eBISS 2024

Knowledge Graph Exploration

Rana İşlek | rana.islek@estudiantat.upc.edu
Berat Furkan Koçak | berat.furkan.kocak@estudiantat.upc.edu

Knowledge Graphs (KGs) are structured representations of information enabling intelligent applications such as data analytics and recommender systems. This poster presents a comprehensive review of ***state-of-the-art methods in KG exploration****, focusing on enhancing accessibility and effectiveness for both expert and non-expert users. The main research question addressed is:* ***How can we improve the accessibility and effectiveness of Knowledge Graph exploration for both expert and non-expert users?***

Introduction

The use of **semantic web technologies** has resulted in the creation of knowledge graphs depicted as triples, enabling exploration across various domains. These graphs facilitate semantic search, the integration of data, and thorough data analysis. Although **semantic links between entities** can help users explore content, they can also result in **a large number of overwhelming choices** that may confuse or frustrate users, especially those unfamiliar with the subject.

We examine **challenges** they identify in the literature, including **complexity**, **heterogeneity**, and **scalability** of KGs, and discuss their **novel proposals** designed to overcome these obstacles. We identified the **most important six different methodology** addressing our research question.

B. Pattern Mining and Connectivity Search (Wu et al., 2023)

Pattern mining and connectivity search **identify relevant structures** within knowledge graphs, handling under-specified queries. **KPRLN**, a deep knowledge preference-aware reinforcement learning network, uses **user-entity relations** to encode preferences as **weights on the KG**. This novel approach achieves state-of-the-art **accuracy** in **recommendations**, outperforming various other algorithms.

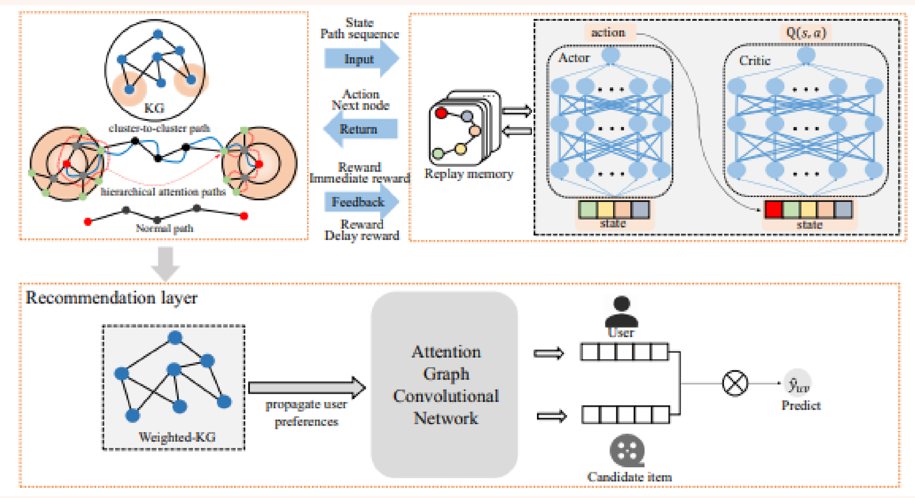


Figure 3. Illustration of the KPRLN framework.

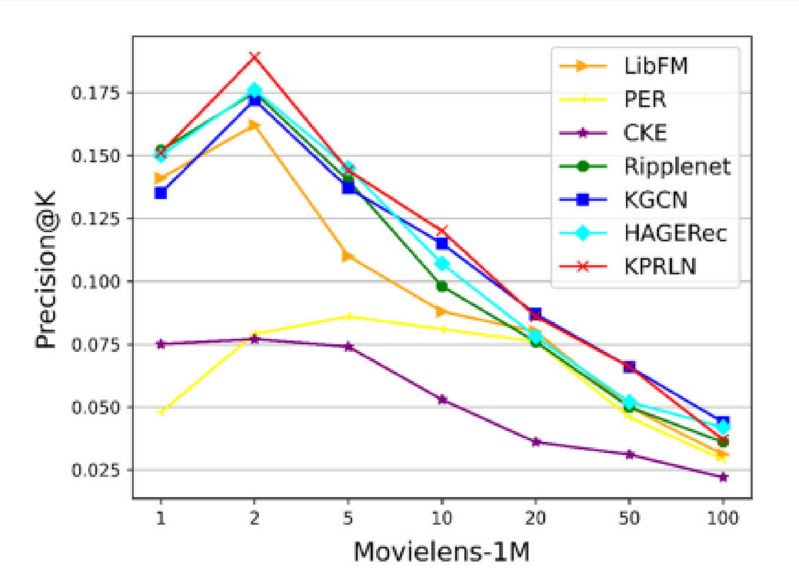


Figure 4. A search for Elizabeth II's residence, comparing the Google search engine and paper's proposal.

