



University of Camerino
SCHOOL OF SCIENCES AND TECHNOLOGY
Master of Science in
Computer Science (LM-18)

**University of Applied Sciences and
Arts Northwestern Switzerland**
SCHOOL OF BUSINESS
Master of Science in
Business Information Systems

HAI-BEMS: A Hybrid Artificial Intelligent Approach for Building Energy Management Systems

Candidate:
Piermichele Rosati

Matricula 124176

FHNW Supervisor:
Dr. Emanuele Laurenzi

Unicam Supervisor:
Prof. Michela Quadrini

Empa Supervisor:
Dr. James Allan

A thesis presented to the School of Sciences and Technology of the University of Camerino and School of Business of the University of Applied Sciences and Arts Northwestern Switzerland in partial fulfillment of the requirements for the degree of Master of Science in Computer Science

A.Y. 2023/2024

Abstract

The Thesis Abstract is written here (and usually kept to just this page)...

Statement of Authenticity

I, Piermichele ROSATI, hereby confirm that this report was performed autonomously using only the sources, aids and assistance stated in the report, and that quotes are readily identifiable as such.



Signed:

Date: 11th June 2024

Contents

Abstract	i
Statement of Authenticity	ii
Table of Contents	iii
1 Introduction	1
1.1 Background	1
1.2 Problem Statement	1
1.3 Thesis Statement	1
1.4 Research Questions	1
2 Literature Review	2
2.1 Knowledge Graphs	2
2.2 Graph Neural Networks	7
2.3 Building Energy Management Systems	7
3 Conclusion	8
Bibliography	8
Glossary	10
Abbreviations	10
List of Figures	10
List of Tables	12
Acknowledgements	13

Chapter 1

Introduction

1.1 Background

1.2 Problem Statement

1.3 Thesis Statement

1.4 Research Questions

Scarselli et al., 2009

Chapter 2

Literature Review

2.1 Knowledge Graphs

TODO: change intro

TODO: reread every section and change (modifying, adding something)

Knowledge Graphs (KGs) have come up as a key technology in the field of data management and Artificial Intelligence (AI), enabling sophisticated data integration, retrieval and analysis, and analysis. This literature review provides an in-depth examination of KGs, their theoretical foundations, practical applications and recent advances.

Theoretical Foundations of Knowledge Graphs

Definition and Structure

KGs are directed graph-based data structures that represent real-world entities and their interrelations, providing a way to model complex domains and their underlying semantics. A KG consists of nodes (entities) and edges (relationships), forming a network of interconnected information. This structure allows KGs to capture rich contextual information and provide a semantic framework for data (Hogan et al., 2021).

A KG refers to a semantic network graph which is consisted of diverse entities, concepts, and relationships in the real world. It is used to formally describe various things and their associations in the real world. KGs are generally represented in triples $KG = \{E, R, F\}$.

- E represents the entity set $\{e_1, e_2, \dots, e_E\}$, and the entity e is the most basic element in the KG, referring to the items that exist objectively and can be distinguished from each other.
- R represents the relation set $\{r_1, r_2, \dots, r_R\}$, and the relation r is an edge in the KG, representing a specific connection between different entities.
- F represents the fact set $\{f_1, f_2, \dots, f_F\}$, and each f is defined as a triple $(h, r, t) \in f$, in which h denotes the head entity, r stands for the relationship, and t indicates the tail entity.

Ontologies and Semantic Web Technologies

An ontology is a formal representation of knowledge in a domain, specifying the concepts, relationships, and constraints that exist within that domain. The term “ontology” can be used to the shared understanding of some domain of interest (Uschold and Gruninger, 1996). Ontologies play a critical role in defining the schema and semantics of KGs. They specify the types of entities, relationships, and constraints, thereby providing a formalized structure for the data. The Semantic Web technologies, particularly the Resource Description Framework (RDF) and the Web Ontology Language (OWL), are fundamental to the development and functioning of KGs (G. (Grigoris) Antoniou and Frank. Van Harmelen, 2008).

- RDF is a standard model for data interchange on the web. It uses triples (subject-predicate-object) to represent information, providing a flexible and extensible framework for creating and managing KGs (W3C CITATION).
- OWL is used to explicitly represent the meaning of terms in vocabularies and the relationships between those terms. It enables more complex and expressive representations compared to RDF Schema (RDFS) (Deborah and van Harmelen Frank, 2004).

Query Languages

SPARQL Protocol and RDF Query Language (SPARQL) is the standard query language for retrieving and manipulating data stored in RDF format. It allows users to write complex queries to extract specific information from a KG, making it a powerful tool for data analysis and knowledge discovery (Pérez et al., 2009).

