

Dipartimento di Ingegneria e Scienza dell’Informazione

– KnowDive Group –

**KGE 2024**

**Knowledge Graph (KG) for Health Facilities in Trentino**

Document Data: November 2024

Reference Persons: **Zehra Deniz Tas, Lydia Assefa Bekele, Cecilia Peccolo**

© 2023 University of Trento

Trento, Italy

KnowDive (internal) reports are for internal only use within the KnowDive Group. They describe preliminary or instrumental work which should not be disclosed outside the group. KnowDive reports cannot be mentioned or cited by documents which are not KnowDive reports. KnowDive reports are the result of the collaborative work of members of the KnowDive group. The people whose names are in this page cannot be taken to be the authors of this report, but only the people who can better provide detailed information about its contents. Official, citable material produced by the KnowDive group may take any of the official Academic forms, for instance: Master and PhD theses, DISI technical reports, papers in conferences and journals, or books.



### **1. Informal Purpose**

"I want to build a Knowledge Graph (KG) that provides comprehensive, accessible, and structured information about health facilities across Trentino. The KG will help residents locate essential healthcare services, such as hospitals, pharmacies, residential care, and semi-residential care, based on type, availability, and capacity. It will support informed healthcare choices, allowing users to find facilities suited to their specific needs, especially in cases requiring urgent or specialized care."

**1.a** **Scenarios:**

1. **Lucia's Search for Memory Care:**Lucia uses the KG to locate nearby residential care facilities that specialize in memory care for her spouse. She filters options by distance, memory care availability, and user reviews, using the KG’s comparative analysis feature to select a suitable facility.
2. **Marco’s Family Health Needs:**Marco queries the KG for nearby pharmacies that offer family health services, such as vaccinations. He filters for extended hours and online appointment options, allowing him to select a facility that meets his needs and fits into his schedule.
3. **Dr. Rossi’s Patient Referrals:**Dr. Rossi needs to refer a patient for orthopedic care. Using the KG, he quickly finds nearby hospitals with relevant specialties and current availability, enabling a timely referral.
4. **Elena’s Research on Healthcare Access:**For her research on healthcare access disparities, Elena uses the KG to gather data on the distribution and types of healthcare facilities across Trentino. The structured data helps her analyze service availability and identify areas where residents may lack adequate access.

**1.b Personas:**

1. **Lucia:** A 65-year-old retired teacher from Trentino seeking a memory care facility for her spouse recently diagnosed with Alzheimer’s. Lucia, who values clear information, feels overwhelmed by the options and needs a way to easily compare nearby facilities by services, availability, and location.
2. **Marco:** A 40-year-old father with a busy schedule, living in suburban Trentino. He needs a pharmacy that offers family services, like vaccinations, with flexible hours and online booking options, to fit around his work and family commitments.
3. **Dr. Rossi:** A 50-year-old general practitioner in Trentino who often refers patients to specialized or urgent care facilities. He needs quick, reliable access to data on local healthcare options to make efficient referral decisions.
4. **Elena:** A 30-year-old public health researcher focusing on healthcare accessibility in Trentino. She requires structured data on healthcare facility distribution and services to analyze regional trends in access and availability for her research.

### **2. Competency Questions (CQs)**

1. What are the nearest hospitals to a specific location?
2. Which pharmacies in Trentino offer home delivery services?
3. How many beds are available in residential care facilities in a specific area?
4. What specialized services are offered by hospitals in Trentino?
5. Which health facilities provide 24/7 emergency care within a certain radius?
6. Are there facilities offering memory care within a specific distance?
7. What are the average wait times for non-emergency services at local hospitals?
8. Are there health services in Trentino that offer online booking?
9. What are the comparative reviews of nearby pharmacies?
10. How does the availability of transportation impact access to healthcare in rural areas?

