Dipartimento di Ingegneria e Scienza dell’Informazione

– KnowDive Group –

Knowledge Graph course 2025 - Project Report

Document Data:	Reference Persons:
January 13, 2026	Mattia Santaniello, Mattia Maci

Index:

Revision History:

Revision	Date	Author	Description of Changes
1	January 13, 2026	Mayukh Bagchi	Document created

1 Introduction

Reusability is one of the main principles in the Knowledge Graph (KG) development process defined by iTelos. The KG project documentation plays an important to enhance the reusabiltiy of the resources handled and produced during the process. A clear description of the resources, the process (and sub processes) developed and evaluation at each step of the process provides a clear understanding of the project, thus serving such an information to external readers for the future exploitations of the project's outcomes.

The current document aims to provide a detailed report of the project developed following the iTelos methodology. The report is structured, to describe:

- Section 2: Definition of the project's purpose and related information gathering.
- Sections 3, 4, 5, 6: The description of the iTelos process phases and their activities, divided by knowledge and data layer activities, as well as the evaluation of the resources produced in terms of fit for the chosen purpose.
- Section 7: The description of the metadata produced for all (and all kind of) the resources handled and generated by the iTelos process, while executing the project.
- Section 8: Conclusion and open issues summary.

2 Project Design

2.1 Purpose

The first step is the definition of the project's objective. We start from an informal definition of our purpose, captured in the following phrase.

> "This project is about building a Knowledge Graph that brings together data from different weather stations to better understand the climate of the Trentino area during the decade 2015-2025. It helps organize and connect meteorological data in a smart way so users can easily explore trends and patterns. Using this graph, you can ask meaningful questions about temperature, rainfall, and changes over time, like:
>
> - what are the most impacted area by the climate change
> - where the major weather stations are located
> - and how the wheater and the temperature distributions evolved"

We can see how our domain of interest includes meteorological observations, spatial information about weather stations, and temporal aspects related to climate evolution over time.

2.2 information gathering

We gathered information from the meteotrentino website, from which we collected data from the meteorological stations on rainfall, temperature, humidity, wind direction, wind speed, atmospheric pressure, and solar radiation, divided by year. Some stations provided only part of this data, while others had missing years or unreliable measurements. TODO see http://storico.meteotrentino.it/web.htm

3 Purpose Definition

In the following section we are taking the informal description of the purpose in the project design and use *iTelos* methodology to extract a formal description of the purpose.

3.1 Domain

The area of knowledge and the field of study interested are climatology and environmental data analysis, with a specific focus on regional climate change. More precisely, the project lies at the intersection of meteorology, geospatial analysis, and semantic data modeling through Knowledge Graphs(KG).

3.2 Context

more specifically we are interested in the context of the Trentino area, during the decade 2015-2025, analyzing local climate variations, weather patterns, and their impact on the environment and communities.

3.3 Scenarios

now we will continue illustrating the possible scenarios:

- **S1: Mario Rossi** is performing a research on the effects of climate change, which could be useful for some supervisor agencies specialized on the theme (e.g ESA). He needs to retrieve some data about the weather of the territory, analyzing how the temperature has evolved over the years and how much it has drifted from the natural trend.
- **S2: Laura Rosa** is planning a long journey on the mountains with her friends, and she needs to know the metheorogical data of the locations they are going to visit, especially the rainfall events and the sunny days of the previous years in order to make decision on the logistics
- **S3: Luigi Verdi** is worried about how the climate events of the last years could affect his farming activities and damage his products.Agriculture is one of the sectors which has been mostly affected negatively by the climate change, so he wants to be prepared

3.4 Personas

Now that the scenarios have been defined it is time to see the profiles of the users mentioned above:

- **P1: Mario Rossi** is a 26 years old researcher at the Trento university, specialized in ecology and enviromental research
- **P2: Laura Rosa** is a 30 years old woman who is also planning a trip on the mountains of the Trentino region, however in her case the main activity is a trekking exursion with her friends
- **P3: Luigi Verdi** is a 56 years old farmer, with a small field nearby Trento. His main products are potatoes and grape, the latter one especially is cultivated in summer and gathered during grape harvest period

