



UNIVERSITY
OF TRENTO - Italy



Dipartimento di Ingegneria e Scienza dell'Informazione

– KnowDive Group –

KGE 2022 - Trentino Education

Document Data:

December 21, 2022

Reference Persons:

Samuele Bortolotti, Erich Robbi

© 2022 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
1.1	Background	1
1.2	Data Availability	2
1.3	Datasets overview	2
1.4	Project structure	3
1.5	Representation Diversity	3
1.5.1	Levels	3
1.6	Representation Diversity Architecture	4
2	Purpose and project's resources	6
2.1	Purpose	6
2.2	Personas	7
2.2.1	Vincente Trevisan	8
2.2.2	Azzurra Gentile	8
2.2.3	Michele Cafaro	9
2.2.4	Simona Tocci	9
2.2.5	Svend Kjærgaard	10
2.2.6	Germano Genovesi	10
2.2.7	Charlie Štěrbová	11
2.2.8	Hermann Schmidt	12
2.2.9	Ettore Candreva	13
2.2.10	Alvise Pisano	14
2.2.11	Giuditta D'Agostino	14
2.3	Scenarios	15
2.3.1	What school provides what I need	15
2.3.2	Last day of school	15
2.3.3	Dissatisfaction	16
2.3.4	Kindergarten for my child	16
2.3.5	Erasmus student	17
2.3.6	Best of education	17
2.3.7	Work and study	18
2.3.8	Education in Europe	18
2.3.9	New schools for Trentino	19
2.3.10	School Competition	19
2.4	Competency Questions	19
2.5	Resources	19
2.6	Organizations	20
2.6.1	Open Data Trentino	20
2.6.2	INVALSI - Servizio Statistico	21
2.6.3	Google Reviews	22

2.6.4	Istituto di Statistica della Provincia di Trento - ISPAT	23
2.6.5	Vivoscuola	23
2.6.6	Data Collection	24
2.7	Metadata	27
3	Inception	29
3.1	Purpose Formalization	29
3.1.1	Domain of Interest	29
3.2	Competency Questions	30
3.3	Resource Collection	36
3.3.1	Data Resources	36
3.3.2	school_list.csv	39
3.3.3	institute_list.csv	39
3.3.4	responsible_list.csv	40
3.3.5	school_courses.csv	40
3.3.6	institute_scrutini.csv	40
3.3.7	institute_ripetenti.csv	41
3.3.8	institute_dropouts.csv	41
3.3.9	school_scores.csv	41
3.3.10	school_reviews.csv	42
3.3.11	unitn_courses.csv	42
3.3.12	unitn_staff.csv	42
3.3.13	departments.csv	42
3.3.14	Knowledge Resources	42
3.4	Ontologies	43
3.5	Data Trimming	43
3.6	Entities defined	44
3.7	Schema modeling	46
3.8	Protégé	47
3.8.1	schools_list.csv	47
3.8.2	institute_list.csv	48
3.8.3	responsible_list.csv	48
3.8.4	school_courses.csv	49
3.8.5	institute_ripetenti.csv	49
3.8.6	unitn_staff.csv	50
3.9	Karma	50
3.9.1	Trentino Commune List.csv	50
3.9.2	courses_offered_in_trentino_school.csv	51
3.9.3	schools_list.csv	52
3.9.4	educational_institutions.csv	53
3.9.5	responsible_list.csv	54
3.9.6	school_courses.csv	54
3.9.7	institute_scrutini.csv	55

3.9.8	institute_scrutini.csv	55
3.9.9	institute_dropouts.csv	56
3.9.10	school_scores.csv	57
3.9.11	school_reviews.csv	57
3.9.12	unitn_courses.csv	58
3.9.13	unitn_staff.csv	58
3.10	Metadata	59
3.10.1	Municipality metadata	59
3.10.2	School Courses metadata	62
3.10.3	Institute metadata	67
3.10.4	Vivo metadata	72
3.10.5	Google Reviews metadata	75
4	Informal Modeling		79
4.1	Purpose Formalization	79
4.2	ER Model	80
4.2.1	Common Etypes	81
4.2.2	Core Etypes	82
4.2.3	Contextual Etypes	83
4.2.4	Teleology OWL	86
4.3	Metadata	93
5	Formal Modeling		101
5.1	ETG generation	101
5.1.1	Ontology Selection	101
5.1.2	Ontology - Teleology alignment	104
5.1.3	Teleontology Model	106
5.2	Schema Alignment	108
5.3	Language Alignment	111
5.4	ETG Model	114
5.5	Data Management	116
5.6	Metadata	118
6	Data Integration		125
6.1	Data Management	125
6.1.1	Entity Alignment	125
6.2	Entity Matching	126
6.3	Evaluation	148
6.4	Metadata	148
7	Outcome Exploitation		154
7.1	Knowledge Graph information	154
7.2	Model application	157
7.2.1	SPARQL	157

7.2.2 Project conclusion and future work	203
8 Issues	207
Bibliography	208

Revision History:

Revision	Date	Author	Description of Changes
0.1	20.04.2020	Fausto Giunchiglia	Document created
0.2	24.10.2022	Samuele Bortolotti, Erich Robbi	Add Introduction and background
0.3	25.10.2022	Samuele Bortolotti, Erich Robbi	Purpose formalization
0.4	29.10.2022	Samuele Bortolotti, Erich Robbi	Scenarios definition
0.4	30.10.2022	Samuele Bortolotti, Erich Robbi	Personas definition
0.5	2.11.2022	Samuele Bortolotti, Erich Robbi	First definitions of the CQs
0.6	9.11.2022	Samuele Bortolotti, Erich Robbi	Correction of CQs
0.7	12.11.2022	Samuele Bortolotti, Erich Robbi	Finished refinement of the scrapers and the scraped data
0.8	14.11.2022	Samuele Bortolotti, Erich Robbi	Finished first draft of inception phase
0.9	20.11.2022	Samuele Bortolotti, Erich Robbi	Typos correction
0.10	21.11.2022	Samuele Bortolotti, Erich Robbi	Start Informal Modeling
0.11	22.11.2022	Samuele Bortolotti, Erich Robbi	ER model correction
0.12	23.11.2022	Samuele Bortolotti, Erich Robbi	Metadata definition
0.13	24.11.2022	Samuele Bortolotti, Erich Robbi	Final revision before the deadline
0.14	26.11.2022	Samuele Bortolotti, Erich Robbi	Choice of the reference ontologies
0.15	28.11.2022	Samuele Bortolotti, Erich Robbi	Schema alignment
0.16	1.12.2022	Samuele Bortolotti, Erich Robbi	ETG creation
0.17	2.12.2022	Samuele Bortolotti, Erich Robbi	Work on transformers for sentiment analysis
0.18	3.12.2022	Samuele Bortolotti, Erich Robbi	Formal modeling metadata
0.19	4.12.2022	Samuele Bortolotti, Erich Robbi	Translating the selected data properties
0.20	5.12.2022	Samuele Bortolotti, Erich Robbi	Final revision before the deadline
0.21	6.12.2022	Samuele Bortolotti, Erich Robbi	Translating the selected data properties
0.22	10.11.2022	Samuele Bortolotti, Erich Robbi	Final revision before the deadline
0.23	11.11.2022	Samuele Bortolotti, Erich Robbi	Data Integration
0.24	12.11.2022	Samuele Bortolotti, Erich Robbi	Entity alignment
0.25	15.11.2022	Samuele Bortolotti, Erich Robbi	Entity matching
0.26	16.11.2022	Samuele Bortolotti, Erich Robbi	Outcome exploitation
0.27	17.11.2022	Samuele Bortolotti, Erich Robbi	Open Issues
0.28	18.11.2022	Samuele Bortolotti, Erich Robbi	ETG evaluation

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 in order to enhance the reusability of the resources handled and produced during the process. The current document aims to provide a detailed report of the KGE project developed following the iTelos methodology.

1.1 Background

Through the employment of new techniques and equipment, companies are producing an unprecedented volume of data. Thanks to such resources, it is possible to resolve complicated issues by gaining access and reusing data appropriately adjusted for secondary needs.

However, the available data is frequently not easily accessible, and even when it is, it cannot be used outside of the purpose for which it was produced.

Indeed, among the resources, we can identify three categories that are strictly connected to the relevance for the purpose itself: Core, Common, and Contextual resources [4].

- **Common resources** are not directly relevant to the user's purpose but are fundamental to support it since they include knowledge that is shared across several domains of interest.
- **Core resources** do not convey knowledge on the most crucial aspects related to the purpose; however, without this knowledge, the service cannot be delivered.
- **Contextual resources** include particular, sometimes exclusive data pertaining to the purpose and for this reason, they are the less reusable ones.

iTelos [5] is a general-purpose methodology that aims at providing a possible solution to the resources re-usability problem by developing independently the data level and the schema level of the resources. To fulfill its objective, *iTelos* uses *ontologies* in order to set a standard schema for the whole application, and *Knowledge Graphs*, graph-structured data models to integrate data.

To support our work in re-usability, we have relied on the *FAIR* [16] guiding principles. FAIR aim to make data *Findable*: data should be easy to find for both humans and computers (meta-data); *Accessible*: once the data is found the user has to know how to access it; *Interoperable*: the data need to interoperate with applications or workflows for analysis, storage, and processing; and finally *Reusable*: there should be well-described metadata in order to replicate data.

The goal of this work is to provide a service that will make it easier to locate schools in the Trentino area, together with information about those schools and the courses they offer based on criteria such as municipality, school type, course length, and instructional activity schedules.



1.2 Data Availability

The project source code can be found on the respective author's GitHub page¹, and all data which has been used can be downloaded from the Open Data Trentino web portal², Wikipedia³, Invalsi Servizio Statistico Cineca⁴, Istituto di Statistica della Provincia di Trento⁵, Vivoscuola⁶ and Digital University⁷.

1.3 Datasets overview

We used eleven data sets for the knowledge graph's development, namely:

- **Municipalities of Trentino** This resource represents a list of municipalities of the autonomous province of Trento, Italy. In particular, it shows the *ISTAT code*⁸, the name, the area, and the estimated population in 2001.
- **Nidi d'Infanzia, Scuole d'Infanzia, Scuole Elementari, Scuole Medie superiori, Scuole Medie inferiori** These resources contains the number of facilities, respectively of Kindergartens, Elementary Schools, Middle and High Schools, which are present in each municipality of the autonomous province of Trento.
- **Insegnamenti e corsi di studio dell'Università di Trento** This dataset provides information about all the courses the University of Trento provides.
- **Istituzioni scolastiche del Trentino** The dataset contains information about all educational facilities presents in Trentino. The relevant information which are listed are the address, the name and the type.
- **Corsi di studio scuole Trentine** The resource concerns the courses each school in Trentino offers.
- **Invalsi Servizio Statistico Cineca** This dataset contains the average score of the *Invalsi* tests⁹ per municipality. Their goal is to assess students' learning levels of Italian, Math, and English at critical points in the academic year.
- **Google Reviews** This resource concerns the reviews that other people leave on points of interest around the world. They can give pretty meaningful information regarding how people have felt about a given school facility, either concerning their education or the ones of others they know.
- **Istituto di Statistica della Provincia di Trento** This website contains all the resources which were collected and made available by the Statistical Institute of the Province of Trento

¹<https://github.com/samuelebortolotti/kge>

²<https://dati.trentino.it/>

³<https://www.wikipedia.org/>

⁴<https://invalsi-serviziostatistico.cineca.it/>

⁵<https://statweb.provincia.tn.it/indicatoriStrutturaliSubPro/>

⁶<https://www.vivoscuola.it/>

⁷<https://webapps.unitn.it/du>

⁸https://en.wikipedia.org/wiki/Italian_National_Institute_of_Statistics

⁹https://it.wikipedia.org/wiki/Prove_nazionali_INVALSI

(ISPAT). It provides several relevant information about Trentino and the cultural level of its inhabitants. In this project, we have employed this service for deepening our knowledge about the number of students per municipality and the information concerning the rate of completion of compulsory schooling per municipality.

- **Unitrento Digital University** Information on the students, faculty, staff, and administration of the University of Trento can be found at the Digital University portal.
- **Vivo Scuola** Vivoscuola provides an integrated variety of services, information, and curiosity about schools to students, teachers, and parents.

1.4 Project structure

The report is organized as follows:

- Section 2: The project purpose, the domain of interest, and the resources involved (both schema and data resources) in the integration process.
- Section 2: The input resources considered by the KGE project.
- Section 4, 5, 6: The integration process along the different iTelos phases, respectively.
- Section 7: How the result of the KGE process (the KG) can be exploited.
- Section 8: Conclusions and open issues summary.

1.5 Representation Diversity

Semantic Heterogeneity arises when there exist different representations of the same real-world phenomena. Every time knowledge graphs or databases are created by independent parties, semantic heterogeneity occurs.

In our definition of *Representation Diversity*, semantic heterogeneity is organized into four components: concepts, language, knowledge, and data [4].

1.5.1 Levels

Representation diversity occurs when:

- different concepts are used to denote the same entity.
- different terms and meanings are used in language.
- different entity types and properties are used.
- different entities and property values are used.

We categorize representation diversity into 4 levels:

1.5.1.1 Conceptual Diversity (L1)

The notion of concepts is well known in the Philosophy of Mind and in Computational Linguistics.

It is necessary to treat concepts as distinct language identities. Multiple hierarchies are used to arrange concepts in terms of "hypernym-hyponym" links. As an example: *Car* or *Vehicle*?

1.5.1.2 Language Diversity (L2)

Languages, in a very broad sense, includes: *Natural languages, Namespaces, and Formal languages*.

Polysemy and synonymy are linguistic phenomena that enable many representations of entities.

Words and concepts can be mapped in a *many to many* mapping in the same language as well as across languages. For instance: the Italian noun *Targa* can be mapped to *Nameplate*, its English counterpart.

1.5.1.3 Knowledge Diversity (L4)

Entity types, or Etypes, namely classes of entities with related properties, are used in order to model knowledge.

Knowledge diversity appears in two different forms, it is the result of the *many-to-many mapping* between Etypes and the properties used to characterize them. As an example: The same Etype *Car* is modeled in a different way, being connected to different groups of properties.

1.5.1.4 Data Diversity (L5)

The data is represented as a set entity, each of them associated with property values, inherited from the entity's Etype.

Data Diversity is the result of the *many-to-many mapping* between entities and the property values used to characterize them. For instance: The same car identified by "*FP372MK*", but with different values associated with the *velocity* property.

1.6 Representation Diversity Architecture

The architecture starts with the language representation layer (L2) and ends with it. L2 enforces the input and the output dependence of the representation of data and the user language. In fact, language is the key enabler of the bidirectional interaction between users and the platform.



In the first phase, the L2 input language is translated into the system-internal L1 conceptual language and the input language is only resumed during the last step when the results of the Data Integration steps are represented back to the user.

In this process, L2 is key in keeping completely distinct the multilingual user-defined data representation and the linguistic system-level data representation.

The management of conceptual diversity (L1) involves the organization of the L1 linguistic concepts, as identified in the first step, into a *Language Entity Graph* (LEG) which codifies the semantic relations across concepts (and, therefore, among, the corresponding L2 input words).

In order to achieve this goal we exploit, as a priori knowledge, a multilingual lexico-semantic resource, called *Universal Knowledge Core* (UKC).

The alignment of meanings across languages and namespaces absorbs a major source of heterogeneity present in the (Semantic) Web.

The net result of this phase is a LEG with the following properties:

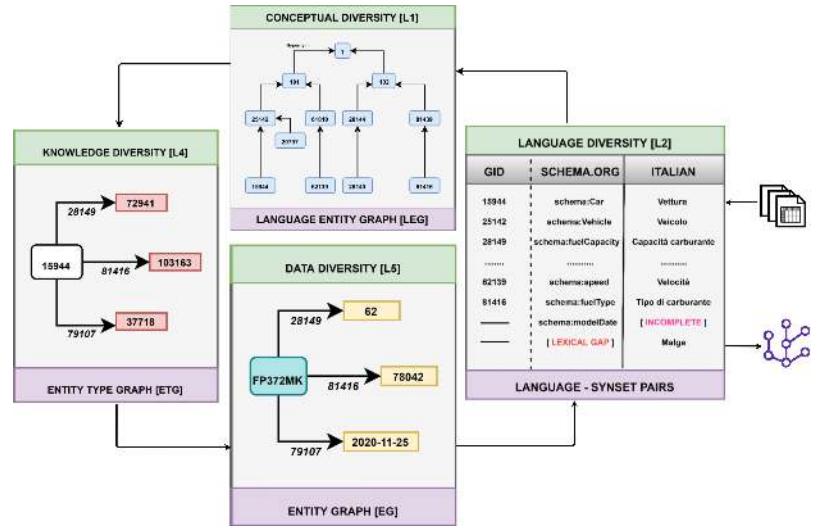
- Concepts identified during the first phase are all and only the nodes in this graph
- Nodes are annotated with the input L2 terms, across languages
- Nodes are organized into a hierarchy that preserves the ordering, across the links of the UKC (in the case of nouns, the *synonym/ hyponym/ hypernym* relations)

Managing knowledge diversity (L4) involves the construction of a (linguistic) *Entity Type Graph* (ETG) encoded using only concepts occurring in the LEG constructed during the previous two phases. In this phase, the first step is to distinguish concepts into Etype(s) and properties (both object properties and datatype properties) while the second step is to organize them into a *subsumption hierarchy*.

In the fourth representation layer (L5), we tackle data diversity via an *Entity Graph* (EG), namely, a *data-level knowledge graph*, by populating the ETG with the entities extracted from the input datasets.

The EG is constituted of a backbone of L1 linguistic id(s), each annotated with the input L2 terms where, for each L2 term, the system remembers the dataset it comes from. This mechanism is implemented via a provenance mechanism that applies to all the input dataset elements, both at the schema and at the data level.

One major advantage of our approach is that the combinatorial explosion deriving from the interaction of the four different types of diversity is avoided and the complexity of the data integration problem reduces to the sum of the complexity of each layer [4].



A schematic illustration of the *Representation Diversity Architecture*

2 Purpose and project's resources

The formalization of the project's objective and the domain is the first and most crucial step in the data integration process, and clear of the entire *iTelos* methodology.

The clarity of the project objective is a key factor in determining how the data integration process turns out. To formulate a purpose, we have defined the project's domain of interest, personas, and scenarios.

This section will discuss the project's purpose, together with the definition of the domain, the personas, and the project scenario: the knowledge resources, the data resources, and the metadata.

2.1 Purpose

The initial, and thus informal, purpose which has been later formalized, is shown below:

"A service that will facilitate the finding of schools, including details about the school and courses offered, in the region of Trentino based on city, commune, school type, course duration, and teaching activities schedules."

From here, we can understand that the project's scope (**Domain of Interest**) includes information regarding the school facilities in Trentino as well as the municipalities of the autonomous region of Trento. Moreover, the courses offered by each educational facility with their associated duration should be included.

This research seeks to provide a complete analysis of the educational facilities in Trentino which comprehend geospatial and temporal domains.

The Trentino area has a wide range of educational institutions, including public and private kindergartens, elementary schools, secondary schools, universities, and other vocational training facilities. The information regarding the school facilities in Trentino as well as the municipalities of the autonomous region of Trento is part of the project domain. Moreover, the courses offered by each educational facility with their associated duration are included.

Initially, the data we managed to collect ranged between 2014 and 2022. We had extremely updated information concerning university courses and general institutional information, and outdated material concerning schools' courses. However, with additional data scraping, we managed to obtain more recent data. We claim that the time boundary is restricted to the current few years (2020 - 2022).

As a possible metric for the quality of the education provided by each municipality, we have decided to rely on the data provided by the *invalsi* test, the data from the schools' Google reviews, the rate of completion of compulsory schooling and the number of students who fail to pass the year per municipality.

The final goal is to develop a knowledge graph that acts as a collective memory of the information regarding the school facilities in Trentino and helps inhabitants or foreigners to find relevant information in the field of education.

The purpose is formalized by means of *Personas* and *Scenarios* described in the following sub-sections.

2.2 Personas

As stated at the beginning of the section, the personas analysis is crucial in order to understand the domain of interest of the project, as well as to formalize the project's purpose. Their main job is to represent users, their needs, their objectives, and their most important traits. If the personas are realistic, they operate as a proxy for the consumers and direct the design and development of the product or service.

In this work, personas are the possible actors which are supposed to exploit the final result of the knowledge graph engineering process. Therefore, the subsection lists different kind of users who acts with the system for different needs. Looking for schools in the region of Trentino is a common need of all types of users of the system, despite having different purposes which might lead to a different result for each of them.

Needs are a type of innate psychological nourishment that is necessary for a person's psychological development, integrity, and well-being [3]. They are directly related to well-being since unmet needs result in a decline in a person's degree of well-being.

The following needs were identified during the persona definition process:

- **Autonomy** [3] This need expresses the wish to take charge, the desire to control one's life, and the conviction that one has a choice.
- **Bridging** That is the need for, and benefit of, having a substantial social network.
- **Competence** This need is met when the person believes he can produce the desired outcomes and consequences through his activities.
- **Curiosity** This need is frequently manifested in the desire to learn and to apply one's skills effectively.
- **Security** [10] The need for security may be described as the desire to feel safe, not in danger, and unconcerned about one's own or others' safety.

The purposes are pleasing states of circumstances that people try to attain by controlling their behavior in different ways. The personas make choices and behave in a certain way based on their purposes.

When creating personas, there are often two different sorts of purposes: those that seek to achieve desired results and those that want to prevent undesirable ones. The first class of purposes applies to most, if not all, of our specified personas.

As a disclaimer: all names and characteristics of the individuals in this section are fictitious. No identification with actual persons (living or deceased), places, buildings, and products is intended or should be inferred. Moreover, the individuals' pictures are taken from *this-person-does-not-exist*¹⁰, a website in which people's faces are generated by the means of a *StyleGAN*.

¹⁰<https://this-person-does-not-exist.com/en>

2.2.1 Vincente Trevisan

2.2.1.1 Differentiator

Vincente is a 14 years old boy who lives in Margone, a small village in Trentino. He has just finished middle school and is interested in attending a high school that allows him to become a software engineer. Despite being a native of Trentino, he does not know which nearby high schools provide such a kind of education.

2.2.1.2 Technological Resources

Internet, Smartphone.

2.2.1.3 Needs

Autonomy, Curiosity.

2.2.1.4 Purposes

Vincente wants to find a school where he can learn to be a software developer.



Vincente Trevisan
Personal Information

ID	1
Age	14
Sex	♂
Occupation	Student
Driver	✗
Education	Middle School

2.2.2 Azzurra Gentile

2.2.2.1 Differentiator

Azzurra is a student who lives in Trento. She is completing middle school, where she has done exceptionally well in every subject throughout the year. She truly enjoyed all of the topics in middle school, but she does not have a clue when it comes time to choose which high school to attend. Therefore, she would like to consult a service that gives an overview of all the possible schools available in Trentino to find the right school for her.

2.2.2.2 Technological Resources

Internet, Smartphone.

2.2.2.3 Needs

Competence.

2.2.2.4 Purposes

Azzurra wants to examine different schools' courses and training programs in order to better understand the alternatives for which school to enroll in.



Azzurra Gentile
Personal Information

ID	2
Age	14
Sex	♀
Occupation	Student
Driver	✗
Education	Middle School

2.2.3 Michele Cafaro

2.2.3.1 Differentiator

Michele is a university student pursuing a bachelor's degree in biomedical engineering in Padova. He has started looking for alternative universities to transfer to since he is dissatisfied with the degree he is currently enrolled in. Michele is investigating the option of relocating to the University of Trento for a number of reasons. Therefore, he wonders if the University of Trento would accept the exams he had passed at his own university.

2.2.3.2 Technological Resources

Internet, PC, Smartphone.

2.2.3.3 Needs

Bridging, Autonomy, Curiosity, Competence.

2.2.3.4 Purposes

Michele wants to browse the courses offered by the University of Trento to see if the classes he has previously taken at the University of Padova are offered at the University of Trento, or are at least comparable.



Michele Cafaro

Personal Information

ID	3
Age	20
Sex	♂
Occupation	Student
Driver	✓
Education	High School

2.2.4 Simona Tocci

2.2.4.1 Differentiator

Simona lives in Pergine Valsugana. She is the mother of a child, and now that her child is old enough, she wants to take the child to a kindergarten because she has to work and therefore cannot leave the child at home alone. Since she has a car, she does not need the kindergarten to be strictly close to Pergine Valsugana.

2.2.4.2 Technological Resources

Internet, PC, Smartphone.

2.2.4.3 Needs

Security.

2.2.4.4 Purposes

Simona wants to find a kindergarten for her son.

She also wants to receive feedback from any previous customers, ensuring other parents that a potential kindergarten is a safe and healthy environment for kids.



Simona Tocci

Personal Information

ID	4
Age	34
Sex	♀
Occupation	Worker
Driver	✓
Education	University

2.2.5 Svend Kjærgaard

2.2.5.1 Differentiator

Svend is a graduate student of Artificial Intelligence Systems at Aarhus University, Denmark. He wants to take part in the Erasmus+ program to complete an academic internship, and the University of Trento is where he would like to do his research internship. The letter of acceptance from the hosting entity must be sent to the online form in order to apply for the European program. For this reason, Svend wants to ask UniTN staff members (office or professors) about any potential opportunities they may offer.

2.2.5.2 Technological Resources

Internet, PC, Smartphone.

2.2.5.3 Needs

Bridging, Competence, Curiosity.

2.2.5.4 Purposes

Svend wants to consult a list of professors who have research areas similar to the topics he learned during his master's in order to ask for an internship.



Svend Kjærgaard

Personal Information

ID	5
Age	22
Sex	♂
Occupation	Student
Driver	✓
Education	University

2.2.6 Germano Genovesi

2.2.6.1 Differentiator

Germano is the father of a son who has just finished elementary school, so now he has to enroll him in a middle school. Since he believes the middle school is crucial for a child's future and that his son is extremely hardworking and scholastically diligent, he would prefer to enroll him in a middle school where he is confident the students are engaged and succeed in school.



Germano Genovesi

Personal Information

ID	6
Age	41
Sex	♂
Occupation	Worker
Driver	✓
Education	High School

2.2.6.2 Technological Resources

Internet, PC.

2.2.6.3 Needs

Security, Competence.

2.2.6.4 Purposes

Germano wants to compare middle schools by quality index, which can be positive reviews and an average of national test results.



2.2.7 Charlie Štěrbová

2.2.7.1 Differentiator

Charlie is 18 years old Czech girl that has moved to Italy for completing her studies as she wants to become a tourist guide in the future. Charlie has always been resourceful and has always wanted to be fully autonomous, despite the fact that her parents are willing to pay for her expenses. Therefore, the first thing she did in Italy was to find a part-time job as a cashier in a supermarket near the city center, in order to provide for herself. She has to work four hours every day with variable shifts, and even if her work takes her a lot of time during the day, she does not want to quit her job since she goes along well with her colleagues.

Since she does not have a driving license and, due to her work shifts she cannot go too far from Trento, she has to find a school in Trento which provides courses that are compatible with her part-time job schedule.

2.2.7.2 Technological Resources

Internet, PC.

2.2.7.3 Needs

Autonomy, Competence.

2.2.7.4 Purposes

Charlie is looking for a school with schedules that work with her part-time employment and which is near Trento.



