

Dipartimento di Ingegneria e Scienza dell'Informazione

– KnowDive Group –

# KGE 2024 - Trentino Climate Change Causes

---

Document Data:

October 29, 2024

Reference Persons:

Lorenzo Fumi, Giacomo Tezza

© 2024 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:

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Purpose Definition</b>	<b>1</b>
2.1	Informal purpose . . . . .	1
2.2	Context definition . . . . .	2
2.3	Domain of Interest (DoI) . . . . .	2
2.4	Scenarios definition . . . . .	2
2.4.1	Scenario 1: Policy Analysis and Environmental Planning . . . . .	3
2.4.2	Scenario 2: Community Awareness and Education . . . . .	3
2.4.3	Scenario 3: Urban and Rural Development . . . . .	3
2.4.4	Scenario 4: Environmental Research and Analysis . . . . .	3
2.4.5	Scenario 5: Infrastructure and Resource Management . . . . .	3
2.5	Personas definition . . . . .	3
2.5.1	Persona 1: Giulia Rossi, Policy Analyst . . . . .	4
2.5.2	Persona 2: Marco Bianchi, Science Teacher . . . . .	4
2.5.3	Persona 3: Sara Verdi, Urban Planner . . . . .	4
2.5.4	Persona 4: Luca Ferro, Environmental Researcher . . . . .	5
2.5.5	Persona 5: Antonio Russo, Resource Management Specialist . . . . .	5
2.6	Competency Question definition . . . . .	6
2.6.1	Policy Analysis and Environmental Planning . . . . .	6
2.6.2	Community Awareness and Education . . . . .	6
2.6.3	Urban and Rural Development . . . . .	6
2.6.4	Environmental Research and Analysis . . . . .	7
2.6.5	Infrastructure and Resource Management . . . . .	7
2.7	Concepts Identification . . . . .	7
2.8	ER modeling . . . . .	9
<b>3</b>	<b>Information Gathering</b>	<b>9</b>
<b>4</b>	<b>Language Definition</b>	<b>10</b>
<b>5</b>	<b>Knowledge Definition</b>	<b>11</b>
<b>6</b>	<b>Entity Definition</b>	<b>11</b>
<b>7</b>	<b>Evaluation</b>	<b>11</b>
<b>8</b>	<b>Metadata Definition</b>	<b>12</b>

Revision History:

Revision	Date	Author	Description of Changes
0.1	October 21, 2024	Lorenzo Fumi, Giacomo Tezza	Document created
0.2	October 29, 2024	Lorenzo Fumi, Giacomo Tezza	Purpose Definition

---

# 1 Introduction

Reusability is one of the main principles in the Knowledge Graph Engineering (KGE) process defined by iTelos. The KGE 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 as well as of the process (and single activities) developed, provides a clear understanding of the project, thus serving such an information to external readers for the future exploitation 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 as follows:

- Section 2: Definition of the project's purpose and its domain of interest.
- Section 3: High level description of the project development, based on the Produce role's objectives.
- Sections 4, 5, 6, 7 and 8: The description of the iTelos process phases and their activities, divided by knowledge and data layer activities.
- Section 9: The description of the evaluation criteria and metrics applied to the project final outcome.
- Section 10: 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 11: Conclusions and open issues summary.

## 2 Purpose Definition

### 2.1 Informal purpose

The aim of the project is to create a Knowledge Graph capable to establish a comprehensive and interconnected representation of the factors contributing to climate change within the autonomous province of Trento. This purpose centers on capturing and linking various types of data (geospatial, meteorological, and demographic) over historical time frames, with the goal of illuminating the relationship between human activities, environmental conditions, and resulting climate impacts in this specific region.