A. Approximate Similarity Search (Pappas et al., 2017)

Approximate similarity search allows users to find nodes or edges similar to a given example using **domain-specific measures**. State-of-the-art methods apply graph theory importance measures **to identify key nodes** in a knowledge graph. The **challenge of connecting these nodes** is addressed by modeling it as a graph **Steiner-tree problem**, using approximations (**SDIST, CHINS, HEUM**) to optimize connections via shortest paths. **Betweenness** adaptation is found to be the **most effective**, with Steiner-Tree algorithms introducing fewer additional nodes to the summary graph.

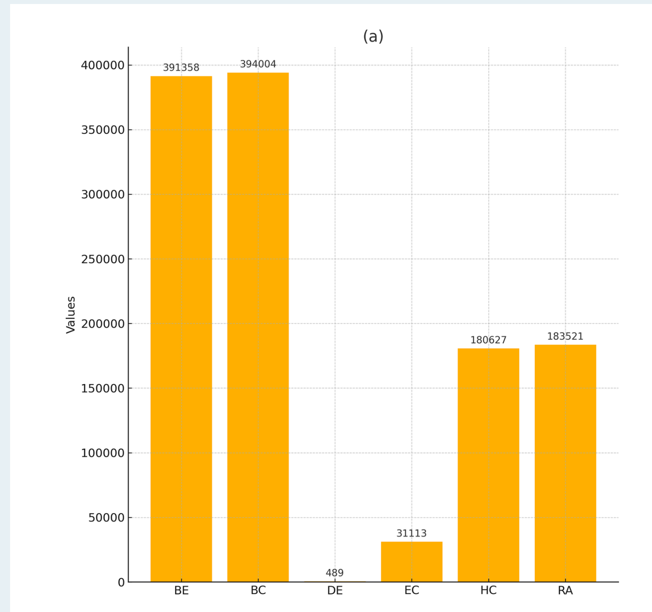


Figure 1: The average execution times of the importance measures (msec)

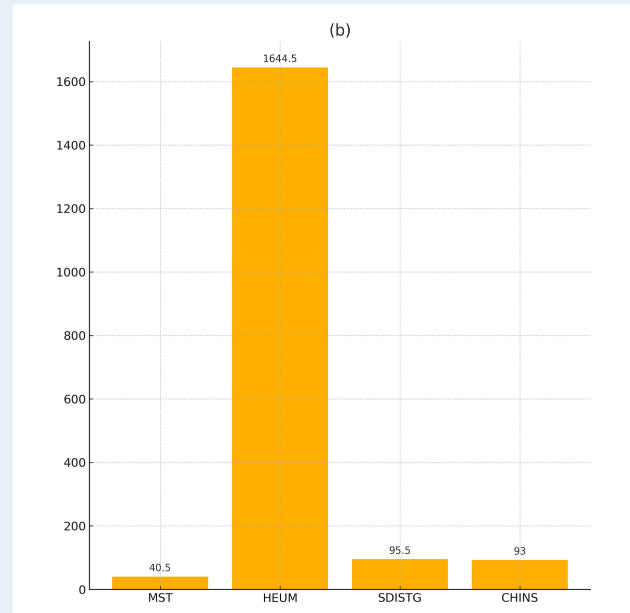
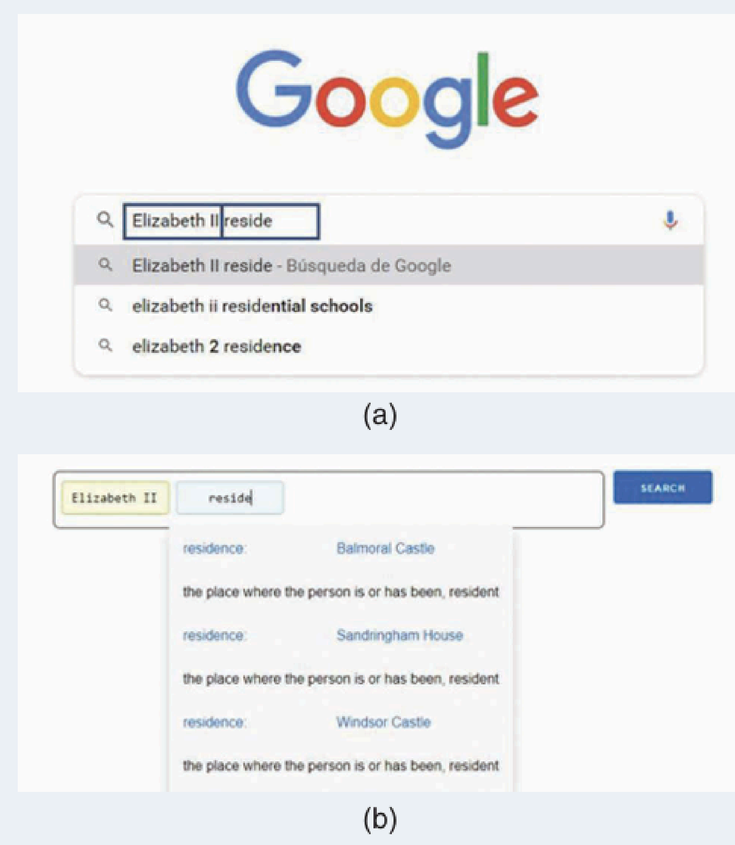