Charlie Štěrbová

Personal Information

ID	7
Age	18
Sex	♀
Occupation	Worker
Driver	X
Education	High School



2.2.8 Hermann Schmidt

2.2.8.1 Differentiator

Hermann is a 55-year-old man from Austria and he works as a senior researcher at Stanford University, United States. His research area is mostly focused on the importance of education and how to improve it during each stage of life. His latest work is carried out in collaboration with the European Schoolnet, which is an organization headquartered in Brussels, with the aim of innovating both education and learning in Europe. Therefore, as the first step for his work he has to collect basic information in regard of the educational activities that take place in all Europe, Trentino included. More specifically, he has to retrieve all details of the schools such as the school name, the country municipality, the courses that the school offers, and a reference contact in case of additional questions.

2.2.8.2 Technological Resources

Internet, PC.

2.2.8.3 Needs

Competence.

2.2.8.4 Purposes

Hermann wants to perform a statistical analysis of all the schools located in the Italian territory.



Hermann Schmidt

Personal Information

ID	8
Age	55
Sex	♂
Occupation	Worker
Driver	✓
Education	PhD



2.2.9 Ettore Candreva

2.2.9.1 Differentiator

Ettore is a 33-year-old man. He is an Italian who lives in Lavis in the autonomous province of Trento. He works in the municipal council and is the provincial councilor for education, university, and culture of the autonomous province of Trento. His job is to guarantee the regularity of school activities and the well-being of students and educational staff.

He, together with the council, has decided to build a new school in Trentino. In order to build an efficient service, they want to place the new facility in areas where there are few schools.

Therefore, to find the places where to build the new facilities, he requires a list of the municipalities which have the fewest schools in the autonomous province of Trento.

2.2.9.2 Technological Resources

Internet, PC.

2.2.9.3 Needs

Competence.

2.2.9.4 Purposes

Ettore needs to have the exact number of schools in a certain Trentino municipality.



Ettore Candreva

Personal Information

ID	9
Age	33
Sex	♂
Occupation	Worker
Driver	✓
Education	University



2.2.10 Alvise Pisano

2.2.10.1 Differentiator

To continue his uncle's construction company, Alvise requires a high school diploma in architecture. For the moment, he is working at a big company's warehouse. Due to his working schedule, he can only attend the lecture in the evening.

Although he wishes to work with his uncle as soon as possible, he is convinced that a good education is required to lead a business to success.

Alvise desire to know which schools provide the education he needs as well as timetables that are compatible with his ideal job. Moreover, he is interested in knowing which schools provide the best service in Trentino, regardless of their position as he owns a car.

2.2.10.2 Technological Resources

Internet, Tablet.

2.2.10.3 Needs

Curiosity, Autonomy.

2.2.10.4 Purposes

Alvise discovers an architectural class at a night school that allows him to continue to work.



Alvise Pisano

Personal Information

ID	10
Age	21
Sex	♂
Occupation	Worker
Driver	✓
Education	Middle School

2.2.11 Giuditta D'Agostino

2.2.11.1 Differentiator

Giuditta is a 61-year-old woman and she is a principal of a school in Trentino. She plans to organize a collaboration with one or more schools to have her students engaged in a collaborative project. To fulfill her objective, Giuditta requires some contact references in order to get in touch with other principals.

2.2.11.2 Technological Resources

Internet, Smartphone, PC.

2.2.11.3 Needs

Security, Competence.

2.2.11.4 Purposes

Giuditta locates schools to arrange meetings with by looking through the local schools' contact information.



Giuditta D'Agostino

Personal Information

ID	11
Age	61
Sex	♀
Occupation	Worker
Driver	✓
Education	University



2.3 Scenarios

A scenario depicts how a person may attempt to execute a task in narrative form, that is, in the form of a tale. In this project, we attempted to describe common events, accidents (i.e. times when things don't go as planned), positive experiences (i.e. times when things went well), and to mimic a typical or ideal scenario (how the consumer should behave).

As a disclaimer: all scenarios in this section are fictitious. No identification with actual persons (living or deceased), places, buildings, and products is intended or should be inferred. Moreover, the pictures are taken from <https://www.pexels.com/> under the Creative Commons Zero license.

2.3.1 What school provides what I need

2.3.1.1 Description

I am a student, and I am a bit worried because I do not know which high school to choose in order to pursue my dream career. I know that I want to stay close to home in Trentino, but I am not sure which high school would give me the education I need. My parents are supportive of my dreams and they decided to help me research the best high school options in the area.



End of the School

Information

Scenario ID 1

2.3.2 Last day of school

2.3.2.1 Description

I am a student in Trentino. I have done exceptionally well in every subject throughout the year, and I am not sure which school to attend. I would like to consult a service that gives an overview of all the possible schools available in Trentino in order to find the right school for me.



Last day of school

Information

Scenario ID 2



2.3.3 Dissatisfaction

2.3.3.1 Description

I want to change schools because I am not happy with my current school. My grades haven't been as good as I want them to be and I feel like I'm not receiving the best education I could get. I'm not sure if it's the school itself or if it's me, but I'm not feeling it. I think it would be a better fit for me to transfer to a different school and see if I can get better grades and have a better education.



**Dissatisfaction
Information**

Scenario ID 3

2.3.4 Kindergarten for my child

2.3.4.1 Description

My child is 3 years old and he is starting kindergarten this fall. I work hard and I don't always have time to take care of him, so I want him to be safe and happy. I think it would be best for him if he attends a kindergarten where the kids are all from similar backgrounds and where the teachers are experienced and know how to handle kids. I want him to socialize with other kids his age.



**Kindergarten
Information**

Scenario ID 4



2.3.5 Erasmus student

2.3.5.1 Description

I am a student at a foreign University. I have been interested in studying abroad for a while, and I have decided that I would like to take part in the Erasmus+ program to complete an academic internship at a foreign University in Europe. I found the University of Trento on the Erasmus+ website, and thus I would like to apply to the program.



**Erasmus student
Information**

Scenario ID 5

However, I am not sure if the University of Trento is the right place for me to do my research internship. I am not sure if the research opportunities are good, so I ask one of the professors at the University of Trento if he knows of any opportunities that I could take part in while I am there. The professor tells me that the University of Trento offers a variety of research internships, and I can look into those opportunities if I am interested.

2.3.6 Best of education

2.3.6.1 Description

I am a parent of a young child and I want my child to attend the best schools possible. I understand the importance of good education and I want to ensure that my child receives the best opportunities possible. I am able to afford to send my child to the best private schools in the area, but I am also open to the idea of sending my child to a public school. I want to make sure that my child has the best chance of achieving their goals and excelling in life.



**Best of education
Information**

Scenario ID 6



2.3.7 Work and study

2.3.7.1 Description

I am a student and I am looking for a school that would allow me not to quit my job. I am currently employed part-time and I cannot afford to quit my job to attend school full-time. I am looking for a school that has an online or evening program that would allow me to continue working while attending school.



Work and study

Information

Scenario ID 7

2.3.8 Education in Europe

2.3.8.1 Description

At an upcoming conference on education in Europe, I want to present my work to the public. I, therefore, need to collect data about education in order to complete my research.

In order to do so, I begin by tracking attendance at educational institutions across Europe. I find that, in many cases, there is a discrepancy between the level of education offered at these institutions and the level of proficiency of students. For example, in Norway, the level of education offered is high, but the students are quite inept in mathematics and science. In contrast, in Greece the level of education offered is high, but the students are quite proficient in mathematics and science.

In order to begin to identify the cause of this discrepancy, I decided to analyze the educational program of each country. Therefore I need to learn more about the educational program in Italy, including Trentino.



Education in Europe

Information

Scenario ID 8

2.3.9 New schools for Trentino

2.3.9.1 Description

I work at the municipal council, and I have just received money from the autonomous province of Trento. My job is to ensure students in Trentino get the right education.

To spend best the funds I want to find the best places where to allocate new schools. I began my search by visiting the municipalities and asking them for a list of the schools which have the fewest students. I am surprised that most of the municipalities do not even know how many schools they have in their jurisdiction.



New schools for Trentino
Information

Scenario ID 9

2.3.10 School Competition

2.3.10.1 Description

I am the principal of a school in Trentino. I plan to organize a collaboration with one or more schools to have my students engaged in a collaborative project. To fulfill my objective, I require some contact references in order to get in touch with other principals.



School Competition
Information

Scenario ID 10

2.4 Competency Questions

Competence questions are the functional requirements of the KGE project.

"These requirements, which we call competency questions, are the basis for a rigorous characterization of the problems that the enterprise model is able to solve". [6]

2.5 Resources

This project aims to answer competence queries by integrating knowledge and data. Additionally, the modeling step would not be complete without linguistic processing. Consequently, the cornerstone of this research is made up of linguistic, knowledge, and data resources.

2.5.0.1 Linguistic Resources

A linguistic resource is a dataset that offers information on languages, such as word relationships and meanings.

Two categories of monolingual and bilingual resources exist:



- Online Dictionaries;
- Wordnet-like resources. Since Wordnets connect the meanings of words in an LKG, they are significantly helpful in data integration [4].

2.5.0.2 Knowledge Resources

A dataset with a KB encoding information about schemas (Etype(s) and properties) is referred to as a knowledge resource.

High-quality KBs are typically referred to as *ontologies*. We will call them *teleologies* (ontologies with metadata that empower their practical use in knowledge and data integration) [4].

2.5.0.3 Data Resources

A dataset that contains data in some format represents a "data resource" (tabular, unstructured, entities, and property values) [4].

2.6 Organizations

2.6.1 Open Data Trentino

2.6.1.1 Description

The Trentino Open Data project was founded by the Autonomous Province of Trento with the involvement of Trentino Digitale, the Consortium of Municipalities, the Bruno Kessler Foundation, and the University of Trento.

The aim is to open its data to the world, not only as a support to the transparency of public service operations but above all as a contribution that the process of enhancing public information capital can lead to economic growth through the development of services. In this sense, the Trentino Open Data project is a tool for the benefit of companies and professionals interested in using the information assets of the Public Administration in new services and new applications subsidiary to the services of the Province.

The Trentino Data Catalog (<http://dati.trentino.it>), provided and managed by Trentino Digitale as a single point of access and exploitation of the Trentino system's open data, now has 6,306 datasets.

The datasets are also published systematically on the national portal <https://dati.gov.it/> and on the European one <https://data.europa.eu/en> [11].



Open Data Trentino

Information

Active:	2012 - Today
Resource Type	Data
Website	https://dati.trentino.it/
Last access	October 29, 2022
Validity	✓
License	CC Attribution v4.0



2.6.2 INVALSI - Servizio Statistico

2.6.2.1 Description

From the year 2021, the data produced by INVALSI have become open data. It is possible to access and download territorial census and sample data easily, without having to follow the procedure, as long as some particular principles linked to this mode of use are respected. On the dedicated page of the INVALSI institutional website, it is possible to download the analytical datasets of census surveys with various territorial details. In fact, the tracks and the complete dataset of the population at the provincial level are currently available, obtained through the census surveys that the INVALSI conducts annually [7].



INVALSI - Servizio Statistico

Information

Active:	2021 - Today
Resource Type	Data
Website	invalsi-serviziostatistico.cineca.it
Last access	October 29, 2022
Validity	✓
License	CC Attribution v4.0

2.6.3 Google Reviews

2.6.3.1 Description

Google Developers (previously Google Code) is Google's site for software development tools and platforms, application programming interfaces (APIs), and technical resources. The site contains documentation on using Google developer tools and APIs - including discussion groups and blogs for developers using Google's developer products.

There are APIs offered for almost all of Google's popular consumer products, like Google Maps, YouTube, Google Apps, and others.

The site also features a variety of developer products and tools built specifically for developers. Google App Engine is a hosting service for web apps. Project Hosting gives users version control for open-source code. Google Web Toolkit (GWT) allows developers to create Ajax applications in the Java programming language (All languages).

The site contains reference information for community-based developer products that Google is involved with like Android from the Open Handset Alliance and OpenSocial from the OpenSocial Foundation [14].



Google Developers

Information

2005 - Today

Data

<https://developers.google.com>

October 29, 2022



CC Attribution v3.0

Active:

Resource Type

Website

Last access

Validity

License



2.6.4 Istituto di Statistica della Provincia di Trento - ISPAT

2.6.4.1 Description

The Statistical Institute of the Province of Trento (ISPAT) promotes and coordinates the development and production of statistics of interest to the provincial administration, as part of the provincial statistical program. designs and develops the information system of statistical knowledge, as a portal for the official statistics of Trentino.

ISPAT collaborates with universities and research institutes in statistical research projects, with priority for the Trentino research system for higher education, also through the funding of grants and research grants in accordance with the provisions of current legislation on the subject [8].



ISPAT

Information

Active:	2014 - Today
Resource Type	Data
Website	statweb.provincia.tn.it
Last access	October 29, 2022
Validity	✓
License	CC Attribution v4.0

2.6.5 Vivoscuola

2.6.5.1 Description

Vivoscuola is the Trentino school site with the goal of supporting the ongoing improvement of the Trentino education system and the contact between schools and families [13].



Vivoscuola

Information

Active:	2009 - Today
Resource Type	Data
Website	vivoscuola.it
Last access	November 5, 2022
Validity	✓
License	CC Attribution v3.0

2.6.6 Data Collection

Municipalities of Trentino

Attributes	Description
<i>ISTAT code</i>	code used by the Italian administration to identify an Italian municipality
<i>italian</i>	Italian name of municipality
<i>german</i>	German name (if exists) of municipality
<i>ladin</i>	Ladin name (if exists) of municipality
<i>area</i>	area that the municipality covers on the territory in km2
<i>population</i>	total population residing under the municipality

Nidi d'Infanzia

Attributes	Description
<i>entity</i>	municipality name
<i>year₁</i>	number of kindergartens in the municipality in year ₁
	...
<i>year_n</i>	number of kindergartens in the municipality in year _n

Scuole d'Infanzia

Attributes	Description
<i>entity</i>	municipality name
<i>year₁</i>	number of pre-schools in the municipality in year ₁
	...
<i>year_n</i>	number of pre-schools in the municipality in year _n

Scuole Elementari

Attributes	Description
<i>entity</i>	municipality name
<i>year₁</i>	number of elementary schools in the municipality in year ₁
	...
<i>year_n</i>	number of elementary schools in the municipality in year _n

Scuole Medie e Superiori

Attributes	Description
<i>entity</i>	municipality name
<i>year₁</i>	number of highschools in the municipality in year ₁
	...
<i>year_n</i>	number of highschools in the municipality in year _n

Scuole Medie Inferiori

Attributes	Description
<i>entity</i>	municipality name
<i>year₁</i>	number of middle schools in the municipality in year ₁
	...
<i>year_n</i>	number of middle schools in the municipality in year _n



Insegnamenti e corsi di studio dell'Università di Trento

Attributes	Description
<i>nomeInsegnamento</i>	course name
<i>descrizioneInsegnamento</i>	course description
<i>corsoStudi</i>	degree course
<i>dipartimento</i>	the academic structure that offers to teach; both the name ("nome" field) and the identification code of the structure ("id" field) are provided
<i>sitoWeb</i>	teaching website within the ESSE3 system
<i>docenti</i>	list of the course's teachers; the name ("nome" field), the surname ("cognome" field), and the identification code of the teacher ("id" field) are provided
<i>assistanti</i>	list of teaching assistants or tutors; the name ("nome" field), the surname ("cognome" field), and the identification code of the person ("id" field) are provided

Istituzioni Scolastiche del Trentino

Attributes	Description
<i>codiceOrigine</i>	internal code of the school entity
<i>codiceProvinciale</i>	municipality code for classification of the school entity
<i>denominazione</i>	name of the school entity
<i>codiceS1p</i>	code of the school entity registered in the Sistema Informativo del Personale (S1P)
<i>denominazioneBreve</i>	short name of the school entity
<i>dataInizio</i>	validity start date of the school entity
<i>telefono</i>	phone number of the school entity
<i>fax</i>	fax number
<i>emailIstituzione</i>	email address of the school entity
<i>emailSegreteria</i>	email address of the administrative office of the school entity
<i>descrizione</i>	school entity's description
<i>codiceMeccanografico</i>	mechanographic code of the school entity
<i>nominativo</i>	name of the supervisor
<i>indirizzo</i>	school entity address
<i>numeroCivico</i>	house number of the facility
<i>gestioneAutonomaMensa</i>	autonomous management of the canteen service
<i>soloEffettuazione</i>	course of study authorized for examination only
<i>terminale</i>	terminal course of study
<i>dataFine</i>	validity end date of the school entity
<i>note</i>	notes
<i>altroTelefono</i>	other phone number
<i>numeroBreve</i>	short phone number
<i>intitolazione</i>	Title
<i>codiceMiur</i>	MIUR code of the school entity
<i>indirizzoInternet</i>	school entity's website
<i>emailDirigenza</i>	management email of the school entity
<i>tipologia</i>	school entity type
<i>codiceEntePat</i>	PAT code

Corsi di studio delle scuole Trentine

Attributes	Description
<i>codiceOrigine</i>	internal code of the school entity
<i>codiceMeccanografico</i>	standard unique code of the element assigned by MIUR
<i>descrizione</i>	description
<i>terminale</i>	binary element, if "si" the study course issues a degree
<i>attivo</i>	activity status of the course (yes/no)
<i>codiceOrigineIndirizzoEsame</i>	exam code assigned by MIUR, it is present if type attribute has value "si"
<i>annoCorsoInizio</i>	Classe di inizio del corso descritto
<i>durataAnni</i>	study course duration (in academic years)
<i>codiceOrigineTitoloStudi</i>	code of the qualification within PAT, it is present if type attribute has value "si"
<i>tipologia</i>	type of qualification released

Invalsi Servizio Statistico Cineca

Attributes	Description
<i>id_comune</i>	code used by the Italian administration to identify an Italian municipality
<i>comune_istituto</i>	name of the main institute's municipality
<i>grado</i>	type of school
<i>materia</i>	subject of the test
<i>anno</i>	number of pre-schools in the municipality in
<i>punteggio_percentuale_medio</i> *	average percentage score - corrected for cheating
<i>dev_std_punteggio_percentuale</i> *	standard deviation of the mean percentage score
<i>punteggio_wle_medio</i>	mean WLE score (estimate of skills according to the Rasch model) on a national scale corrected for cheating
<i>dev_std_punteggio_wle</i>	standard deviation of the mean WLE score
<i>perc_compertura_stu</i>	percentage of participation in the corresponding test, i.e. students who took the test on expected students

* from 2017-18, with the introduction of CBT tests, this statistic is no longer calculated for grades 8, 10 and 13

Google Reviews

Attributes	Description
<i>reviewID</i>	identifier of the review
<i>username</i>	name of the user who writes the review
<i>content</i>	text data of the review
<i>score</i>	score of the review (stars score)
<i>thumbs_up_count</i>	how many likes the review has

Unitrento Digital University

Attributes	Description
<i>id</i>	identifier of the person
<i>nome</i>	name of the person
<i>cognome</i>	surname of the person
<i>telefono</i>	phone number of the person
<i>cun</i>	national university council area
<i>ssd</i>	disciplinary scientific sector
<i>ruolo</i>	role covered by the person
<i>nomeStruttura</i>	department name
<i>idStruttura</i>	department id

Vivoscuola - Schools list

Attributes	Description
<i>istituto principale</i>	name of the group of schools in the area
<i>indirizzo</i>	school address
<i>scuola</i>	school name
<i>comune</i>	municipality of the school
<i>telefono</i>	school telephone number

Vivoscuola - School Info

Attributes	Description
<i>school name</i>	name of the school
<i>tipo istituto</i>	type of school
<i>indirizzo</i>	school address
<i>telefono</i>	school telephone number
<i>fax</i>	school fax number
<i>email istituto</i>	institute's email
<i>email dirigenza</i>	management's email
<i>email segreteria</i>	administrative office email
<i>sito web</i>	school web site
<i>codice MIUR</i>	school MIUR code
<i>indirizzo di studio₁</i>	school's field of study ₁
...	
<i>indirizzo di studio_n</i>	school's field of study _n
<i>indirizzo di studio_x articolazione₁</i>	school's articulation ₁ of study _x
...	
<i>indirizzo di studio_x articolazione_n</i>	school's articulation _n of study _x

2.7 Metadata

Metadata is structured information that describes, explains, locates or otherwise makes it easier to retrieve, use or manage an information resource.

Thus, metadata, in general, has three main purposes:

- information resource description;

-
- information resource organization;
 - information resource discovery.

In the context of *iTelos* methodology, the information resources are data resources and technologies. We discuss the scope of metadata concerning quality and reusability.

3 Inception

This section aims to report the KGE sub-process performed during the inception phase, by describing each activity both in the schema and data layer.

This section reports on and describes the following:

- The project's goal, including the objectives description and the project's domain, personas, and scenarios;
- Knowledge resources: The reference teleologies initially collected to satisfy the purpose along the integration process;
- Data resources: The datasets that were initially gathered to fulfill the requirements throughout the integration procedure.
- Metadata: The metadata that has been specified for the informational and data sources listed in the items above.

3.1 Purpose Formalization

The formalization of the project's objective and the domain is the first and most crucial step in the data integration process and the inception phase. The clarity of the project objective is a key factor in determining how the data integration process turns out.

Since the *iTelos* Data Integration technique is purpose-driven and is also known as the Purpose Driven Data Integration Methodology, we adopted it. In order to formulate a purpose, one must first define the project's domain of interest, personas, and scenarios involving those personas.

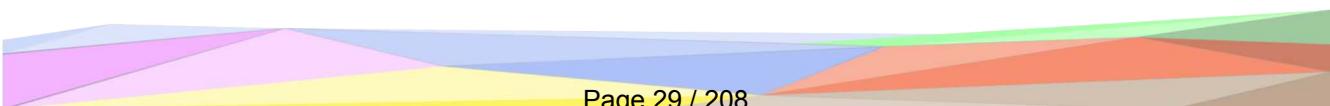
The project's end product (KG) will provide locals and foreigners with some practical and effective methods to know more about education in Trentino. We have identified the domain of interest, and built 11 distinct personas and 10 various scenarios with a variety of features, with the subsequent definition of the competency questions.

The whole project's aim is centered on determining the most effective answers to these concerns.

3.1.1 Domain of Interest

This research seeks to provide a complete analysis of the educational facilities in Trentino which comprehend geospatial and temporal domains. Since it is possible to combine many domains to produce a newly composed domain that serves a specific purpose, domain composition has become an important step that was undertaken to identify the project's aim.

Setting the parameters under which the project's material would be taken into consideration was crucial, since Trentino's educational system included a wide range of topics. Trentino has a wide range of educational institutions, including public and private kindergartens, elementary schools, secondary schools, universities, and other vocational training facilities.



Moreover, as mentioned in Subsection 2, the data we were able to gather spans the years 2014 through 2022. However, we have sufficient recent data to assert that our temporal domain refers to the last few years (2020 - 2022).

Furthermore, a crucial feature we want our system to provide is to define metrics in order to assess the quality of the education provided by an educational facility, such as the quality of the reviews, the rate of completion of compulsory schooling, the number of students who fail to pass the year per municipality.

As a result, after learning about the type of data that was already accessible, the project's original goal was established as follows:

"A service that will help parents and students to find schools, including details about schools and courses offered, in the region of Trentino based on city, municipality, school type, course duration, teaching activities schedules, contact information, and education statistics".

3.2 Competency Questions

The collection of functional user requirements for the whole project is comprised of competency questions, and these questions will be used to assess the project's final outputs. Therefore, the CQs serve as guidelines on how to build the project and aid in determining the significance of various components. The following QCs are posed:

Scenarios	Personas	Competency Questions	Common Entities	Core Entities	Contextual Entities
1	1	Give me how many schools have courses on Software Engineering	Institute, School		Study Course
1	1	Give me schools with courses on Software Engineering	Educational Institute, School	Subregional Academic Division	Study Course
1	1	Give me a list of courses at the University of Trento I can consult when I finish High School	Educational Institute, School		Study Course

1	1	Give me schools that prepare me to become a Software Engineer	Educational Institute, School	Subregional Academic Division	Study Course
1	2	Give me popular schools	Educational Institute, School	Subregional Academic Division, Institute Contact Information	Education Quality, School Statistics
1	2	Give me schools that have courses in the multi-disciplinary area	Educational Institute, School		Study Course
1	2	Give me the schools contact information	Contact Information, Educational Institute, School	Institute Contact Information	
2	2	Give me study programs of the schools in Trentino	Educational Institute, School		Study Course
3	3	Give me the University of Trento's study programs similar to mine (biology)	Educational Institute, School		Study Course
3	3	Give me information about the university's biology department	Educational Institute, Contact Information, School		Study Course
2	3	Get me all study programs the University of Trento is proposing	Educational Institute, School		Study Course

4	4	Give me Kindergartens near Pergine Valsugana	Educational Institute, School	Subregional Academic Division	
4	4	Give me Kindergartens in Trentino	Educational Institute, School	Subregional Academic Division	
4	4	Give me contact informations about this particular Kindergarten	Educational Institute, School	Institute Contact Information, Subregional Academic Division	
5	5	Give me the list of professors of University of Trento	Educational Institute, School	Professor	
5	5	Give me the courses a professor teaches	Educational Institute, School	Professor	Study Course
5	5	Give me contact information about a particular professor of the University of Trento	Contact Information	Professor	
5	5	Give me ratings about ITT Buonarroti school	Educational Institute, School		Review, Education Quality
6	4	Give me the top rated kindergartens in Trentino	Educational Institute, School	Subregional Academic Division	Education Quality, Review
6	4	Give me the top rated kindergartens near Pergine Valsugana	Educational Institute, School	Subregional Academic Division	Education quality, Review

6	6	Give me the top rated schools in Trentino near my home	Educational Institute, School	Subregional Academic Division	Education Quality, Review
6	6	Give me % of students admitted to the next class for a particular school	Educational Institute, School		School Statistics
6	6	Give me % of student that failed the year for a institute	Educational Institute, School		School Statistics
6	6	Give me the invalsi score of a particular school	Educational Institute, School		Invalsi Score
6	6	Give me % of student that abandoned a particular institute	Educational Institute, School		School Statistics
7	7	Give me schools that offer evening programs in Trentino	Educational Institute, School	Subregional Academic Division	
7	7	Give me courses provided by evening schools	Educational Institute, School		Study Course
7	7	Give me contact information about a particular evening school	Educational Institute, School	Institute Contact Information	

7	10	Give me % of students admitted to the next class for a particular night school	Educational Institute, School		School Statistics
7	10	Give me courses provided by evening schools	Educational Institute, School		Study Course
7	10	Give me contact information about a particular evening school	Educational Institute, School	Institute Contact Information	
8	8	Give me detailed information about schools in Trentino	Educational Institute, School	Institute Contact Information	
8	8	Give me the percentage of students admitted to the next class for a particular school	Educational Institute, School		School Statistics
8	8	Give me the percentage of students that failed the year for an institute	Educational Institute, School		School Statistics
8	8	Give me the invalsi score of a particular school	Educational Institute, School		Invalsi Score
8	8	Give me courses provided by schools	Educational Institute, School		Study Course

8	8	Give me the number of enrolled students of a particular school	Educational Institute, School		School Statistics
8	8	Give me the graduation rate of a particular municipality in Trentino	Institute	Subregional Academic Division	
8	8	Give me the number of people who didn't finish school in a certain municipality in Trentino		Subregional Academic Division	
9	9	Give me a list of number of school per municipality	Educational Institute	Subregional Academic Division	
9	9	Give me a list of number of students per municipality	Educational Institute, School	Subregional Academic Division	
10	11	Give me a list of contact information of schools in Trentino	Educational Institute, School	Institute Contact Information	
10	11	Give me a list of schools of Trentino	Educational Institute, School	Subregional Academic Division	

Competency Questions

3.3 Resource Collection

As for the second input of the inception phase of the *iTelos* methodology, the data sources are required. With the term data source, we are referring to all those data and knowledge resources that will be refined in order to build the final KG.