By integrating geospatial data, such as that from OpenStreetMap, with historical meteorological records, including temperature trends, air quality metrics, and other environmental indicators, the Knowledge Graph aims to reflect the environmental transformations that have occurred over the years. Additionally, historical demographic data, such as population density, energy consumption, gas usage, water consumption, and waste production, will provide insight into local patterns of resource utilization and waste generation. These variables will collectively inform an

---

understanding of how these human and environmental factors interlink to drive climate change in Trento.

Ultimately, the Knowledge Graph will serve as a tool to clarify the underlying causes of climate change within the context of this region, allowing users to explore the interconnected nature of socio-environmental impacts and aiding in the development of targeted actions or policies to address these issues.

## **2.2 Context definition**

The Context of this Knowledge Graph project encompasses geographical, temporal, and domain-specific boundaries that shape and constrain the analysis of climate change in the autonomous province of Trento, Italy. This project primarily reflects a Reference Context Purpose, aimed at examining the broader environmental and socio-economic factors in Trento without a focus on individual user perspectives. The context, however, implicitly considers the interaction between human activities and the reference environment, making it relevant for understanding how individuals and communities are affected by and contribute to climate change within this regional setting.

## **2.3 Domain of Interest (DoI)**

The Domain of Interest for this Knowledge Graph project is climate change within the autonomous province of Trento. This domain encompasses both geographical and temporal dimensions that are essential to understanding the climate dynamics and human interactions in the region.

Spatially, the domain is limited to the province of Trento, situated in Northern Italy and characterized by diverse topography that includes mountainous landscapes, forests, rivers, and valleys. These geographic features play a critical role in shaping the local climate and contribute to the unique environmental challenges faced by the area. Trento's natural landscape, alongside its urban areas and agricultural zones, is integral to the Knowledge Graph, as it provides the physical context in which climate-related variables and human activities interact.

Temporally, the project focuses on historical data extending over the past several decades, capturing long-term changes and trends in climate, population, resource consumption, and environmental quality. This temporal scope is chosen to reveal patterns and variations that highlight how climate change has evolved in Trento over time, influenced by both natural and human factors. By studying these historical trends, we aim to identify and understand the temporal progression of climate effects and contributing factors specific to this region.

## **2.4 Scenarios definition**

These scenarios illustrate how users might interact with the Knowledge Graph to analyze, interpret, and respond to climate-related information, providing a foundation for further definition of user personas in the next section.

---

#### **2.4.1 Scenario 1: Policy Analysis and Environmental Planning**

In this scenario, the Knowledge Graph supports policy analysts and environmental planners seeking to assess the impact of regional policies on climate change factors, such as emissions, energy consumption, and waste production. By exploring the relationships between policy implementations and climate indicators over time, users can evaluate the effectiveness of past policies and shape future strategies for mitigating climate impacts.

#### **2.4.2 Scenario 2: Community Awareness and Education**

This scenario focuses on enhancing community awareness about climate change in Trento. Educators, community leaders, and local activists can use the Knowledge Graph to access and visualize data on local climate conditions and human activities, creating educational materials or interactive sessions for residents. This scenario helps raise awareness of the impact of individual and collective actions on the environment, fostering informed community discussions and behavior change.

#### **2.4.3 Scenario 3: Urban and Rural Development**

In this scenario, urban planners and development specialists leverage the Knowledge Graph to analyze the effects of urbanization and rural land use on climate change in Trento. They can investigate factors like population density, energy use, and waste production within urban and rural areas to assess environmental impacts and guide sustainable development practices that align with climate objectives.

#### **2.4.4 Scenario 4: Environmental Research and Analysis**

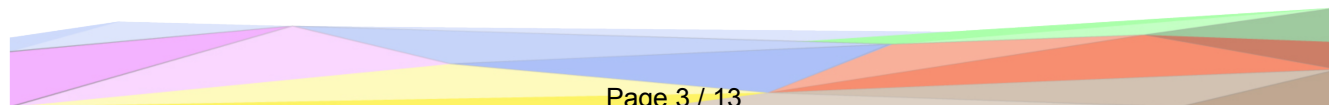
Researchers and scientists studying climate change can use the Knowledge Graph to analyze historical climate and socio-economic data in Trento, identifying correlations and trends within this specific regional context. This scenario enables deep analysis of local climate change phenomena, serving as a foundation for scientific studies and publications that contribute to understanding and addressing climate change at the regional level.

