# Vincenzo Schiano Di Cola*

Tutor: Francesco Piccialli, Nicola Mazzocca
Company (DATABOOZ) Tutor: Paolo Benedusi

XXIX Cycle - II year presentation

Data Science: Knowledge Graph applications

*e-mail: vincenzo.schianodicola@unina.it

Nowadays, research organizations are getting interested in scenarios that require exploiting diverse collection of data, like IoT devices. The ultimate goal is to retrieve large amounts of data from various sources and provide added value to this information. *Machine Learning* (ML) can provide new perspectives to identify hidden patterns and classes within the Data Science process.
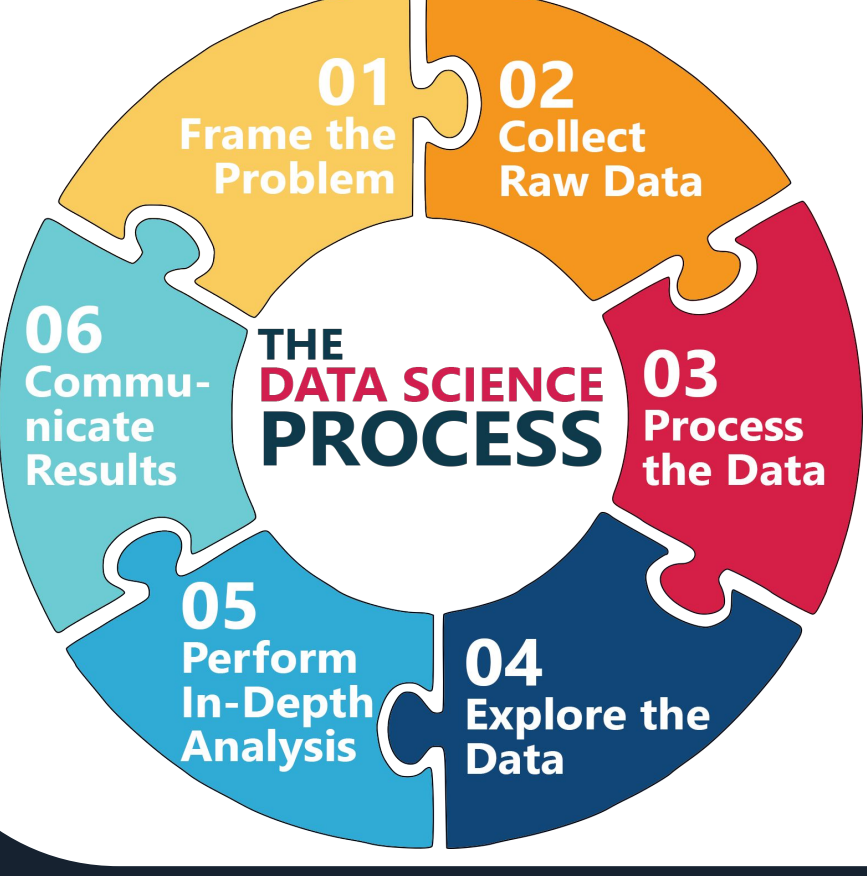
Moreover, *Knowledge Graph* (KG) technologies can improve services based on Machine Learning and data analytics. Applying KG to real-world problems represents a great challenge since it requires modeling, reasoning, and ultimately, advanced mathematical algorithms and computational techniques.

The research group is involved in multiple Data Science project. The second year research application was based on a project in collaboration with DATABOOZ that uses IoT data in a Cultural Heritage framework, with data from the MANN and Castel Nuovo collected through BBBs (Beaglebone Black IoT boards) and mobile application on tablets


