Dipartimento di Ingegneria e Scienza dell'Informazione

– KnowDive Group –

KGE 2024 Project
Sport Facilities & Events in Trentino

Index:

Revision History:

Revision	Date	Author	Description of Changes
0.1	October 16, 2024	Christian Sassi, Pietro Bologna	Document created
1.1	October 25, 2024	Christian Sassi, Pietro Bologna	Purpose definition phase
1.2	October 29, 2024	Christian Sassi, Pietro Bologna	Revision of purpose definition phase

1   Introduction

Access to sports facilities and related events is essential for enhancing the quality of life in modern cities and regions. In areas like Trentino Province, providing convenient access to a range of sports infrastructure and events is increasingly important for both residents and visitors. Promoting an active lifestyle and supporting community-driven initiatives are key to strengthening the livability of the region.

To address these needs, we present a project for the Knowledge Graph Engineering course, integrating data on sports facilities and events in Trentino. This Knowledge Graph will offer citizens, tourists, and local authorities a comprehensive, interconnected view of available sports facilities, and related sport events. By combining data on facilities and events, the Knowledge Graph will empower users to make informed decisions that enhance community participation, public health, and the region's sporting culture.

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.

You can access the GitHub repository, which contains all the materials used during the project's development, via this link.

2 Purpose Definition

Access to sports facilities and events has a significant impact on the quality of life in regions like Trentino Province. Reliable information on the availability and accessibility of these resources can inspire a more active lifestyle and enhance community engagement. The **purpose** of this project is to develop a Knowledge Graph that consolidates information on sports facilities and events in Trentino, creating a unified resource for residents, tourists, and local authorities. By providing a comprehensive overview, the Knowledge Graph supports informed decision-making, promotes community involvement, enhances public health, and cultivates a vibrant sports culture in the region.

2.1 Informal Purpose

We want to build a Knowledge Graph that brings together all the information on sports facilities and events across Trentino. The goal is to make it easy for people to find sports venues, check out upcoming events, and get involved in physical activities. By offering details on what facilities are available, when events are happening, and what kinds of sports are offered, this project aims to promote an active lifestyle. Ultimately, we want this Knowledge Graph to be a go-to resource for anyone looking to make informed decisions about sports and activities in the region.
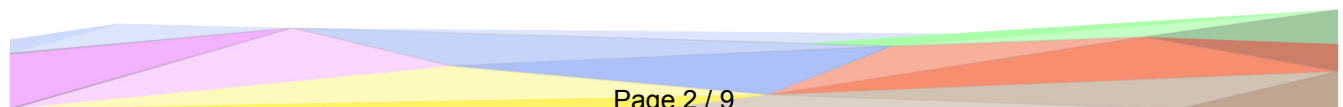
2.2 Domain of Interest (DoI)

After analyzing the purpose, the next step is to define the Domain of Interest (DoI). The DoI provides details about the geographical area and time frame relevant to the project purpose. The Domain of Interest for this project is as follows:

- The **geographical space** for this project is defined by the administrative boundaries of the Trentino Province. We ensure that only sports facilities and centers located within this region are included in our dataset, encompassing both urban and rural locations where sports facilities, such as soccer fields, tennis courts, basketball courts, and other venues, are situated. Additionally, this geographical space applies to the sports events data, ensuring that all events included are those taking place within the Trentino Province. Many of these events are based on those organized for the Festival dello Sport 2024, which brings a concentrated focus on sports culture and activities within one of the main city of the region.

- The **temporal domain** for this project is focused on the year 2024. The data on sports facilities and events reflects the information available for this specific period, capturing the current landscape of sports resources and scheduled events within Trentino Province throughout the year.

2.3 Scenarios definition

Here we define a set of usage scenarios, describing the multiple aspects considered by the project purpose. Four key scenarios are considered:

1. **Weekday**: Across the Trentino, on a typical weekday.

2. **Weekend**: In the province of Trento, on a weekend, when sports facilities may see an higher activity as locals and tourists alike participate in and spectate at community sports events.

3. **Holidays**: In Trentino, during the holiday periods (e.g. Christmas, Summer, etc.), when a high influx of tourists is expected.

4. **Festival dello Sport:** In Trento, during the *Festival dello Sport 2024*. During this time, a variety of sports events and exhibitions are scheduled, attracting both residents and visitors.

2.4   Personas

Now, we define a set of real users acting within the scenarios defined above. Each Persona is defined over a specific features included in the main purpose.

1. **Luca** is a 35-year-old engineer working in Trento. He is passionate about outdoor sports, especially padel.