#### **2.4.5 Scenario 5: Infrastructure and Resource Management**

This scenario addresses the needs of infrastructure managers and resource planners who require insight into the demand and usage patterns of resources like energy and water. By using the Knowledge Graph, they can assess how climate change and demographic shifts affect resource consumption, helping to optimize infrastructure planning and ensure the resilience of critical resources in the face of environmental change.

### **2.5 Personas definition**

These semi-fictional characters represent key groups of end-users who will interact with the Knowledge Graph within the various scenarios defined. Each persona embodies distinct goals,



---

motivations, and challenges, aligning with the core purpose of the Knowledge Graph project and reflecting broader user needs within the climate change domain. These personas, along with their associated use cases, provide a robust framework to ensure that the Knowledge Graph design addresses a range of user perspectives.

### 2.5.1 Persona 1: Giulia Rossi, Policy Analyst

**Background:** Giulia is a 34-year-old environmental policy analyst working for the provincial government in Trento. She has a master's degree in environmental science and specializes in evaluating the effectiveness of local and regional environmental policies.

**Goals:** Giulia wants to assess the impact of Trento's climate policies on reducing emissions, improving air quality, and managing resources. She's also interested in identifying data-supported recommendations for future policy adjustments.

**Challenges:** Her main challenge is accessing up-to-date, interconnected data that shows how past policies have influenced climate indicators. She needs reliable, actionable insights to back her policy recommendations, especially when presenting findings to decision-makers.

**Use Case:** Giulia uses the Knowledge Graph within the Policy Analysis and Environmental Planning scenario 2.4.1, accessing historical data on climate indicators and policy impacts. She generates reports that link policy decisions to measurable outcomes, allowing her to advocate for evidence-based policy adjustments.

### 2.5.2 Persona 2: Marco Bianchi, Science Teacher

**Background:** Marco, 45, teaches high school science in Trento. Passionate about climate education, he seeks engaging ways to make complex climate concepts accessible and relatable for his students.

**Goals:** Marco aims to help students understand the local effects of climate change and encourage eco-conscious behaviors. He wants to create interactive lessons and demonstrations using real-world data from their region.

**Challenges:** Marco struggles to find localized data that visually represents climate change impacts in an easily understandable way for teenagers. He needs a tool that translates raw data into a story that resonates with his students.

**Use Case:** In the Community Awareness and Education scenario 2.4.2, Marco uses the Knowledge Graph to pull localized climate data and display visualizations in his classroom. He develops lesson plans that show the impact of specific actions, like waste reduction, on the environment, inspiring students to participate in local environmental efforts.

### 2.5.3 Persona 3: Sara Verdi, Urban Planner

**Background:** Sara is a 39-year-old urban planner with the municipality of Trento, specializing in sustainable development. She has experience in balancing urban growth with environmental conservation and climate resilience.

---

**Goals:** Sara is focused on planning urban and rural areas that are environmentally sustainable. She aims to understand the interplay between population density, resource consumption, and climate impact to develop long-term plans that align with climate goals.

**Challenges:** Sara faces the challenge of accessing integrated datasets that link urbanization patterns to climate data, such as emissions or energy consumption, over time. She needs insights that help her balance growth and sustainability effectively.

**Use Case:** Within the Urban and Rural Development scenario 2.4.3, Sara uses the Knowledge Graph to analyze relationships between urban density and resource consumption, identifying areas where sustainable practices can be implemented. This enables her to propose zoning and infrastructure plans that support climate resilience.

