

PP10: Introducing Graph Structure in Real-World Data Analyses

Kountouris Nikolaos (Nikolaos.Kountouris@pfizer.com)
Panagiotakis Georgios (Georgios.Panagiotakis@pfizer.com)
Puri Maria (Maria.Puri@pfizer.com)
Pfizer Inc.

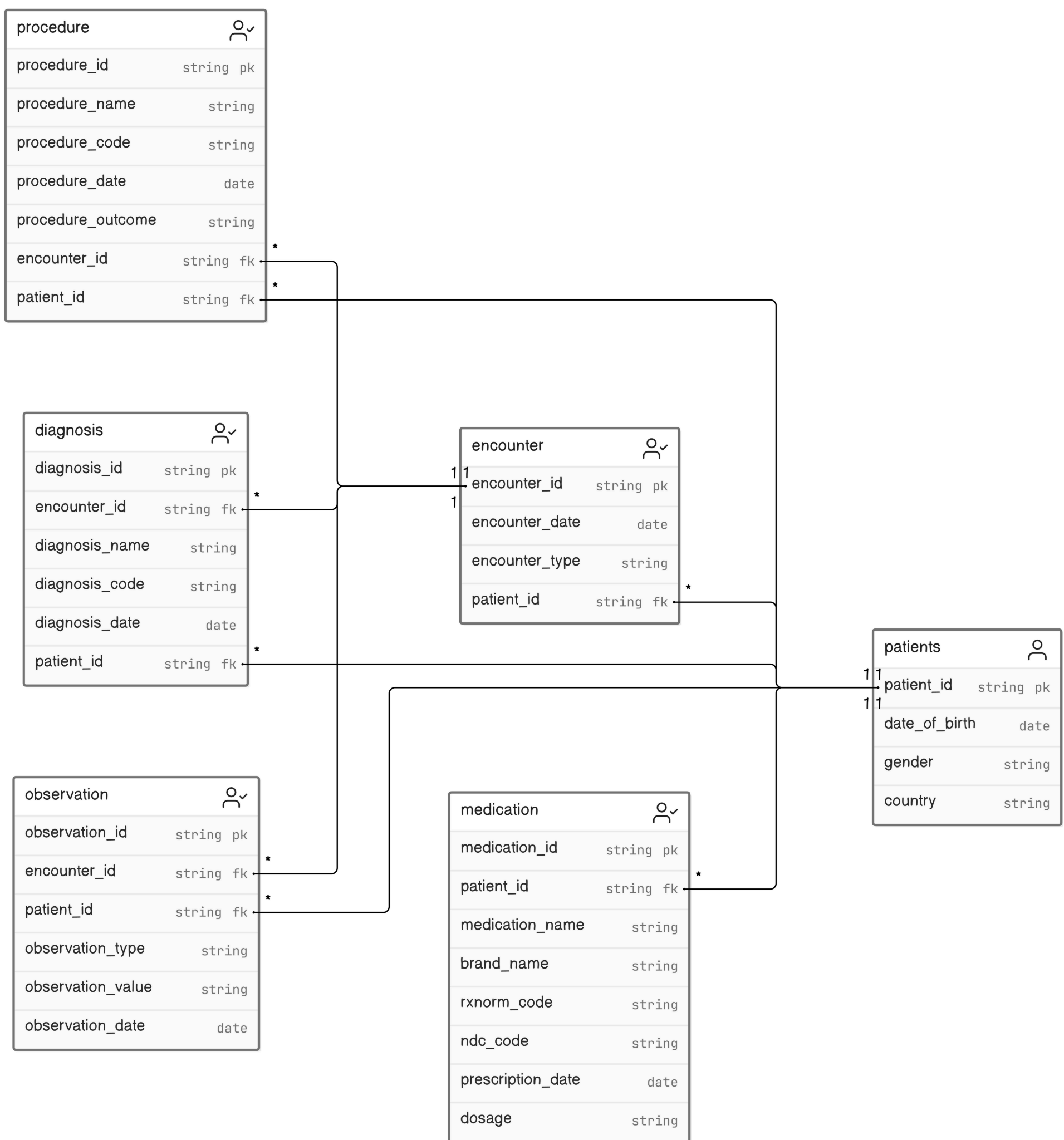
Introduction

Graphs serve as a robust means to depict and investigate the intricate relationships hidden in data. With nodes symbolizing entities and edges illustrating the links between them, graph theory and network analysis empower researchers to reveal obscured patterns, forecast results, and formulate potent strategies. Employing graph-based approaches (both databases and algorithms) in data analysis not only deepens our insights into complex systems but also bolsters decision-making capabilities, especially in fields like medical research where they are extremely beneficial.