### **3. Formal Purpose**

The goal of this project is to develop a Knowledge Graph (KG) that aggregates, structures, and makes accessible comprehensive information on health facilities across Trentino. This KG will enable residents to locate healthcare services—including hospitals, pharmacies, residential care, and semi-residential care facilities—according to their type, availability, capacity, and specific services offered. By enhancing residents' ability to make informed choices, particularly in urgent or specialized care situations, the KG aims to improve community health awareness and accessibility. In addition, the KG will be designed to support data reusability, ensuring that the information remains accessible and applicable across various user contexts and scenarios.

**Contextualization:**

* **Domain of Interest:** The domain of the KG is health services within the geographic scope of the Trentino region, focusing on both public and private facilities offering a wide range of health and care services.
* **Geographical Boundaries:** The KG will cover the entire autonomous province of Trentino, addressing healthcare needs in both urban and rural areas.
* **Temporal Boundaries:** The KG will be updated with real-time data to reflect current facility availability and services, with regular revisions to maintain accuracy.
* **Domain Boundaries:** The KG focuses exclusively on healthcare facilities, categorizing them by type (e.g., hospitals, pharmacies) and specific capabilities (e.g., emergency services, specialized care).

### **3.a Concept Identification**

Based on the CQs and the formal purpose, key concepts and entity types are identified, organized by Focus Levels:

#### **Core Concepts**

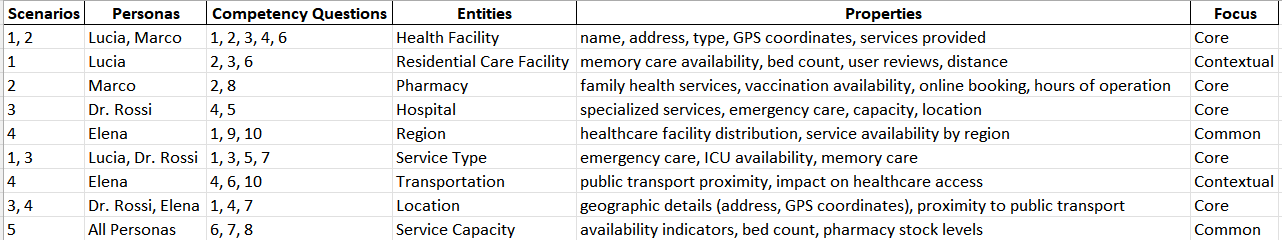
* Health Facility: Represents various healthcare providers, including hospitals, pharmacies, residential and semi-residential care facilities.
* Service Type: Different types of healthcare services provided at each facility (e.g., ICU, memory care, pharmacy services).
* Location: Geographic details such as address and GPS coordinates of each facility.
* Capacity: Availability indicators like bed count, memory care units, or pharmacy stock levels.

#### **Contextual Concepts**

* Availability: Operational hours and open/closed status, especially relevant for pharmacies and urgent care centers.
* Service Capacity: Specific to services like ICU availability or specialized care units, indicating the facility’s ability to accommodate additional patients.

#### **Common Concepts**

* Region: Subdivisions within Trentino to group healthcare resources by area.
* Facility Type Classification: A standardized classification (e.g., Hospital, Pharmacy) to facilitate querying by facility type.



### **3.b. ER Modeling**

### The ER model provides a structured foundation for implementing the KG, defining core entities and relationships based on focus levels and competency questions.

### **Entity Classification**

#### **Strong Entities**

* **HEALTH\_FACILITIES**: Independent entity representing various health facilities. It serves as a central link to other entities based on facility type and location.
* **PHARMACIES**: Independent entity with unique location attributes that can exist without any associated entities.
* **HOSPITAL\_FACILITIES**: An independent entity with unique location and facility-specific details.
* **SPECIALIST\_SERVICES\_PRIVATE** and **SPECIALIST\_SERVICES\_PUBLIC**: These entities can exist independently but are related by service code to allow mapping between public and private services.
* **DRG\_HOSPITAL\_ADMISSIONS**: Represents hospital admissions data, independent of any other entities.