#### 2.5.4 Persona 4: Luca Ferro, Environmental Researcher

**Background:** Luca, 29, is an environmental researcher pursuing a Ph.D. in climate science with a focus on regional climate impacts. He is keen on studying how climate change affects ecosystems and communities in specific regions, including Trento.

**Goals:** Luca aims to conduct in-depth research into the correlation between human activities and climate changes in Trento. He wants access to comprehensive datasets to explore cause-effect relationships between environmental changes and socio-economic factors.

**Challenges:** Luca's research requires high-quality, localized datasets that are often fragmented. He needs a consolidated data source that supports complex queries and data exploration to support his academic research.

**Use Case:** In the Environmental Research and Analysis scenario 2.4.4, Luca utilizes the Knowledge Graph to access interlinked environmental and socio-economic data over time. This allows him to explore and test hypotheses on the impacts of human activities, contributing to publications that inform the scientific community.

#### 2.5.5 Persona 5: Antonio Russo, Resource Management Specialist

**Background:** Antonio, 52, is a senior infrastructure and resource manager at a public utility company. He oversees resource planning, including water and energy supply, and ensures resilience against climate impacts.

**Goals:** Antonio's goal is to optimize resource management strategies in response to evolving climate patterns and demographic changes. He needs accurate, predictive data to assess future resource demands.

**Challenges:** He often contends with incomplete data, making it difficult to accurately project future resource requirements. He requires data that factors in both current usage trends and anticipated climate impacts to effectively plan for the long term.

**Use Case:** Within the Infrastructure and Resource Management scenario 2.4.5, Antonio leverages the Knowledge Graph to study historical trends in resource consumption linked to climate patterns, allowing him to prepare adaptive strategies that anticipate resource needs under various climate scenarios.



---

## 2.6 Competency Question definition

These critical questions defines what the Knowledge Graph (KG) must answer to meet the needs of the personas within each scenario. These CQs represent the functional requirements of the KG, providing a guide to the entities and relationships that should be modeled. Each question reflects a specific need of a persona in a scenario and indicates the diverse information paths the KG must support.

### 2.6.1 Policy Analysis and Environmental Planning

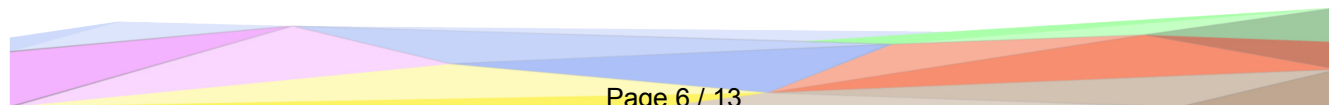
- CQ1 How have specific policies influenced emission levels in Trento over the past decade? (2.5.1, 2.4.1)
- CQ2 Which regions within Trento have seen the most significant reductions in air pollutants following policy implementations? (2.5.1, 2.4.1)
- CQ3 What relationships exist between waste reduction initiatives and air quality metrics over the years? (2.5.1, 2.4.1)
- CQ4 What are the policy measures that have shown the most significant impact on reducing energy consumption in urban versus rural areas? (2.5.1, 2.4.1)

### 2.6.2 Community Awareness and Education

- CQ5 What is the historical trend of energy consumption in Trento, and how does it correlate with population density? (2.5.2, 2.4.2)
- CQ6 How does waste production in Trento vary by neighborhood, and what seasonal patterns are observed? (2.5.2, 2.4.2)
- CQ7 Which areas have experienced the most pronounced changes in air quality, and what are the potential causes? (2.5.2, 2.4.2)
- CQ8 What are the projected environmental impacts if current trends in waste production continue? (2.5.2, 2.4.2)

### 2.6.3 Urban and Rural Development

- CQ9 How does resource consumption (water, gas, energy) differ between urban and rural areas of Trento? (2.5.3, 2.4.3)
- CQ10 What is the relationship between urban density and waste production over time in Trento? (2.5.3, 2.4.3)
- CQ11 Which urban areas are at the greatest risk of environmental degradation due to current resource consumption rates? (2.5.3, 2.4.3)
- CQ12 How have infrastructure improvements influenced energy efficiency and climate resilience in different parts of Trento? (2.5.3, 2.4.3)