The Inception phase aims to collect and semi-formalize the resources (knowledge and data) used to build the final KG.

Due to the heterogeneity of the sources provided in input, the execution of this phase can be different source by source, or even dataset by dataset.

As a consequence, an iterative execution over the source list is considered for this phase [4]. Collecting data, in general, aims to achieve the following results:

- increases the number of entities (E-types);
- increases the number of entity attributes (data properties).

Most of the datasets were downloaded manually from Open Data Trentino [11], INVALSI [7], and ISPAT [8].

For what concerns Google Reviews [14] and Vivoscuola [13] data, we have provided a scraper written in Python which collects the reviews for each school. In order to do that, we have relied on *Selenium*¹¹ and *BeautifulSoup*¹².

The first was used in order to simulate the browser search of all the schools collected from *Vivoscuola* [13], while the latter was in order to retrieve the information contained in the reviews, such as the rating and each review's text.

In the following sections, we describe the data resources collected as well as the knowledge resources of the data we have managed to retrieve.

3.3.1 Data Resources

As for the data resources, along with those listed in Section 1.6, we have employed the semi-formal resources provided by Ali Hamza and Tecla Venturelli¹³ of the last year data integration project. The code for the scraping process can be found at the following GitHub repository: <https://github.com/samuelebortolotti/kge>.

The reason why we have adopted such data is that they comprise the same information conveyed in *Istituzioni Scolastiche del Trentino*, *Corsi di studio delle scuole Trentine* and *Municipalities of Trentino* but they have kept information which is useful in order to fulfill our purpose.

For instance, the *german* and *ladin* name and the *population* amount in the *Municipalities* resources were dropped since the only information preserved is the code, the Italian name and the area. While, for the resource *Corsi di studio delle scuole Trentine*: *codiceOrigineTitoloStudi*, *tipologia* were deleted. Finally, for the *Istituzioni Scolastiche del Trentino* dataset, they have created two resources with cleaned data. All of them comply with our purpose, therefore it was not necessary to perform other data processing.

Last year project's data, cleaned and formatted, which we have employed are:

¹¹<https://www.selenium.dev/>

¹²<https://beautiful-soup-4.readthedocs.io/en/latest/>

¹³<https://github.com/alihamzaunitn/kdi-educationtrentino/tree/master>

- Trentino Commune List.csv
<https://github.com/alihamzaunitn/kdi-educationtrentino/blob/master/Datasets/Data%20Integration/Trentino%20Commune%20List.csv>, which collects data of the municipalities;
- courses_offered_in_trentino_schools.csv
https://github.com/alihamzaunitn/kdi-educationtrentino/blob/master/Datasets/Data%20Integration/courses_offered_in_trentino_school.csv, which collects data about the courses, Trentino schools provide;
- educational_institutions.csv
https://github.com/alihamzaunitn/kdi-educationtrentino/blob/master/Datasets/Data%20Integration/educational_institutions.csv, which collects data concerning educational institutions in Trentino;
- responsible_person_educational_institutions.csv
https://github.com/alihamzaunitn/kdi-educationtrentino/blob/master/Datasets/Data%20Integration/responsible_person_educational_institutions.csv, which collects data concerning the people responsible for the educational institutions in Trentino;
- schools_in_trentino.csv
https://github.com/alihamzaunitn/kdi-educationtrentino/blob/master/Datasets/Data%20Integration/schools_in_trentino.csv, which collects data concerning the schools present in Trentino territory.

A comprehensive overview of the datasets is shown here:

3.3.1.1 Trentino Commune List.csv

Trentino Commune List

Attributes	Description
comune-code	identifier of the municipality
comune-name	name of the municipality
area	area dimension, expressed in square meters

3.3.1.2 courses_offered_in_trentino_school.csv

courses_offered_in_trentino_school

Attributes	Description
codiceOrigine	identifier of the course
codiceMeccanografico	mechanical code
descrizione	description of the course

3.3.1.3 educational_institutions.csv

educational_institutions.csv

Attributes	Description
<i>institute_name</i>	name of the institute
<i>institute_phone</i>	mechanical code
<i>institute_fax</i>	description of the course
<i>institute_website</i>	website of the course

3.3.1.4 responsible_person_educational_institutions.csv

responsible_person_educational_institutions

Attributes	Description
<i>institute_name</i>	name of the institute
<i>responsible_person</i>	name of the responsible person
<i>responsible_person_role</i>	role of the responsible person

3.3.1.5 schools_in_trentino.csv

schools_in_trentino

Attributes	Description
<i>institute_name</i>	name of the institute
<i>school_email</i>	email of the school
<i>school_type</i>	whether the school is public, private, religious or secular
<i>school_responsible_person_name</i>	name of the school responsible person's role
<i>school_name</i>	name of the school

Concerning the scraped data from Vivoscuola and Google Reviews, the following section describes each field of the resources.

3.3.2 school_list.csv

schools_list

Attributes	Description
<i>school_type</i>	type of the school, it could be Prinviciale or Autonoma or more
<i>address</i>	address of the school
<i>phone_number</i>	phone number of the institute (presumably, as all schools of the same institute have the same values)
<i>fax_number</i>	fax number of the institute (presumably, all schools of the same institute have same values)
<i>school_email</i>	email of the institute (presumably, all schools of the same institute have same values) in which the school belongs to (all schools of the same institute have same values)
<i>management_email</i>	email of the institute's management section (presumably, all schools of the same institute have same values) in which the school belongs to (all schools of the same institute have same values) in which the school belongs to
<i>office_email</i>	email of the institute's office section (presumably, all schools of the same institute have same values) in which the school belongs to (all schools of the same institute have same values) in which the school belongs to
<i>website</i>	website of the institute presumably (all schools of the same institute have same values)
<i>miurcode</i>	MIUR code of the school
<i>school_name</i>	name of the school
<i>school_id</i>	identifier of the school
<i>institute_id</i>	identifier of the institute (defined by processing the URL of the institute) in which the school belongs to
<i>institute_id_number</i>	identifier as a number of the institute in which the school belongs to

3.3.3 institute_list.csv

institute_list

Attributes	Description
<i>institute_type</i>	institute type
<i>indirizzo</i>	institute's address
<i>telefono</i>	institute's phone number
<i>fax</i>	institute's fax number
<i>email_istituto</i>	email address of the institute
<i>email_dirigenza</i>	email address of the institute's management section
<i>email_segreteria</i>	email address of the institute's office section
<i>sito_web</i>	institute's web site
<i>miur_code</i>	institute's MIUR code
<i>numero_plessi_scuole</i>	number of schools assigned to the institute
<i>nome_istituto</i>	institute name
<i>id_istituto</i>	institute's id as a string (taken from the institute url)
<i>id_number</i>	institute's processed id as a number

3.3.4 responsible_list.csv

responsible_list

Attributes	Description
<i>name</i>	responsible's name
<i>surname</i>	responsible's surname
<i>title</i>	title of the responsible (could be: sorastant, principal, director, pedagogican coordinator, etc...)
<i>institute_id</i>	institute's identifier

3.3.5 school_courses.csv

school_courses

Attributes	Description
<i>teaching_unit</i>	information needed for fetching data from the API
<i>date_for_find_type_of_courses_of_study</i>	information needed for fetching data from the API
<i>title</i>	name of the title the student will achieve when completing the studies
<i>indirizzo</i>	description of the field of study offered by the school
<i>school_id</i>	id of the school providing the study course

3.3.6 institute_scrutini.csv

institute_scrutini

Attributes	Description
<i>ord_scol</i>	school's stage, it could be "Primaria", Secondaria di primo grado, (elementary school, middle school) etc.
<i>tipo_orario</i>	Timetable type, it could be "DIURNO" or "SERALE" (daylight lectures and evening lectures respectively)
<i>num_anni</i>	duration of the school's program
<i>etichetta</i>	information about the school year
<i>percentuale_ammessi</i>	admitted students to the next year as percentage
<i>anno</i>	school year of enrollment reference as a number
<i>miur_code</i>	school's MIUR code

3.3.7 institute_ripetenti.csv

institute_ripetenti

Attributes	Description
<i>data_inserimento</i>	time when the record was submitted to the server
<i>ord_scol_cap</i>	school's stage, it could be "Primaria", Secondaria di primo grado, (elementary school, middle school) etc.
<i>ord_scol_ord</i>	school's stage as a number
<i>strutt_liv_1_cap</i>	institute's name
<i>strutt_liv_1_cod_miur</i>	institute's MIUR code, presumably (sometimes is missing)
<i>strutt_liv_1_cod_pat</i>	institute's PAT code (used for fetching data from an API endpoint)
<i>tipo_ora_id</i>	school's stage, it could be "Primaria", Secondaria di primo grado, (elementary school, middle school) etc.
<i>valore</i>	repeating students as a percentage
<i>valore_d</i>	total number of students
<i>valore_n</i>	number of repeating students
<i>anno_scol</i>	school year
<i>miur_code</i>	institute's MIUR code

3.3.8 institute_dropouts.csv

institute_dropouts

Attributes	Description
<i>data_inserimento</i>	time when the record was submitted to the server
<i>ord_scol_cap</i>	school's stage, it could be "Primaria", Secondaria di primo grado, (elementary school, middle school) etc.
<i>ord_scol_ord</i>	school's stage as a number
<i>strutt_liv_1_cap</i>	institute's name
<i>strutt_liv_1_cod_miur</i>	institute's MIUR code, presumably (sometime is missing)
<i>strutt_liv_1_cod_pat</i>	institute's PAT code (used for fetching data from an API endpoint)
<i>tipo_ora_id</i>	school's stage, it could be "Primaria", Secondaria di primo grado, (elementary school, middle school) etc.
<i>valore</i>	dropout students as a percentage
<i>valore_d</i>	total number of students
<i>valore_n</i>	number of dropout students
<i>anno_scol</i>	school year
<i>miur_code</i>	institute's MIUR code

3.3.9 school_scores.csv

school_scores

Attributes	Description
<i>school_id</i>	school id
<i>miur_code</i>	school's MIUR code
<i>score</i>	school's google score
<i>school_name</i>	school's name

3.3.10 school_reviews.csv

school_reviews

Attributes	Description
<i>school_id</i>	school id
<i>miur_code</i>	school's MIUR code
<i>text_review</i>	school's google review
<i>school_name</i>	school's name

While regarding the digital university data, the data cleaning and formatting process have brought the following result, where we have kept only the most information relevant to our purpose.

3.3.11 unitn_courses.csv

unitn_courses

Attributes	Description
<i>id_course</i>	course's id
<i>course_name</i>	name of the course
<i>degree_program</i>	degree in which the course is registered to

3.3.12 unitn_staff.csv

unitn_staff

Attributes	Description
<i>id_employee</i>	employee's id
<i>name</i>	name of the employee
<i>surname</i>	surname of the employee
<i>CUN area</i>	CUN area of the employee (could be missing if the person is not a professor)
<i>SSD</i>	disciplinary scientific sector of the employee (could be missing if the person is not a professor)

3.3.13 departments.csv

departments

Attributes	Description
<i>id_department</i>	id of the department
<i>name</i>	name of the department

3.3.14 Knowledge Resources

Concerning the Knowledge resources, we have constructed the metadata representing the ontology of the collected data.

To perform such an operation, we have followed the tools compliant to the *iTelos* methodology, such as *Protégé*¹⁴ and *Karma*¹⁵.

¹⁴<https://protege.stanford.edu/>

¹⁵<https://usc-isi-i2.github.io/karma/>

All of them, along with the ontologies (OWL files) and the cleaned CSV/JSON files, are available in *RDF* format in the project's official repository <https://github.com/samuelebortolotti/Education-Trentino>.

3.3.14.1 Protégé

Protégé is a framework and a free open-source ontology editor for creating intelligent systems. It allows modeling an ontology compliant with the *OWL*¹⁶ standard.

During the inception phase, we modeled an ontology on top of the extracted and cleaned dataset. Protégé is the tool we have employed in order to realize each dataset ontology.

3.3.14.2 Karma

Karma is a framework and a free open-source ontology editor for creating intelligent systems. It allows modeling an ontology compliant with the *OWL*¹⁷ standard. Karma automates most of the integration process since users can integrate information and model it in accordance with an ontology of their choice.

In order to map each dataset with the corresponding ontology, we have relied on Karma.

3.4 Ontologies

In order to generate the ontologies of the dataset we have collected, we have taken inspiration from the entities <https://schema.org/School> and <https://schema.org/EducationalOrganization> from the schema.org knowledge resource. However, we did not have all the information required, for instance, *openingHours* is missing in the definition of an *EducationalInstitute*, but we had the necessity to adapt the schemas and to understand the data by referring to the metadata provided along with the resources.

Moreover, in the schema.org definition of school, it can be noticed that *e-mail* can be a property of a school, but in our case, we do not have such data for schools in Trentino. From our data we can see how schools belonging to a certain institute have identical values for *Phone Number*, *Fax Number*, *School Email*, *Management Email*, *Office Email*. As a consequence, we decided not to add such properties in the *School* entity but instead add them in the *Institute* entity.

Finally, some other data does not fit schema.org entities, such as the *Invalsi test*, since they contain contextual information.

3.5 Data Trimming

As a result of the data scraping, we ended up with a huge amount of information. In order to reduce the complexity of the product, it was decided to remove some columns in the csv files.

For instance, if we look at Table 3, we can see that depending on the type and location of an institute, the latter can have persons with different roles as a coordinator, respectively *Sorastant*, *Pedagogican Coordinator*, *President of Managing Body*, *Principal*, *Director*. Given

¹⁶<https://www.w3.org/OWL/>

¹⁷<https://www.w3.org/OWL/>

these circumstances, it was decided to drop all those columns and to create instead a formal entity *Responsible Authority*.

Another logical justification for reducing the properties in the final entities was that certain information was only needed to gather other critical data for a particular entity: in Table 4 there are the properties *teaching_unit* and *date_for_find_type_of_courses_of_study*; those were discarded in the definition phase of the ontologies as they provided no useful information whatsoever in terms of our purpose: their use was to fetch school courses through an API endpoint.

3.6 Entities defined

As for the entities, we have employed *Protégé*, along with *Karma* in order to produce an *RD-F/OWL* file.

Below we list the entities we have defined up to now, on top of the retrieved data in the inception phase.

School

Name
Type
Stage
Address
MIUR code
Timetable type
Number of Years

Where review is the following entity:

Review

Text
Likes
Username
Date

From the data, we have retrieved the following ontology. Since there were several pieces of information concerning a school year, we have decided to create a separate entity *School Statistics*, which is in charge of dealing with all the statistics of different school years.

School Statistics

Year
Dropout Students
Dropout Percentage
Number of Students
Failed Number of Students
Failed Percentage
Admitted Percentage

A *School Statistics* comprise all the statistics which were issued for a school according to a specific academic year. This entity was issued since there was the need to couple the statistics to a school year by year. For this reason, it was not possible to include them in the school entity.

Institute

Name
Type
Website
Address
Number of Schools
MIUR code
PAT code

Institute Contact Information

Phone Number
Fax Number
Institute Mail
Management Mail
Office Mail

Study Course

Title
Articulation

Invalsi Score

Subject
Average Score
Average Score Percentage
Average WLE
Coverage Percentage STU

This entity represents an Invalsi test instance. In this case, we have defined a very specific entity that adapts perfectly to the Invalsi test but may not suit other types of tests.

Professor

Name
Surname
CUN
SSD

Professor is a specialized entity that includes information such as the phone number and the CUN (*Consiglio Universitario Nazionale*) they belong to. Moreover, professors have courses they can teach, and of course, all the information related to a hypothetical *Person* entity.

Responsible Authority

Name
Surname
Title

Responsible Authority is a specialized functionality of *Person*. Usually, the response is the school principal, but in general, is any person which has an administrative role in the school facility.

Contact Information

Phone
Email
Office Address

Education Quality

User Score

Subregional Academic Division

Name
Number of Students Enrolled in Kindergarten
Number of Students Enrolled in Primary School
Number of Students Enrolled in Middle School
Number of Students Enrolled in High school
Number of Students Enrolled in CF
School Completion Rate
Graduate Rate

Region

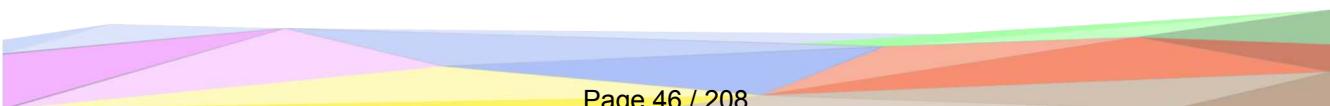
Lat
Lon
StartYear
EndYear

3.7 Schema modeling

Another crucial phase in the *iTelos* methodology is schema modeling.

In order to produce Semi-Formal resources the dataset collected, cleaned, and formatted have to be associated with a schema representing the information they are carrying.

Such a schema has to be formally defined in RDF-OWL, using a specific tool suggested by *iTelos* (*Protègè*). Once the ontology for each dataset has been created, the objective of the activity is to create, for each pair composed of a dataset and its relative schema, a single object representing a semi-formal information model.

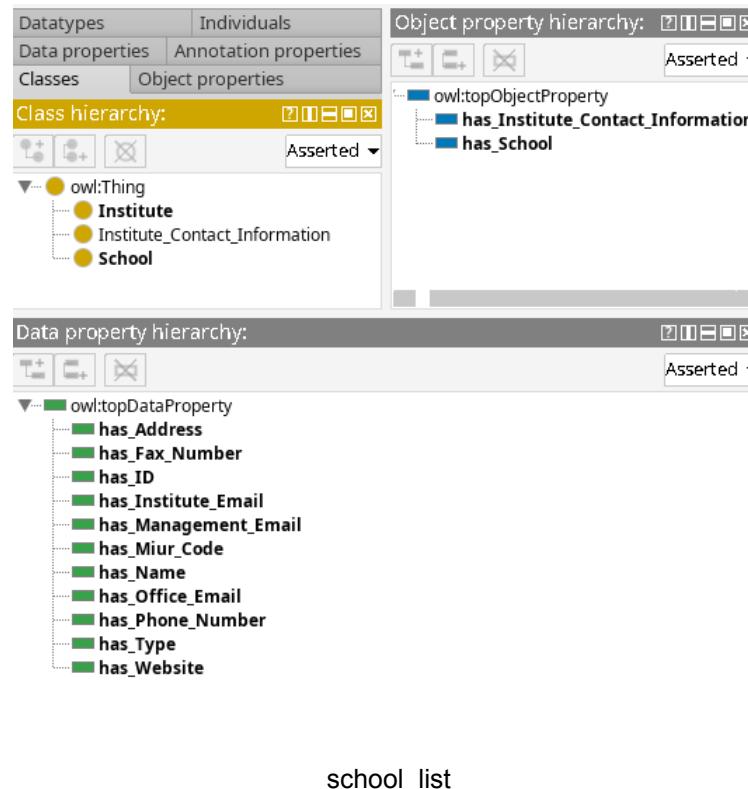


Such an object is represented by a (dataset-specific) KG. To achieve this result a specific tool is offered by iTelos (Karma) used to map each dataset over its own schema, thus merging the data and knowledge layer of a KG.

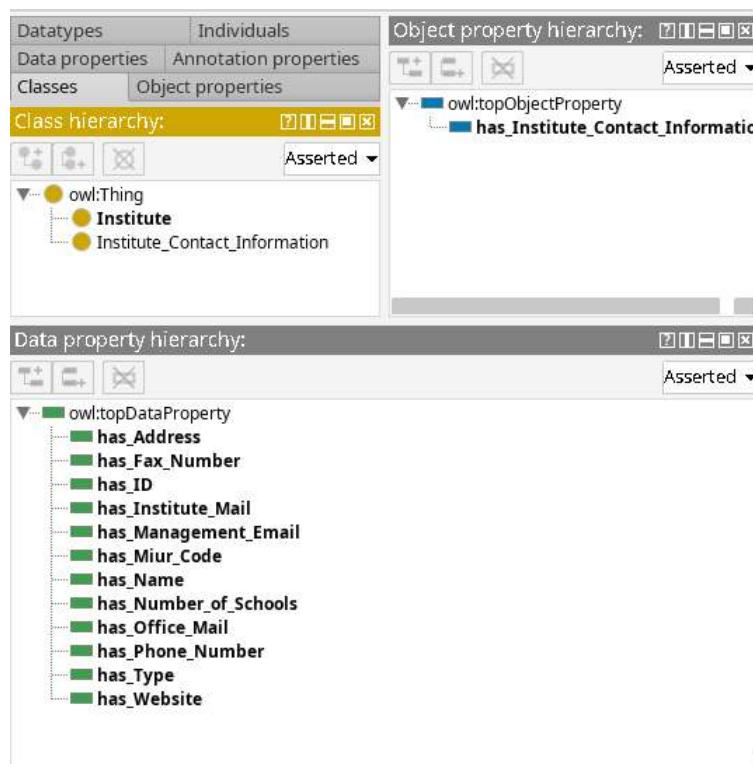
3.8 Protégé

Here, we list some of the steps we have performed in order to model the schema of the semi-formal resources we have collected.

3.8.1 schools_list.csv

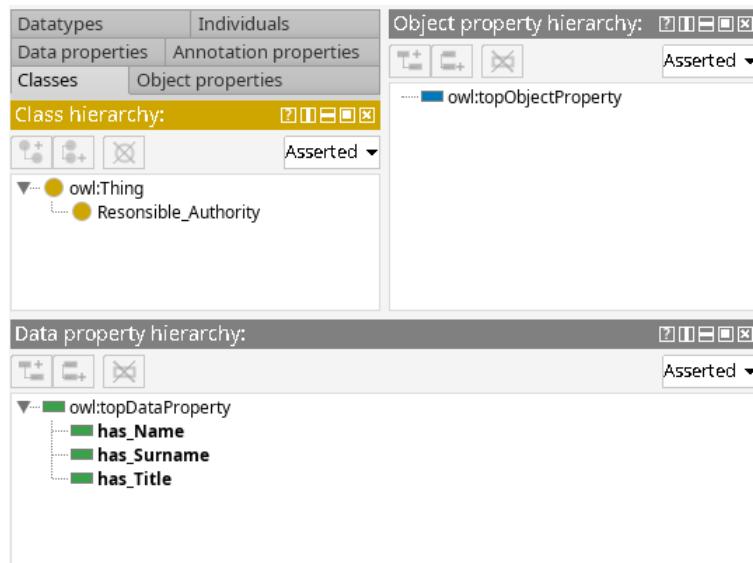


3.8.2 institute_list.csv



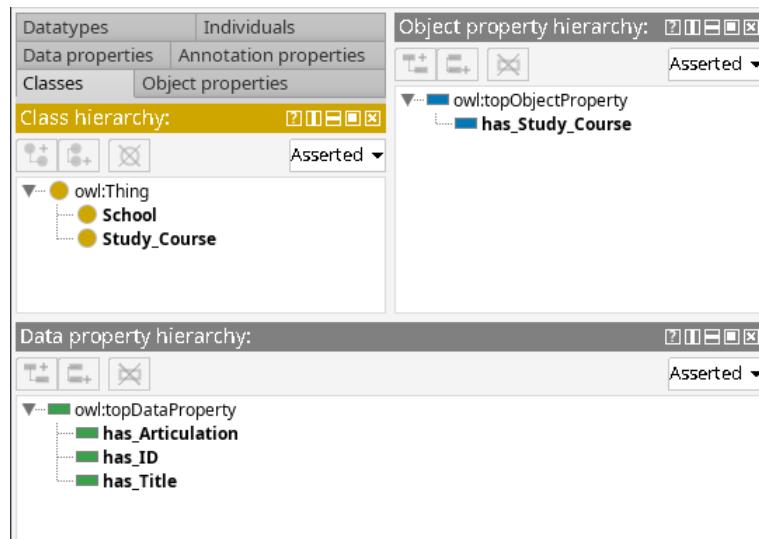
institute_list

3.8.3 responsible_list.csv



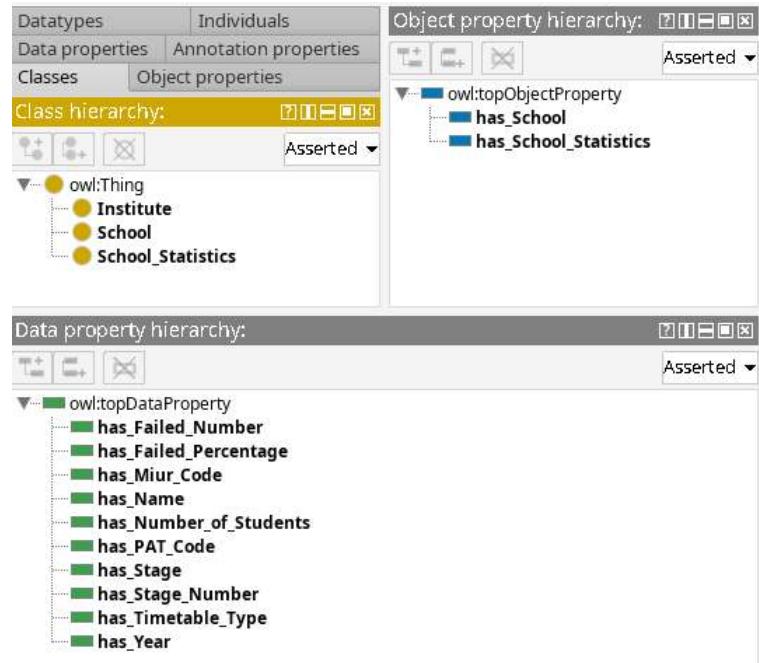
responsible_list

3.8.4 school_courses.csv



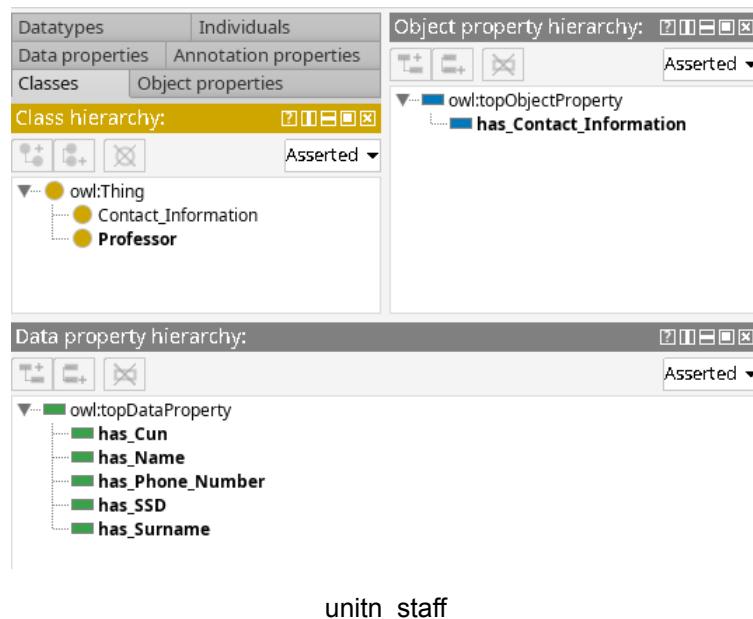
school_courses

3.8.5 institute_ripetenti.csv



instiute_ripetenti

3.8.6 unitn_staff.csv



3.9 Karma

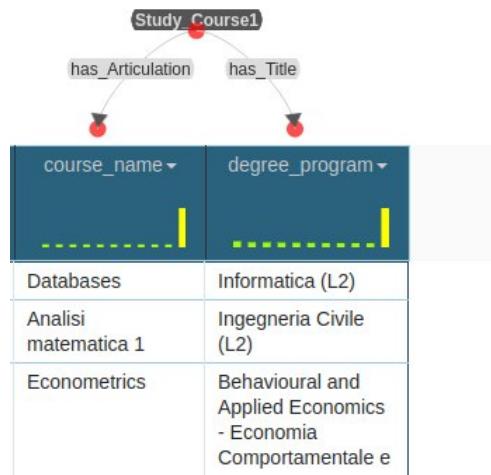
Instead, in this section, we list the merging operation between the knowledge layer and the data layer by means of Karma.