MANN

Castel Nuovo
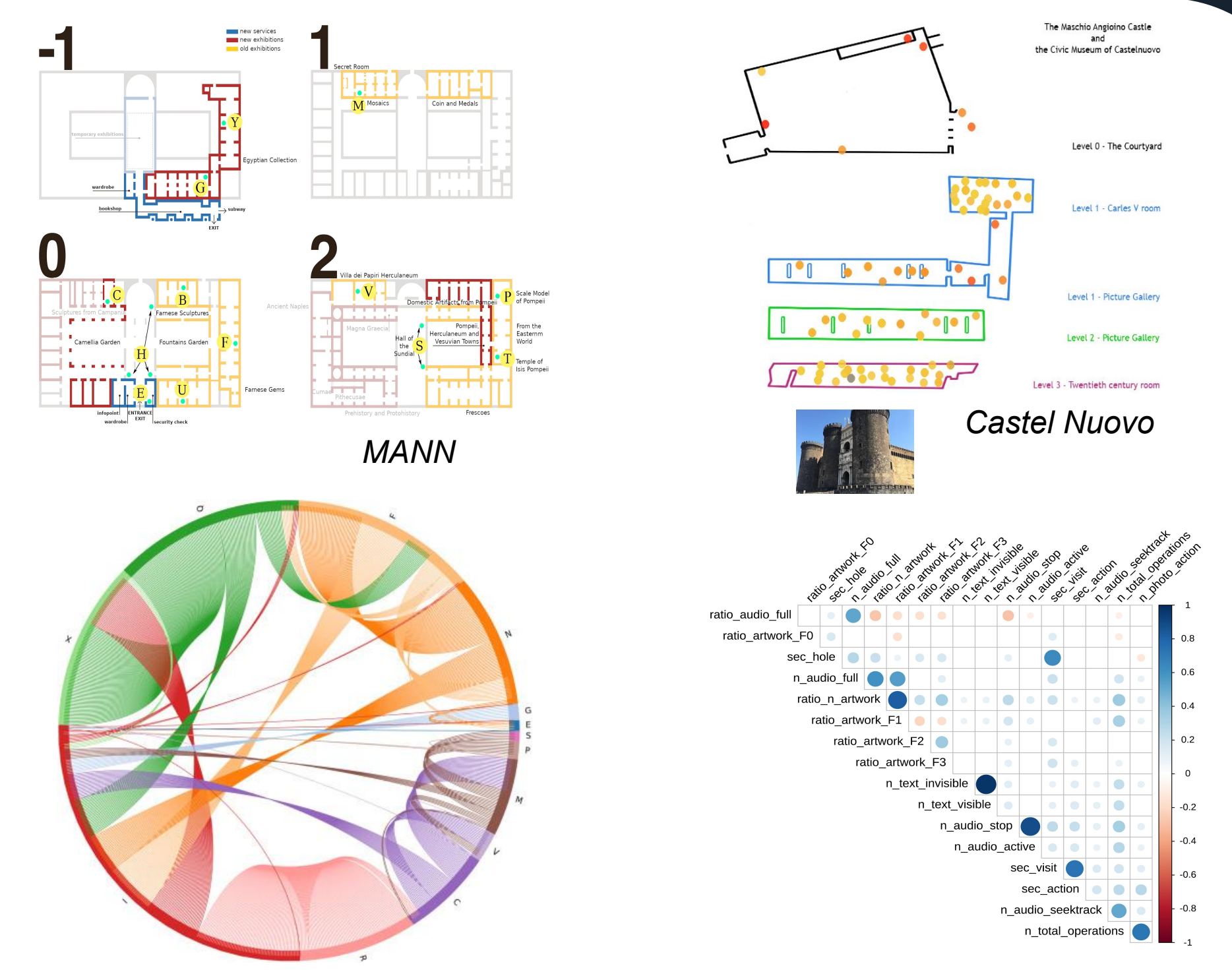
Organizing data through a KG can represent a suitable way to discover information quickly, easily, and directly. The first step is to model the schema in a fourth normal form so as to query a NoSQL database and present the first layer of results.
A future step is to apply knowledge graph embedding which arranges the