2. **Anna** is a 21-year-old university student from Verona, studying in Trento. She enjoys playing volleyball with her friends.

3. **Matteo** is a 42-year-old tourist visiting Trento during the Christmas holidays. He has a passion for winter sports.

4. **Camilla** is a 24-years-old a volunteer student at the annual "Festival dello Sport" in Trento and assist visitors discovering the wide range of activities available throughout the city.

2.5   Competency Questions (CQs)

From the above information, here we have the list of Competency Questions (CQs) created considering the personas in the defined scenarios.

1. **CQ1**: Luca inquires about available padel courts in Trento after 7 PM.

2. **CQ2**: With uncertain weather, Luca wonders if there are any indoor padel courts available.

3. **CQ3**: While on a business trip to Levico Terme, Luca wants to know if there are any padel courts available after 6 PM.

4. **CQ4**: Luca also asks if there are any padel events during the weekend of the Festival dello Sport.

5. **CQ5**: Anna wants to know if there are volleyball courts available in Trento after her class at 3 PM.

6. **CQ6**: Spending the weekend with friends in Folgaria, Anna asks if there are any volleyball or beach volleyball courts available throughout the day.

7. **CQ7**: As a volleyball enthusiast, Anna would like to know if there will be any volleyball events during the summer of 2024.

8. **CQ8**: During the winter holidays, Matteo asks if there are any ski facilities within Stelvio National Park.

9. **CQ9**: Matteo also wants to know if there are any skiing events held in Stelvio National Park during the winter season.

10. **CQ10**: Not finding what he's looking for, Matteo asks if there are any skiing events during his vacation in Trentino.

11. **CQ11**: A visitor asks Camilla about the events happening today, October 10, at the Festival dello Sport.

12. **CQ12**: After attending the suggested climbing event, the visitor asks Camilla if there are any climbing gyms in Trento.

13. **CQ13**: While volunteering at the Festival dello Sport, Camilla becomes interested in tennis and wants to know if there are any tennis-related events and if tennis courts are available when she returns to Molveno for the weekend.

2.6 Concepts identification

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

Scenarios	Personas	Competency Questions	Entities	Properties	Focus
1	1	CQ1, CQ2, CQ3	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			SportFacility	id, name, location, open_hours, phone_number, email, website, price, sports, location	Core
			PadelCourt	id, name, covered	Contextual
2, 4	1	CQ4	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			Event	id, name, date, sports, location, organization, guests	Core
			Organization	id, name	Core
			Guest	id, name	Contextual
1	2	CQ5	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			SportFacility	id, name, location, open_hours, phone_number, email, website, price, sports, location	Core
			VolleyballCourt		Contextual
2	2	CQ6	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			SportFacility	id, name, location, open_hours, phone_number, email, website, price, sports, location	Core
			VolleyballCourt	id, name, surface, lit	Contextual
3	2	CQ7	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			Event	id, name, date, sports, location, organization, guests	Core
			Organization	id, name	Core
			Guest	id, name	Contextual
3	3	CQ8	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			SportFacility	id, name, location, open_hours, phone_number, email, website, price, sports, location	Core
			SkiArea	id, name	Contextual
3	3	CQ9, CQ10	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			Event	id, name, date, sports, location, organization, guests	Core
			Organization	id, name	Core
			Guest	id, name	Contextual
4	4	CQ11	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			Event	id, name, date, sports, location, organization, guests	Core
			Organization	id, name	Core
			Guest	id, name	Contextual
4	4	CQ12	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			SportFacility	id, name, location, open_hours, phone_number, email, website, price, sports, location	Core
			ClimbingArea	id, name, surface, lit, covered	Contextual
2, 4	4	CQ13	Region	id, name	Common
			Location	id, name, address, city, region	Core
			Sport	id, name	Core
			Event	id, name, date, sports, location, organization, guests	Core
			Organization	id, name	Core
			Guest	id, name	Core
			SportFacility	id, name, location, open_hours, phone_number, email, website, price, sports, location	Core
			TennisCourt	id, name	Contextual