3.9.1 Trentino Commune List.csv



Trentino Commune List

3.9.2 courses_offered_in_trentino_school.csv



`courses_offered_in_trentino_school`

3.9.3 schools_list.csv

school_type	address	phone_number	fax_number	school_email	management_email	office_email	website	miur_code	school_name
Provinciale	Piazza Municipio, 5 - 38050 NOVALEDO	+39 0461 764581	+39 0461 771046	ic.centrovalsgan...	dir.ic.centrovalsg...	segri.ic.centrovalsg...	http://www.iccentr...	TNEE813064	SCUOLA PRIMARIA "C. CORRADI" NOVALEDO
Provinciale	Via Baldassari, 15 - 38050 RONCEGN TERME	+39 0461 764581	+39 0461 771046	ic.centrovalsgan...	dir.ic.centrovalsg...	segri.ic.centrovalsg...	http://www.iccentr...	TNEE813075	SCUOLA PRIMARIA "P. MARTINELLI" RONCEGNO

school_list

3.9.4 educational_institutions.csv

institute_name	institute_phone	institute_fax	institute_email	institute_websit	institute_educational_email	institute_type	responsible_person	responsible_surname	responsible_name	responsible_person_role	institute_address	institute_comune
CEO "A. Maffei" RIVA DEL GARDA	0464/552316	0464/552316	maffei@pec...	https://e.lice...	dir.liceo.maf...	Pubblico	ZAMBONI ANTONIA	ZAMBONI	ANTONIA	Dirigente scolastico	VIALE F.A. LUTTI	RIVA DEL GARDA
STITUTO COMPRENSIVO RONCEGNO	0461/764581	0461/771046			dir.ircneg...	Pubblico	FORENZASALVATORE	FORENZA	FORENZASALVATORE	Dirigente scolastico	VIA F. MEGGIO	RONCEGNO TERME
COOPERAZIONE SOCIALE A.R.L. Z.VERONESI	0464/434047	0464/458179	www.cfpgeve...	scuolaveron...		Laico	OLIVIA OLIVO	OLIVIA	OLIVO	Dirigente scolastico	VIA SAIBANTI	ROVERETO

educational_institutions

3.9.5 responsible_list.csv

Resonsi...Authority1		
has_Name	has_Surname	has_Title
nome ▾	cognome ▾	titolo ▾
.....
BRUNO	GENTILINI	DIRIGENTE SCOLASTICO
MARCO	FELICETTI	DIRIGENTE SCOLASTICO
ENRICA	RIGOTTI	DIRIGENTE SCOLASTICO

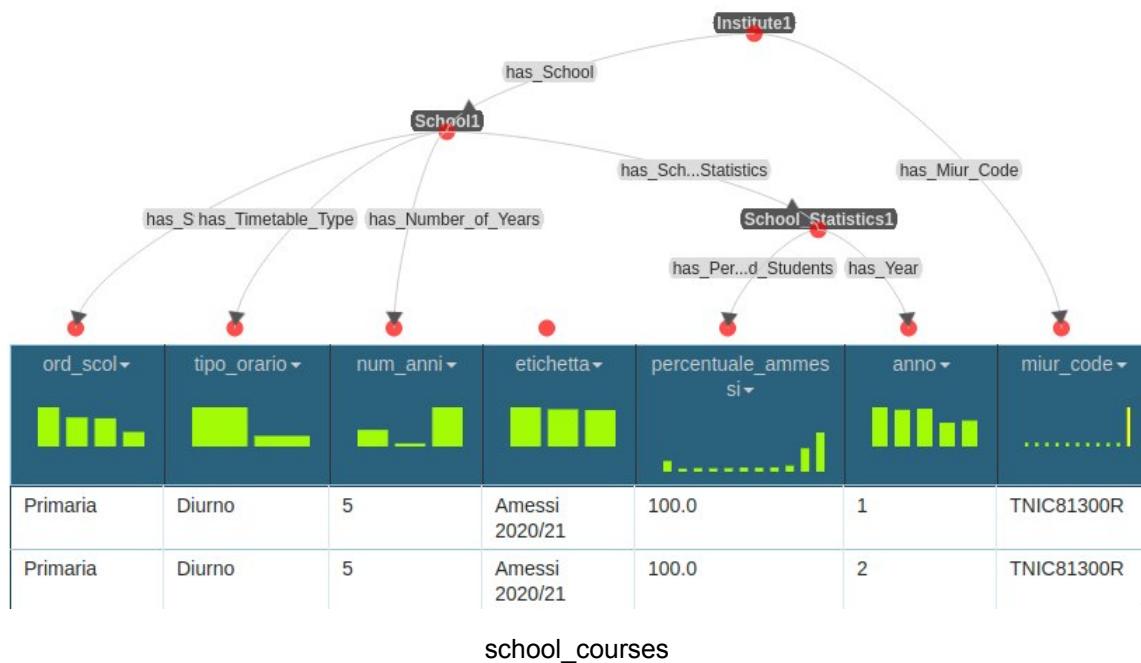
responsible_list

3.9.6 school_courses.csv

Study_Course1	
has_Title	has_Articulation
title ▾	indirizzo ▾
.....
DIPLOMA ESAME DI STATO CONCLUSIVO DEL SECONDO	Diploma di istituto tecnico settore economico indirizzo "amministrazione,
DIPLOMA ESAME DI STATO CONCLUSIVO DEL SECONDO	Diploma di istituto tecnico settore economico indirizzo "amministrazione,

school_courses

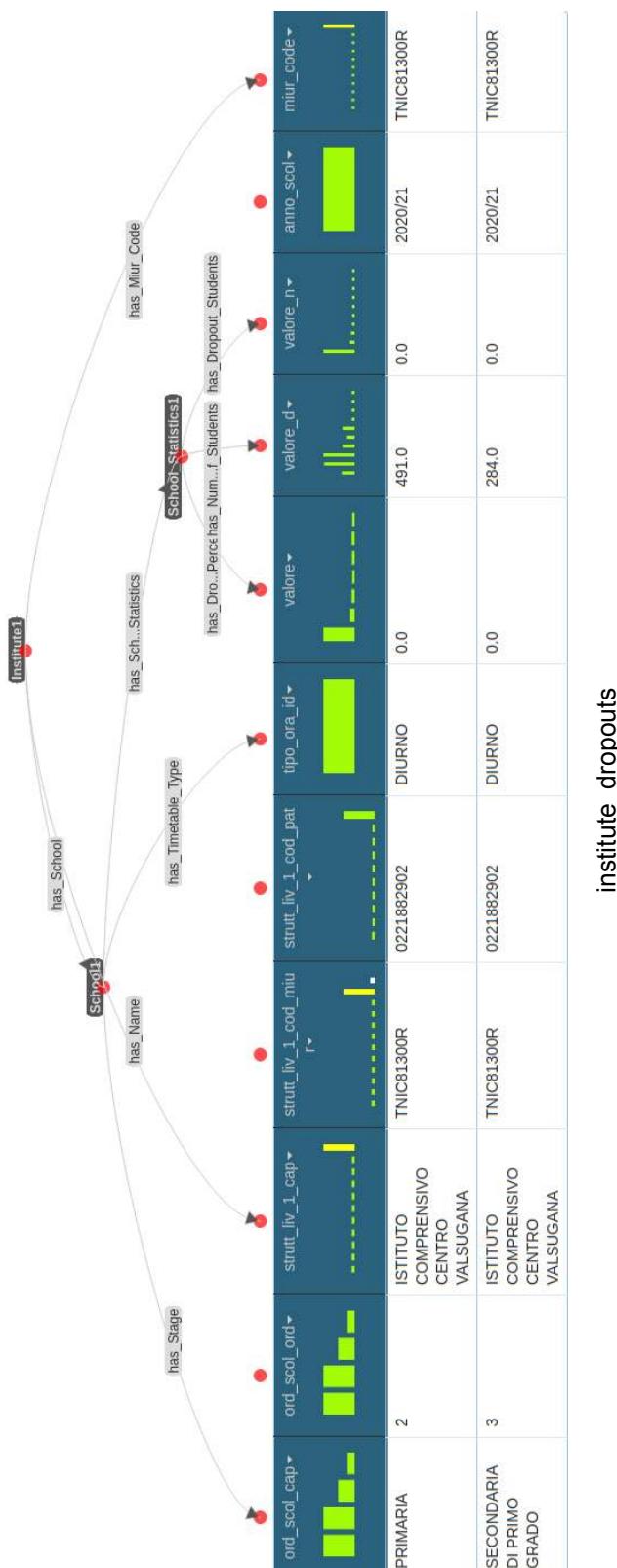
3.9.7 institute_scrutini.csv



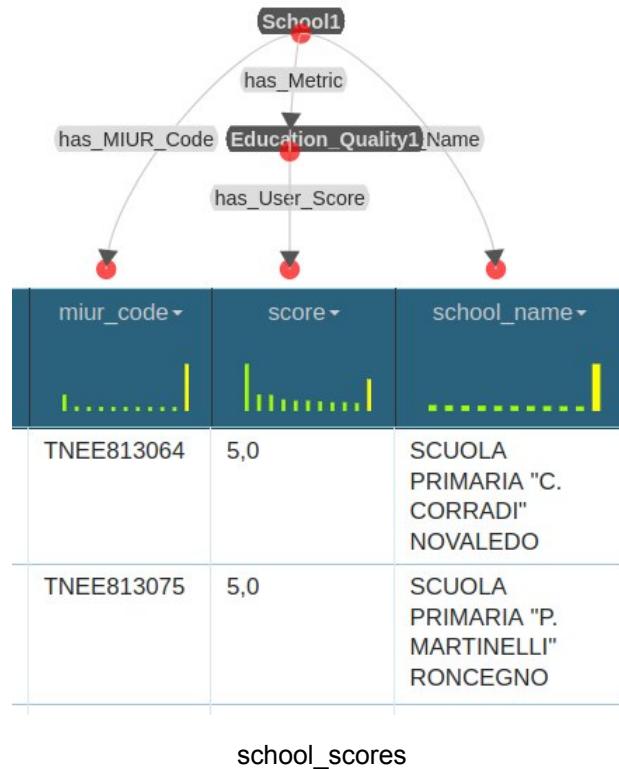
3.9.8 institute_scrutini.csv



3.9.9 institute_dropouts.csv

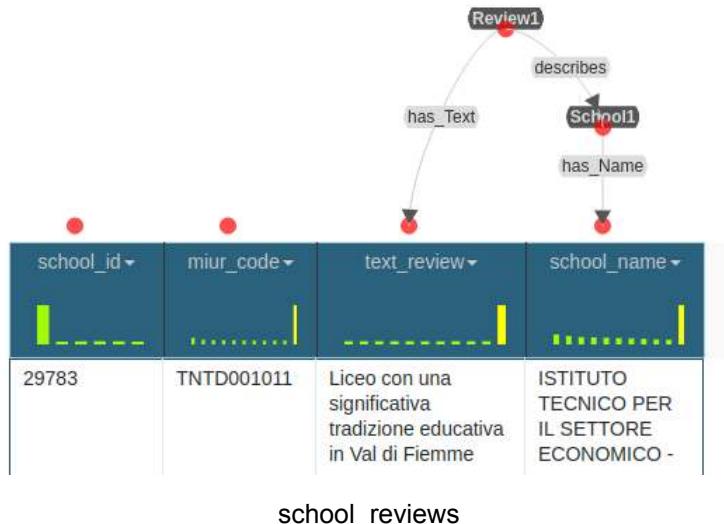


3.9.10 school_scores.csv



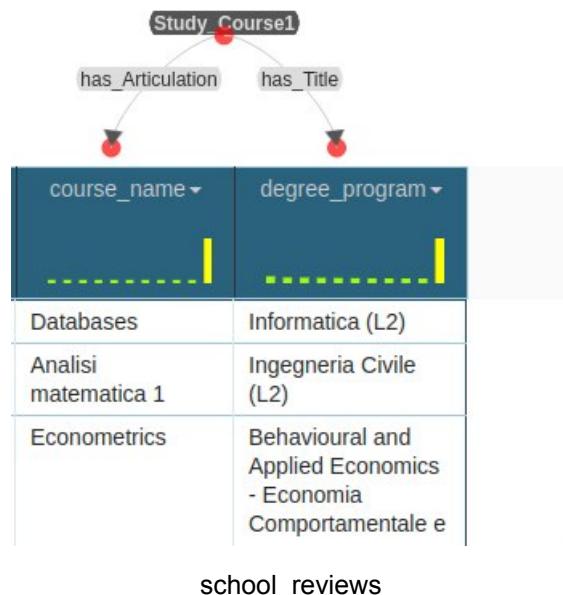
`school_scores`

3.9.11 school_reviews.csv

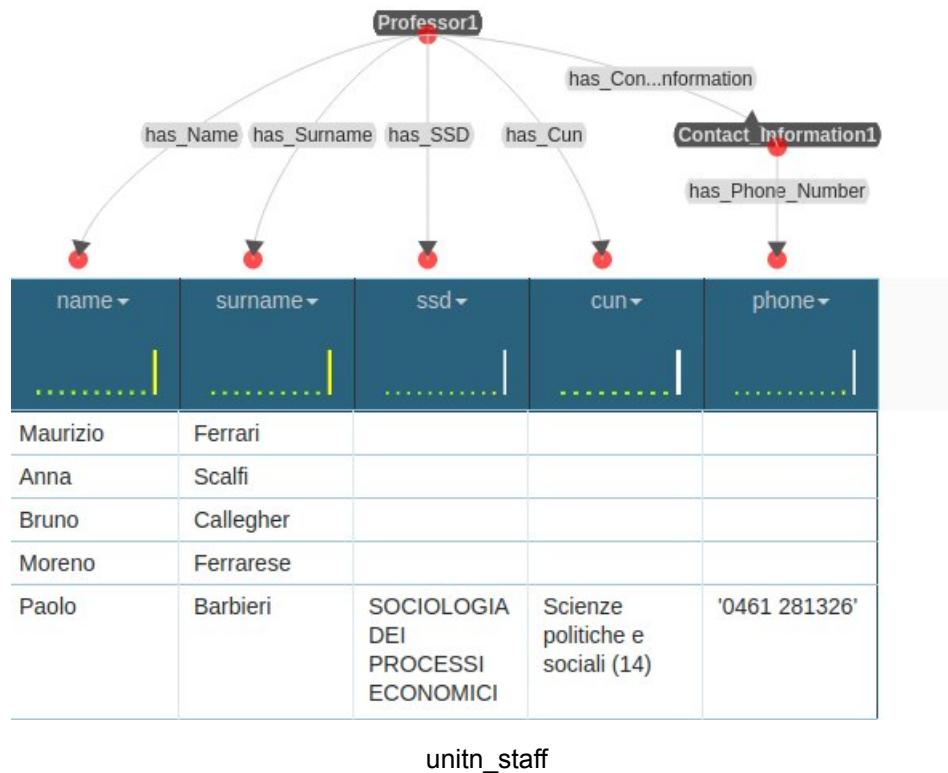


`school_reviews`

3.9.12 unitn_courses.csv



3.9.13 unitn_staff.csv

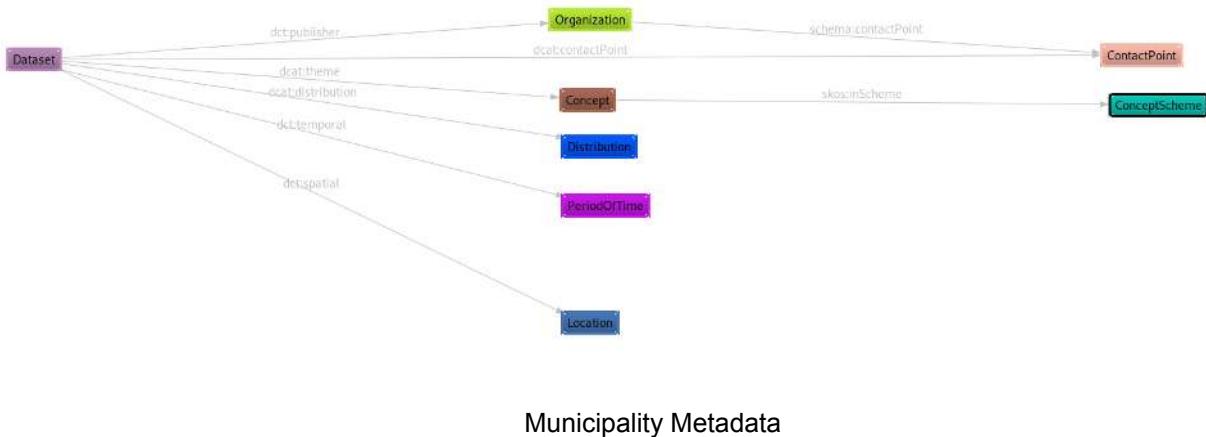


3.10 Metadata

RDF is also a standard paradigm for data transfer on the Internet. It contains capabilities that make data merging easier even when the underlying schemas differ, and it explicitly facilitates schema evolution over time without needing all data consumers to be updated. We have adopted the tool *SHAPEness metadata editor* in order to build the metadata associated with the resources we have gathered.

Thus, we present our dataset information in RDF format here, with particular graphs and tables presented below.

3.10.1 Municipality metadata



```
@prefix : <https://www.epos-eu.org/epos-dcat-ap#> .  
@prefix schema: <http://schema.org/> .  
@prefix spdx: <http://spdx.org/rdf/terms#> .  
@prefix owl: <http://www.w3.org/2002/07/owl#> .  
@prefix gsp: <http://www.opengis.net/ont/geosparql#> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .  
@prefix dqv: <http://www.w3.org/ns/dqv#> .  
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .  
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .  
@prefix oa: <http://www.w3.org/ns/oa#> .  
@prefix dct: <http://purl.org/dc/terms/> .  
@prefix sh: <http://www.w3.org/ns/shacl#> .  
@prefix dcat: <http://www.w3.org/ns/dcat#> .  
@prefix locn: <http://www.w3.org/ns/locn#> .  
@prefix foaf: <http://xmlns.com/foaf/0.1/> .  
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .  
@prefix adms: <http://www.w3.org/ns/adms#> .
```

```
@prefix org: <http://www.w3.org/ns/org#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix http: <http://www.w3.org/2006/http#> .
@prefix dash: <http://datashapes.org/dash#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636>
    rdf:type skos:ConceptScheme ;
    dct:description "Concept of the Trentino Study Course Dataset" ;
    dct:title "Dataset concept" .

<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365>
    rdf:type schema:ContactPoint ;
    schema:availableLanguage "it-IT" ;
    schema:contactType "None" .

<https://www.epos-eu.org/epos-dcat-ap#Distribution/2dee49b8-e959-4ea-a3d1-48342784a61d>
    rdf:type dcat:Distribution ;
    dct:conformsTo "HTML" ;
    dct:description "List of municipalities of the autonomous province of Trento, Italy" ;
    dct:format "HTML" ;
    dct:identifier "https://en.wikipedia.org/wiki/Municipalities\_of\_Trentino"^^xsd:anyURI ;
    dct:issued "2011-09-10T16:27:17Z"^^xsd:dateTime ;
    dct:language "en-EN" ;
    dct:license "https://creativecommons.org/licenses/by-sa/3.0/"^^xsd:anyURI ;
    dct:modified "2022-09-05T16:27:17Z"^^xsd:dateTime ;
    dct:title "Municipalities_of_Trentino" ;
    dct:type "dcat:Distribution"^^xsd:anyURI ;
    dcat:accessURL "https://en.wikipedia.org/wiki/Municipalities\_of\_Trentino"^^xsd:anyURI ;
    dcat:downloadURL "https://en.wikipedia.org/wiki/Municipalities\_of\_Trentino"^^xsd:anyURI ;
    dcat:mediaType "text/html" .

<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/3781753d-6d8a-460b>
```

```
-ba7d-11386779cb23>
    rdf:type            dct:PeriodOfTime ;
    schema:endDate     "2011-04-25T19:37:51Z"^^xsd:dateTime ;
    schema:startDate   "2022-09-05T19:37:51Z"^^xsd:dateTime .

<https://www.epos-eu.org/epos-dcat-ap#Organization/e892ce35-d4e1-468c
-9df5-c0733faa444c>
    rdf:type            schema:Organization ;
    schema:contactPoint <https://www.epos-eu.org/epos-dcat-ap#
                           ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;
    schema:identifier   "https://en.wikipedia.org/wiki/
                           Municipalities_of_Trentino"^^xsd:anyURI ;
    schema:legalName    "The list shows the municipalities (
                           comuni) of the autonomous province of Trento, Italy" ;
    schema:leiCode      "https://en.wikipedia.org/wiki/
                           Municipalities_of_Trentino" ;
    schema:url         "https://en.wikipedia.org/wiki/
                           Municipalities_of_Trentino"^^xsd:anyURI .

<https://www.epos-eu.org/epos-dcat-ap#Dataset/8ab7280e-0772-4617-a052
-fb38cab23de0>
    rdf:type            dcat:Dataset ;
    dct:accessRights   "free" ;
    dct:conformsTo     "http://www.w3.org/2001/XMLSchema-instance
                         " ;
    dct:created        "2011-04-25T19:36:18Z"^^xsd:dateTime ;
    dct:description    "This dataset shows the list of
                           municipalities (comuni) of the autonomous province of
                           Trento, Italy" ;
    dct:identifier     "https://en.wikipedia.org/wiki/
                           Municipalities_of_Trentino"^^xsd:anyURI ;
    dct:issued         "2011-04-25T19:36:19Z"^^xsd:dateTime ;
    dct:language       "en-EN" ;
    dct:modified       "2022-09-05T19:36:19Z"^^xsd:dateTime ;
    dct:provenance     "Wikipedia contributors" ;
    dct:publisher      <https://www.epos-eu.org/epos-dcat-ap#
                           Organization/e892ce35-d4e1-468c-9df5-c0733faa444c> ;
    dct:spatial        <https://www.epos-eu.org/epos-dcat-ap#
                           Location/81c713e2-ed14-4e72-aa78-14b31823c8e6> ;
    dct:temporal       <https://www.epos-eu.org/epos-dcat-ap#
                           PeriodOfTime/3781753d-6d8a-460b-ba7d-11386779cb23> ;
    dct:title          "Municipalities of Trentino" ;
    dct:type           "dct:Dataset"^^xsd:anyURI ;
    dcat:contactPoint <https://www.epos-eu.org/epos-dcat-ap#
```

```

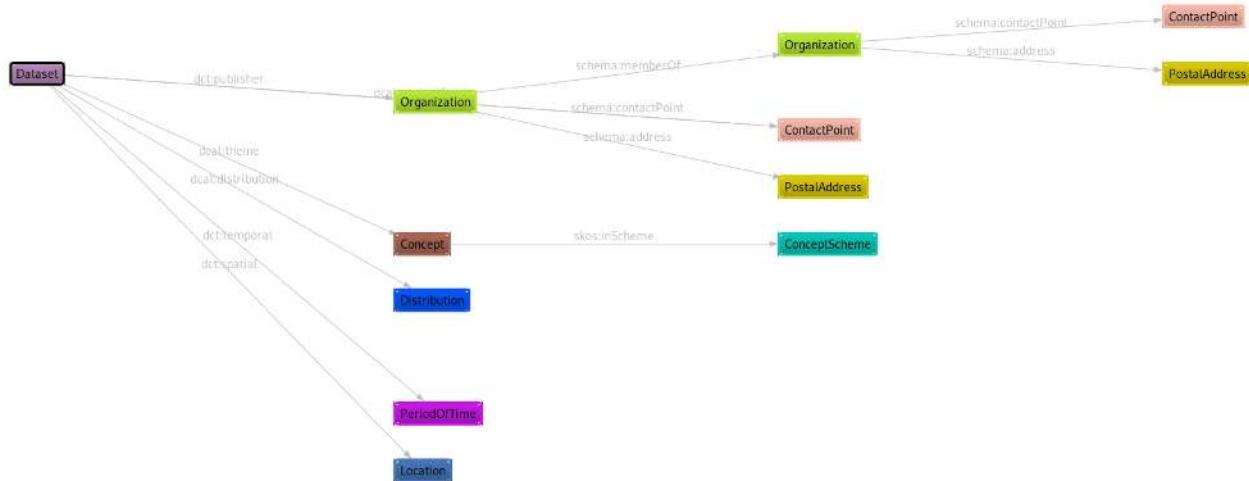
        ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;
        dcat:distribution <https://www.epos-eu.org/epos-dcat-ap#Distribution/2dee49b8-e959-4eaa-a3d1-48342784a61d> ;
        dcat:keyword "wikipedia, municipalities, Trentino, Italy" ;
        dcat:theme <https://www.epos-eu.org/epos-dcat-ap#Concept/9e6a61c3-4e43-4e6c-a53c-d75a2dc7bb> .

<https://www.epos-eu.org/epos-dcat-ap#Location/81c713e2-ed14-4e72-aa78-14b31823c8e6>
    rdf:type dct:Location .

<https://www.epos-eu.org/epos-dcat-ap#Concept/9e6a61c3-4e43-4e6c-a53c-d75a2dc7bb>
    rdf:type skos:Concept ;
    skos:definition "Trentino, Location, Space" ;
    skos:inScheme <https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636> ;
    skos:prefLabel "Space" .

```

3.10.2 School Courses metadata



School Courses Metadata

```

@prefix : <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix schema: <http://schema.org/> .
@prefix spdx: <http://spdx.org/rdf/terms#> .

```

```

@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix gsp: <http://www.opengis.net/ont/geosparql#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dqv: <http://www.w3.org/ns/dqv#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix sh: <http://www.w3.org/ns/shacl#> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix locn: <http://www.w3.org/ns/locn#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix adms: <http://www.w3.org/ns/adms#> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix http: <http://www.w3.org/2006/http#> .
@prefix dash: <http://datashapes.org/dash#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

```

<<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/91072b16-c630-4093-b376-138acb4b2da5>>

```

    rdf:type                  schema:ContactPoint ;
    schema:availableLanguage "it-IT" ;
    schema:contactType      "email" ;
    schema:email              "gruppoportale@provincia.tn.it" ;
    schema:telephone          "0461495111" .

```

<<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636>>

```

    rdf:type            skos:ConceptScheme ;
    dct:description   "Concept of the Trentino Study Course
Dataset" ;
    dct:title         "Dataset concept" .