Cypher is another query language for graph databases, such as Neo4j, that allows users to interact with graph data using a pattern-matching syntax. Cypher queries are used

to traverse the graph, retrieve specific patterns, and perform operations on the data (Francis et al., 2018).

Applications of Knowledge Graphs

General Applications

KGs have been adopted across various domains due to their ability to integrate heterogeneous data sources, provide semantic context, and enable advanced querying and reasoning.

- In healthcare, KGs are used to integrate patient records, clinical trials, research data, and medical ontologies, enabling personalized medicine and decision support systems. They help in identifying relationships between diseases, treatments, and patient outcomes (Kapanipathi et al., 2020).
- Financial institutions leverage KGs to connect data from various sources, such as market data, regulatory information, and customer transactions. This integration facilitates risk management, fraud detection, and compliance monitoring (Tchechmedjiev et al., 2019).
- In e-commerce, KGs enhance product recommendation systems by linking customer preferences, purchase history, and product information. They enable more personalized and relevant recommendations, improving customer satisfaction and sales (Zhang et al., 2021).

Enterprise Knowledge Management

Within enterprises, KGs are used to manage and utilize internal knowledge effectively. They integrate data from different departments, such as human resources, finance, and operations, providing a unified view of the organization's information. This integration supports decision-making, collaboration, and innovation (Pujara et al., 2013).

Search and Information Retrieval

KGs significantly enhance search engines by providing semantic search capabilities. They enable the understanding of user queries in context, allowing for more accurate and relevant search results. Google's Knowledge Graph is a prominent example, enhancing search results with information about entities and their relationships (Singhal et al., 2012).

Recent Advancements in Knowledge Graphs

Integration with Machine Learning

Recent research has focused on integrating KGs with Machine Learning (ML) and Deep Learning (DL) techniques to enhance their capabilities and applications. These integrations have led to significant advancements in various areas, including Natural Language Processing (NLP), recommendation systems, and predictive analytics.

- **Knowledge Graph Embeddings (KGEs):** KGE techniques represent entities and relationships in a continuous vector space, enabling the use of machine learning algorithms for tasks such as link prediction, entity classification, and clustering. Popular methods include TransE, TransH, and TransR, each providing different ways to model relationships in the embedding space (Wang et al., 2017).
- **Graph Neural Networks (GNNs):** GNNs are DL models designed to operate on graph-structured data. They leverage the relational nature of graphs to perform tasks such as node classification, link prediction, and graph classification. GNNs have been successfully applied to enhance the capabilities of KGs in various domains (Wu et al., 2021).

GNNs are described in Sec. 2.2.

Natural Language Processing and Question Answering

KGs have been instrumental in advancing NLP applications, particularly in question answering systems. By providing structured and semantically rich information, KGs enable systems to understand and generate human language more effectively.

- **Question Answering Systems:** KGs support question answering systems by enabling them to retrieve and reason over structured data. These systems can answer complex queries by traversing the graph and applying logical inferences based on the relationships between entities (Yasunaga et al., 2021).
- **Semantic Search and Text Analysis:** KGs enhance text analysis and semantic search by providing contextual information about entities mentioned in the text. This contextual understanding improves the accuracy of information retrieval and the relevance of search results (Fernández et al., 2011).

Challenges and Future Directions

Scalability and Performance

As KGs grow in size and complexity, scalability and performance become critical challenges. Efficient storage, querying, and updating of large KGs require advanced techniques and architectures. Research in distributed computing, graph databases, and parallel processing is ongoing to address these issues (CITATION).

Data Quality and Integration

Ensuring the accuracy and consistency of data in KGs is essential for their reliability. Data quality issues, such as inconsistencies, duplications, and inaccuracies, can significantly impact the performance of applications relying on KGs. Developing methods for automatic data cleaning, validation, and integration is an active area of research (Paulheim, 2017).

Privacy and Security

The integration of sensitive data into KGs raises concerns about privacy and security. Protecting personal and confidential information while allowing for meaningful data analysis is a significant challenge. Research is focusing on developing techniques for secure data sharing, access control, and anonymization within KGs (Bonatti et al., 2017).

Interoperability and Standardization

Interoperability and standardization are crucial for the widespread adoption of KGs. Ensuring that different KGs can work together seamlessly and that their data can be easily integrated requires the development of common standards and protocols. Efforts such as the Linked Open Data (LOD) initiative and W3C standards aim to address these challenges (Bizer et al., 2023).

Conclusion

Knowledge Graphs represent a transformative technology for data integration, retrieval, and analysis, offering significant benefits across various domains. Their ability to provide

semantic context and capture complex relationships makes them invaluable for applications in healthcare, finance, e-commerce, and enterprise knowledge management. Recent advancements in machine learning, particularly in the integration with GNNs, have further enhanced the capabilities of KGs, opening new avenues for research and application. However, challenges related to scalability, data quality, privacy, and interoperability remain and must be addressed to fully realize the potential of KGs. Continued research and development in these areas will be crucial for the future evolution and adoption of KGs.