3.5 Competency Question

Each Persona has to deal with different scenarios, which could be incapsuleted in the definition of Competency Questions(CQ). Some of them could be formalized in the following table

Person	Question
Mario Rossi	**Q1:** What is the trend of the temperature values in Trentino during the decade 2015-2025?
Mario Rossi	**Q2:** How the metheorogical facilities are distributed in the territory?
Mario Rossi	**Q3:** Which areas in Trentino have seen the highest increase in temperature over time?
Laura Rosa	**Q4:** How was the temperature on the Trentino mountains during the winter of the last three years?
Laura Rosa	**Q5:** What are the locations where rainfalls most frequently happen?
Laura Rosa	**Q6:** What are the periods of the year where there are sunny days?
Luigi Verdi	**Q7:** Did the number of extreme climate events increased in the area around my field?
Luigi Verdi	**Q8:** What is the trend of humidity in the summer of the last three years?
Luigi Verdi	**Q9:** Did the number of rainfalls decreased over the last five years?

Table 1: Competency Questions

3.6 PurposeFormalization

now from the scenarios, the personas, and the competency questions we extract a list of usefull concepts, During the concept identification phase, we focused on concepts that are directly re-

quired to answer the defined competency questions and that can be grounded in the available data sources.
Based on the available datasets, the core entities required to answer the defined competency questions are **MeteorologicalFacility, MeteorologicalMeasurement and GeographicLocation**.

Scenarios	Personas	Competency Questions	ID	Entities	Properties	Focus
S1,S2,S3	P1,P2,P3	Q1,Q2,Q3,Q4,Q4,Q5,Q6,Q7,Q8,Q9	Q190107	Meteorological facility	Stazione, Tavoletta n., Coordinate Est/Nord, Latitudine, Longitudine, Note	Core
S1,S2,S3	P1,P2,P3	Q1,Q2,Q3,Q4,Q4,Q5,Q6,Q7,Q8,Q9	UNITN2026KGMMC02	Meteorological measurement	ID, Date, Air temperature, Min air temperature, Max air temperature, Rain in ml, Tot rain in ml, Air relative humidity, Avg wind direction, Avg wind velocity, Atmospheric pressure, Tot solar radiation	Contextual
S1,S2,S3	P1,P2,P3	Q2,Q3,Q4,Q4,Q5,Q7	Q2221906	Geographic location	ID, nome, latitudine, longitudine	Core
S1,S2,S3	P1,P2,P3	Q1,Q3,Q4,Q4,Q5,Q6,Q7,Q8,Q9	UNITN2026KGMMC04	Meteorological varable	Name, unit, value	Common
S1,S2,S3	P1,P2,P3	Q1,Q2,Q3,Q4,Q4,Q5,Q6,Q7,Q8,Q9	Q11471	Time	Year, month, timestamp	Common
			UNITN2026KGMMC06	Climate variation		Common
			Q1277161	Extreme events		Common
			UNITN2026KGMMC08	Trend		Core
			UNITN2026KGMMC09	Climate change indicator		Common
			Q125928	Climate change		Common

Figure 1: Purpose formalization overview

Structural concepts such as **Meteorological Facility** and **Meteorological Measurement** are associated with explicit properties, as they directly correspond to entities present in the data sources. Other concepts, including **temperature trend**, **climate variation** and **climate change indicators**, are considered derived analytical notions are inferred through queries and statistical analyses over the collected measurements.