---

#### 2.6.4 Environmental Research and Analysis

- CQ13 What are the trends in average temperature changes across Trento over the last 50 years? (2.5.4, 2.4.4)
- CQ14 How does population density correlate with air quality indicators in urban and rural zones of Trento? (2.5.4, 2.4.4)
- CQ15 What are the statistical correlations between gas consumption and emission levels? (2.5.4, 2.4.4)
- CQ16 What patterns emerge when comparing climate indicators (temperature, precipitation) with demographic changes? (2.5.4, 2.4.4)
- CQ17 What seasonal variations are observed in air quality, and how do these relate to energy consumption patterns? (2.5.4, 2.4.4)

#### 2.6.5 Infrastructure and Resource Management

- CQ18 How do seasonal shifts impact water and energy demand across Trento? (2.5.5, 2.4.5)
- CQ19 What is the projected resource demand if population growth continues at the current rate in Trento? (2.5.5, 2.4.5)
- CQ20 Which regions are projected to face resource scarcity, and how can management practices be adapted? (2.5.5, 2.4.5)
- CQ21 How does residential versus industrial gas consumption vary by region, and what trends are projected? (2.5.5, 2.4.5)
- CQ22 What historical data on resource consumption align with trends in urban development? (2.5.5, 2.4.5)

### 2.7 Concepts Identification

This section organizes the essential entities (ETypes) and their properties necessary to satisfy the Competency Questions (CQs) and meet the purpose of the Knowledge Graph (KG) project; categorizing concepts into three levels of importance: Common, Core, and Contextual, which indicate the relevance and potential reusability of each entity and property within the KG. This categorization not only shapes the KG for this specific purpose but also enhances its reusability and interoperability for broader applications.

The Focus Categories are defined as follows:

- Common: Entities and properties broadly applicable to the purpose and potentially reusable across multiple contexts.
- Core: Entities and properties essential to the purpose and central to addressing the specified CQs, though possibly more domain-specific.

- Contextual: Highly specific entities and properties that address niche or unique aspects of the purpose, likely less reusable outside this context.

Scenarios	Personas	CQs	Entities	Properties	Focus
all	all	CQ2, CQ4, CQ6, CQ7, CQ9, CQ11, CQ20, CQ21, CQ22	Location	location_id, name, latitude, longitude, altitude, type	Common
S1, S2, S4	P1, P2, P4	CQ1, CQ2, CQ7, CQ13, CQ14, CQ15, CQ16, CQ17	Weather Station	station_id, location_id, name, type	Core
all	all	all	Measurement	measurement_id, station_id, timestamp	Core
S1, S2, S4	P1, P2, P4	CQ3, CQ7, CQ14, CQ17	Air Quality	air_quality_id, value, timestamp	Core
S1, S4	P1, P4	CQ1, CQ2, CQ15	Pollution	pollution_id, station_id, polluting, value, measurement_unit, timestamp	Core
S1, S2, S3	P1, P2, P3	CQ3, CQ6, CQ8, CQ10	Waste	waste_id, location_id, differentiated, undifferentiated, timestamp	Core
S1, S2, S3, S4, S5	P1, P2, P3, P4, P5	CQ4, CQ5, CQ9, CQ11, CQ12, CQ15, CQ17, CQ18, CQ19, CQ20, CQ21, CQ22	Consumption	consumption_id, value, timestamp	Core
S2, S3, S4	P2, P3, P4	CQ8, CQ11, CQ12, CQ13, CQ16	Weather	weather_id, station_id, temperature_min, temperature_max, temperature_avg, rainfall, rain_mm, snowfall, humidity, visibility_avg, wind_avg, timestamp	Common
S2, S3, S4, S5	P2, P3, P4, P5	CQ5, CQ10, CQ14, CQ16, CQ19	Demographic	demographic_id	Common

Table 1: Entity Types (ETypes) and Properties

## 2.8 ER modeling

The competency questions are used to design the ER model, ensuring that entities have the right properties to analyze climate change data in the right way.