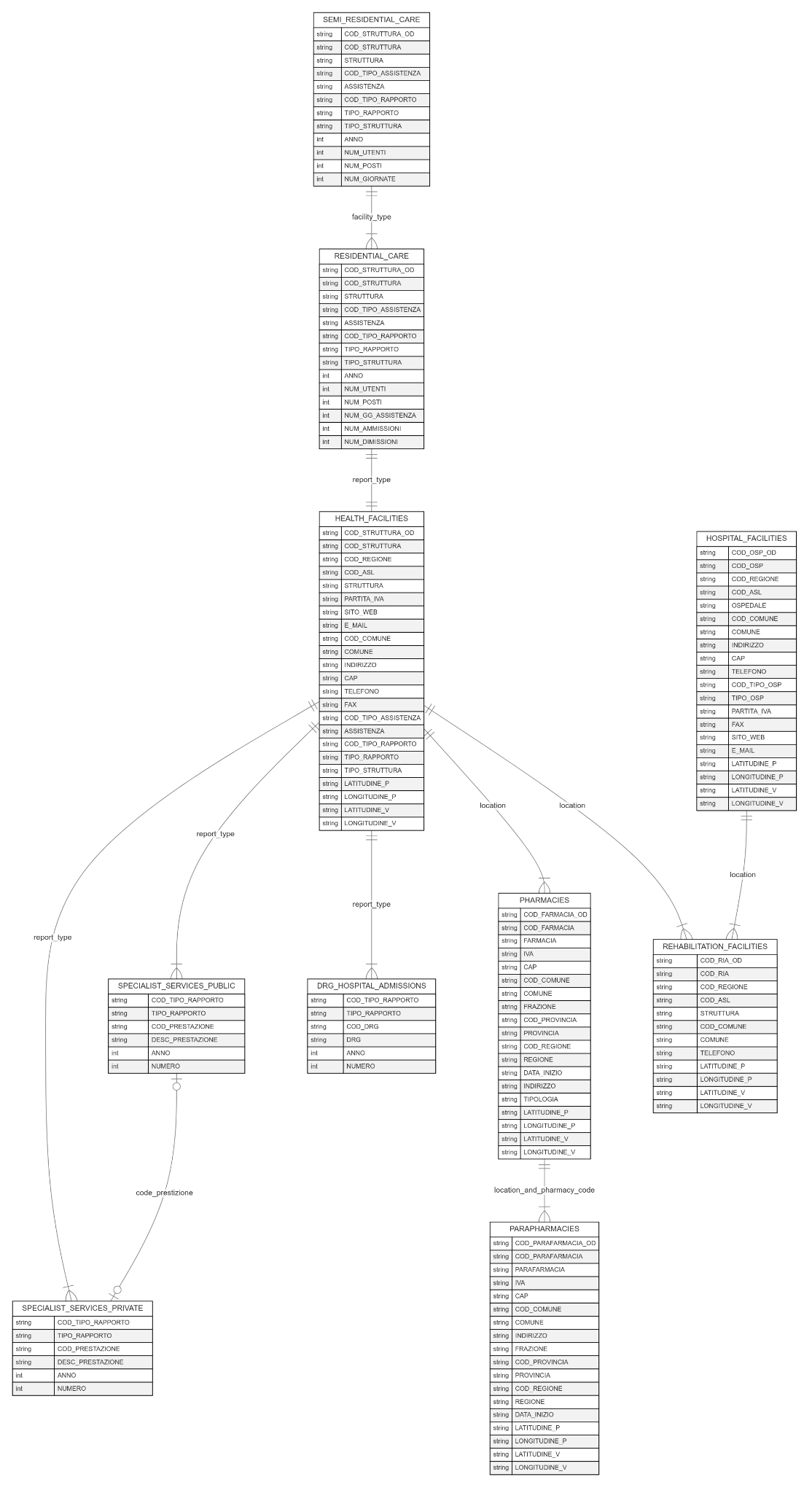
#### 

#### **Weak Entities**

* **SEMI\_RESIDENTIAL\_CARE**: Depends on **RESIDENTIAL\_CARE** as it uses the same facility ID. It cannot exist independently without a related residential care facility.
* **RESIDENTIAL\_CARE**: Although semi-dependent, it is meaningful only in the context of a **HEALTH\_FACILITY** that provides residential care.
* **PARAPHARMACIES**: Dependent on **PHARMACIES** for the location information.
* **REHABILITATION\_FACILITIES**: Dependent on **HEALTH\_FACILITIES** or **HOSPITAL\_FACILITIES** for the location, indicating its dependence on a hospital or healthcare facility.

### **Cardinalities**

* **One-to-Many** (||--|{): Indicates that one instance of the first entity can relate to multiple instances of the second entity.
* **Many-to-Many** (|o--o|): Indicates that multiple instances of both entities can relate to each other.
* **One-to-One** (||--||): Indicates that one instance can relate to one.



**4. Information Gathering**

**Objective**The Information Gathering phase focuses on the collection and preparation of resources essential to building a comprehensive and structured Knowledge Graph (KG) of healthcare facilities across Trentino. This KG will help residents find suitable healthcare services based on facility type, availability, and capacity.

The phase has three core aims:

1. **Identify and Access** high-quality sources and datasets relevant to the healthcare facilities in Trentino.
2. **Organize Resources** along two main process dimensions (Producer and Consumer) and two layers (Data and Knowledge) to support informed and flexible access to healthcare data.
3. **Enhance Data Quality** by removing irrelevant or incomplete information, ensuring consistency with standards for data interoperability and reusability.

**4.a Producer Activities**

**4.a.a. Knowledge Layer**

* **Sources Description:**  
  description of the website where we got the data
* **Informal Resource Collection:**  
  how we got the data.
* **Informal Resources Classification**:  
  why we selected specific data

**4.a.b Data Layer**

* **Sources Description:**  
  detailed description of the dataset and the columns
* **Resources Collection and Scraping:**  