2.2 Graph Neural Networks

2.3 Building Energy Management Systems

Chapter 3

Conclusion

Bibliography

- Bizer, C., Heath, T., & Berners-Lee, T. (2023). Linked data - the story so far. In *Linking the world's information: Essays on tim berners-lee's invention of the world wide web* (1st ed., pp. 115–143). Association for Computing Machinery. <https://doi.org/10.1145/3591366.3591378>
- Bonatti, P., Kirrane, S., Polleres, A., & Wenning, R. (2017). Transparent personal data processing: The road ahead. In S. Tonetta, E. Schoitsch & F. Bitsch (Eds.), *Computer safety, reliability, and security* (pp. 337–349). Springer International Publishing.
- Deborah, L. M., & van Harmelen Frank. (2004). Owl web ontology language overview. <http://www.w3.org/TR/2003/PR-owl-features-20031215/>
- Fernández, M., Cantador, I., López, V., Vallet, D., Castells, P., & Motta, E. (2011). Semantically enhanced information retrieval: An ontology-based approach. *Journal of Web Semantics*, 9, 434–452. <https://doi.org/10.1016/j.websem.2010.11.003>
- Francis, N., Green, A., Guagliardo, P., Libkin, L., Lindaaker, T., Marsault, V., Plantikow, S., Rydberg, M., Selmer, P., & Taylor, A. (2018). Cypher: An evolving query language for property graphs. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 1433–1445. <https://doi.org/10.1145/3183713.3190657>
- G. (Grigoris) Antoniou and Frank. Van Harmelen. (2008). *A semantic web primer*. MIT Press.
- Hogan, A., Blomqvist, E., Cochez, M., D’Amato, C., Melo, G. D., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., Ngomo, A. C. N., Polleres, A., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys*, 54. <https://doi.org/10.1145/3447772>
- Kapanipathi, P., Abdelaziz, I., Ravishankar, S., Roukos, S., Gray, A., Astudillo, R., Chang, M., Cornelio, C., Dana, S., Fokoue, A., Garg, D., Gliozzo, A., Gurajada, S., Karanam, H., Khan, N., Khandelwal, D., Lee, Y.-S., Li, Y., Luus, F., ...

- Yu, M. (2020). Leveraging abstract meaning representation for knowledge base question answering. <http://arxiv.org/abs/2012.01707>
- Paulheim, H. (2017). Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web*, 8, 489–508. <https://doi.org/10.3233/SW-160218>
- Pérez, J., Arenas, M., & Gutierrez, C. (2009). Semantics and complexity of sparql. *ACM Transactions on Database Systems*, 34. <https://doi.org/10.1145/1567274.1567278>
- Pujara, J., Miao, H., Getoor, L., & Cohen, W. (2013). Knowledge graph identification. *The Semantic Web–ISWC 2013: 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21–25, 2013, Proceedings, Part I 12*, 542–557.
- Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009). The graph neural network model. *IEEE Transactions on Neural Networks*, 20, 61–80. <https://doi.org/10.1109/TNN.2008.2005605>
- Singhal, A., et al. (2012). Introducing the knowledge graph: Things, not strings. *Official google blog*, 5(16), 3.
- Tchechmedjiev, A., Fafalios, P., Boland, K., Gasquet, M., Zloch, M., Zapilko, B., Dietze, S., Todorov, K., & Zloch, M. (2019). Claimskg: A knowledge graph of fact-checked claims, 309–324. https://doi.org/10.1007/978-3-030-30796-7_20
- Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications.
- Wang, Q., Mao, Z., Wang, B., & Guo, L. (2017). Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering*, 29, 2724–2743. <https://doi.org/10.1109/TKDE.2017.2754499>
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32, 4–24. <https://doi.org/10.1109/TNNLS.2020.2978386>
- Yasunaga, M., Ren, H., Bosselut, A., Liang, P., & Leskovec, J. (2021). Qa-gnn: Reasoning with language models and knowledge graphs for question answering. <http://arxiv.org/abs/2104.06378>
- Zhang, W., Deng, S., Chen, M., Wang, L., Chen, Q., Xiong, F., Liu, X., & Chen, H. (2021). Knowledge graph embedding in e-commerce applications: Attentive reasoning, explanations, and transferable rules. *ACM International Conference Proceeding Series*, 71–79. <https://doi.org/10.1145/3502223.3502232>

List of Figures

List of Tables

The acknowledgements and the people to thank go here, don't forget to include your project advisor...