entities, from the KG, within a vector space, and then apply ML techniques.
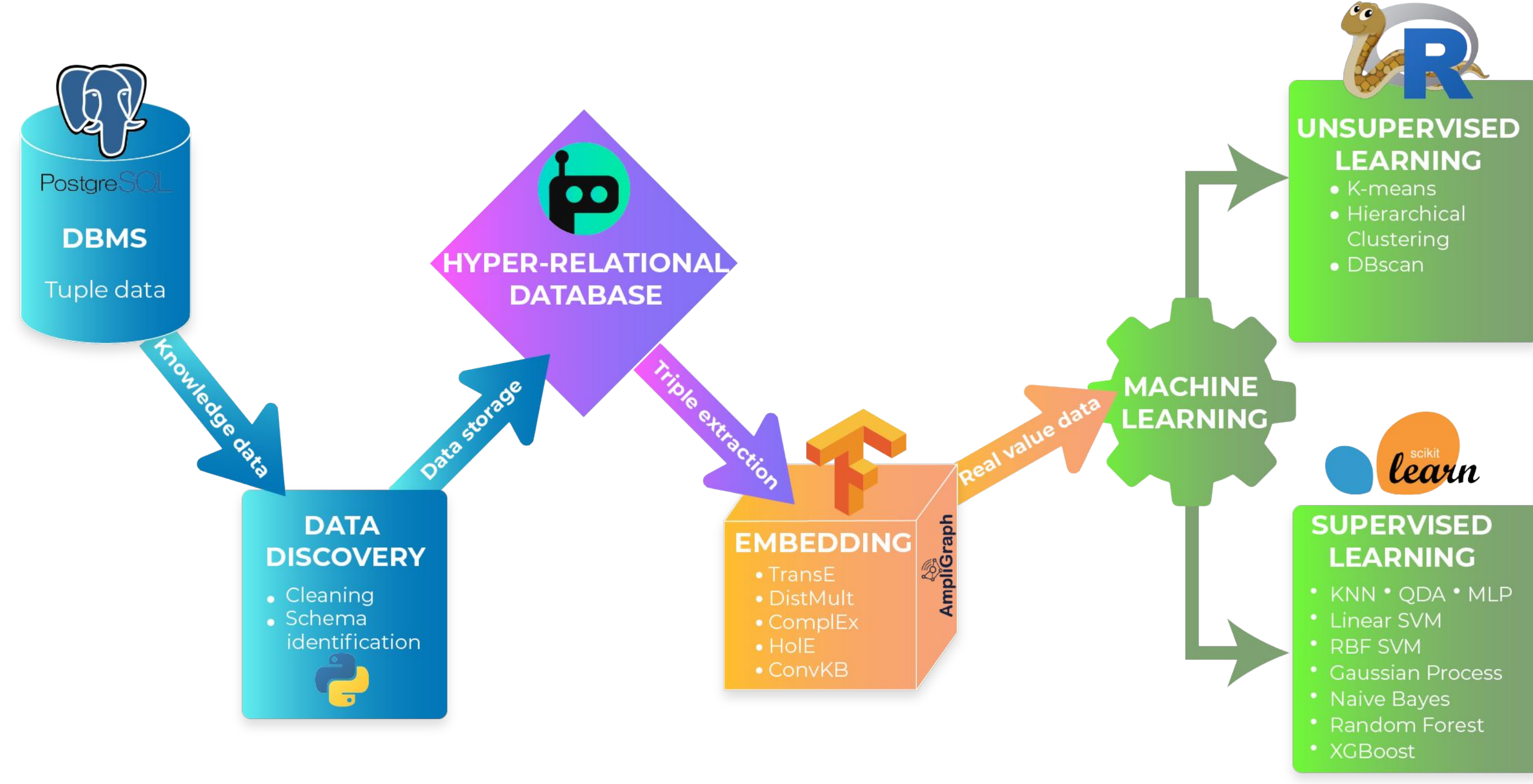
The final outcome is to identify aggregate behaviours.
From a technical point of view, this framework relies on *GRAKN*, a novel and intelligent graph database able to model complex datasets, and *Amplighrph*, a graph embedding system for multiple embeddings.