  methods used to pull in the data, write for example that we scraped the timetable of the facilities from the web
* **Resources Classification:**  how this data is sorted into categories, as in the Knowledge Layer (common, core, and contextual).

**4.b Consumer Activities**

Consumer activities focus on gathering structured, high-quality resources that are already standardized and meet quality guidelines, facilitating easier integration into the KG.

#### **4.b.a. Knowledge Layer**

To enrich the healthcare knowledge graph, we draw from the **FHIR ontology** to better define and structure our healthcare entities. The FHIR ontology provides a structured framework for healthcare data, including entities, relationships, and attributes that align well with our project’s needs for representing healthcare facilities, services, and operational data. By leveraging specific FHIR entities and definitions, we ensure a standardized approach, enabling clearer representation and interoperability across systems.

We selected resources within the FHIR ontology that closely match our primary entities, aiding in the accurate definition of our knowledge graph components. The following are particularly useful for our project:

* **Location**: This FHIR entity helps define attributes for healthcare facility locations, which aligns with our **HEALTH\_FACILITIES** entity, specifying aspects like address, geo-coordinates, and type.
* **Organization**: Used for healthcare providers, it enables us to define key attributes for **PHARMACIES**, **HOSPITAL\_FACILITIES**, and other facility types, covering aspects like facility name, operational details, and contact information.
* **ServiceRequest**: This entity is relevant for representing **SPECIALIST\_SERVICES** in both public and private facilities, defining services based on type, location, and provider, which helps map the services offered in Trentino.
* **Encounter**: This can be used to represent **DRG\_HOSPITAL\_ADMISSIONS**, offering a framework for capturing admission details, patient interactions, and discharge information, which aligns well with the data needs in our graph.

**4.b.b. Data Layer**

**..**

**4.c Dataset cleaning**

**4.c Dataset Standardization**

**4.d Resource Formatting**

*…………ontology screenshot*