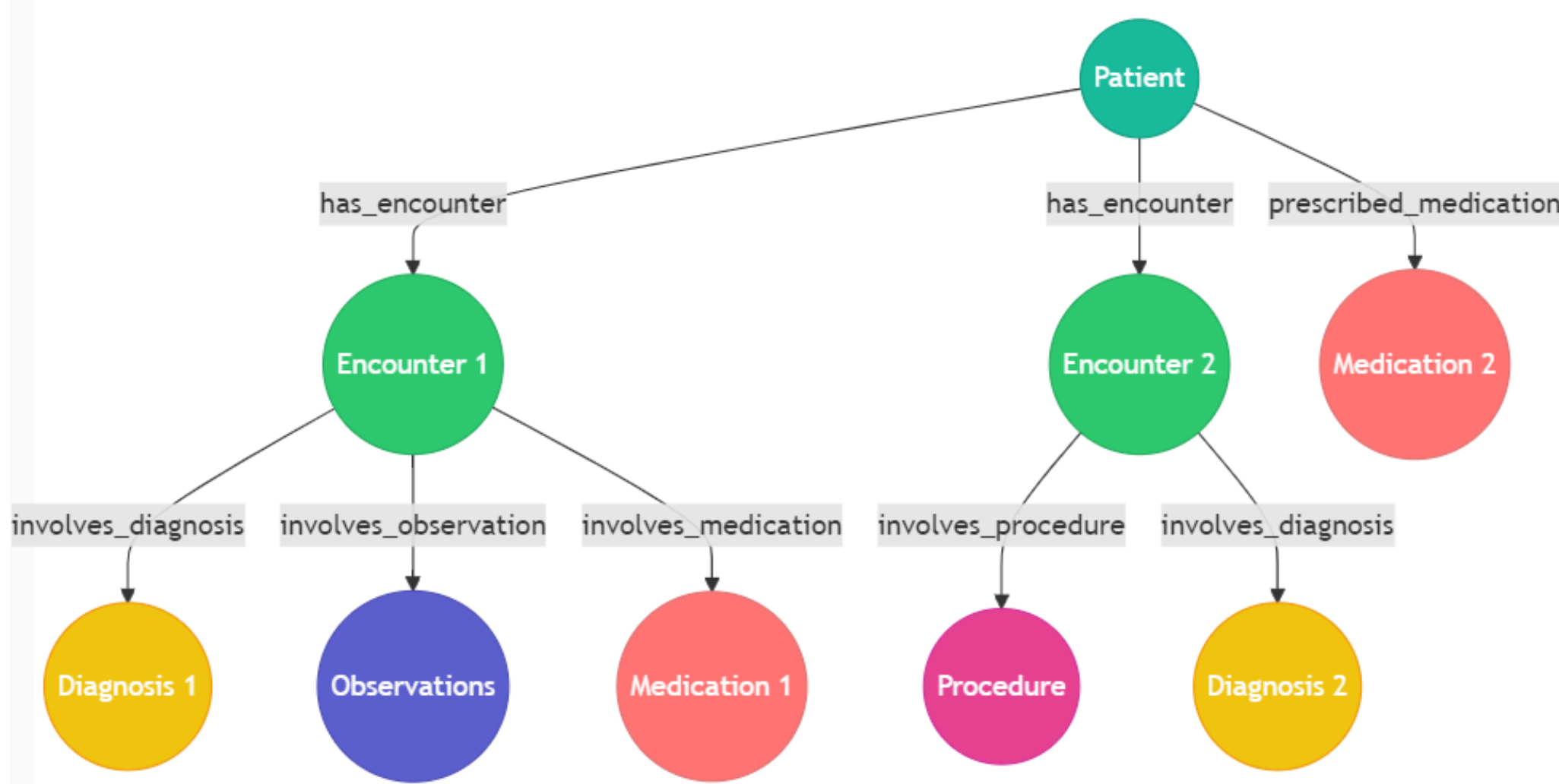
Database Comparison

Aspect	Graph DBs	Relational DBs
Efficient Relationship Handling	Efficient with complex, interconnected data.	Performance issues with deeply nested or many-to-many relationships.
Flexibility and Schema-less Nature	Schema-less, allowing dynamic and flexible data models.	Require predefined schemas, making adaptation to changing data structures challenging.
Scalability	Handles growing amounts of interconnected data efficiently, often better scalability for certain applications (e.g. Disease Outbreak and Transmission Tracking).	Can face performance and complexity challenges as data volume and relationships grow.
Standardization and Tooling	Less mature, with fewer standardized tools and frameworks, lacking a single standard query language.	Highly mature, with a wide range of tools and the well-established SQL standard ensuring consistency and compatibility.
Integration	More challenging to integrate with existing systems and workflows designed around relational databases.	Easy integration with a broad ecosystem of tools, applications, and legacy systems.

Visualization

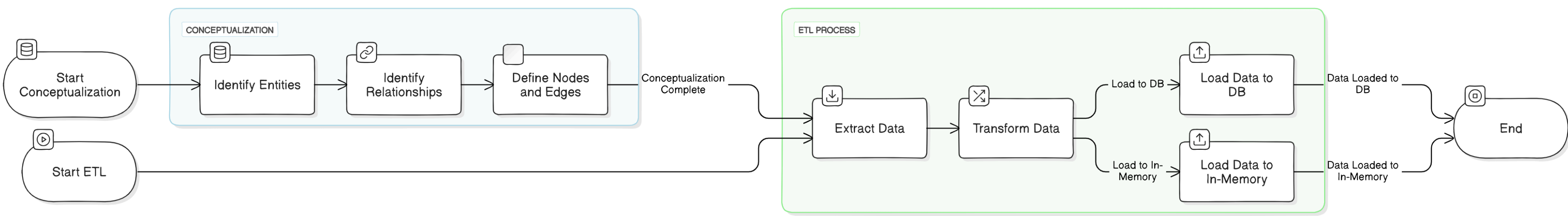


Relational Database Schema



Graph Database Structure (Patient)

Relational to Graphs



To use graph technology with relational data:

- 1) Conceptualize data as nodes and edges.