```

<<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365>>

```

    rdf:type                  schema:ContactPoint ;
    schema:availableLanguage "it-IT" ;
    schema:contactType      "email" ;

```

```
schema:email           "serv.formazione@provincia.tn.it" .  
  
<https://www.epos-eu.org/epos-dcat-ap#Distribution/2dee49b8-e959-4eaa  
-a3d1-48342784a61d>  
    rdf:type          dcat:Distribution ;  
    dct:conformsTo   "http://www.w3.org/2001/XMLSchema-instance"  
                      ;  
    dct:description  "List of courses of study of Trentino  
                       schools along with description and status of activity" ;  
    dct:format        "http://www.w3.org/2001/XMLSchema-instance  
                       ^^^xsd:anyURI" ;  
    dct:identifier    "https://dati.trentino.it/dataset/823309b7-  
                       d59e-4bdc-ae14-9a3611c12b27/resource/23d28ec8-c2cc-41cc-80  
                       af-6f9443bbbca2/download/exportcorsi10092014.xml" ^^^xsd:  
                       anyURI ;  
    dct:issued        "2011-09-10T16:27:17Z" ^^^xsd:dateTime ;  
    dct:language      "it-IT" ;  
    dct:license       "https://creativecommons.org/publicdomain/  
                       zero/1.0/deed.it" ^^^xsd:anyURI ;  
    dct:modified      "2014-09-10T16:27:17Z" ^^^xsd:dateTime ;  
    dct:title         "exportcorsi10092014.xml" ;  
    dct:type          "dcat:Distribution" ^^^xsd:anyURI ;  
    dcat:accessURL   "https://dati.trentino.it/dataset/823309b7-  
                       d59e-4bdc-ae14-9a3611c12b27/resource/23d28ec8-c2cc-41cc-80  
                       af-6f9443bbbca2/download/exportcorsi10092014.xml" ^^^xsd:  
                       anyURI ;  
    dcat:downloadURL "https://dati.trentino.it/dataset/823309b7-  
                       d59e-4bdc-ae14-9a3611c12b27/resource/23d28ec8-c2cc-41cc-80  
                       af-6f9443bbbca2/download/exportcorsi10092014.xml" ^^^xsd:  
                       anyURI ;  
    dcat:mediaType    "application/xhtml+xml" .  
  
<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/3781753d-6d8a-460b  
-ba7d-11386779cb23>  
    rdf:type          dct:PeriodOfTime ;  
    schema:endDate   "2014-09-10T19:37:51Z" ^^^xsd:dateTime ;  
    schema:startDate  "2014-09-10T19:37:51Z" ^^^xsd:dateTime .  
  
<https://www.epos-eu.org/epos-dcat-ap#Organization/e892ce35-d4e1-468c  
-9df5-c0733faa444c>  
    rdf:type          schema:Organization ;  
    schema:address   <https://www.epos-eu.org/epos-dcat-ap#  
                      PostalAddress/985e9420-0d39-42db-836d-e4bd0b236a0c> ;  
    schema:contactPoint  <https://www.epos-eu.org/epos-dcat-ap#
```

ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;
schema:email "serv.formazione@provincia.tn.it" ;
schema:identifier "W05601" ;
schema:legalName "Servizio formazione professionale,
formazione terziaria e funzioni di sistema" ;
schema:leiCode "W05601" ;
schema:memberOf <<https://www.epos-eu.org/epos-dcat-ap#Organization/a7057281-4a43-4cd2-8556-372e40f8d710>> ;
schema:telephone "0461491377" ;
schema:url "<https://www.provincia.tn.it/Amministrazione/Strutture-organizzative/Servizio-formazione-prof.le-formazione-terziaria-e-funz.-di-sistema>"^^xsd:anyURI .

<<https://www.epos-eu.org/epos-dcat-ap#Dataset/8ab7280e-0772-4617-a052-fb38cab23de0>>
rdf:type dcat:Dataset ;
dct:accessRights "free" ;
dct:conformsTo "http://www.w3.org/2001/XMLSchema-instance" ;
dct:created "2014-09-10T19:36:18Z"^^xsd:dateTime ;
dct:description "List of study courses of Trentino schools
with description and status of activity." ;
dct:identifier "p_TN:90e4dd5d-f967-4a0a-854a-a8bcef48114b" ;
dct:issued "2014-09-10T19:36:19Z"^^xsd:dateTime ;
dct:language "it-IT" ;
dct:modified "2014-09-10T19:36:19Z"^^xsd:dateTime ;
dct:provenance "Servizio formazione professionale,
formazione terziaria e funzioni di sistema" ;
dct:publisher <<https://www.epos-eu.org/epos-dcat-ap#Organization/e892ce35-d4e1-468c-9df5-c0733faa444c>> ;
dct:spatial <<https://www.epos-eu.org/epos-dcat-ap#Location/81c713e2-ed14-4e72-aa78-14b31823c8e6>> ;
dct:temporal <<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/3781753d-6d8a-460b-ba7d-11386779cb23>> ;
dct:title "Corsi di studio delle scuole Trentine" ;
dct:type "dct:Dataset"^^xsd:anyURI ;
dcat:contactPoint <<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365>> ;
dcat:distribution <<https://www.epos-eu.org/epos-dcat-ap#Distribution/2dee49b8-e959-4eaa-a3d1-48342784a61d>> ;
dcat:keyword "education, istitution, school
organization" ;

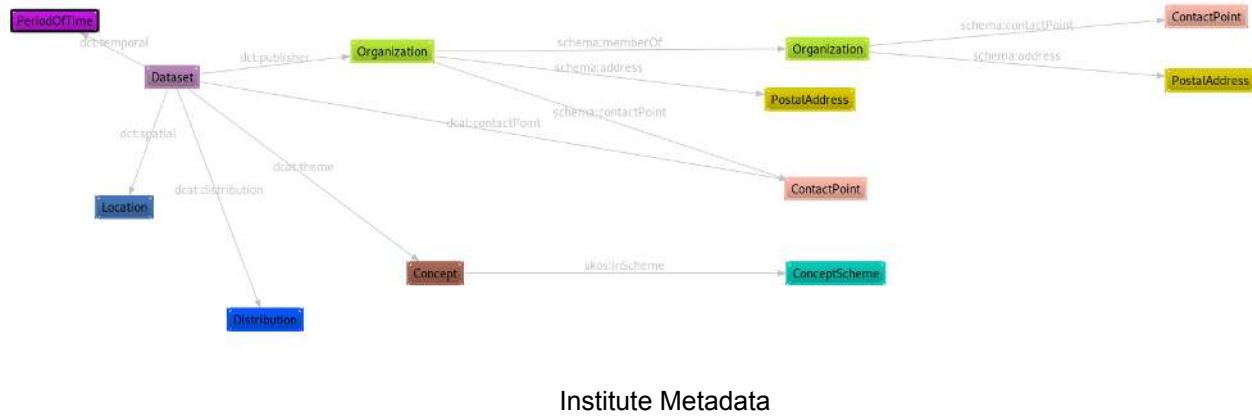
```
    dcat:theme <https://www.epos-eu.org/epos-dcat-ap#Concept/9e6a61c3-4e43-4e6c-a53c-d75a2dc7bb> .  
  
<https://www.epos-eu.org/epos-dcat-ap#Location/81c713e2-ed14-4e72-aa78-14b31823c8e6>  
    rdf:type dct:Location .  
  
<https://www.epos-eu.org/epos-dcat-ap#Organization/a7057281-4a43-4cd2-8556-372e40f8d710>  
    rdf:type schema:Organization ;  
    schema:address <https://www.epos-eu.org/epos-dcat-ap#PostalAddress/41881aae-6d13-468e-a1f6-1ad528a069e5> ;  
    schema:contactPoint <https://www.epos-eu.org/epos-dcat-ap#ContactPoint/91072b16-c630-4093-b376-138acb4b2da5> ;  
    schema:email "gruppoportale@provincia.tn.it" ;  
    schema:identifier "https://www.provincia.tn.it/  
        Amministrazione"^^xsd:anyURI ;  
    schema:legalName "Provincia Autonoma di Trento" ;  
    schema:leiCode "00337460224" ;  
    schema:telephone "0461495111" ;  
    schema:url "https://www.provincia.tn.it/  
        Amministrazione"^^xsd:anyURI .  
  
<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/985e9420-0d39-42db-836d-e4bd0b236a0c>  
    rdf:type schema:PostalAddress ;  
    schema:addressCountry "IT" ;  
    schema:addressLocality "Trento" ;  
    schema:postalCode "16121" ;  
    schema:streetAddress "15 Piazza Dante" .  
  
<https://www.epos-eu.org/epos-dcat-ap#Concept/9e6a61c3-4e43-4e6c-a53c-d75a2dc7bb>  
    rdf:type skos:Concept ;  
    skos:definition "Education, culture and sport" ;  
    skos:inScheme <https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636> ;  
    skos:prefLabel "Culture" .  
  
<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/41881aae-6d13-468e-a1f6-1ad528a069e5>  
    rdf:type schema:PostalAddress ;  
    schema:addressCountry "IT" ;  
    schema:addressLocality "Trento" ;
```

```

    schema:postalCode      "16121" ;
    schema:streetAddress   "15 Piazza Dante" .

```

3.10.3 Institute metadata



```

@prefix : <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix schema: <http://schema.org/> .
@prefix spdx: <http://spdx.org/rdf/terms#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix gsp: <http://www.opengis.net/ont/geosparql#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dqv: <http://www.w3.org/ns/dqv#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix sh: <http://www.w3.org/ns/shacl#> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix locn: <http://www.w3.org/ns/locn#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix adms: <http://www.w3.org/ns/adms#> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix http: <http://www.w3.org/2006/http#> .
@prefix dash: <http://datashapes.org/dash#> .

```

```
@prefix dc: <http://purl.org/dc/elements/1.1/> .  
  
<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/91072b16-c630-4093-b376-138acb4b2da5>  
    rdf:type schema:ContactPoint ;  
    schema:availableLanguage "it-IT" ;  
    schema:contactType "email" ;  
    schema:email "gruppoportale@provincia.tn.it" ;  
    schema:telephone "0461495111" .  
  
<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636>  
    rdf:type skos:ConceptScheme ;  
    dct:description "Concept of the School institutions in Trentino" ;  
    dct:title "Dataset concept" .  
  
<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365>  
    rdf:type schema:ContactPoint ;  
    schema:availableLanguage "it-IT" ;  
    schema:contactType "email" ;  
    schema:email "serv.formazione@provincia.tn.it" .  
  
<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/aef057f0-b812-4bf9-a03c-a769acac6a1d>  
    rdf:type dct:PeriodOfTime ;  
    schema:endDate "2014-09-10T19:57:19Z"^^xsd:dateTime ;  
    schema:startDate "2014-09-10T19:57:19Z"^^xsd:dateTime .  
  
<https://www.epos-eu.org/epos-dcat-ap#Distribution/2dee49b8-e959-4ea-a3d1-48342784a61d>  
    rdf:type dcat:Distribution ;  
    dct:conformsTo "http://www.w3.org/2001/XMLSchema-instance" ;  
    dct:description "List of educational institutions in Trentino, with information such as: Type of institution, Location, Educational offer, ID" ;  
    dct:format "http://www.w3.org/2001/XMLSchema-instance"^^xsd:anyURI ;  
    dct:identifier "https://dati.trentino.it/dataset/istituzioni-scolastiche-trentino"^^xsd:anyURI ;  
    dct:issued "2011-09-10T16:27:17Z"^^xsd:dateTime ;  
    dct:language "it-IT" ;
```

```
dct:license      "https://creativecommons.org/publicdomain/
    zero/1.0/deed.it"^^xsd:anyURI ;
dct:modified     "2014-09-10T16:27:17Z"^^xsd:dateTime ;
dct:rights       "free" ;
dct:title        "exportesco10092014.xml" ;
dct:type         "dcat:Distribution"^^xsd:anyURI ;
dcat:accessURL   "https://dati.trentino.it/dataset/18034fef
    -88e5-4d4d-906d-5fdc381303cc/resource/1a04a591-14d7-49c9-
    a011-d54bc427fa0f/download/exportesco10092014.xml"^^xsd:
    anyURI ;
dcat:downloadURL "https://dati.trentino.it/dataset/18034fef
    -88e5-4d4d-906d-5fdc381303cc/resource/1a04a591-14d7-49c9-
    a011-d54bc427fa0f/download/exportesco10092014.xml"^^xsd:
    anyURI ;
dcat:mediaType    " application/xhtml+xml" .

<https://www.epos-eu.org/epos-dcat-ap#Dataset/6173b3d1-3706-460f
-8525-a0cd33cd4d0b>
    rdf:type dcat:Dataset .

<https://www.epos-eu.org/epos-dcat-ap#Organization/e892ce35-d4e1-468c
-9df5-c0733faa444c>
    rdf:type             schema:Organization ;
    schema:address       <https://www.epos-eu.org/epos-dcat-ap#
    PostalAddress/985e9420-0d39-42db-836d-e4bd0b236a0c> ;
    schema:contactPoint <https://www.epos-eu.org/epos-dcat-ap#
    ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;
    schema:email          "serv.formazione@provincia.tn.it" ;
    schema:identifier     " W05601" ;
    schema:legalName      "Servizio formazione professionale,
    formazione terziaria e funzioni di sistema" ;
    schema:leiCode         " W05601" ;
    schema:memberOf        <https://www.epos-eu.org/epos-dcat-ap#
    Organization/a7057281-4a43-4cd2-8556-372e40f8d710> ;
    schema:telephone        "0461491377" ;
    schema:url              "https://www.provincia.tn.it/
    Amministrazione/Strutture-organizzative/Servizio-
    formazione-prof.le-formazione-terziaria-e-funz.-di-sistema
    "^^xsd:anyURI .

<https://www.epos-eu.org/epos-dcat-ap#Dataset/8ab7280e-0772-4617-a052
-fb38cab23de0>
    rdf:type dcat:Dataset ;
    dct:accessRights "free" ;
```

```
dct:conformsTo      "http://www.w3.org/2001/XMLSchema-instance"
"
dct:created         "2014-09-10T16:56:45Z"^^xsd:dateTime ;
dct:description     "List of educational institutions in
Trentino, with information such as: Type of institution,
webite, study plan, ID (Origin Code)" ;
dct:identifier      "p_TN:653ca277-0643-4834-badc-dc27ecc8e99e
"
;
dct:issued          "2014-09-10T19:48:43Z"^^xsd:dateTime ;
dct:language        "it-IT" ;
dct:modified        "2014-09-10T19:48:43Z"^^xsd:dateTime ;
dct:provenance      "Servizio formazione professionale,
formazione terziaria e funzioni di sistema" ;
dct:publisher       <https://www.epos-eu.org/epos-dcat-ap#
Organization/e892ce35-d4e1-468c-9df5-c0733faa444c> ;
dct:source          <https://www.epos-eu.org/epos-dcat-ap#
Dataset/6173b3d1-3706-460f-8525-a0cd33cd4d0b> ;
dct:spatial         <https://www.epos-eu.org/epos-dcat-ap#
Location/1d74288c-892a-4e85-bbb6-3d9c0b7aa0fe> ;
dct:temporal        <https://www.epos-eu.org/epos-dcat-ap#
PeriodOfTime/aef057f0-b812-4bf9-a03c-a769acac6a1d> ;
dct:title           "Istituzioni scolastiche del Trentino" ;
dct:type            "dct:Dataset"^^xsd:anyURI ;
dcat:contactPoint  <https://www.epos-eu.org/epos-dcat-ap#
ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;
dcat:distribution   <https://www.epos-eu.org/epos-dcat-ap#
Distribution/2dee49b8-e959-4eaa-a3d1-48342784a61d> ;
dcat:keyword        "education, istitution, school
organization" ;
dcat:theme          <https://www.epos-eu.org/epos-dcat-ap#
Concept/9e6a61c3-4e43-4e6c-a53c-d75a2dc7bb> .

<https://www.epos-eu.org/epos-dcat-ap#Organization/a7057281-4a43-4cd2
-8556-372e40f8d710>
rdf:type           schema:Organization ;
schema:address     <https://www.epos-eu.org/epos-dcat-ap#
PostalAddress/41881aae-6d13-468e-a1f6-1ad528a069e5> ;
schema:contactPoint <https://www.epos-eu.org/epos-dcat-ap#
ContactPoint/91072b16-c630-4093-b376-138acb4b2da5> ;
schema:email        "gruppoportale@provincia.tn.it" ;
schema:identifier   "https://www.provincia.tn.it/
Amministrazione"^^xsd:anyURI ;
schema:legalName    "Provincia Autonoma di Trento" ;
schema:leiCode      "00337460224" ;
```

```
    schema:telephone      "0461495111" ;
    schema:url           "https://www.provincia.tn.it/
                           Amministrazione"^^xsd:anyURI .

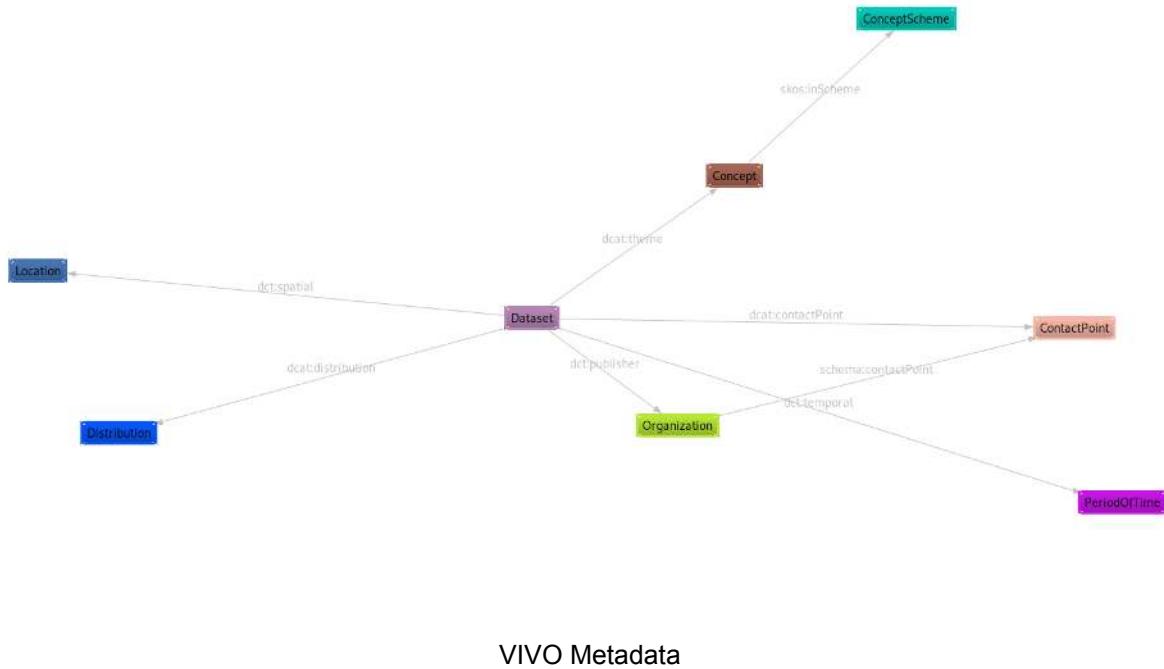
<https://www.epos-eu.org/epos-dcat-ap#Location/1d74288c-892a-4e85-
bbb6-3d9c0b7aa0fe>
    rdf:type          dct:Location ;
    locn:geometry   "POINT(11.116667 46.066666)"^^gsp:wktLiteral .

<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/985e9420-0d39-42
db-836d-e4bd0b236a0c>
    rdf:type          schema:PostalAddress ;
    schema:addressCountry "IT" ;
    schema:addressLocality "Trento" ;
    schema:postalCode     "16121" ;
    schema:streetAddress  "15 Piazza Dante" .

<https://www.epos-eu.org/epos-dcat-ap#Concept/9e6a61c3-4e43-4e6c-a53c
-d75a2dc7bb>
    rdf:type          skos:Concept ;
    skos:definition  "Education, culture and sport" ;
    skos:inScheme    <https://www.epos-eu.org/epos-dcat-ap#
                      ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636> ;
    skos:prefLabel   "Culture" .

<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/41881aae-6d13-468
e-a1f6-1ad528a069e5>
    rdf:type          schema:PostalAddress ;
    schema:addressCountry "IT" ;
    schema:addressLocality "Trento" ;
    schema:postalCode     "16121" ;
    schema:streetAddress  "15 Piazza Dante" .
```

3.10.4 Vivo metadata



VIVO Metadata

```
@prefix : <https://www.epos-eu.org/epos-dcat-ap#> .  
@prefix schema: <http://schema.org/> .  
@prefix spdx: <http://spdx.org/rdf/terms#> .  
@prefix owl: <http://www.w3.org/2002/07/owl#> .  
@prefix gsp: <http://www.opengis.net/ont/geosparql#> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .  
@prefix dqv: <http://www.w3.org/ns/dqv#> .  
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .  
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .  
@prefix oa: <http://www.w3.org/ns/oa#> .  
@prefix dct: <http://purl.org/dc/terms/> .  
@prefix sh: <http://www.w3.org/ns/shacl#> .  
@prefix dcat: <http://www.w3.org/ns/dcat#> .  
@prefix locn: <http://www.w3.org/ns/locn#> .  
@prefix foaf: <http://xmlns.com/foaf/0.1/> .  
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .  
@prefix adms: <http://www.w3.org/ns/adms#> .  
@prefix org: <http://www.w3.org/ns/org#> .  
@prefix cnt: <http://www.w3.org/2011/content#> .  
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix http: <http://www.w3.org/2006/http#> .
@prefix dash: <http://datashapes.org/dash#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636>
    rdf:type skos:ConceptScheme ;
    dct:description "Concept of the Vivoscuola Dataset" ;
    dct:title "Vivoscuola concept" .

<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365>
    rdf:type schema>ContactPoint ;
    schema:availableLanguage "it-IT" ;
    schema:contactType "None" .

<https://www.epos-eu.org/epos-dcat-ap#Distribution/2dee49b8-e959-4eaa-a3d1-48342784a61d>
    rdf:type dcat:Distribution ;
    dct:conformsTo "HTML" ;
    dct:description "List of school information" ;
    dct:format "HTML" ;
    dct:identifier "https://www.vivoscuola.it/Scuole2"^^xsd:anyURI ;
    dct:issued "2011-09-10T16:27:17Z"^^xsd:dateTime ;
    dct:language "en-EN" ;
    dct:modified "2022-09-05T16:27:17Z"^^xsd:dateTime ;
    dct:title "Vivoscuola, il portale della scuola in Trentino" ;
    dct:type "dcat:Distribution"^^xsd:anyURI ;
    dcat:accessURL "https://www.vivoscuola.it/Scuole2"^^xsd:anyURI ;
    dcat:downloadURL "https://www.vivoscuola.it/Scuole2"^^xsd:anyURI ;
    dcat:mediaType "text/html" .

<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/3781753d-6d8a-460b-ba7d-11386779cb23>
    rdf:type dct:PeriodOfTime ;
    schema:endDate "2022-09-05T19:37:51Z"^^xsd:dateTime ;
    schema:startDate "2009-01-01T19:37:51Z"^^xsd:dateTime .

<https://www.epos-eu.org/epos-dcat-ap#Organization/e892ce35-d4e1-468c>
```

-9df5-c0733faa444c>
 rdf:type schema:Organization ;
 schema:contactPoint <https://www.epos-eu.org/epos-dcat-ap#ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;
 schema:identifier "https://en.wikipedia.org/wiki/Municipalities_of_Trentino"^^xsd:anyURI ;
 schema:legalName "The list shows the municipalities (comuni) of the autonomous province of Trento, Italy" ;
 schema:leiCode "https://en.wikipedia.org/wiki/Municipalities_of_Trentino" ;
 schema:url "https://en.wikipedia.org/wiki/Municipalities_of_Trentino"^^xsd:anyURI .

<https://www.epos-eu.org/epos-dcat-ap#Dataset/8ab7280e-0772-4617-a052-fb38cab23de0>
 rdf:type dcat:Dataset ;
 dct:accessRights "free" ;
 dct:created "2009-01-01T00:00:00Z"^^xsd:dateTime ;
 dct:description "Dataset about schools in Trentino" ;
 dct:identifier "https://www.vivoscuola.it/Scuole2"^^xsd:anyURI ;
 dct:issued "2011-04-25T19:36:19Z"^^xsd:dateTime ;
 dct:language "en-EN" ;
 dct:modified "2022-09-05T19:36:19Z"^^xsd:dateTime ;
 dct:provenance "Vivoscuola Trentino" ;
 dct:publisher <https://www.epos-eu.org/epos-dcat-ap#Organization/e892ce35-d4e1-468c-9df5-c0733faa444c> ;
 dct:spatial <https://www.epos-eu.org/epos-dcat-ap#Location/81c713e2-ed14-4e72-aa78-14b31823c8e6> ;
 dct:temporal <https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/3781753d-6d8a-460b-ba7d-11386779cb23> ;
 dct:title "Vivoscuola, il portale della scuola in Trentino" ;
 dct:type "dct:Dataset"^^xsd:anyURI ;
 dcat:contactPoint <https://www.epos-eu.org/epos-dcat-ap#ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;
 dcat:distribution <https://www.epos-eu.org/epos-dcat-ap#Distribution/2dee49b8-e959-4eaa-a3d1-48342784a61d> ;
 dcat:keyword "schools, Trentino, Italy" ;
 dcat:theme <https://www.epos-eu.org/epos-dcat-ap#Concept/9e6a61c3-4e43-4e6c-a53c-d75a2dc7bb> .

<https://www.epos-eu.org/epos-dcat-ap#Location/81c713e2-ed14-4e72-aa78-14b31823c8e6>

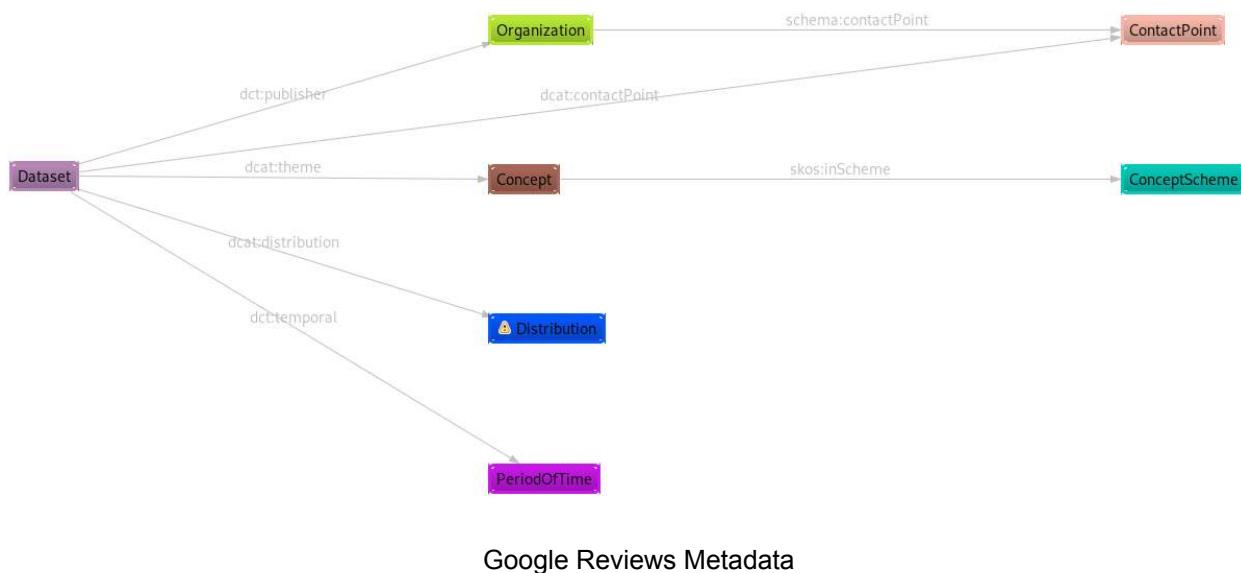
```

    rdf:type      dct:Location ;
    locn:geometry "POINT(11.116667 46.066666)"^^gsp:wktLiteral .

<https://www.epos-eu.org/epos-dcat-ap#Concept/9e6a61c3-4e43-4e6c-a53c
-d75a2dc7bb>
    rdf:type      skos:Concept ;
    skos:definition "Trentino, Italy, School, Vivoscuola" ;
    skos:inScheme <https://www.epos-eu.org/epos-dcat-ap#
    ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636> ;
    skos:prefLabel "School" .

