



UNIVERSITY
OF TRENTO - Italy



Dipartimento di Ingegneria e Scienza dell'Informazione

– KnowDive Group –

Knowledge Graph course 2025 - Project Report

Document Data:

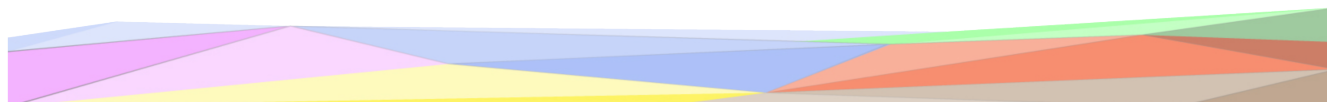
January 10, 2026

Reference Persons:

Mattia Santaniello, Mattia Maci

© 2026 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.



Index:

1	Introduction	1
2	Project Design	1
2.1	Purpose	1
2.2	information gathering	1
3	Purpose Definition	2
3.1	Personas	2
3.2	Competency Question	2
3.3	ER Diagram	3
3.4	Information Gathering	5
3.5	Evaluation - Purpose Definition	5
4	Language Definition	5
5	Knowledge Definition	5
6	Entity Definition	5
7	Evaluation	5
8	Metadata Definition	5
9	Open Issues	5

Revision History:

Revision	Date	Author	Description of Changes
1	January 10, 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 role to enhance the reusability 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”

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

From the Project design it is possible to derive a formal description of the situation by first illustrating the possible scenarios:

- **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.
- **Laura Rosa** is planning a long journey on the mountains with her friends, and she needs to know the meteorological 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
- **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

3.1 Personas

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

- **Mario Rossi** is a 26 years old researcher at the Trento university, specialized in ecology and environmental research
- **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 excursion with her friends
- **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

3.2 Competency Question

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

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

Table 1: Competency Questions

3.3 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
Location	location_ID ,gps_coordinates,name
Coordinates	latitude,longitude
Weather	temperature,humidity,date, wind speed,is_sunny
MetheorologicalFacility	facility_ID ,name,gps_coordinates, inauguration_date,still_active
ClimateEvent	event_ID ,location,weather_status,event_date

Table 2: Informal definition of the entities

As we can see, in the table 3.3 there are some entities with a more general use, like *Location* and *Coordinates*, which represents respectively the place where a specific event happens and the gps coordinates of a certain position. Next we find the weather conditions associated to a location, which includes the temperature and the humidity, the wind speed, wheter the day was sunny or not and in the end the date timestamp of those specific metheorological configuration. Last we have the metheorological facility, which include its foundation date togheter with its position, and the dangerous climate event happened at a specific location and it is associated with the weather conditions of that day. Notice how in this latter case we also store a further date timestamp since we want to keep track of when the catastrophe happened, which does not

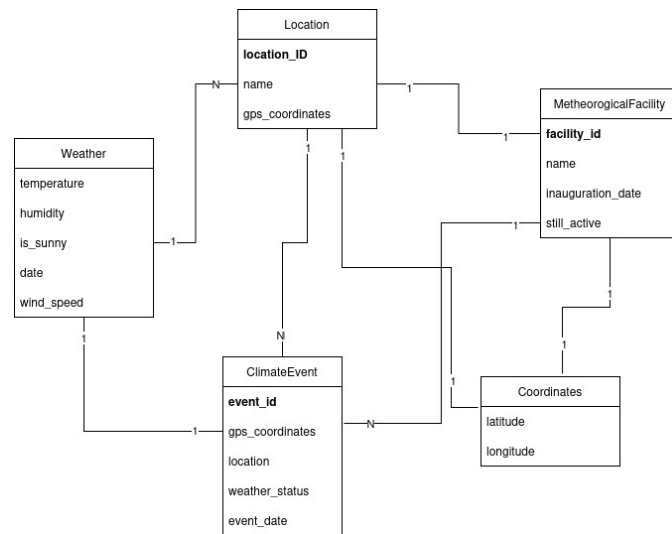


Figure 1: ER diagram

necessary match the time memorized in the Weather entity (maybe because the weather conditions are sampled every 24 hours, and the climate event happen in the meantime). With the definition of the entities (although informal) done it is possible now to define the ER diagram that declares the relationship between the subjects involved, as in figure 1. It is interesting to notice how there is a one-to-one relationship between MeteorologicalFacility entity and Coordinates while it could be sufficient to identify a facility through the location. The main reason why such decision has been made is because it is possible that a meteorological station is not exactly placed inside a specific location but it could be in the nearby.

-
- 3.4 Information Gathering**
 - 3.5 Evaluation - Purpose Definition**
 - 4 Language Definition**
 - 5 Knowledge Definition**
 - 6 Entity Definition**
 - 7 Evaluation**
 - 8 Metadata Definition**
 - 9 Open Issues**