PhD Candidate: Andrea Colombo

Advisor: Prof. Stefano Ceri

“Knowledge Management with Graphs in Finance and Legislation”

The transformative role of Knowledge Graphs in modern knowledge management systems has been unlocking novel opportunities and research directions both in the scientific and industrial worlds, especially in the wake of generative Artificial Intelligence tools. The flexibility of representing interconnected knowledge in an easily explorable graph, combined with rich semantics, is compelling for creating more effective pipelines to knowledge-intensive applications that manage and consume knowledge coming from multiple sources. Among these, automated reasoning is a powerful tool for deriving novel, relevant knowledge by applying inference rules, possibly in a recursive fashion, leveraging the graph's semantics and interconnected structure. There is no strict definition of Knowledge Graphs, although the scientific literature is converging on a paradigm for modeling highly interconnected knowledge, enhancing the role of relationships among data as first-class citizens.

This thesis contributes to advancing knowledge management with graphs from two perspectives; on one side, inspired by the resurgence of logic-based languages for expressing intuitive business rules traversing knowledge natively expressed through graphs, it demonstrates the power of such paradigms in expressing complex business rules and definitions, modeling complex financial domains that lacked a declarative implementation, thus resulting obscure for non-technical users. Following this line, it also demonstrates how declarative models can synergize with generative AI to fully unlock transparency in reasoning processes by providing natural language explanations of inferred facts.

While reasoning over graphs can be effectively executed through logic-based paradigms, the underlying data are stored within graph databases, whose most native data model is the Property Graph (PG) model. This data model has recently been empowered with an international standard for queries, the Graph Query Language (GQL), and proposals of common graph database tools, such as schemas, keys, and triggers are being developed. Following this line, this thesis contributes to these efforts by proposing an approach for injecting reasoning capabilities within graph databases, by replicating the behavior of logic-based inference with PG tools.

In the second part of the thesis, domain-specific graph applications are considered to demonstrate the transformative power of graphs and the adoption of the Property Graph data model as a paradigm to reshape knowledge management. In particular, the thesis focuses on the possibility of capturing the complexities of the legislative system within highly interconnected graphs, going beyond the unstructured nature of normative acts. Such a proposal enhances both knowledge management and discovery, especially when considering recently introduced AI-based applications and agents, for tasks as navigation of the intricacies of normative acts, allowing users to receive more informative and contextualized answers.

The results of this thesis are part of a broad vision that demonstrates how adopting a graph data model is beneficial in multiple directions: it allows an easy implementation of reasoning applications, enabling higher levels of transparency in automated decision processes, providing, at the same time, an effective tool to enhance knowledge discovery especially in the wake of novel AI tools, whose power can be unlocked by the adoption of graph solutions, as for the legal domain.