```

3.10.5 Google Reviews metadata



```

@prefix : <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix schema: <http://schema.org/> .
@prefix spdx: <http://spdx.org/rdf/terms#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix gsp: <http://www.opengis.net/ont/geosparql#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dqv: <http://www.w3.org/ns/dqv#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix dct: <http://purl.org/dc/terms/> .

```

```
@prefix sh: <http://www.w3.org/ns/shacl#> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix locn: <http://www.w3.org/ns/locn#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix adms: <http://www.w3.org/ns/adms#> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix http: <http://www.w3.org/2006/http#> .
@prefix dash: <http://datashapes.org/dash#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636>
    rdf:type skos:ConceptScheme ;
    dct:description "Concept of the Google Reviews Dataset" ;
    dct:title "Google maps reviews concept" .

<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365>
    rdf:type schema>ContactPoint ;
    schema:availableLanguage "en-EN" ;
    schema:contactType "website" .

<https://www.epos-eu.org/epos-dcat-ap#Distribution/2dee49b8-e959-4ea-a3d1-48342784a61d>
    rdf:type dcat:Distribution ;
    dct:description "List of Google Maps reviews for the schools" ;
    dct:identifier "https://support.google.com/business/answer/3474122?hl=en"^^xsd:anyURI ;
    dct:issued "2022-11-10T16:27:17Z"^^xsd:dateTime ;
    dct:language "it-IT" ;
    dct:modified "2022-11-10T16:27:17Z"^^xsd:dateTime ;
    dct:title "Google reviews html text" ;
    dct:type "dcat:Distribution"^^xsd:anyURI ;
    dcat:accessURL "https://support.google.com/business/answer/3474122?hl=en"^^xsd:anyURI ;
    dcat:downloadURL "https://support.google.com/business/answer/3474122?hl=en"^^xsd:anyURI ;
    dcat:mediaType "text/html" .
```

```
<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/3781753d-6d8a-460b  
-ba7d-11386779cb23>  
    rdf:type          dct:PeriodOfTime ;  
    schema:endDate   "2022-11-10T19:37:51Z"^^xsd:dateTime ;  
    schema:startDate "2022-11-10T19:37:51Z"^^xsd:dateTime .  
  
<https://www.epos-eu.org/epos-dcat-ap#Organization/e892ce35-d4e1-468c  
-9df5-c0733faa444c>  
    rdf:type          schema:Organization ;  
    schema:contactPoint <https://www.epos-eu.org/epos-dcat-ap#  
                          ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;  
    schema:identifier   "W05601" ;  
    schema:legalName    "Google" ;  
    schema:leiCode      "W05601" ;  
    schema:url         "https://www.google.com/business/"^^xsd:  
                        anyURI .  
  
<https://www.epos-eu.org/epos-dcat-ap#Dataset/8ab7280e-0772-4617-a052  
-fb38cab23de0>  
    rdf:type          dcat:Dataset ;  
    dct:accessRights  "free" ;  
    dct:created        "2022-11-10T19:36:18Z"^^xsd:dateTime ;  
    dct:description    "Google maps reviews" ;  
    dct:identifier     "p_TN:90e4dd5d-f967-4a0a-854a-a8bcef48114b  
                        " ;  
    dct:issued         "2014-09-10T19:36:19Z"^^xsd:dateTime ;  
    dct:language        "it-IT" ;  
    dct:modified        "2022-11-10T19:36:19Z"^^xsd:dateTime ;  
    dct:provenance      "Google reviews users" ;  
    dct:publisher       <https://www.epos-eu.org/epos-dcat-ap#  
                          Organization/e892ce35-d4e1-468c-9df5-c0733faa444c> ;  
    dct:temporal        <https://www.epos-eu.org/epos-dcat-ap#  
                          PeriodOfTime/3781753d-6d8a-460b-ba7d-11386779cb23> ;  
    dct:title           "Google Maps reviews" ;  
    dct:type            "dct:Dataset"^^xsd:anyURI ;  
    dcat:contactPoint  <https://www.epos-eu.org/epos-dcat-ap#  
                          ContactPoint/d3a0d642-84d5-4ebc-9ed6-e520f3340365> ;  
    dcat:distribution   <https://www.epos-eu.org/epos-dcat-ap#  
                          Distribution/2dee49b8-e959-4eaa-a3d1-48342784a61d> ;  
    dcat:keyword         "education, istitution, school  
                        organization, reviews, google maps, google, school" ;  
    dcat:theme           <https://www.epos-eu.org/epos-dcat-ap#  
                          Concept/9e6a61c3-4e43-4e6c-a53c-d75a2dc7bb> .
```

```
<https://www.epos-eu.org/epos-dcat-ap#Concept/9e6a61c3-4e43-4e6c-a53c  
-d75a2dc7bb>  
    rdf:type      skos:Concept ;  
    skos:definition "Google maps reviews text" ;  
    skos:inScheme  <https://www.epos-eu.org/epos-dcat-ap#  
                    ConceptScheme/7cb615e2-9ecd-43d0-b917-17eebbd50636> ;  
    skos:prefLabel "Reviews" .
```

4 Informal Modeling

The informal modeling process is described in detail in this section. Similar to the previous part, this one tries to highlight the many sub-activities carried out by every team member as well as the results of the phase.

The following actions are described in further depth in this section:

- Purpose formalization updated for the Informal Modeling part of the project;
- Entity-Relationship Model definition.

4.1 Purpose Formalization

The informal modeling step of *iTelos* is a crucial linking stage that allows us to create the fundamental Etypes and draws conclusions about their related attributes from competency questions. The ER model is built using the retrieved Etypes and characteristics. There are two sections of the entire operation: the first one relates to the knowledge layer, while the other is concerned with the data layer. The knowledge layer's critical step in this phase is intended to transform analyzed competency questions into classes and attributes.

The definition of the Competency Questions allowed us to define the following Entities:

- School
- School Statistics
- Region
- Institute
- Institute Contact Information
- Subregional Academic Division
- Responsible Authority
- Invalsi Score
- Study Course
- Contact Information
- Review
- Education Quality
- Professor

During the inception phase, we noticed that looking at the datasets we gathered, there is an important relationship between an *Institute* and a *School*. In the Italian legal system, institutes are public or private organizations tasked with the responsibility of imparting education at the many stages into which it is split (elementary and secondary schools for instance). Some of them



have a distinctive name, while others incorporate the name of the institution in their denomination. As a result, an institution can have numerous schools of various stages (for example, an institute can include an elementary and middle school) or it can just have one school, such as the Buonarroti technical institute¹⁸, which is considered as an institute but it is only composed by a single high school. Furthermore, many of the characteristics that the school would hold are instead owned by an institution: for example, schools that are part of an institution share the same school principals. Moreover, it has been noticed that an institution may have many people in charge of various functions, namely the School Principal, President of the Managing Body, Director, Pedagogical Coordinator, and in a few cases, Sorastant.

Another critical point to emphasize regards the dataset holding information on the percentage of students admitted to the following year. The dataset has the following structure:

- Stage of the School
- Timetable type of the School
- Number of Years of the School
- Percentage of admitted students
- School Year
- MIUR code

We have observed that if an institution has numerous schools that share the same stage, the percentage of admitted students present in the dataset is an average of those schools. Nonetheless, we have opted to attribute that value to every school at a specific stage rather than assign the property to the institute to which they belong. The latter choice would have implied creating an entity that collects information about the institute per year, which, according to our purpose, would not be relevant.

4.2 ER Model

We can determine the necessary Etype(s), and their object attributes as well as data properties based on our Competency Questions and the acquired datasets. The key pillars of the ER model are the definition of Etype(s) specification and the development of various Etype(s) relationships.

The classes of the Etypes are also made clear. School, Institute, Region, and Contact Information are the common Etypes, while Subregional Academic Division, Institute Contact Information, Responsible Authority, and Professor are the core Etypes; instead, Invalsi Score, Review, School Statistics, Study Course and Education Quality are marked as contextual Etypes as it is shown in Figure 27.

Furthermore, the data type we have employed are only those supported by Protégé. As a matter of fact, this has prevented us from using enumerations and type constraints which are fundamental for a database in order to guarantee data integrity. Regardless, we enforced such constraints when we were dealing with the data. The same applies to dates, which have been modeled in the ontology as strings.

¹⁸<https://www.buonarroti.tn.it/>

In the following subsections, we describe the entity we have introduced, divided into common, core, and contextual. Moreover, we motivate some of the design choices we have performed to build the ER Model.

4.2.1 Common Etypes

- **Region:** a common type entity describing the domain of interest, both in spatial and temporal terms. In our case, this entity complies with the *Dol* we have defined in Section 2.7 and refers to the Trentino Region in the temporal domain ranging from 2020 to 2022. We have identified the following attributes:
 - **Lat:** the region latitude;
 - **Lon:** the region longitude;
 - **StartYear:** the starting year of our temporal domain;
 - **EndYear:** the ending year of our temporal domain.
- **School:** a common type entity describing any school associated with a building. As attributes we have identified:
 - **Name:** names of the school;
 - **Address:** address of the school building;
 - **Type:** attribute which identifies the type of the school. It could be "Provinciale", "Delega F.P.S.M.", "Gestione Autonoma" or "Delega Co.E.S.I";
 - **MIUR code:** attribute whose means is to identify the school, issued by *Ministero dell'istruzione, dell'università e della ricerca*;
 - **Stage:** attribute which identifies the stage of the school, it could be "Primaria", "Secondaria di Primo Grado", etc;
 - **Timetable type:** if it is "SERALE", it specifies that the school offers courses in the evening. Otherwise, if it is "DIURNO", it means that the lectures take place during the morning or during the afternoon;
 - **Number of Years:** attribute which specifies the total number of years that has a school (for instance, a primary school has *Number of Years* equal to 5).
- **Contact Information:** a common type entity describing a generic contact type information of two types: email and phone number. As attributes we have identified:
 - **Phone Number:** attribute which identifies the phone number of the contact information;
 - **Email:** attribute which identifies the email of the contact information.
 - **Office address**
- **Institute:** a common type entity that represents an institute. As attributes we have identified:
 - **Name:** name of the institute;

-
- **Type**: whether the institute is religious, secular, public, or private;
 - **Website**: website URL of the institute;
 - **Address**: address of the institute;
 - **Number of schools**: the number of schools which are included in the institute;
 - **MIUR code**: MIUR code of the institute;
 - **PAT code**: PAT code of the institute, namely a code assigned by the autonomous province of Trento.

We have decided not to put the specialization of School, such as High School, Middle School, and Kindergarten since, according to the data we have, they will not stand as standalone entities as they do not have either specific data properties or object properties. For this reason, we have decided to model this concept as an attribute in the *School* entity.

Moreover, as it will be clear from the next subsection, Subregional Academic Division cannot be modeled as a common entity, because it has information that is strictly related to the purpose of our project.

4.2.2 Core Etypes

- **Professor**: a core type entity that depicts a Professor. Along with the base information that a person can have, such as name and surname, Professor contains the data needed to identify a teacher in the university context. As attributes we have identified:
 - **CUN**: specifies the area of the *Consiglio Universitario Nazionale* associated with the professor;
 - **SSD**: indicates the area in which the professor works;
 - **Name**: refers to the name of the professor;
 - **Surname**: refers to the surname of the professor.
- **Responsible Authority**: a core type entity which depicts a Responsible Authority of an Institute. This entity is a specialization of a Person. Along with the base information that a person can have, such as name and surname, Responsible Authority contains the title, which is employed in order to identify the Responsible Authority mansion in the institutional context. As attributes we have identified:
 - **Title**: specifies the occupation of the Responsible Authority in the Institute context;
 - **Name**: refers to the name of the Responsible Authority;
 - **Surname**: refers to the surname of the Responsible Authority.
- **Subregional Academic Division**: a core type entity that represents a location in Trentino that hosts a single or more school. In our scenario, Subregional Academic Division is more than a location since it embeds data that is associated with the context, such as the graduate rate or the number of students. As attributes we have identified:
 - **Name**: indicates the name of the location;

- **Number of Students Enrolled in Kindergarten:** property that indicates how many students are enrolled in Kindergartens;
 - **Number of Students Enrolled in Primary School:** property that indicates how many students are enrolled in Primary Schools;
 - **Number of Students Enrolled in Middle School:** property that indicates how many students are enrolled in Middle Schools;
 - **Number of Students Enrolled in High school:** property that indicates how many students are enrolled in High schools;
 - **Number of Students Enrolled in CF:** property that indicates how many students are enrolled in Centri di Formazione (Professional Schools);
 - **School Completion Rate:** represents the school completion rate of that location;
 - **Graduate Rate:** represents the graduate rate of that location.
- **Institute Contact Information:** a core type entity that represents a set of specialized contact information belonging to the institute authorities. As attributes we have identified:
 - **Phone Number:** phone number of the school;
 - **Fax Number:** fax number of the institute;
 - **Institute Mail:** main email address of the institute;
 - **Management Mail:** email address of the management authorities;
 - **Office Mail:** email address of the administration office;
 - **institute mail:** institute mail of the institute.

Concerning the core entities, we have decided to do not to include the specialization of each Responsible Authority as standalone Etypes. The reason behind our choice is that we do not have enough data to properly characterize those entities, and in order to satisfy our purpose, it is sufficient to simply list their role.

The reason why we have modeled entities depicting contact information is strictly related to how we have defined the competency questions. In particular, all the personas and the associated scenarios always refer to a way of contacting the institute, however, they never define the means.

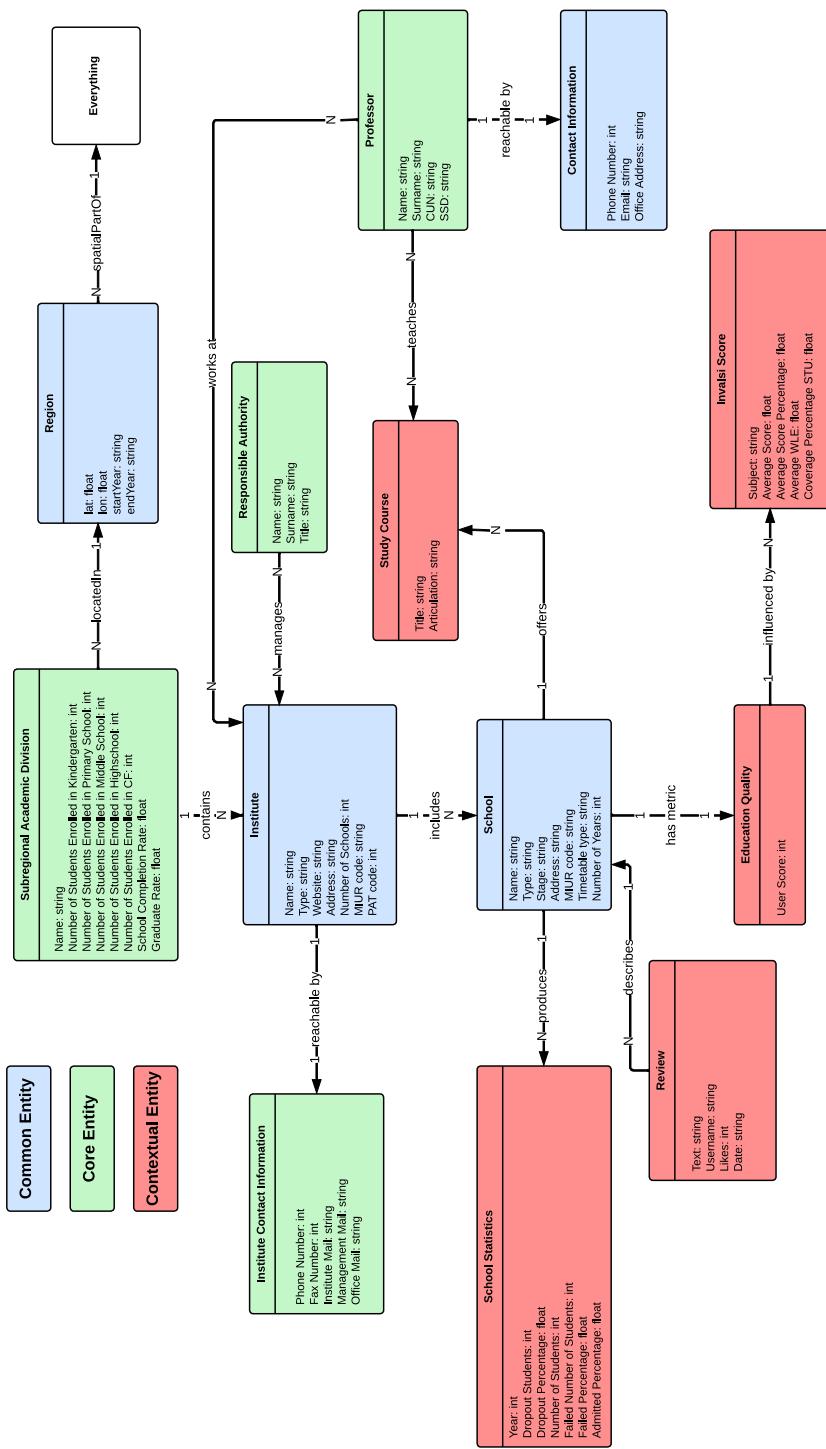
4.2.3 Contextual Etypes

- **Review:** a contextual type entity that represents a Review, which is associated with a school. As attributes we have identified:
 - **Text:** the content of the review;
 - **Username:** the username of the creator of the review;
 - **Likes:** the number of likes associated to that review;
 - **Date:** the date of the review creation.

- **School Statistics:** this contextual entity represents all the statistics associated with the school during an academic year. This entity has the following list of properties:
 - **Year:** indicates the academic year of the associated statistics;
 - **Dropout Students:** indicates how many students drop out from school during that specific academic year;
 - **Dropout Percentage:** indicates the percentage of dropout students during that specific academic year;
 - **Number of Students:** indicates how many students there are in a particular year;
 - **Failed Number of Students:** indicates how many students failed the school year;
 - **Failed Percentage:** indicates the percentage of failed students during that specific year;
 - **Admitted Percentage:** indicates the percentage of admitted students during the reference year.
- **Invalsi Score:** a contextual entity that describes the score of the school associated with the Invalsi test during 2022. its properties are:
 - **Subject:** subject of the test, such as *Math* or *Italian*;
 - **Average Score:** the average score of the students in that test;
 - **Average Score Percentage:** the average score in percentage in that test;
 - **Average WLE:** average WLE score, which is an ability estimation according to the Rasch model.
 - **Coverage Percentage STU:** Percentage of students who have taken part in the exam.
- **Study Course:** this entity represents a generic course, either a school or a university course. Its properties are:
 - **Title:** indicates the name of the course.
 - **Articulation:** indicates the articulation of the course, namely the area of interest.
- **Education Quality:** this contextual entity represents the abstract concept "Education Quality". Quality has different meanings among different individuals: for instance, for some, a school can provide a quality education if its students perform well on national tests, while for others it may be more relevant to look for students' opinions. For this reason, according to the data we were able to collect, we have defined a notion of quality that deals with quantitative data. We are aware of the fact that we are modeling a qualitative and subjective concept, therefore we expect that this concept can be integrated with more information.
 - **User Score:** indicates the Google Reviews rating (from 1 to 5) assigned to the institute by the Google Maps user.

To conclude, according to the output of the Inception phase, namely the OWL RDF/XML files of entity types produced, along with the properties informed by the competency questions, we have built the visual diagram of the Teleology as Figure 27 shows.



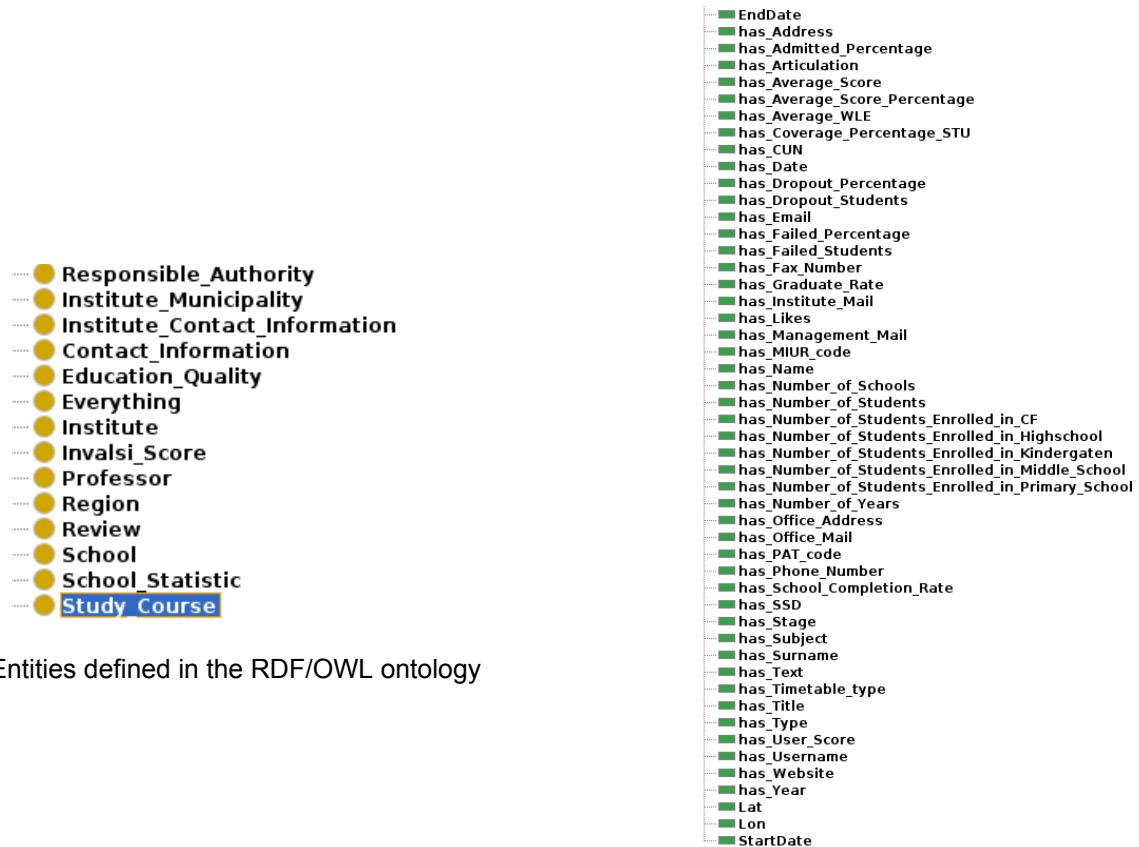


ER Model

4.2.4 Teleology OWL

Once the teleology UML has been modeled, we have realized the RDF/OWL ontology thanks to the software *Protégé*. In this section, you can find some screenshots documenting the teleology construction as well as some comments.

Figure 29 and Figure 28 show the work that has been done in *Protégé* in order to define the ontology of the previously shown teleology.



Entities defined in the RDF/OWL ontology

A subset of data properties defined in the RDF/OWL ontology

Furthermore, we employed object restrictors in order to enforce indirectly an association between Etypes (one to one, one to many, many to one, many to many). We have not used the cardinality constraints due to the following reasons:

- Cardinality constraints will make the logical axiomatization unnecessarily to higher-level flavors of OWL;
- cardinality can produce problems in integrating data.

The images here show the modeling in *Protégé*.

File Edit View Reasoner Tools Refactor Window Help

< > urn:webprotege:ontology:a3b24373-383b-4086-b82b-ba985401fa6b (urn:webprotege:ontology:a3b24373-383b-4086-b82b-ba985401fa6b)
Subregional_Academic_Division

Active ontology x Entities x Individuals by class x DL Query x

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Subregional_Academic_Division

Annotations Usage

Annotations: Subregional_Academic_Division

Annotations +
rdfs:label [language: en-us]
Subregional Academic Subdivision

Description: Subregional_Academic_Division

Equivalent To +

SubClass Of +
contains some Institute
locatedIn some Region

owl:Thing

- Subregional Academic Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School
- School_Statistic
- Study_Course

Asserted

File Edit View Reasoner Tools Refactor Window Help

< > urn:webprotege:ontology:a3b24373-383b-4086-b82b-ba985401fa6b (urn:webprotege:ontology:a3b24373-383b-4086-b82b-ba985401fa6b)
Contact_Information

Active ontology x Entities x Individuals by class x DL Query x

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Contact_Information

Annotations Usage

Annotations: Contact_Information

Annotations +
rdfs:label [language: en-us]
Contact Information

Description: Contact_Information

Equivalent To +

SubClass Of +
reachable_by some Professor

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School
- School_Statistic
- Study_Course

Asserted

Class hierarchy: Education_Quality

Annotations Usage

Annotations: Education_Quality

rdfs:label [language: en-us]
Education Quality

Description: Education_Quality

Equivalent To +

SubClass Of +

- has_metric some School
- influenced_by some Invalsi_Score

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Everything

Annotations Usage

Annotations: Everything

Annotations +

Description: Everything

Equivalent To +

SubClass Of +

- spatialPartOf some Region

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Institute

Annotations Usage

Annotations: Institute

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School
- School_Statistic
- Study_Course

Asserted

Annotations +

rdfs:label [language: en-us]
Institute

Description: Institute

Equivalent To +

SubClass Of +

- contains some Subregional_Academic_Division
- includes some School
- manages some Responsible_Authority
- reachable_by some Institute_Contact_Information
- works_at some Professor

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Institute_Contact_Information

Annotations Usage

Annotations: Institute_Contact_Information

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School
- School_Statistic
- Study_Course

Asserted

Annotations +

rdfs:label [language: en-us]
Institute Information

Description: Institute_Contact_Information

Equivalent To +

SubClass Of +

- reachable_by some Institute

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Invalsi_Score

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School
- School_Statistic
- Study_Course

Annotations Usage

Invalsi_Score

Annotations +

rdfs:label [language: en-us]
Invalsi Score

Description: Invalsi_Score

Equivalent To +

SubClass Of +

influenced_by some Education_Quality

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Professor

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School
- School_Statistic
- Study_Course

Annotations Usage

Professor

Annotations +

rdfs:label [language: en-us]
Professor

Description: Professor

Equivalent To +

SubClass Of +

reachable_by some Contact_Information
teaches some Study_Course
works_at some Institute

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Region

Asserted

Annotations Usage

Annotations: Region

Annotations + rdfs:label [language: en-us] Region

owl:Thing Subregional_Academic_Division Contact_Information Education_Quality Everything Institute Institute_Contact_Information Invalsi_Score Professor Region Responsible_Authority Review School School_Statistic Study_Course

Description: Region

Equivalent To +

SubClass Of + locatedIn some Subregional_Academic_Division spatialPartOf some Everything

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Responsible_Authority

Asserted

Annotations Usage

Annotations: Responsible_Authority

Annotations + rdfs:label [language: en-us] Responsible

owl:Thing Subregional_Academic_Division Contact_Information Education_Quality Everything Institute Institute_Contact_Information Invalsi_Score Professor Region Responsible_Authority Review School School_Statistic Study_Course

Description: Responsible_Authority

Equivalent To +

SubClass Of + manages some Institute

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Review

Annotations Usage

Annotations: Review

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review**
- School
- School_Statistic
- Study_Course

Annotations +

rdfs:label [language: en-us] Review

Description: Review

Equivalent To +

SubClass Of +

describes some School

Active ontology x Entities x Individuals by class x DL Query x

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: School

Annotations Usage

Annotations: School

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School**
- School_Statistic
- Study_Course

Annotations +

rdfs:label [language: en-us] School

Description: School

Equivalent To +

SubClass Of +

- describes some Review**
- has_metric some Education_Quality**
- includes some Institute**
- offers some Study_Course**
- produces some School_Statistic**

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: School_Statistic

Annotations Usage

Annotations: School_Statistic

Annotations +

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School
- School Statistic**
- Study_Course

Asserted

Description: School_Statistic

Equivalent To +

SubClass Of +

produces some School

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: Study_Course

Annotations Usage

Annotations: Study_Course

Annotations +

owl:Thing

- Subregional_Academic_Division
- Contact_Information
- Education_Quality
- Everything
- Institute
- Institute_Contact_Information
- Invalsi_Score
- Professor
- Region
- Responsible_Authority
- Review
- School
- School_Statistic
- Study Course**

Asserted

Description: Study_Course

Equivalent To +

SubClass Of +

offers some School

teaches some Professor

4.3 Metadata

Along with the ER model, we have produced the associated Metadata compliant with the DCAT [2] vocabulary, thanks to the SHAPENess [9] metadata editor.