Figure 2: The algorithms for linking those nodes (msec)

C. Theme-Based and Keyword Search (Yagüe et al., 2024)

Theme-based and keyword search approaches improve knowledge graph exploration through **intuitive methods** like **theme identification** and keyword searches. For example, a proposed **search box system** for **Wikidata** allows **non-expert users** to perform exploratory searches without navigating through multiple pages. This system simplifies user interaction by providing dynamic, **real-time suggestions** based on the text input, significantly reducing the time and complexity of searches. It leverages the SPARQL query language via the **Wikidata Query Service (WDQS)** to fetch relevant entities and relationships, **significantly reducing search time and complexity.**



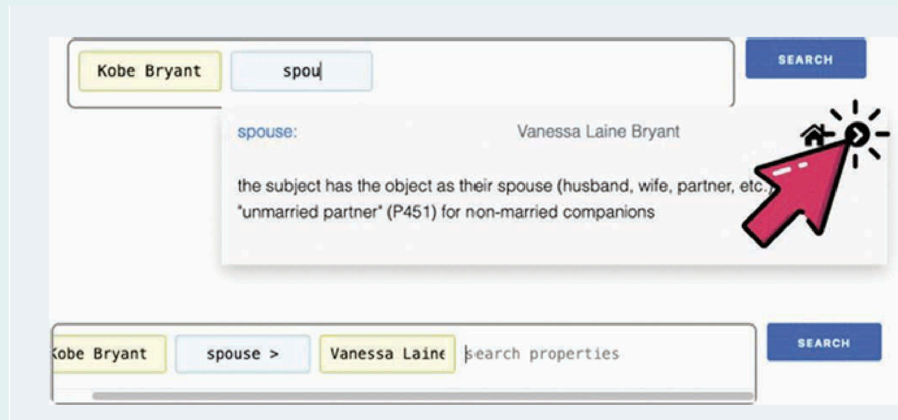


Figure 6. The search by predicate and dynamic filtering based on a previous entity.

Figure 5. A search for Elizabeth II's residence, comparing the Google search engine and paper's proposal.

D. Knowledge Graph Embeddings & Interactive Visualization Tools (Yuan et al., 2021)

Knowledge graph embeddings represent entities and relations **as vectors in a continuous space**, capturing semantic information to guide exploration. The **KGScope** tool supports interactive visual exploration with **embedding-based guidance**. It features the Embedding View (**EBView**), which displays the semantic context of entities during exploration.

In Figure 7:

- **1st Group:** Contains entities primarily linked to musical roles and specific awards like the Grammy Lifetime Achievement Award.
- **2nd Group:** Contains entities related to TV, including Emmy Awards and Golden Globe Awards winners.
- **3rd Group:** Contains entities associated with politicians and non-entertainment roles

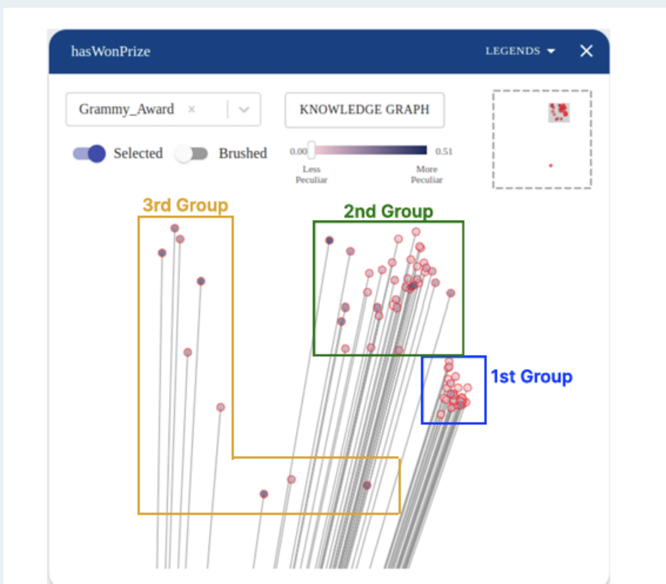


Figure 7. En example of the distributions of Grammy Award awardees in the EBView.

