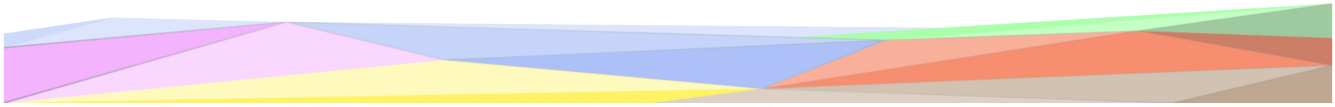
UNIVERSITY
OF TRENTO - Italy

Dipartimento di Ingegneria e Scienza dell'Informazione

– KnowDive Group –

KGE 2023 - Trentino Territory and Transportation

Document Data:	Reference Persons:
October 16, 2023	Rubens Rissi Onzi, Ferrari Eugenio



Index:

Revision History:

Revision	Date	Author	Description of Changes
0.1	October 16, 2023	Rubens Rissi Onzi, Ferrari Eugenio	Phase 1 - 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 sub processes) developed, provides a clear understanding of the project, thus serving such an information to external readers for the future exploitations of the project's outcomes.

This project aims at producing useful data for applications that intend to tell it's users about possible delays using the data we provide. The data will feature details mainly about bus stops locations using it's territory data. With that benefiting Trento inhabitants of urban areas.
The current document has the objective 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 its domain of interest.
- Section 3: Description of the project development, based on the two main sub process considered by iTelos: producer and consumer.
- Sections 4: Purpose formalization step of the iTelos methodology. This step aims to formalize the purpose, by extracting the functional requirement.

2 Purpose and Domain of Interest (DoI)

2.1 Project Purpose

The goal of this project is to provide data to applications and services that needs information suitable for predicting bus delays. To do so, we want to incorporate data that could affect transportation delays in urban areas, more in detail we will consider facilities, population density, number of right and left turns and number of traffic signs [1]. The data shall be integrated in knowledge graphs (KGs), using as base bus transportation and territorial data available. This release is intended to enable applications to use our KGs, created by the available data, to predict possible transportation delays in Trento, city of Italy.

2.2 Domain of Interest

The project will focus on bus transportation and will use data available from September until June 2023 of bus transportation in Trento urban areas, territorial and facilities data up to February 2023 from Trentino OSM Places and population data from 2018.

[1] Relevant features to predict delays: `https://journals.sagepub.com/doi/abs/10.3141/1666-12`

3 Project Development

For this project, we intend to integrate and produce data for applications aimed at informing users about possible delays using our designed data.

3.1 Data Production

To pursuit our project purpose we want data about Trento urban area territory, transportation and population. In details, we need to collect and create data about the bus stops, bus lines and their timetables. For each bus stop we also want data about the number of facilities and amenities located in that area as well as the population density for that area of the Trento city and the bus line path (number of turns and traffic signs).

3.2 Data Composition

To fulfil the purpose of a possible application, the data will need to be aggregated in a way to facilitate the inference about delays. The data will be gathered by known resources as *Trentino* OSM places, KGE22 - *Trentino* Urban Transportation and *Comune di Trento - "Dati statistici nelle Circoscrizioni di Trento"*.

4 Purpose Formalization

The Purpose of the KG knowledge layers can be formalized by the scenarios, personas and Competency questions. For this Project they are as it follows:

4.1 Scenarios

1. A day in Trento on a weekday.
2. A day in Trento on a weekend.
3. A day in Trento on rush hours.
4. A day in Trento on nighttime.

4.2 Personas

1. Giovanni, 19, is a college student that lives in the city centre, even though he studies far, in Mesiano.
2. Isabella, 83, is a senior citizen and lives in the outskirts of Trento, Cassoti di Povo, she often goes with her husband for groceries in weekdays.
3. Lily, 21, is a waiting staff worker in a hotel in the city center, she lives with roommates in a flat in Madonna Bianca. She also likes to go to parties and events in the city.

4.3 Competency Questions

1. Isabella, after lunch, wants to reach the city center, where she can find lot of shops to buy a gift to her daughter. She wants to know how much time it is gonna take to reach the center, and arrive home for dinner.
2. On Tuesday, Giovanni ends his lectures at 19:30 in Mesiano. He's curious about his arrival time at the city center, considering that 19:30 falls during rush hour when many people are heading home from work.
3. Isabella wants to visit her daughter in the weekend. As her daughter is available only on Sunday mornings, she have to be aware at witch time she should take the bus, as in the weekend there aren't many available in her part of the city. Her daughter also lives in the other side of the city, Gardolo. She also needs to change bus in between.
4. Lily don't want to go too early to work, but as she works sometimes in the night shift starting at 18:30, or in the day shift starting at 7:00. She usually have to go early to not get late. The bus that she rides in the afternoon is always full, by the time that she needs to leave. The hotel she works is well placed, having many amenities like museum and attractions nearby.
5. After enjoying a dinner with his friends, Giovanni decide to head to one of his friend's houses in Martignano. They are fortunate to be right on schedule for the last bus. The buses to Martignano are usually punctual, as the area had fewer residents compared to the city center, resulting in less traffic. They want to confirm if the bus is running on time so that they can catch it.