Some concepts refer to concep already modeled in wikidata, therefore for this concept we linked the ID to the wikidata one

3.7 ER Diagram

At this point the next step of the formalization procedure is to produce a first bone structure which allows us to define the involved entities more clearly. Lets first define their schemas:

Etype	**Properties**
MeteorologicalFacility	***Tavoletta n.***,*Stazione, Coordinate Est/Nord, Latitudine, Longitudine, Note*
MeteorologicalMeasurement	***ID***, *Date, Air temperature, Min air temperature, Max air temperature, Rain in ml, Tot rain in ml, Air relative humidity, Avg wind direction, Avg wind velocity, Atmospheric pressure, Tot solar radiation*
GeographicLocation	***ID***,*nome, latitudine, longitudine*

Table 2: Informal definition of the entities

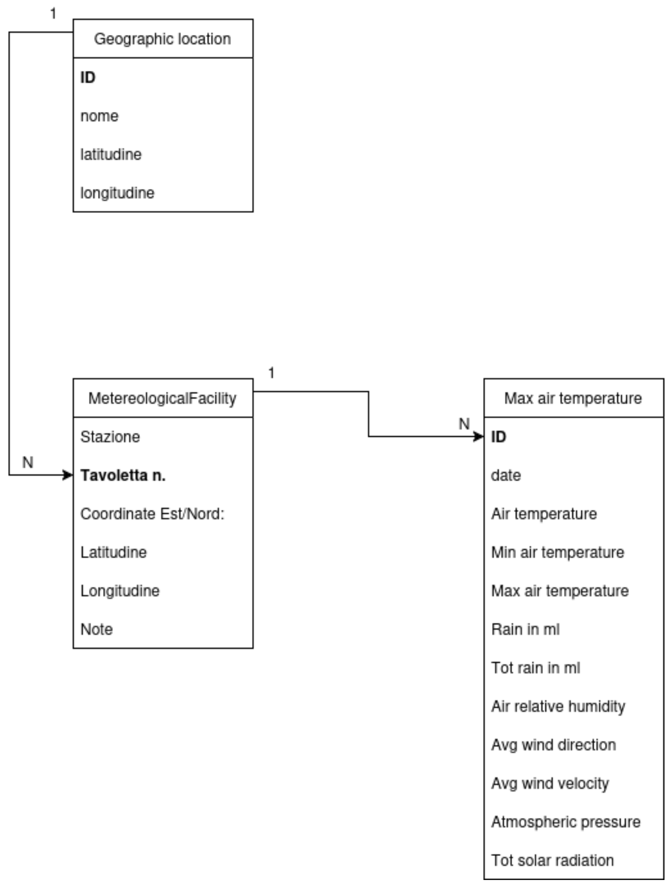

Figure 2: ER diagram

As shown in Table 3.7, the core entities of the Knowledge Graph represent the fundamental components required to describe and analyze meteorological data at a regional scale.

In particular, the entity *GeographicLocation* represents a generic spatial reference within the Trentino area and is characterized by a name and geographical coordinates, which allow the association of measurements and facilities to specific areas of interest.

The entity *MeteorologicalFacility* models the physical weather stations deployed in the territory. Each facility is uniquely identified and associated with a precise geographical location through latitude and longitude attributes. Additional descriptive properties, such as the station name and technical identifiers, are included to support spatial analysis and the study of the distribution of meteorological facilities across the region.

Meteorological observations are modeled through the *MeteorologicalMeasurement* entity, which represents a single measurement recorded by a meteorological facility at a given point in time. This entity includes several meteorological variables, such as air temperature, minimum and maximum temperature, rainfall, relative humidity, wind characteristics, atmospheric pressure, and solar radiation, together with a temporal reference.

Each meteorological measurement is associated with exactly one meteorological facility, while each facility can be related to multiple measurements collected over time, resulting in a one-to-many relationship.

Based on the definition of these entities, the Entity–Relationship diagram shown in Figure 2 illustrates the structural organization of the Knowledge Graph. The separation between meteorological facilities, measurements, and geographic locations allows the model to clearly distinguish between spatial entities, observation sources, and temporal data. This design supports the analysis of long-term climate trends and spatial variations by enabling aggregation and comparison of measurements across different locations and time periods.

3.8 Information Gathering

3.9 Evaluation - Purpose Definition

4 Language Definition

5 Knowledge Definition

6 Entity Definition

7 Evaluation

8 Metadata Definition

9 Open Issues