A KG graph approach is presented to support museum stakeholders services through the extraction of knowledge and new insights.
The KG represents a unified model for both data of MANN and Castel Nuovo.
The model included parameters of distances among Point Of Interests of a museum, calculated trough R, and statistical informations about users behaviours were presented with PowerBi.
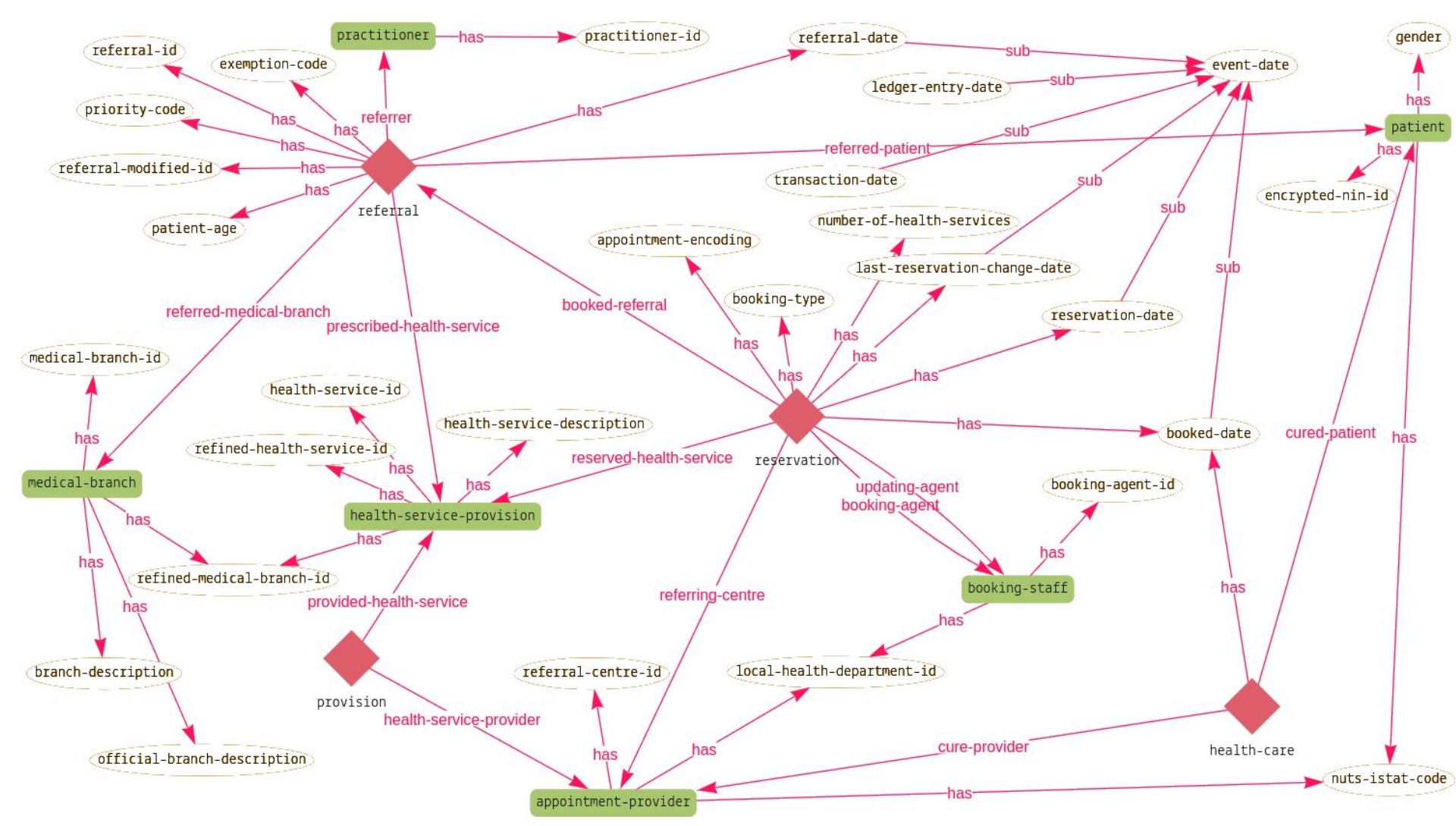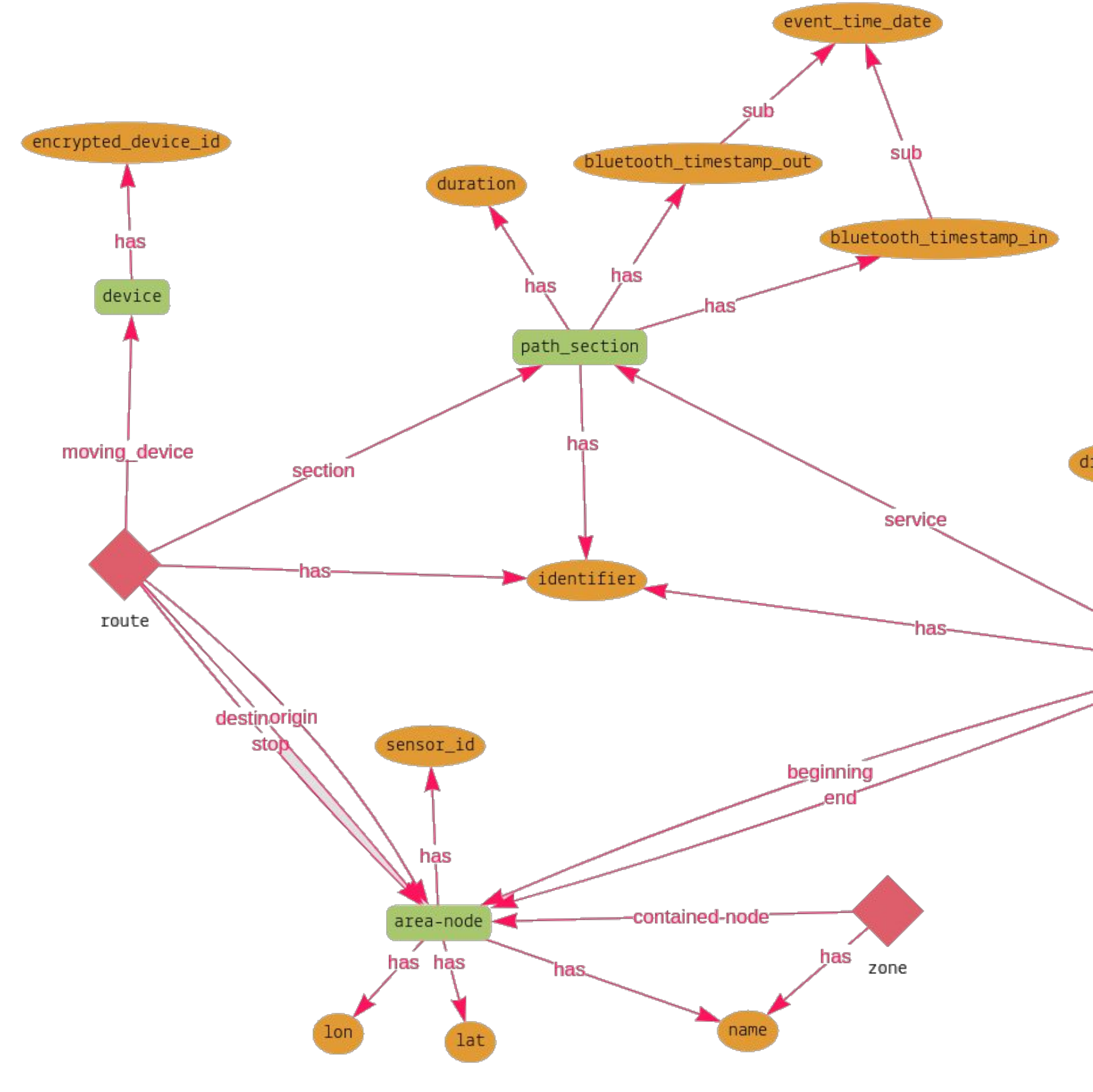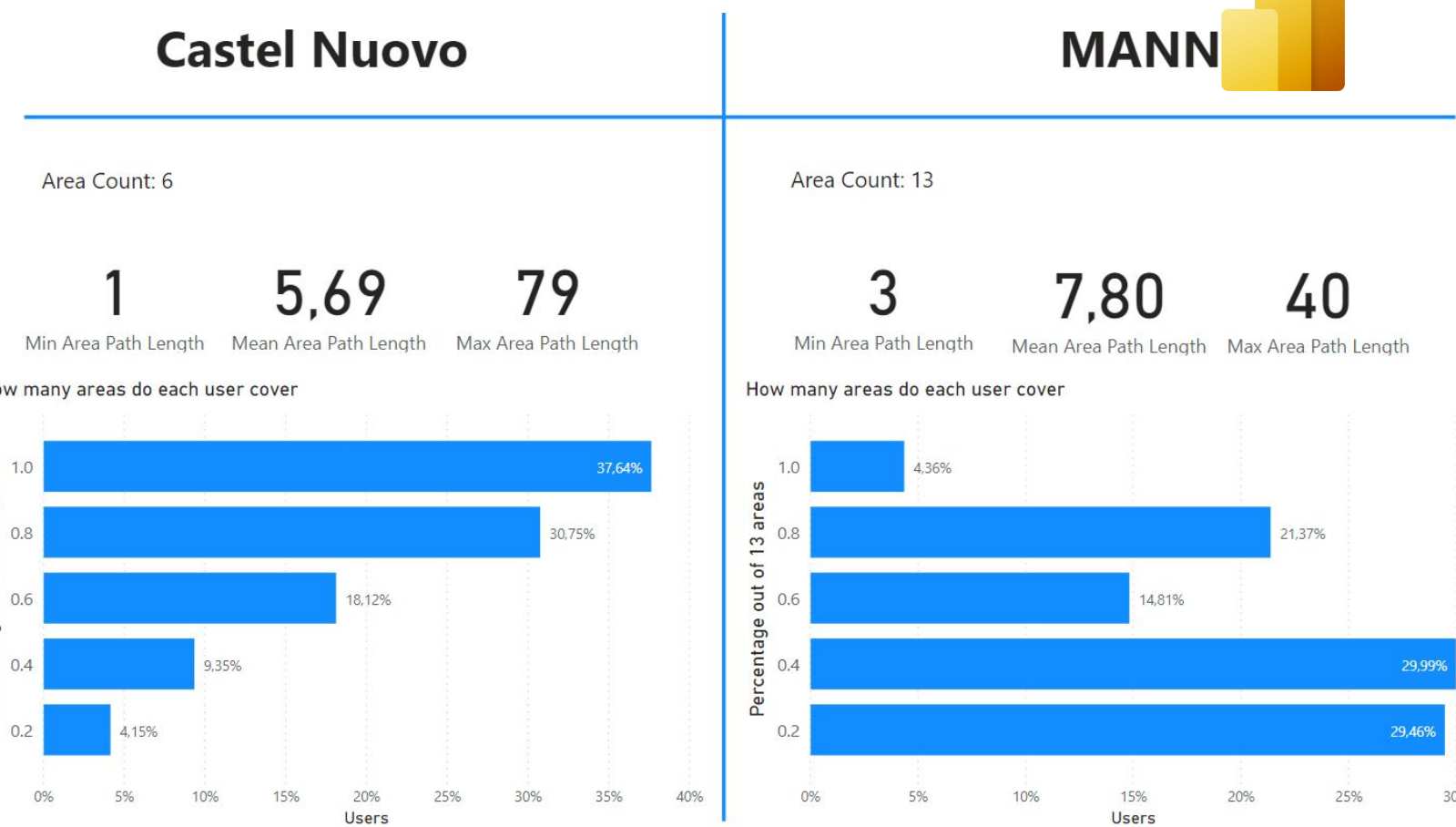
Similar process was also applied to booking data of the Local Health Department of Naples, and presented at the Grakn Conference.

Future developments:
The third year will focus on collaboration with foreign universities, and on other applications in Data Science for Predictive Analysis.
Extend embedding result on Cultural Heritage data:
- spatially model the nodes of the museum
- identification of groups in the museum
- cluster of typical of days

KG research: Node classification, Community detection and Link prediction

Challenging topics:
- Explainable AI (XAI)
- Generative Adversarial Networks (GAN)
- Reinforcement Learning (RL)

Testing all this technologies on different application domains

UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

Ph.D
itee
INFORMATION TECHNOLOGY
ELECTRICAL ENGINEERING