Graph of the ER Model Metadata

```

@prefix : <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix schema: <http://schema.org/> .
@prefix spdx: <http://spdx.org/rdf/terms#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix gsp: <http://www.opengis.net/ont/geosparql#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dqv: <http://www.w3.org/ns/dqv#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix sh: <http://www.w3.org/ns/shacl#> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix locn: <http://www.w3.org/ns/locn#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix adms: <http://www.w3.org/ns/adms#> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix http: <http://www.w3.org/2006/http#> .
@prefix dash: <http://datashapes.org/dash#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

```

```
<https://www.epos-eu.org/epos-dcat-ap#Agent/828bf6b9-9f4b-4355-bcd0-d68b188811cb>
    rdf:type    foaf:Agent ;
    dct:type    <https://www.epos-eu.org/epos-dcat-ap#Concept/74cfe6d8-39b1-4491-a4ca-dba88fdccf10> ;
    foaf:name   "Erich Robbi" .

<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/84efa52e-3879-4d09-b0b9-36362e915866>
    rdf:type          dct:PeriodOfTime ;
    schema:endDate   "2022-11-23T00:00:00Z"^^xsd:dateTime ;
    schema:startDate "2022-11-23T00:00:00Z"^^xsd:dateTime .

<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/a21146bd-a38c-49fe-a02f-f8575d1d73e2>
    rdf:type          skos:ConceptScheme ;
    dct:description  "The position or rank of someone or something when compared to others in a society, organization, group, etc." ;
    dct:title        "Status" .

<https://www.epos-eu.org/epos-dcat-ap#Dataset/3a2e0c02-aebe-44d1-b688-a5e21d81d088>
    rdf:type          dcat:Dataset ;
    dct:accessRights "Public" ;
    dct:accrualPeriodicity "Once"^^xsd:anyURI ;
    dct:conformsTo   "iTilos" ;
    dct:created      "2022-11-23T00:00:00Z"^^xsd:dateTime ;
    dct:description  "This research seeks to provide a complete analysis of the educational facilities in Trentino\nwhich comprehend geospatial and temporal domains.\n\nThe Trentino area has a wide range of educational institutions, including public and private\nkindergartens, elementary schools, secondary schools, universities, and other vocational training facilities. The information regarding the school facilities in Trentino as well as the municipalities of the autonomous region of Trento is part of the project domain. Moreover, the courses\noffered by each educational facility with their associated duration are included.\nInitially, the data we have managed to collect ranges between 2014 and 2022. We had extremely updated information concerning university
```

courses and general institutional information,\\nand outdated material concerning 'schools courses. However, with additional data scraping, we\\nmanaged to obtain more recent data. We claim that the time boundary is restricted to the current\\nfew years (2020 - 2022).", "Samuele Bortolotti (<https://github.com/samuelebortolotti>) and Erich Robbi (<https://github.com/erich-r>) are the members assigned to the realization and finalization.", "For more information see the following websites: <https://github.com/samuelebortolotti/Education-Trentino>", "ER Model Metadata Regarding Education in Trentino" ;

dct:identifier "https://github.com/samuelebortolotti/Education-Trentino/tree/main/Teleologies/Informal%20Modeling"^^xsd:anyURI ;

dct:issued "2022-11-23T00:00:00Z"^^xsd:dateTime ;

dct:language "English" ;

dct:modified "2022-11-23T00:00:00Z"^^xsd:dateTime ;

dct:publisher <<https://www.epos-eu.org/epos-dcat-ap#Agent/f9cfb89e-3bb9-41ba-b263-e0386fce30ed>> , <<https://www.epos-eu.org/epos-dcat-ap#Organization/2dad489f-4ce3-4e0d-9122-b70df55b8ef2>> , <<https://www.epos-eu.org/epos-dcat-ap#Agent/828bf6b9-9f4b-4355-bcd0-d68b188811cb>> ;

dct:temporal <<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/84efa52e-3879-4d09-b0b9-36362e915866>> ;

dct:title "Entity Relational Model Metadata" ;

dct:type "Collection"^^xsd:anyURI ;

dcat:contactPoint <<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/3a8a3044-3e7a-463d-94d9-15c4f4822938>> , <<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/ca07b831-220f-4c49-bc3b-6e8215eb3238>> ;

dcat:distribution <<https://www.epos-eu.org/epos-dcat-ap#Distribution/2d8e5529-0e68-4e0d-af09-49a96f1c6c25>> ;

dcat:keyword "Knowledge and Data integration" , "Erich Robbi" , "UnitN" , "University of Trento" , "Trentino" , "Samuele Bortolotti" , "iTilos" , "Knowledge and Graph Engineering" ;

dcat:landingPage "<https://github.com/samuelebortolotti/Education-Trentino>" ;

dcat:theme <<https://www.epos-eu.org/epos-dcat-ap#Concept/4040fa1c-755a-4706-9fcdb2e9d6fb755>> , <<https://www.epos-eu.org/epos-dcat-ap#Concept/520bf2b3-6fcc-45ad-971a-b17df5fe98f0>> ;

```
foaf:page           "https://github.com/samuelebortolotti
/Education-Trentino" .
```

<<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/b0bb325d-f408-4dac-80e0-2389d8efb97f>>

```
    rdf:type          schema:PostalAddress ;
    schema:addressCountry "IT" ;
    schema:addressLocality "Povo (TN)" ;
    schema:postalCode    "I-38123" ;
    schema:streetAddress "Via Sommarive 9" .
```

<<https://www.epos-eu.org/epos-dcat-ap#Concept/4040fa1c-755a-4706-9fcdb2e9d6fb755>>

```
    rdf:type          skos:Concept ;
    skos:definition  "Education Institute is the facilities where
                      people learn and improve themselves through practice and
                      theory." ;
    skos:inScheme     <https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/f913bd20-8e7b-4d82-9688-ca745a432f24> ;
    skos:prefLabel    "Educational Institute" .
```

<<https://www.epos-eu.org/epos-dcat-ap#Concept/74cf6d8-39b1-4491-a4ca-dba88fdccf10>>

```
    rdf:type          skos:Concept ;
    skos:definition  "Member" ;
    skos:inScheme     <https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/17ebceb0-a50b-4801-82df-cc0d484ffa90> ;
    skos:prefLabel    "Project Member" .
```

<<https://www.epos-eu.org/epos-dcat-ap#Concept/b9645f82-9500-414c-a51d-255ff657c69e>>

```
    rdf:type          skos:Concept ;
    skos:definition  "Containing all the necessary parts, answers
                      , or information." ;
    skos:inScheme     <https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/a21146bd-a38c-49fe-a02f-f8575d1d73e2> ;
    skos:prefLabel    "Completed" .
```

<<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/cd879ddb-55b6-4620-a0e4-c85a2df9f483>>

```
    rdf:type          schema:PostalAddress ;
    schema:addressCountry "IT" ;
    schema:addressLocality "Trento" ;
    schema:postalCode    "38122" ;
```

```
    schema:streetAddress      "Via Calepina 14" .  
  
<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/f913bd20-8e7b-4d82-9688-ca745a432f24>  
    rdf:type          skos:ConceptScheme ;  
    dct:description  "Education is a purposeful activity directed  
                      at achieving certain aims, such as transmitting knowledge  
                      or fostering skills and character traits." ;  
    dct:title        "Education" .  
  
<https://www.epos-eu.org/epos-dcat-ap#Distribution/2d8e5529-0e68-4e0d-af09-49a96f1c6c25>  
    rdf:type          dcat:Distribution ;  
    dct:conformsTo   "UML" ;  
    dct:description  "Entity Relational Model regarding Education  
                      in Trentino" ;  
    dct:format       "application/pdf" ;  
    dct:identifier   "https://github.com/samuelebortolotti/Education-Trentino/tree/main/Teleologies/Informal%20Modeling"^^xsd:anyURI ;  
    dct:issued        "2022-11-23T00:00:00Z"^^xsd:dateTime ;  
    dct:language     "English" ;  
    dct:license      "https://opensource.org/licenses/MIT"^^xsd:  
                      anyURI ;  
    dct:modified      "2022-11-23T00:00:00Z"^^xsd:dateTime ;  
    dct:rights        "https://opensource.org/licenses/MIT" ;  
    dct:title         "Entity Relational Model" ;  
    dct:type          "Document"^^xsd:anyURI ;  
    adms:status       <https://www.epos-eu.org/epos-dcat-ap#Concept/b9645f82-9500-414c-a51d-255ff657c69e> ;  
    dcat:accessURL   "https://github.com/samuelebortolotti/Education-Trentino/tree/main/Teleologies/Informal%20Modeling"^^xsd:anyURI ;  
    dcat:mediaType    "image/pdf" ;  
    foaf:page         "https://github.com/samuelebortolotti/Education-Trentino" .  
  
<https://www.epos-eu.org/epos-dcat-ap#Organization/2dad489f-4ce3-4e0d-9122-b70df55b8ef2>  
    rdf:type          schema:Organization ;  
    schema:address   <https://www.epos-eu.org/epos-dcat-ap#PostalAddress/b0bb325d-f408-4dac-80e0-2389d8efb97f> ;  
    schema:email      "knowdive@disi.unitn.it" ;
```

```
    schema:identifier "http://knowdive.disi.unitn.it"^^xsd:-
        anyURI ;
    schema:legalName "Knowdive" ;
    schema:leiCode "http://knowdive.disi.unitn.it" ;
    schema:logo "http://knowdive.disi.unitn.it/wp-content/
        uploads/knowdive-new-logo.png"^^xsd:anyURI ;
    schema:memberOf <https://www.epos-eu.org/epos-dcat-ap#
        Organization/78574deb-e40e-4c05-a5c6-40beeeef01aa0> ;
    schema:url "http://knowdive.disi.unitn.it"^^xsd:-
        anyURI .

<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/17ebceb0-a50b
-4801-82df-cc0d484ffa90>
    rdf:type skos:ConceptScheme ;
    dct:description "Member of the Education in Trentino Project
        \t" ;
    dct:title "Member" .

<https://www.epos-eu.org/epos-dcat-ap#Agent/f9cfb89e-3bb9-41ba-b263-
e0386fce30ed>
    rdf:type foaf:Agent ;
    dct:type <https://www.epos-eu.org/epos-dcat-ap#Concept/74
        cfe6d8-39b1-4491-a4ca-dba88fdccf10> ;
    foaf:name "Samuele Bortolotti, Erich Robbi" .

<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/ca07b831-220f-4c49
-bc3b-6e8215eb3238>
    rdf:type schema>ContactPoint ;
    schema:availableLanguage "en-US" , "it-IT" ;
    schema:contactType "Member" ;
    schema:email "samuele.bortolotti@studenti.unitn.
        it" .

<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/5edb028b-96af
-4035-8d62-c253acd5a3c5>
    rdf:type skos:ConceptScheme ;
    dct:description "Domain of Interest" ;
    dct:title "iTilos Principles" .

<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/3a8a3044-3e7a-463d
-94d9-15c4f4822938>
    rdf:type schema>ContactPoint ;
    schema:availableLanguage "it-IT" , "en-US" ;
    schema:contactType "Member" ;
```

```
    schema:email           "erich.robbi@studenti.unitn.it" .  
  
<https://www.epos-eu.org/epos-dcat-ap#Organization/78574deb-e40e-4c05-a5c6-40beef01aa0>  
    rdf:type              schema:Organization ;  
    schema:address        <https://www.epos-eu.org/epos-dcat-ap#PostalAddress/cd879ddb-55b6-4620-a0e4-c85a2df9f483> ;  
    schema:email          "ateneo@unitn.it" , "ateneo@pec.unitn.it"  
    ;  
    schema:identifier     "www.unitn.it"^^xsd:anyURI ;  
    schema:legalName      "Università degli Studi di Trento" ;  
    schema:leiCode        "00340520220" ;  
    schema:logo           "https://static-cdn.unitn.it/sites/www.unitn.it/themes/unitn\_theme/images/newlogo\_unitn\_en.png"^^xsd:anyURI ;  
    schema:telephone       "0461281111" ;  
    schema:url            "www.unitn.it"^^xsd:anyURI .  
  
<https://www.epos-eu.org/epos-dcat-ap#Concept/520bf2b3-6fcc-45ad-971a-b17df5fe98f0>  
    rdf:type              skos:Concept ;  
    skos:definition      "Inception is the second iTelos phase:\r\n    nInputs:\r\n    - Classified Competency Questions (CQ)\r\n    - Datasets\r\n    - Reference teleologies\r\n    nOutputs:\r\n    - ER Model\r\n    - Selected Datasets" ;  
    skos:inScheme         <https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/5edb028b-96af-4035-8d62-c253acd5a3c5> ;  
    skos:prefLabel        "Informal Modeling Phase" .
```

5 Formal Modeling

The third iTelos phase is Formal Modeling. In this phase, we are asked to build the ETG model and to syntactically align the datasets for knowledge and data level using the ER model, the selected datasets, and reference ontologies.

As a result, this section concentrates on the following activities:

- ETG generation;
- Schema Alignment;
- Language Alignment;
- ETG Model;
- Data Management;

5.1 ETG generation

At this level, we aim to generate a reusable and feasible ETG by reusing reference ontologies provided by the knowledge source we have selected. The ETG generating activity includes three fundamental sub-activities: ontology selection, language alignment, and schema development.

5.1.1 Ontology Selection

The goal of ontology selection is to identify relevant ontologies for reference, which represent use-case scenarios that may be reused to model the ER.

To be more explicit, the initial stage in the ETG generation activity is to reuse ontology from well-known knowledge resources that are semantically equivalent to concepts that are present in the ER model defined in the previous phase. The most important element is to determine which concepts such as objects, functions, and actions from existing ontologies may be connected and aligned to the ER model.

We opted to explore the VIVO portal in order to find a first reference teleontology, expecting to discover any specified contextual entity as we are modeling entities related to education, academia, and school and university courses. Although we identified various items related to our teleology, we still had to use [schema.org](#) to find additional specific entities that were previously not found. The entities discovered on [schema.org](#) and VIVO may be differentiated in the figure via the hue of the latter.

Finally, as a remark, some of the ontologies either on [schema.org](#) or on VIVO do not have any properties. For example, the [schema.org Instructor](#) entity is visible on the site but lacks its attributes; in their place, there is a high-level description that specifies what an instructor is, from which entity attributes should be inherited (i.e. the *Person* entity), and in which circumstances the entity should be utilized (for example *CourseInstance*).

With the above statements in mind, we then proceed with the alignment of the teleology, which can be seen in Figure 45 and in Table 5.



Name	Description	Knowledge Resource
Anything		
Anything Agent	Agents are things that do stuff	vivo
Anything Agent Person	An instance of a human being (Homo sapiens)	vivo
Anything Agent Person Non-Academic	A person holding a position that is not considered to be an academic appointment.	vivo
Anything Agent Person <i>Faculty Member</i>	A person with at least one academic appointment to a specific faculty of a university or institution of higher learning.	vivo
Anything Agent Organization	A kind of Agent corresponding to social institutions such as companies, societies etc.	vivo
Anything Spatial Region		vivo
Anything Spatial Region Location	Top level of all location classes.	vivo
Anything Spatial Region Location Geographic Location	A location having coordinates in geographic space.	vivo
Anything Spatial Region Location Geographic Location Geographic Region	A location having coordinates in geographic space.	vivo

Anything Spatial Region Location Geographic Location Geographic Region <i>Subnational Region</i>	Smaller administrative division into which a country may be divided.	vivo
Anything <i>Process</i>		vivo
Anything Process <i>Event</i>	Something that happens at a given place and time.	vivo
Anything Process Event <i>Course</i>	A course as taught in one time period by one or more instructors, normally but not always for credit.	vivo
Anything <i>Intangible</i>	A utility class that serves as the umbrella for a number of 'intangible' things such as quantities, structured values, etc.	schema.org
Anything Intangible <i>StructuredValue</i>	Structured values are used when the value of a property has a more complex structure than simply being a textual value or a reference to another thing.	schema.org
Anything Intangible StructuredValue <i>ContactPoint</i>	A contact point—for example, a Customer Complaints department.	schema.org
Anything Intangible <i>Rating</i>	A rating is an evaluation on a numeric scale, such as 1 to 5 stars.	schema.org
Anything <i>Review</i>	A review of an item - for example, of a restaurant, movie, or store.	schema.org

Anything Review <i>UserReview</i>	A review created by an end-user (e.g. consumer, purchaser, attendee etc.).	schema.org
Anything <i>CreativeWork</i>	The most generic kind of creative work, including books, movies, photographs, software programs, etc.	schema.org
Anything CreativeWork <i>LearningResource</i>	The LearningResource type can be used to indicate CreativeWorks (whether physical or digital) that have a particular and explicit orientation towards learning, education, skill acquisition, and other educational purposes.	schema.org
Anything CreativeWork LearningResource <i>Quiz</i>	Quiz: A test of knowledge, skills, and abilities.	schema.org

5.1.2 Ontology - Teleology alignment

In the following table, we illustrate the reasons why we have aligned our teleology with the selected reference ontologies.

Teleology Entity	Ontology Entity	Motivation
<i>Invalsi Score</i>	<i>Quiz</i>	Invalsi Score is the score which refers to the Invalsi test, a multidisciplinary test that is taken by Italian students in order to extract statistics regarding Italian education, divided by the municipality. A Quiz is a more general entity since it refers to a general test of knowledge and skill.

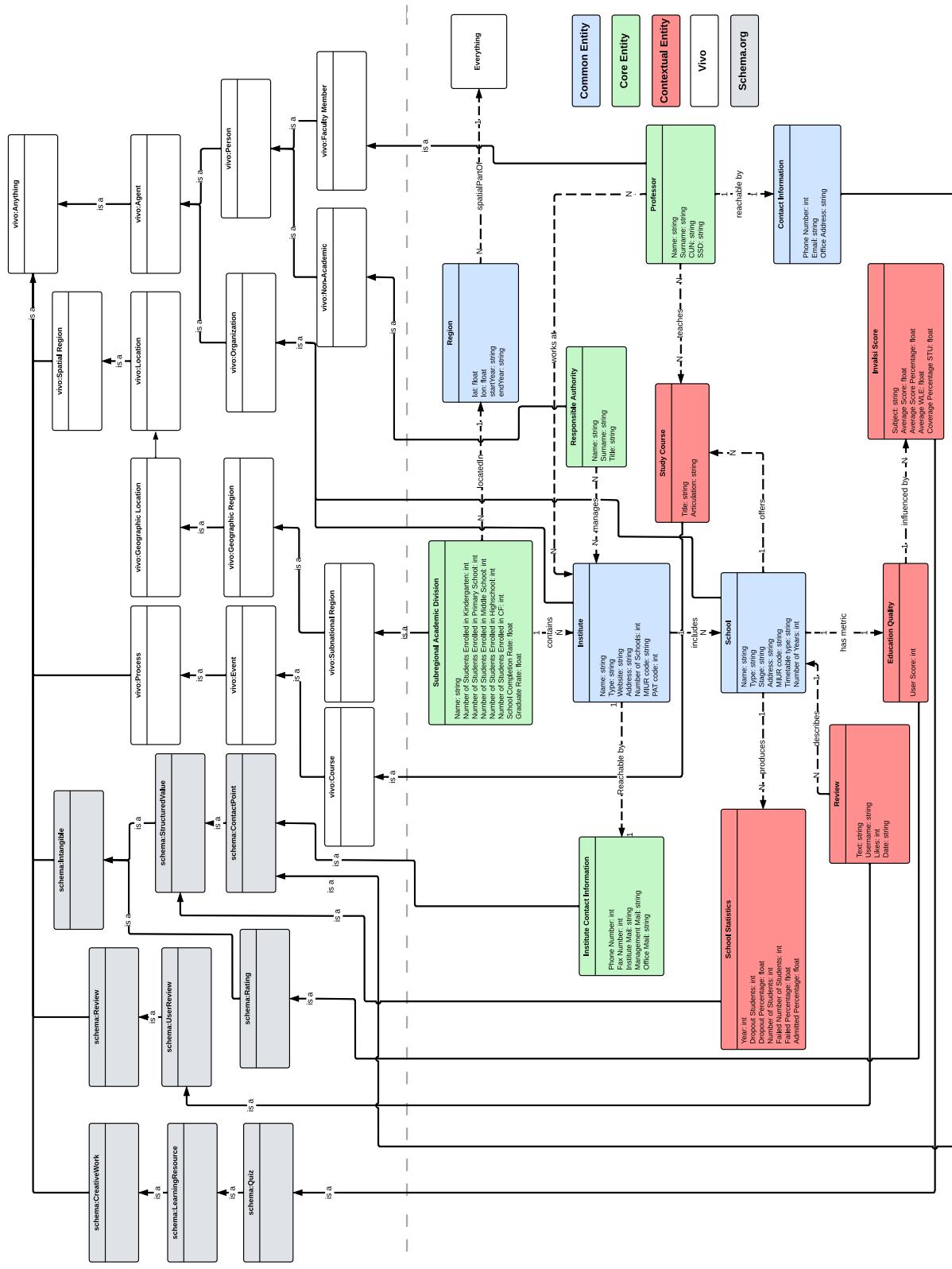
<i>Review</i>	<i>UserReview</i>	In our teleology, review refers to a review created by an end user, which is exactly the concept covered by the UserReview ontology.
<i>Education Quality</i>	<i>Rating</i>	Education quality for us is a collection of metrics that can be employed in order to classify the quality of the education provided by a school in quantitative terms. We have chosen Rating as reference ontology since, likewise our defined Education Quality, a rating is an evaluation on a numeric scale.
<i>School Statistics</i>	<i>StructuredValue</i>	A school statistic according to our purpose is a collection of data that is produced by a school and can be used by others. Since the statistics are complex data types, they comply with the Schema.org definition of StructuredValue.
<i>Contact Information</i>	<i>ContactPoint</i>	In our idea, Contact Information refers to all the data which can be used in order to reach a person. We have chosen ContactPoint since it refers to a general contact point. Moreover, it is the superclass of PostalAddress and contains information such as email, phone, and fax numbers, which are present in our original teleology.
<i>Institute Contact Information</i>	<i>ContactPoint</i>	In our idea, the Institute Contact Information refers to all the data which can be used in order to reach some institute authorities. We have chosen ContactPoint for the same reason as for Contact Information.

<i>Study Course</i>	<i>Course</i>	In our definition, a course is taught in an institute in a specific time period of time.
<i>Subregional Academic Division</i>	<i>Subnational Region</i>	Subregional Academic Division is a specialization of a municipality since it contains general statistics regarding the institute contained in the municipality. Subnational Regional refers to a smaller administrative division into which a country may be divided. Hence, it refers to a more general entity with respect to the Subregional Academic Division.
<i>Institute</i>	<i>Organization</i>	Institute is an organization which comprises several schools.
<i>School</i>	<i>Organization</i>	A school is an organization that offers study courses that are taught by professors.
<i>Responsible Authority</i>	<i>Non-Academic</i>	In our definition, a Responsible Authority is a person with a specific position in an institution decoupled from positions which are academic appointments.
<i>Professor</i>	<i>Faculty Member</i>	In our definition a professor is a person who is an academic appointment to a specific University.

5.1.3 Teleontology Model

The result of the combination of the ontologies (VIVO and Schema.org) with the underlined teleology is shown in this ER model.

Teleontology with Ontologies



5.2 Schema Alignment

Once the teleology and ontology have been mapped, we noticed that some of the properties of our teleology did not match the data sources we were able to gather.

Indeed, we have deleted the `date`, `likes`, and `username` from `Review`, since those data do not comply with the resources we were able to collect. However, from the information, we have gathered another relevant property that could be added to the `Review` entity, which is `sentiment`. This data property represents the sentiment associated with each review, which could be beneficial to answer all the competence questions regarding the quality of education provided by a certain school or institution.

Moreover, `Region` (which is the entity that represents our geospatial location) has been replaced with the `Geographical Region` entity of VIVO since, by reading its description, they map the very same concept we intended. In this scenario, of course, we have moved the object and data properties of `Region` to the ones of `Geographical Region`.

As a further consideration, VIVO does not have attributes that are linked to the entities (the `Range` field in Protégé has missing items). As a matter of fact, we have decided to keep the entities and directly use our data properties.

On the other hand, `schema.org` comes with predefined data properties. However, not all of them are useful for the purpose of our project. Moreover, a crucial part of this phase is to re-use data properties of the reference ontologies when they represent the same concept we have in our teleontology. Indeed, we have decided to remove the `faxNumber` and `phoneNumber` properties of our `Institute Contact Information` and employ the ones provided by `ContactPoint`. Additionally, we have removed the `Contact Information` entity since it is basically equivalent to the `ContactPoint` entity.