Figure 1: Concepts identification

2.7 ER model definition

The ER diagram, which graphically represents the knowledge gathered in the prior stages. This ER diagram provides detailed information for a technician to delve deeper into the project. The diagram is based on the entity types (ETypes) and attributes identified before.
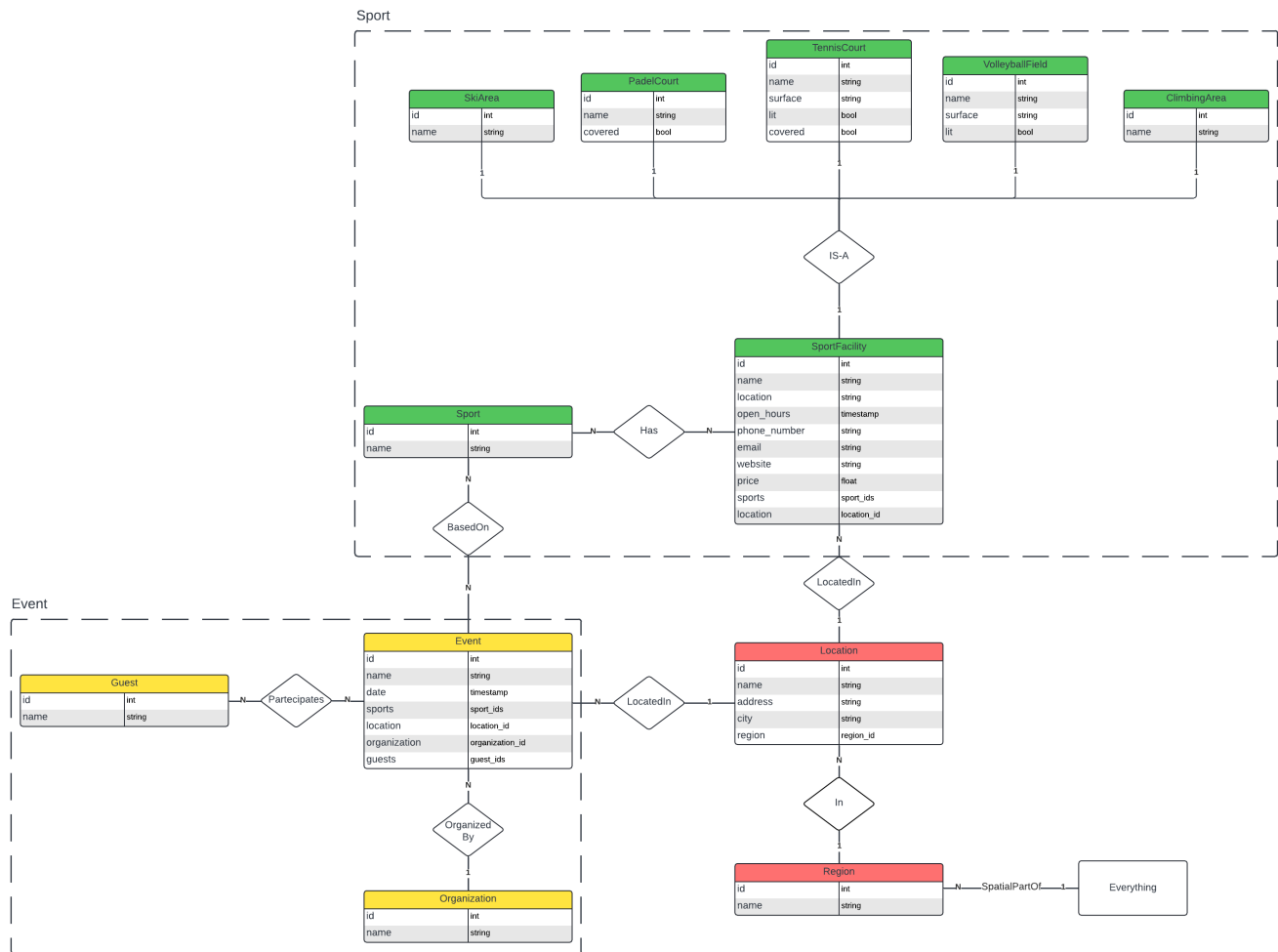
Figure 2: ER model

The following ETypes shown in Figure 2 have been identified to illustrate how the different entities interact within the model, providing a clear and coherent structure for representing information about sports facilities and events in Trentino.

1. **SportFacility**: Represents a facility dedicated to sports, where individuals can participate in various physical activities.

    - `id`: Unique identifier for each sports facility.
    - `name`: Name of the sports facility.
    - `location`: Address or general location information.
    - `open_hours`: Operating hours for the facility.
    - `phone_number`: Contact phone number for the facility.
    - `email`: Contact email address for the facility.
    - `website`: Website link for more information.
    - `price`: Cost to use the facility, if applicable.

- **sports**: List of sports available at the facility, linked to sport IDs.
- **location_id**: Foreign key linking to a specific location.

2. **SkiArea**: Represents a specialized area dedicated to skiing activities.