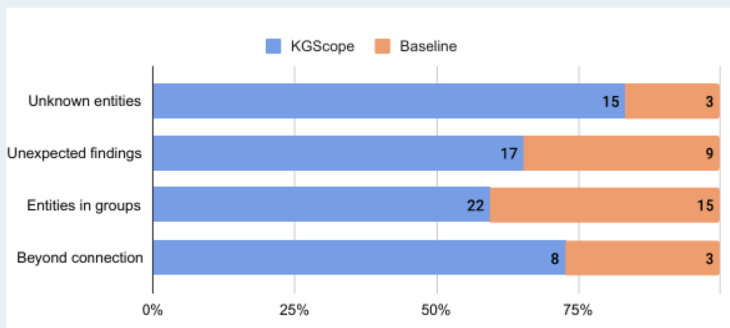


Figure 8. Comparison of KGScope and Baseline: The number of insights found by participants related to the four categories.

E. Query Engine and Interaction Flow (Lissandrini et al., 2020)

Re2xOLAP enhances user interaction by simplifying queries and providing **real-time search** capabilities without needing domain expertise. This system creates a virtual graph from a triplestore, mapping queries to an **in-memory graph**. Users can refine results by **disaggregating**, **subsetting**, or **finding similar items**. The ReOLAP procedure's time complexity is **independent of** observation numbers, thanks to virtual **graph-based matching**. Re2xOLAP is the first solution offering example-driven exploratory analysis on statistical KGs, **eliminating the need for SPARQL queries.**

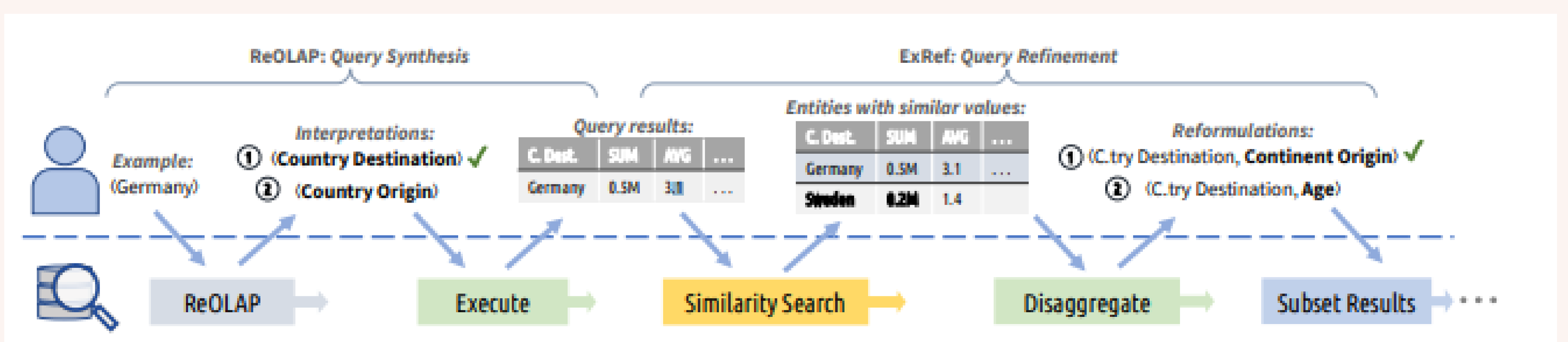


Figure 9. Re2xOLAP: example query synthesis and refinement steps, an interaction step is mapped to a pair of arrows

F. Progressive and Incremental Results (Jung et al., 2020)

Progressive and incremental results provide initial outcomes **quickly** and refine them **progressively**. **AttnIO**, a novel approach in this field, uses a **Graph Neural Network** (GNN) with attention mechanisms to navigate knowledge paths within a KG. It **dynamically** updates attention distribution based on **dialog context**, ensuring resilience to sub-optimal paths and enhancing **knowledge retrieval accuracy.**