This model maintains historical data in various environmental measurements to support tracking climate change and understanding its impact within Trentino.

At the top level, the Location entity represents the geographic area associated with each measurement.

Measurement is the super-type for data measurement like air quality level, pollution levels, temperature, and so on.

The Weather Station entity links each station to its Location, with additional details like altitude and station type, specifying whether it measures weather, air quality, or pollution.

The model allows measuring climate data across time and spatial domains.

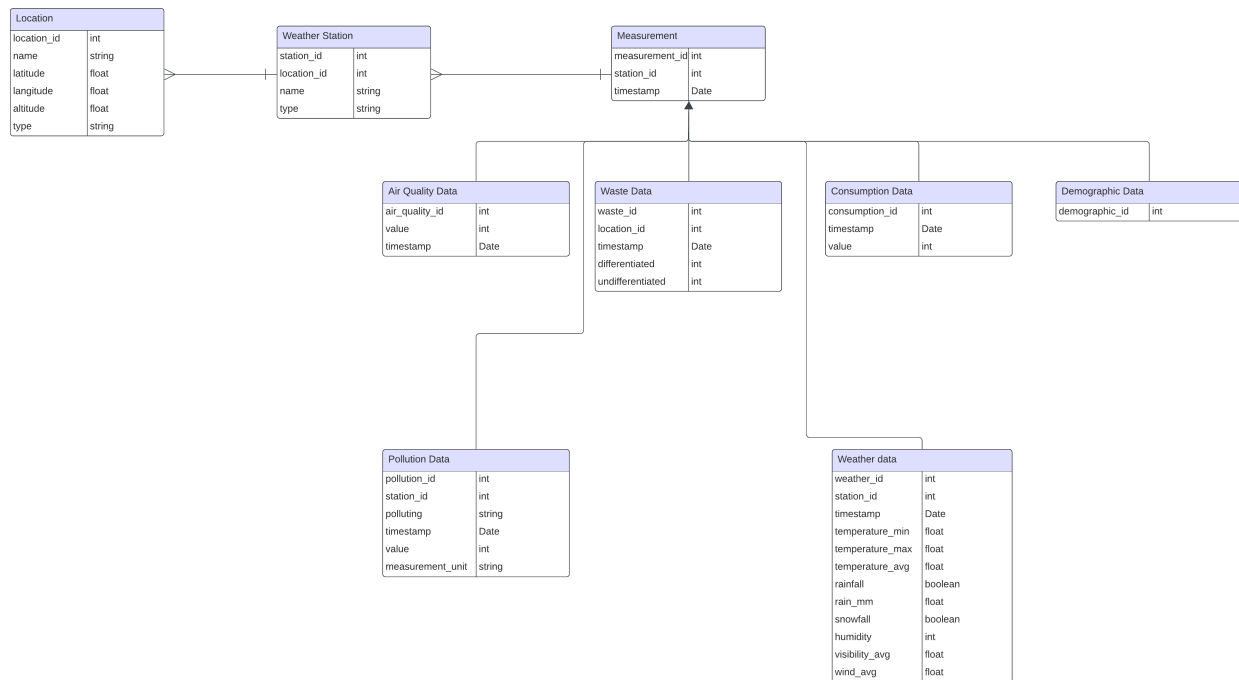


Figure 1: ER diagram

## 3 Information Gathering

In this section the second main input for the project is described, namely the data source list (if available). The resources (language, schema and data values) available as input for projects, has to be properly described. More in details for each resource has to be reported:

- 
- The name, and the description of the information the resource is carrying.
  - Type of resource. If it is a language, schema or data value dataset.
  - The source from which such resource can be collected.
  - If the resources is diversity-aware (thus already produced by iTelos) or needs to be improved in terms of diversity (i.e., data coming from low quality sources).