- 2) Extract and transform the data into a graph format using ETL tools or custom scripts.
- 3) Load the transformed data into a graph database or an in-memory library, then apply graph algorithms for analysis.

Use cases Examples

- Analyzing Medication Use and Interactions
- Analyzing Patient's Health Journey
- Detecting Adverse Reactions to Drugs
- Identifying Co-treatments and Comorbidities
- Tracking the Progression of Diseases

Conclusion

Graphs excel at handling complex interconnected relationships in medical information, providing valuable insights about patient and treatment outcomes. By integrating various data sources, graph-based approaches create a comprehensive view of patient health and can identify patterns in epidemiological data. As healthcare evolves towards data-driven decision making, these technologies are becoming crucial for maximizing the value of health data, leading to better patient care and more efficient healthcare systems.

[1] I. Robinson, J. Webber, and E. Eifrem, "Graph Databases," 2nd ed. Sebastopol, CA: O'Reilly Media, 2015. [Online]. Available: <https://www.oreilly.com/library/view/graph-databases-2nd/9781491930885/>

[2] K. Hänsel, S. N. Dudgeon, K. H. Cheung, T. J. S. Durant, and W. L. Schulz, "From Data to Wisdom: Biomedical Knowledge Graphs for Real-World Data Insights," J. Med. Syst., vol. 47, no. 1, p. 65, May 2023. doi: 10.1007/s10916-023-01951-2. PMID: 37195430; PMCID: PMC10191934.