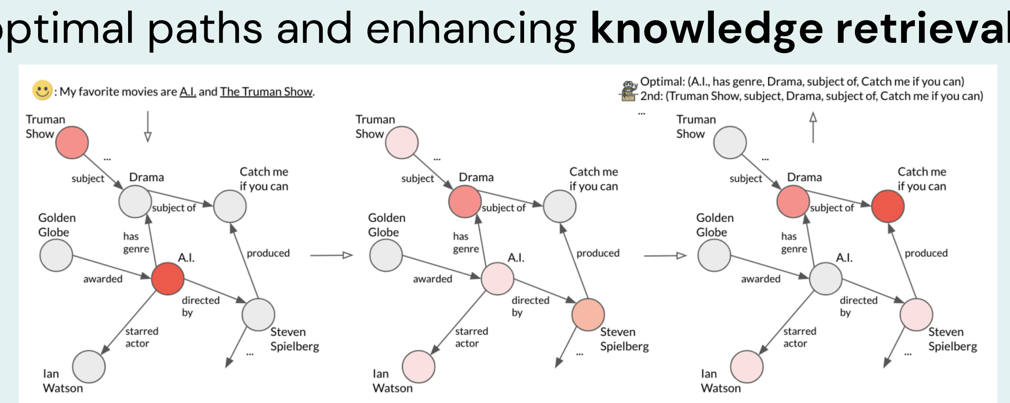


Figure 10: Path decoding process of AttnIO

G. Discussion & Conclusion

Efficiency and Time Complexity: *Re2xOLAP* and *Theme-Based and Keyword Search* provide significant efficiency for non-expert users by simplifying interactions and reducing query time (require minimal computational resources during runtime, although Re2xOLAP's initial bootstrapping can be time-intensive).
Accuracy and Adaptability: *AttnIO* and *KPRLN* demonstrate superior accuracy in their respective tasks (dialogue systems and recommendations) by dynamically learning and adapting to user preferences and dialog contexts (comes at the cost of higher computational complexity, particularly during model training and real-time updates).
Complexity Handling: Methods like *Approximate Similarity Search* and *Pattern Mining* address the complexity of KGs by using advanced graph theory and reinforcement learning techniques (ideal for high-precision tasks requiring detailed analysis but may be overkill for simpler queries and explorations).

References

1. Al-Tawil, M., Dimitrova, V., Thakker, D., & Abu-Salih, B. (2023). Emerging Exploration Strategies of Knowledge Graphs. *IEEE Access, 11*, 94713-94731.
2. Chen, Y., Min, S., Chen, D., Wang, M., & Ren, X. (2020). Improving Commonsense Question Answering by Graph-based Iterative Retrieval over Multiple Knowledge Sources.
3. Jung, S., Jang, Y., Kim, J., Kim, K., Hwang, W., & Song, Y. (2020). AttnIO: Knowledge Graph Exploration with In-and-Out Attention Flow for Knowledge-Grounded Dialogue.
4. Lissandrini, M., Mottin, D., Hose, K., & Pedersen, T. B. (2022). Knowledge Graph Exploration Systems: Are We Lost? *CIDR 2022*.
5. Lissandrini, M., Hose, K., & Pedersen, T. B. Example-Driven Exploratory Analytics over Knowledge Graphs.
6. Pappas, A., Troullinou, G., Roussakis, G., Kondylakis, H., & Plexousakis, D. (2017). Exploring Importance Measures for Summarizing RDF/S KBs.
7. Troullinou, G., Kondylakis, H., Daskalaki, E., & Plexousakis, D. (2017). Ontology Understanding Without Tears: The Summarization Approach. *Semantic Web Journal, IOS Press*.
8. Wang, X., & Cheng, G. (2023). A Survey on Extractive Knowledge Graph Summarization: Applications, Approaches, Evaluation, and Future Directions.
9. Wu, D., Tang, M., Zhang, S., et al. (2023). KPRLN: Deep Knowledge Preference-Aware Reinforcement Learning Network for Recommendation. *Complex Intell. Syst., 9*, 6645-6659.
10. Yagüe París, J., Espada, J. P., Martínez, J. S., & Crespo, R. G. (2024). Search Box System to Improve User Interaction in Knowledge Graph Searches: A Solution for Users Without Expert Skills.
11. Yuan, H., Yu, T.-W., Pan, J.-Y., & Lin, W. C. (2021). KGScope: Interactive Visual Exploration of Knowledge Graphs with Embedding-Based Guidance.