   - **id**: Unique identifier for the ski area.
   - **name**: Name of the ski area.

3. **PadelCourt**: A court dedicated to padel, a racquet sport.

   - **id**: Unique identifier for the padel court.
   - **name**: Name of the padel court.
   - **covered**: Indicates if the court is covered.

4. **TennisCourt**: A court dedicated to tennis.

   - **id**: Unique identifier for the tennis court.
   - **name**: Name of the tennis court.
   - **surface**: Type of surface on the court (e.g., clay, grass).
   - **lit**: Indicates if the court has lighting.
   - **covered**: Indicates if the court is covered.

5. **VolleyballField**: A field dedicated to volleyball.

   - **id**: Unique identifier for the volleyball field.
   - **name**: Name of the volleyball field.
   - **surface**: Type of surface on the field (e.g., sand, grass).
   - **lit**: Indicates if the field has lighting.

6. **ClimbingArea**: An area designated for climbing activities.

   - **id**: Unique identifier for the climbing area.
   - **name**: Name of the climbing area.

7. **Sport**: Represents a specific type of sport.

   - **id**: Unique identifier for each sport.
   - **name**: Name of the sport.

8. **Event**: Represents an organized sports event, including details on participation and scheduling.

   - **id**: Unique identifier for each event.
   - **name**: Name of the event.
   - **date**: Date and time when the event takes place.
   - **sports**: Sports included in the event, linked to sport IDs.

- **location**: Location where the event is held, linked to location ID.
- **organization**: Organization responsible for the event, linked to organization ID.
- **guests**: List of guest participants, linked to guest IDs.

9. **Guest**: Represents a guest or participant involved in a sports event.

   - **id**: Unique identifier for each guest.
   - **name**: Name of the guest.

10. **Organization**: Represents an organization responsible for hosting or coordinating sports events.

    - **id**: Unique identifier for each organization.
    - **name**: Name of the organization.

11. **Location**: Represents a physical location where sports facilities and events are held.

    - **id**: Unique identifier for each location.
    - **name**: Name of the location.
    - **address**: Full address of the location.
    - **city**: City in which the location is situated.
    - **region**: Region within which the location falls, linked to region ID.

12. **Region**: Represents a geographical area or administrative region that can contain multiple locations.

    - **id**: Unique identifier for each region.
    - **name**: Name of the region.

This ER model is also structured around several relationships:

1. **LocatedIn**: It contextualizes where the facilities and events are physically situated, allowing users to identify precise addresses or areas for sports facilities and event venues.

2. **Has**: It defines the kinds of sports activities or facilities available within each SportFacility, supporting users in finding facilities based on specific sports or equipment.

3. **IS-A**: It allows the model to recognize these entities as specific kinds of SportFacility, enabling queries about both general facilities and specific facility types, based on shared attributes.

4. **OrganizedBy**: It identifies the entity responsible for organizing an event, allowing users to find events hosted by specific organizations or understand the event's affiliation.

5. **Participates**: It enables user queries about the guest lists.

Due to the broad accessibility of the sports world, we decided not to focus solely on the university domain; instead, we envisioned this service as accessible to everyone in Trentino. Building on this idea, our goal became developing a service capable of helping citizens find relevant sports events and facilities throughout Trentino.

From a temporal perspective, we chose to support requests specific to the year 2024, while from a geographical perspective, we based our service on the entire Trentino region.

Both these temporal and geographical choices have their strengths and weaknesses. Regarding the temporal choice, the main limitation is that we are restricted to the year 2024, as sports events data from previous years is often incomplete or inconsistent, with some years missing data entirely. Looking forward, another limitation is that our service remains fixed to 2024, as we cannot guarantee that future data providers will address these critical gaps in coverage. By focusing on 2024, however, we ensure that our service operates reliably, leveraging a dataset we know to be comprehensive—this is a major strength of our approach.

As for the geographical choice, as outlined in our purpose definition, our service is designed to function across the entire Trentino region, making it one of the service's greatest strengths. However, it is important to note that for some entities, especially sports facilities, certain fields may be missing, such as contact information—particularly for less documented facilities. This also applies to specific types of facilities, such as volleyball fields, where some sports facilities may have incomplete data. Despite these gaps, we chose to model these properties because they align with the project's purpose: providing a complete and unified view of sports facilities and events in the region. By including these attributes in the model, we establish a structure that can accommodate future data as it becomes accessible. This approach ensures that the Knowledge Graph can be expanded and refined over time, allowing it to better serve the needs of residents, tourists, and local authorities seeking information about sports opportunities in Trentino.