4.4 Concepts Identification and Categorization

From the scenarios, personas and CQs we extract the following entities with their properties:

Scenarios	Personas	CQs	Entities	Properties	Focus	Popularity
1-2-3	2	1	bus_stop	id: int, name: string, coordinates: string, schedule: glob	Contextual	Common
1-2-3	2	1	bus_line	id: int, name: string, length: float, left_turns: int, right_turns: int, traffic_signals: int	Contextual	Common
1-2-3	2	1	facility	id: int, name: string, coordinates: string, type: enum	Common	Common
1-3	1	2	bus_stop	id: int, name: string, coordinates: string, schedule: glob	Contextual	Common
1-3	1	2	bus_line	id: int, name: string, length: float, left_turns: int, right_turns: int, traffic_signals: int	Contextual	Common
1-3	1	2	facility	id: int, name: string, coordinates: string, type: enum	Common	Common
2	2	3	bus_stop	id: int, name: string, coordinates: string, schedule: glob	Contextual	Common
2	2	3	bus_line	id: int, name: string, length: float, left_turns: int, right_turns: int, traffic_signals: int	Contextual	Common
2	2	3	city_area	id: int, name: string, coordinates: string, population_density: float	Core	Core
4	3	4	bus_stop	id: int, name: string, coordinates: string, schedule: glob	Contextual	Common
4	3	4	bus_line	id: int, name: string, length: float, left_turns: int, right_turns: int, traffic_signals: int	Contextual	Common
4	3	4	facility	id: int, name: string, coordinates: string, type: enum	Common	Common
4	1	5	bus_stop	id: int, name: string, coordinates: string, schedule: glob	Contextual	Common
4	1	5	bus_line	id: int, name: string, length: float, left_turns: int, right_turns: int, traffic_signals: int	Contextual	Common
4	1	5	city_area	id: int, name: string, coordinates: string, population_density: float	Core	Core

4.5 ER Modeling

Given the entities and property identified in the step above, we can design the purpose ER model as in Figure 1.

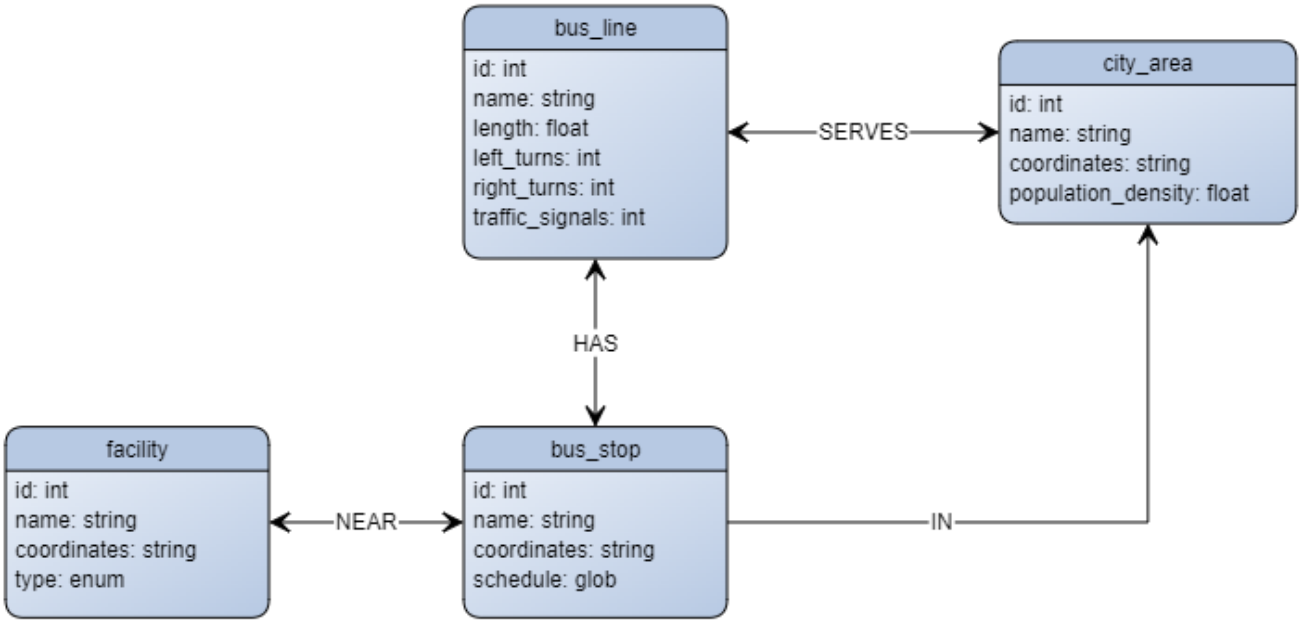


Figure 1: ER model

In our ER model we have 4 entities. Primarily, we created ”bus_line” and ”bus_stop” that represent the information about the bus transportation in Trento, city of Italy. We decided to include ”city_area” as an entity because it allows us to model population density for the Trento urban areas, since we don’t have more detailed data specific to each bus stop. We also introduced ”facility” as an entity because it contributes to understanding factors that could lead to delays.