the front end, enabling its visualization. The presentation of theknowledge graph on the user interface allows users to interact with and explore the complex relationships and entities itcontains. This visualization not only aids in understanding the intricate connections between different pieces of information but also is a powerful tool for data analysis and decision- making. Meanwhile, the system will automatically generate engineering project reports based on the collected information,and an illustrative example is shown in Fig 5.

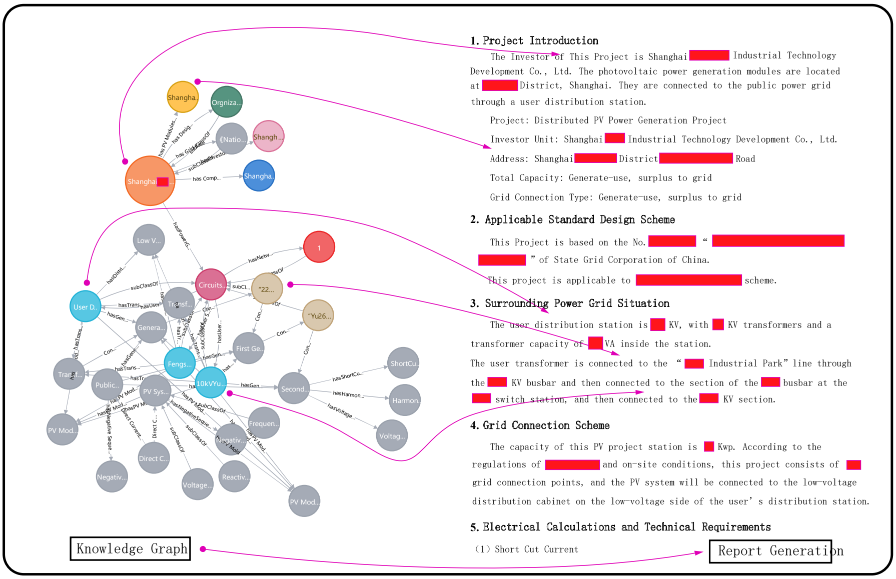

Fig. 5. Knowledge Graph Driven Report Generation Example

5 Evaluation

In the presented system, a knowledge graph (graph database) is innovatively utilized as the backend to automate the fusion of multi-source heterogeneous data, leveraging ontologies and other technologies. This approach facilitates com- plex data query functionalities, showcasing seamless linkage from project information to engineering circuit details, thus aiding decision-making processes. The novelty lies in the pro- vision of sophisticated data relationship querying capabilities. Traditionally, querying complex relationships across different tables (ranging from project information to circuit technical data) poses significant challenges, with tabular methods falling short in efficiently managing the intricate associations among disparate data sources. By integrating this information within a unified space, the system enables comprehensive queries, such as calculating the total photovoltaic (PV) capacity of projects invested by a specific investor shown in Algorithm 1. The code aggregates PV component capacities across circuits to identify the project with the maximum total PV capacity.

$Algorithm$ 1: $Neo4j\ Cypher\ Code\ Example$
$Input$:
$MATCH\ (Project: Photovaoltaic) - [*] \rightarrow (pvModule: PV\ module)$
$WITH\ project, SUM\left(toFloat(replace(pvModule.capacity, \text{kWp},))\right) AS\ totalCapacity$
$RETURN\ project, totalCapacity\ AS\ Total\ capacity$
$ORDER\ BY\ totalCapacity\ DESC$
$LIMIT$ 1

This method addresses the challenge of fusing multi-source heterogeneous data, significantly simplifying complex queriesthat were previously cumbersome, if not impossible, to con- duct using conventional tabular approaches. When executed, this code not only retrieves data about the entity but also allowsfor the exploration of interconnected information, thereby illuminating the context of the entity within the broader knowledge network. This capability is exemplified by a hypothetical interactive visualization, where a user can click on an entity within a generated image to access all related information. Such a feature demonstrates the dynamic and interconnected nature of the knowledge graph, facilitating an intuitive and in-depth exploration of the data, shown in Fig 6.