Moreover, this section aims at reporting the execution of the activities involved in the Information Gathering iTelos phase.

Information Gathering sub activities:

- Sources identification
- Datasets collection
- Datasets cleaning
- Datasets standardisation

The report of the work done during the first phase of the methodology, has to includes also the description of the different choices made, with their strong and weak points. In other words the report should provide to the reader, a clear description of the reasoning conducted by all the different team members.

## 4 Language Definition

This section is dedicated to the description of the Language Definition phase. Like in the previous section, it aims to describe the different sub activities performed by all the team members, as well as the phase outcomes produced.

Language Definition sub activities:

- Concept identification
- Dataset filtering

The report of the work done during this phase of the methodology, has to includes also the description of the different choices made, with their strong and weak points. In other words the report should provide to the reader, a clear description of the reasoning conducted by all the different team members.

---

## 5 Knowledge Definition

This section is dedicated to the description of the Knowledge Definition phase. Like in the previous section, it aims to describe the different sub activities performed by all the team members, as well as the phase outcomes produced.

Knowledge Definition sub activities:

- KTelos
  - Teleology definition
  - Teleontology definition
- Dataset cleaning and formatting

The report of the work done during this phase of the methodology, has to includes also the description of the different choices made, with their strong and weak points. In other words the report should provide to the reader, a clear description of the reasoning conducted by all the different team members.

## 6 Entity Definition

This section is dedicated to the description of the Entity Definition phase. Like in the previous section, it aims to describe the different sub activities performed by all the team members, as well as the phase outcomes produced.

Entity Definition sub activities:

- Entity matching
- Entity identification
- Data mapping

The report of the work done during this phase of the methodology, has to includes also the description of the different choices made, with their strong and weak points. In other words the report should provide to the reader, a clear description of the reasoning conducted by all the different team members.

## 7 Evaluation

This section aims at describing the evaluation performed at the end of the whole process over the final outcome of the iTelos methodology. More in details, this section as to report:

- the final Knowledge Graph information statistics (like, number of etypes and properties, number of entities for each etype, and so on).

- 
- Knowledge layer evaluation: the results of the application of the evaluation metrics applied over the knowledge layer of the final KG.
  - Data layer evaluation: the results of the application of the evaluation metrics applied over the data layer of the final KG.
  - Query execution: the description of the competency queries executed over the final KG in order to test the suitability of the KG to satisfy the project purpose.

## 8 Metadata Definition

In this section the report collects the definitions of all the metadata defined for the different resources produced along the whole process. The metadata defined in this phase describes both the final outcome of the project, and the intermediate outcome of each phase (language, schema, and data source standardised values).

The definition of the metadata, is crucial to enable the distribution (sharing) of the resource produced, through the data catalogs. For this reason it is important to describe also where such metadata will be published to distribute the resources it describes (for example the DataScientia catalogs).

In particular the structure of this section is organized as follows, with the objective to describe the metadata relative to all the type of resources produced by the project.

- Project metadata description
- Language resources metadata description
- Knowledge resources metadata description
- Data resources metadata description

## 9 Open Issues

This section concludes the current document with final conclusions regarding the quality of the process and final outcome, and the description of the issues that (for lack of time or any other cause) remained open.

- Did the project respect the scheduling expected in the beginning ?
- Are the final results able to satisfy the initial Purpose ?
  - If no, or not entirely, why ? which parts of the Purpose have not been covered ?

---

Moreover, this section aims to summarize the most relevant issues/problems remained open along the iTelos process. The description of open issues has to provide a clear explanation about the problems, the approaches adopted while trying to solve them and, eventually, any proposed solution that has not been applied.

- which are the issues remained open at the end of the project ?