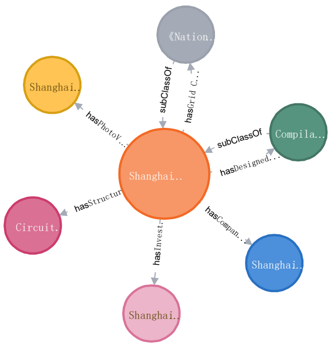

Fig. 6. Result of Query

The comparative analysis highlights that the developed system surpasses similar projects in functionality, particularly excelling in real-time data processing, leveraging Natural Language Processing (NLP) for the extraction of unstructured data, and the automatic generation of professional engineering reports, shown in Table 1. These capabilities not only ensure the up-to-dateness and comprehensiveness of the knowledge graph but also significantly enhance the efficiency and ac- curacy of information synthesis and reporting in PV power projects. Consequently, this positions the system as a superior solution within its domain, offering distinct advantages in knowledge management and project documentation.

Table 1. COMPARISON OF SYSTEM FUNCTIONS

System Functions	Systems		
	This System	Paper 1 [1]	Paper 2 [2]
KG Query	√	√	√
Visualization	√	√	√
Report Generation	√	×	×
Real Time Data	√	×	√
Data Extraction	NLP	Manual	Manual

Additionally, the evaluation phase of the project encompassed a comprehensive satisfaction survey involving 572 respondents from the stakeholders of this developed system, including operational managers, power engineers, maintenance technicians and procurement manager. This survey aimed to assess the adaptability and satisfaction level of the developed system across these varied industries. The findings, graphically represented in Table 2, underscore the versatile applicability of this system and high satisfaction rates among professionals from these fields. The results validate not only the broad utility of the system but also highlight its potential to revolutionize knowledge management and project documentation practices across a wide spectrum of industries.

Table 2. SATISFACTION SURVEY

Different Stakeholders	Occupations			
	Operations Manager	Power Engineer	Maintenance Technicians	Procurement Manager
Satisfaction	87%	94%	90%	88%

In conclusion, it is a distinguished performance of the developed system. Unlike its counterparts, this system not only automates the process of extracting valuable insights from both structured and unstructured data sources through sophisticated algorithms but also seamlessly translates these insights into comprehensive, industry-standard engineering re- ports. This dual functionality marks a significant leap towards enhancing the efficiency and accuracy of project management in the domain of PV power projects. Furthermore, akin to other projects in this field, the developed system boasts a robust visualization of the knowledge graph, allowing users to intuitively navigate through complex datasets. The query functionality of the system extends beyond mere data retrieval, enabling exploration of intricate relationships and interconnected data points within the graph. This facilitates not only a deeper understanding of the project scope but also aids in identifying potential opportunities and challenges early in the project lifecycle. Overall, the exceptional capabilities of the system in automated data extraction, professional report generation, and comprehensive knowledge graph visualization and querying set a new benchmark for technological solutions in the realm of PV power project management.

6 Conclusion

This paper introduced a systematic solution designed to tackle the complex challenges of managing heterogeneous and multi-source data within PV power station projects. By harnessing the power of automated knowledge graphs for data organization and analysis, the system efficiently processes unstructured data to generate professional engineering reports, offering a significant leap in knowledge management. The application of advanced ontological techniques for converting unstructured data into a structured RDF format underpins a robust solution that improves data integrity and accessibility while dramatically reducing the manual labor associated with data handling. The system is capable of maintaining intricate data relationships within a graph database, showcasing a critical enhancement over traditional tabular data storage methods in project management. This advancement facilitates seamless data visualization and querying, enabling effortless navigation through the knowledge graph to pinpoint specific details and understand their interconnections. Moreover, the automated creation of detailed and precise project engineering reports from structured data and knowledge graph insights significantly enriches communication and documentation standards in PV project management. Such capabilities prove satisfactory for a wide range of stakeholders needing quick, comprehensive project insights to guide decision-making processes.

Looking ahead, there are several avenues for further expansion of this system. Integration with emerging technologies such as deep generative models could offer predictive analytics capabilities, forecasting potential issues and optimizing system performance proactively. Expanding the adaptability of this system to accommodate data from a broader spectrum of renewable energy sources would also enhance its utility, making it a more versatile tool for integrated energy management. Lastly, exploring interoperability with existing industry standards and platforms could streamline data exchange and collaboration across different segments of the energy sector.

References

1. J. Zhou and X. Yang, "The construction method of knowledge graphon photovoltaic industry chain," in *Knowledge and Systems Sciences*, J. Chen, V.-N. Huynh, X. Tang, and J. Wu, Eds. Singapore: Springer Nature Singapore, 2023, pp. 274–284.
2. H. Huang, Z. Hong, H. Zhou, J. Wu, and N. Jin, "Knowledge graph construction and application of power grid equipment," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–10, 2020.
3. J. Wu, Q. Li, Q. Chen, G. Peng, J. Wang, Q. Fu, and B. Yang, "Evaluation, analysis and diagnosis for hvdc transmission system faults via knowledge graph under new energy systems construction: A critical review," *Energies*, vol. 15, no. 21, 2022. [Online]. Available: https://www.mdpi.com/1996 1073/15/21/8031
4. S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: Representation, acquisition, and applications," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 2, pp. 494–514, 2022.
5. L. Otero-Cerdeira, F. J. Rodríguez-Martínez, and A. Gómez-Rodríguez, "Ontology matching: A literature review," *Expert Systems with Applica-tions*, vol. 42, no. 2, pp. 949–971, 2015.
6. M. T. Özsu, "A survey of rdf data management systems," *Frontiers of Computer Science*, vol. 10, pp. 418–432, 2016.
7. Y. Tang, T. Liu, M. He, Q. Wang, H. Zhang, G. Liu, and R. Dai, "Graphdatabase based knowledge graph storage and query for power equipmentmanagement," in *2020 12th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 2020, pp. 1–5.

8. Y. Duan, L. Shao, G. Hu, Z. Zhou, Q. Zou, and Z. Lin, “Specifying architecture of knowledge graph with data graph, information graph, knowledge graph and wisdom graph,” in *2017 IEEE 15th International Conference on Software Engineering Research, Management and Appli-cations (SERA)*, 2017, pp. 327–332.
9. X. Chen, S. Jia, and Y. Xiang, “A review: Knowledge reasoning over knowledge graph,” *Expert Systems with Applications*, vol. 141, p. 112948, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417419306669
10. C. Xiong, R. Power, and J. Callan, “Explicit semantic ranking for academic search via knowledge graph embedding,” in *Proceedings of the 26th International Conference on World Wide Web*, ser. WWW’17. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2017, p. 1271–1279. [Online]. Available: https://doi.org/10.1145/3038912.3052558
11. S. Ji, S. Pan, E. Cambria, P. Marttinen, and S. Y. Philip, “A survey on knowledge graphs: Representation, acquisition, and applications,” *IEEE transactions on neural networks and learning systems*, vol. 33, no. 2,pp. 494–514, 2021.
12. F.-L. Li, H. Chen, G. Xu, T. Qiu, F. Ji, J. Zhang, and H. Chen, “Alimekg: Domain knowledge graph construction and application in e- commerce,” in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 2020, pp. 2581–2588.
13. M. C. Silva, P. Eugénio, D. Faria, and C. Pesquita, “Ontologies and knowledge graphs in oncology research,” *Cancers*, vol. 14, no. 8, 2022. [Online]. Available: https://www.mdpi.com/2072-6694/14/8/1906
14. H. L. Nguyen, D. T. Vu, and J. J. Jung, “Knowledge graph fusion for smart systems: A survey,” *Information Fusion*, vol. 61, pp. 56–70, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253519307729
15. M. A. Wajid and A. Zafar, “Multimodal fusion: a review, taxonomy, open challenges, research roadmap and future directions,” *Neutrosophic Sets and Systems*, vol. 45, no. 1, p. 8, 2021.
16. B. Yang and Y.-m. Liao, “Research on enterprise risk knowledge graph based on multi-source data fusion,” *Neural Computing and Applications*,vol. 34, no. 4, pp. 2569–2582, 2022.
17. M. Dipali, “Learning graph databases: Neo4j an overview.” *International Journal of Engineering Applied Sciences and Technology*, vol. 8, pp. 216–219, 2024. DOI: 10.7753/IJCATR1212.1007.
18. F. M. S. López and E. G. S. De La Cruz, “Literature review about neo4j graph database as a feasible alternative for replacing rdbms,” *Industrial Data*, vol. 18, no. 2, pp. 135–139, 2015.
19. A. B. Sai, A. K. Mohankumar, and M. M. Khapra, “A survey of evaluation metrics used for nlg systems,” *ACM Computing Surveys(CSUR)*, vol. 55, no. 2, pp. 1–39, 2022.
20. C. C. Osuji, T. C. Ferreira, and B. Davis, “A systematic review of data- to-text nlg,” 2024.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.