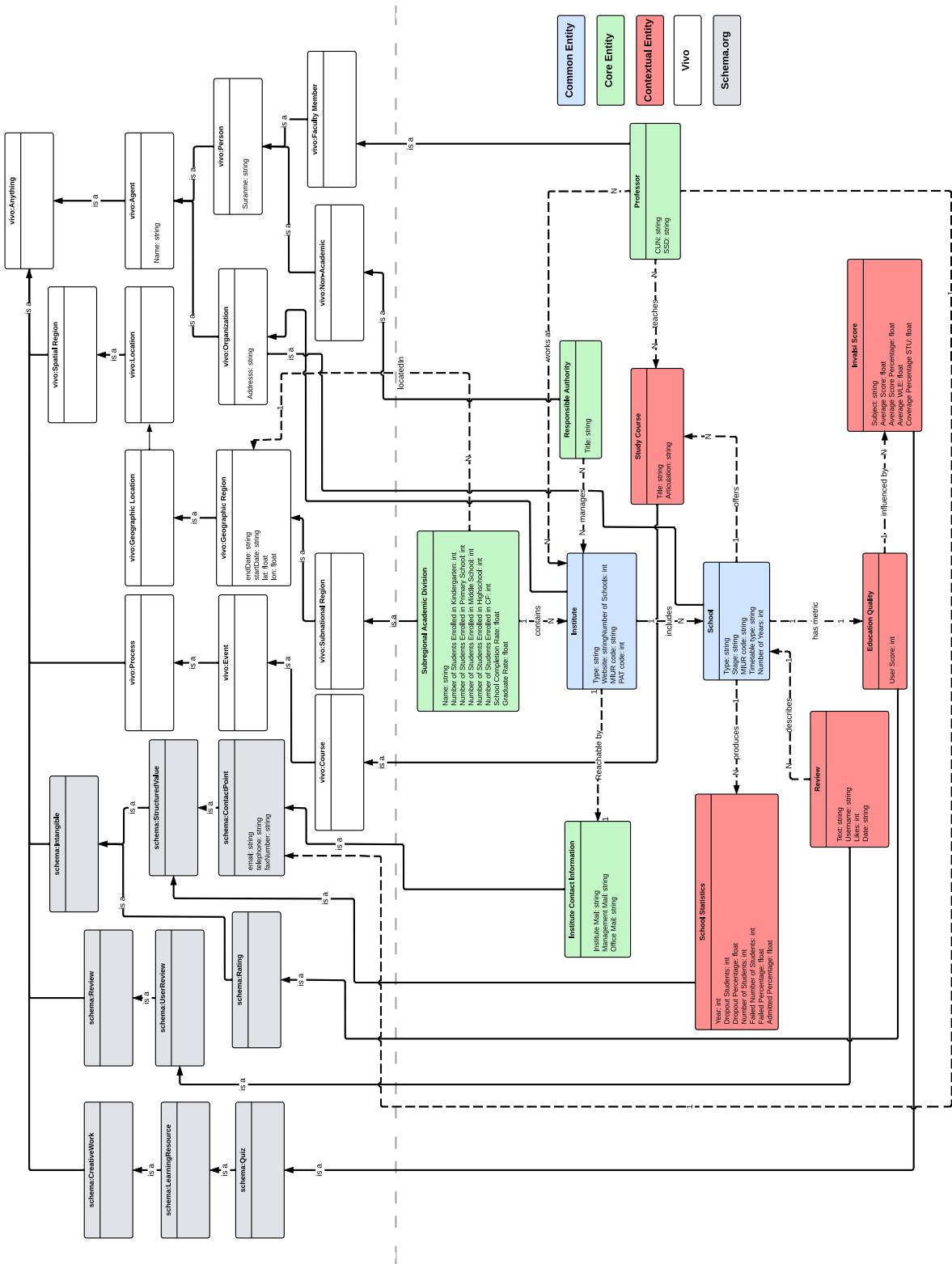
Furthermore, we have noticed that there are some data properties that do not comply with those ones of `schema.org`. For instance, the `User Score` data properties is a core attribution of our teleology and refer to the star rating a user can give to a school by means of a Google Review, which does not represent the same meaning of `ratingValue` property of the `schema.org Rating` entity.

As an effective representation, `name` was mapped to `Agent` as it is shared by both the organizations which in our case are `School` and `Institute`, and by the `Person` entities, such as `Professor` and `Responsible Authority`. Furthermore, we have noticed that `email`, `telephone`, `faxNumber` of the `schema.org ContactPoint` entity are reusable since they share the same meaning we intended for `Institute Contact Information`.

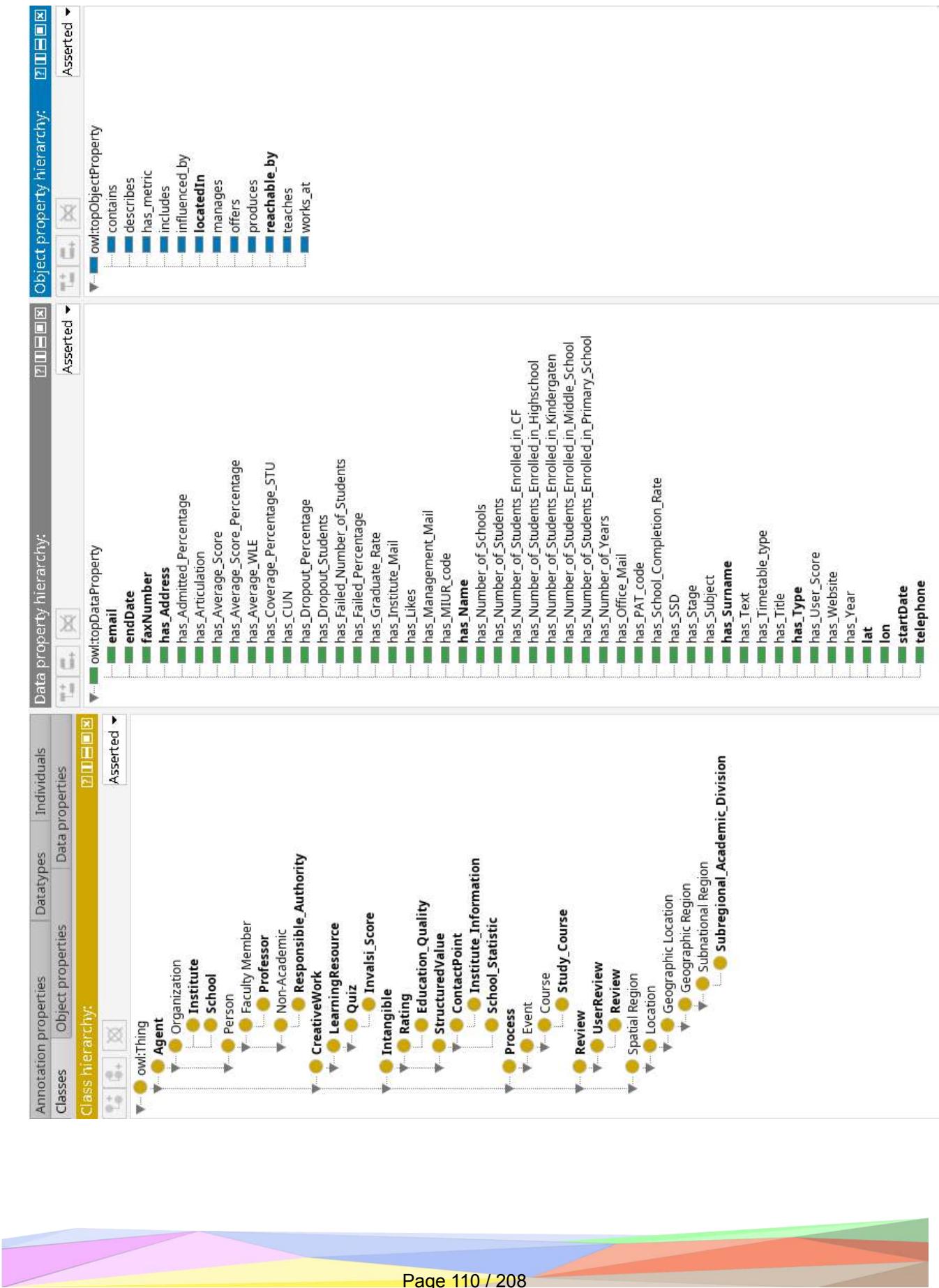
Additionally, the `Contact Information` entity, which was linked to `Professor`, has been removed as it produces redundancy with the aforementioned `ContactPoint` entity.

Therefore, the final schema after the attribute and entity alignment is the following:

Aligned Teleontology



Here, we list some screenshots of the operations we have performed on Protégé in order to build the final alignment operation:



5.3 Language Alignment

Since the ER model still has the problem of conceptual diversity (L1) and language diversity (L2), it can be seen as an intermediate version of ETG. Conceptual diversity means that there are concepts that are not unique in the model. Language diversity suggests that the same concepts are expressed in a different language.

UKC is a specialized tool to address problems related to conceptual and language diversity. The majority of the concepts in our project can be found in UKC CC, and the global IDs for each concept were matched. To obtain the new GID, it is necessary to define the other portion of concepts that are absent from UKC first. When it comes to the definition of this concept, we first decided to look for them in VIVO or Schema.org, and if we still couldn't find the precise description, we defined them ourselves.

Unfortunately, most of the time the concepts were too specific for our purpose so we were not able to find them in the UKC. As an effective representation, all the data concerning the Invalsi tests cannot be generalized to a generic evaluation. The same reasoning applies to the data contained in the Subregional Academic Division.

Here we list some screenshots depicting the procedure we have followed in order to map the different concepts:

The screenshot shows the UKC interface in the 'Formal Modeling' step (step 4). The top navigation bar includes icons for Opening, Inception, Informal Modeling, Formal Modeling (highlighted in blue), Data Integration, and Closing. Below the navigation is a toolbar with Open issues, Closed issues, and Actions buttons, and a Save Changelog button. The main area is titled 'Open issues' and lists the following items:

- influenced_by
- teaches
- produces
- includes
- has_Average_WLE
- has_User_Score
- has_Website
- has_Failed_Percentage
- Lat
- has_Timetable_type

influenced_by

^

What do you mean by *influenced_by*?

e.g. tree *
influence|

Noun (Synset 30624)

A cognitive factor that tends to have an effect on what you do

Noun (Synset 53888)

The effect of one thing (or person) on another

Verb (Synset 96386)

Induce into action by using one's charm

Verb (Synset 105129)

Have and exert influence or effect

Noun (Synset 28090)

A power to affect persons or events especially power based on prestige etc

Noun (Synset 51835)

One having power to influence another

Verb (Synset 96029)

Shape or influence; give direction to

Noun (Synset 686)

Causing something without any direct or apparent effort

None of the above

Solve

What do you mean by *has_Number_of_Students_Enrolled_in_Primary_School*?

e.g. tree *
number

- Noun (Synset 36273)
A short theatrical performance that is part of a longer program
- Noun (Synset 27604)
A clothing measurement
- Noun (Synset 44948)
A select company of people
- Noun (Synset 27734)
The property possessed by a sum or total or indefinite quantity of units or individuals
- Noun (Synset 20632)
An item of merchandise offered for sale
- Noun (Synset 33992)
A numeral or string of numerals that is used for identification
- Noun (Synset 33997)
The number is used in calling a particular telephone
- Noun (Synset 33470)
The grammatical category for the forms of nouns and pronouns and verbs that are used depending on the number of entities involved (singular or dual or plural)
- Noun (Synset 34718)
One of a series published periodically
- Noun (Synset 35823)
A symbol used to represent a number
- None of the above

Insert a brand new concept

has_Number_of_Students_	Part of speech NOUN
Description Number of students enrolled in primary schools of a given location or institutions	
	

Parent concept *

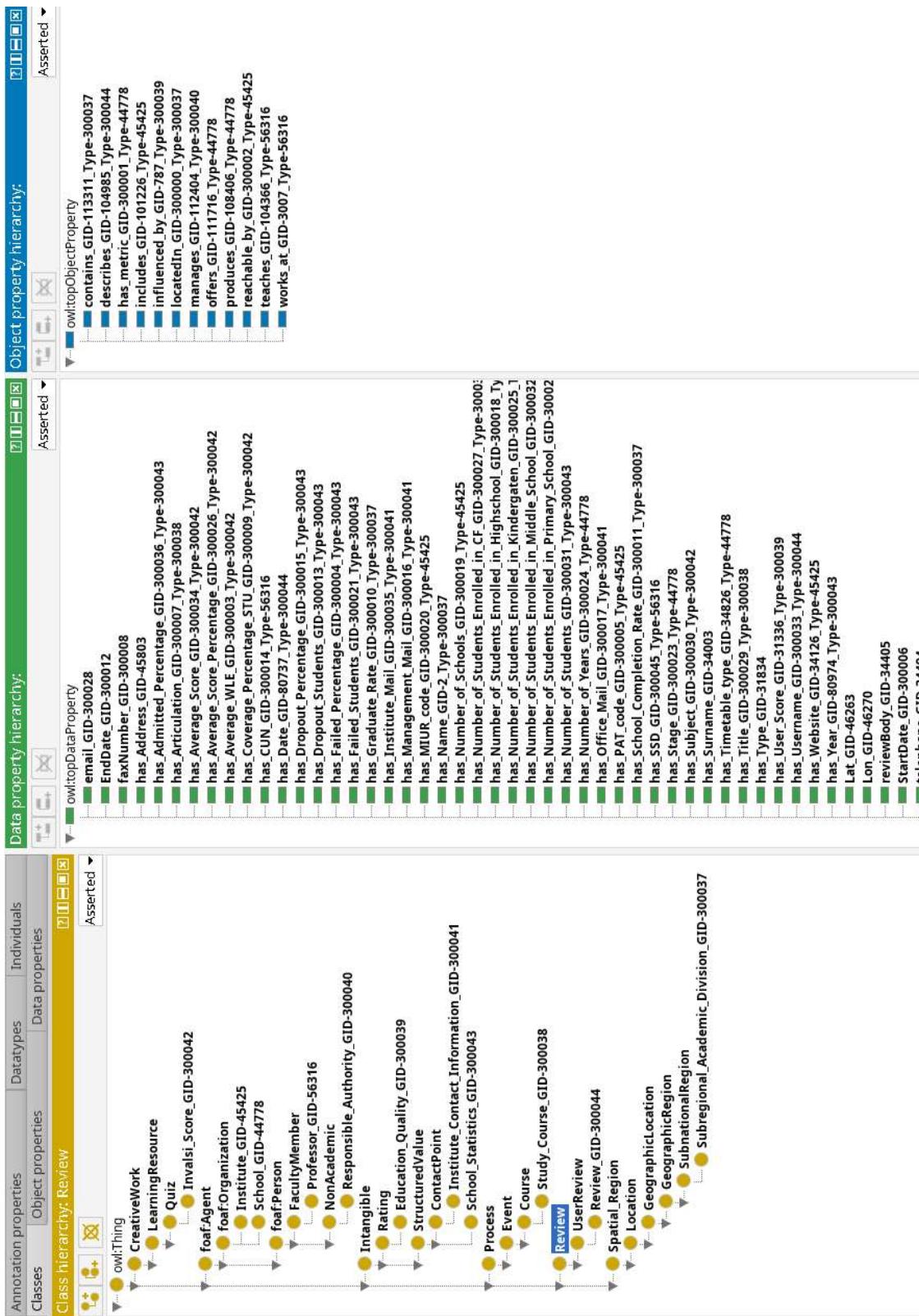
Number

- Number
 - a numeral or string of numerals that is used for identification*
- Number
 - a clothing measurement*
- Phone number
 - the number is used in calling a particular telephone*
- Number
 - the property possessed by a sum or total or indefinite quantity of units or individuals*
- Numeral

5.4 ETG Model

The result of the language alignment is an ontology comprehending some identifiers for specific concepts.

In the image below we show the final result of the language alignment as the first sketch for our final ETG.



We have used Protégé in order to generate the classes (Etype(s) in our project), objects, and data properties of our final ontology.

The Protégé result is accessible in the GitHub repository as an OWL RDF/XML file.

5.5 Data Management

We have decided to filter out the remaining unnecessary data concerning the final step of our teleontology building phase.

All datasets are first transformed in CSV format, then the value formats in various datasets were compared to our ETG model and we determined whether the same characteristics belonged to the same entity across various datasets and if they had the same values.

In most of our data sources, the values are expressed in the Italian language, so we examined the data properties defined with the goal of determining which ones might be translated. Since all the Italian data fields refer to single words or a bunch of concepts, we have employed Wordnet within the `nltk` Python package¹⁹ to align them to the English term. The table below lists the data properties that were subject to translation.

EType	Data Property	Original Value	New Value
<i>Invalsi Score</i>	Subject	Italiano	Italian
<i>Invalsi Score</i>	Subject	Matematica	Mathematics
<i>Institute</i>	Type	Istituzione scolastica e formativa	Educational and training institution
<i>Institute</i>	Type	Scuola dell'infanzia	Kindergarten
<i>Institute</i>	Type	Circolo di coordinamento pedagogico	Pedagogical coordination circle
<i>School</i>	Type	Provinciale	Provincial
<i>School</i>	Type	Gestione autonoma	Autonomous Management
<i>School</i>	Type	Delega F.P.S.M.	F.P.S.M. Delegation
<i>School</i>	Type	Delega Co.E.S.I.	Co.E.S.I. Delegation
<i>School</i>	Timetable type	DIURNO	Morning
<i>School</i>	Timetable type	SERALE	Evening
<i>School</i>	Stage	PRIMARIA	Primary School
<i>School</i>	Stage	SECONDARIA DI PRIMO GRADO	Middle School
<i>School</i>	Stage	SECONDARIA DI SECONDO GRADO	High School
<i>School</i>	Stage	FORMAZIONE PROFESSIONALE	Professional School
<i>Responsible</i>	Title	DIRIGENTE SCOLASTICO	Principal
<i>Responsible</i>	Title	PRESIDENTE ENTE GESTORE	President of Managing Body
<i>Responsible</i>	Title	DIRETTORE	Director
<i>Responsible</i>	Title	COORDINATORE PEDAGOGICO	Pedagogic Coordinator

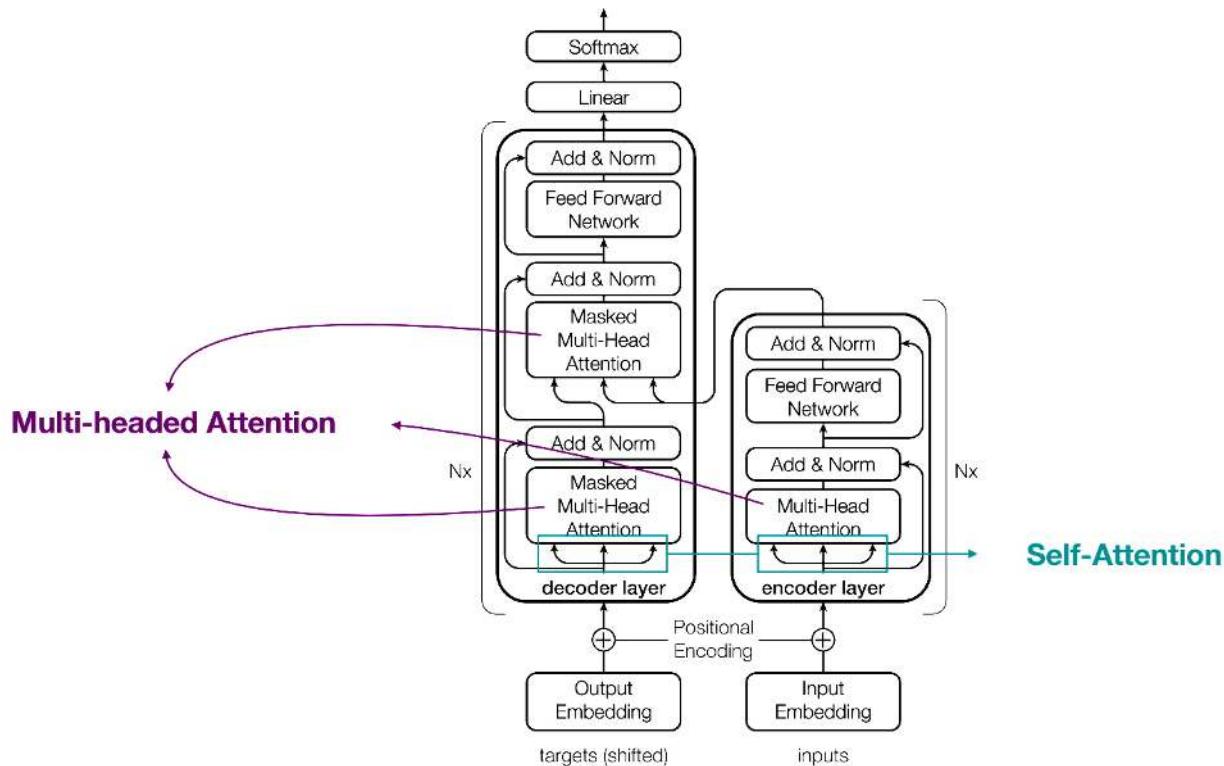
As an additional consideration considering the language translation, we have decided to do not to translate the name of the course. This choice was made to remark on the fact that the course is taught in Italian.

Additionally, there was a necessity to gather sentiment from the reviews we obtained during the Inception phase. By concentrating on our purpose, we have determined that we are interested in understanding whether a review depicts the school to which it relates positively or negatively. To achieve this ideal condition, we used a pre-trained transformer that evaluates the

¹⁹<https://www.nltk.org/howto/wordnet.html>

sentiment of a text given as an input, since assessing all of the reviews by hand would take too much time.

Just to provide a brief introduction about the model we have employed, a transformer [12] is a state-of-art in Natural Language Processing (NLP) technology. BERT (Bidirectional Encoder Representations from Transformers), distilBERT (a smaller variant of BERT), GPT (Generative Pre-trained Transformer), and T5 are some of the most well-liked pre-trained transformers in the literature. The transformer's foundations are composed of three elements: attention, transfer learning, and encoder-decoder.



Courtesy of Nvidia <https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model/>

Transformer models are essentially massive encoder/decoder blocks that handle data, similar to the majority of neural networks.

Positional encoders are used by transformers to tag data items entering and leaving the network. These tags are followed by attention units, which calculate an algebraic map showing how each element connects to the others.

In what is known as "multi-headed attention", attention requests are often processed in parallel by computing a matrix of equations. These resources enable computers to recognize the same patterns that humans see.

The first attempt was made considering FEEL-IT: Emotion and Sentiment Classification for the Italian Language [1]²⁰, a transformer trained for the task of sentiment analysis which

²⁰<https://huggingface.co/MilaNLProc/feel-it-italian-sentiment>

classifies a given sentence either positive or negative. The first approach has brought a considerable number of false positives which may drive the user to the wrong conclusion.

Here we list some of the misleading results we have retrieved from the model:

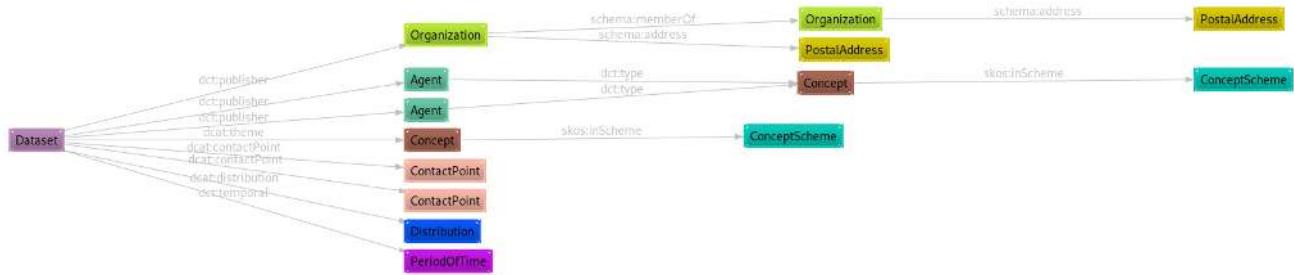
Cultura ed educazione al primo posto!	ISTITUTO TECN positive
Medie e superiori, i ricordi più belli	ISTITUTO TECN positive
Interessante	ISTITUTO TECN negative
La mia adolescenza in ottima compagnia	ISTITUTO TECN positive
Splendidi ricordi, ottimi insegnanti	ISTITUTO TECN positive
 Splendidi ricordi, ottimi insegnanti	ISTITUTO TECN positive
Ricco di servizi	ISTITUTO TECN negative
Una scuola importante	ISTITUTO TECN positive
 mia stima alla Professoressa oggettoDi questo atto vandalico. Coraggiosi!! ...Altro	SETTORE SERV negative
Bello	SETTORE SERV negative
Questa è la scuola in cui svolgo la mia attività come bidella e che dà delle ottime opportunità di inserimento nel mondo del lavoro ai nostri giovani che hanno voglia di studiare e di trovare un	
 Piaciuto tantissimo	SETTORE INDU negative
Bello	SETTORE INDU negative
Scuola qualificata per un ottima formazione!	SETTORE SERV positive
 Scuola superiore e alta formazione con metrologie innovative e sguardo umano	SETTORE INDU positive
Bella scuola (sarà che la frequento eh?)	SETTORE INDU negative
Scuola grafica innovativa e ben strutturata	SETTORE INDU positive

The second trial was made considering Italian BERT Sentiment model²¹. This model, instead, has produced better results compared to the first one. Moreover, it introduces *neutral*, which is another possible outcome for a sentence sentiment, which reflects situations in which the context is neither positive nor negative.

5.6 Metadata

Along with the final ETG, we have produced the associated Metadata compliant with the DCAT [2] vocabulary, thanks to the SHAPENess [9] metadata editor.

²¹<https://huggingface.co/neuraly/bert-base-italian-cased-sentiment>



ETG Metadata

```

@prefix : <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix schema: <http://schema.org/> .
@prefix spdx: <http://spdx.org/rdf/terms#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix gsp: <http://www.opengis.net/ont/geosparql#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dqv: <http://www.w3.org/ns/dqv#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix sh: <http://www.w3.org/ns/shacl#> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix locn: <http://www.w3.org/ns/locn#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix adms: <http://www.w3.org/ns/adms#> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix http: <http://www.w3.org/2006/http#> .
@prefix dash: <http://datashapes.org/dash#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

<https://www.epos-eu.org/epos-dcat-ap#Organization/2dad489f-4ce3-4e0d-9122-b70df55b8ef2>
    rdf:type schema:Organization ;
    schema:address <https://www.epos-eu.org/epos-dcat-ap#PostalAddress/b0bb325d-f408-4dac-80e0-2389d8efb97f> ;

```

```
    schema:email          "knowdive@disi.unitn.it" ;
    schema:identifier     "http://knowdive.disi.unitn.it"^^xsd:-
        anyURI ;
    schema:legalName      "Knowdive" ;
    schema:leiCode        "http://knowdive.disi.unitn.it" ;
    schema:logo           "http://knowdive.disi.unitn.it/wp-content/
        uploads/knowdive-new-logo.png"^^xsd:anyURI ;
    schema:memberOf       <https://www.epos-eu.org/epos-dcat-ap#
        Organization/78574deb-e40e-4c05-a5c6-40beef01aa0> ;
    schema:url            "http://knowdive.disi.unitn.it"^^xsd:-
        anyURI .
```

<<https://www.epos-eu.org/epos-dcat-ap#Concept/c7958238-a526-46b6-8a22-7b8b90f4b917>>

```
    rdf:type              skos:Concept ;
    skos:definition      "Formal modeling phase" ;
    skos:inScheme         <https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/038c1555-1840-4b3c-a08a-a7dfd0378614> ;
    skos:prefLabel        "KGE" .
```

<<https://www.epos-eu.org/epos-dcat-ap#Organization/78574deb-e40e-4c05-a5c6-40beef01aa0>>

```
    rdf:type              schema:Organization ;
    schema:address        <https://www.epos-eu.org/epos-dcat-ap#PostalAddress/cd879ddb-55b6-4620-a0e4-c85a2df9f483> ;
    schema:email           "ateneo@unitn.it" , "ateneo@pec.unitn.it"
    ;
    schema:identifier      "www.unitn.it"^^xsd:anyURI ;
    schema:legalName        "Università degli Studi di Trento" ;
    schema:leiCode          "00340520220" ;
    schema:logo             "https://static-cdn.unitn.it/sites/www.
        unitn.it/themes/unitn_theme/images/newlogo_unitn_en.png"^^
        xsd:anyURI ;
    schema:telephone        "0461281111" ;
    schema:url              "www.unitn.it"^^xsd:anyURI .
```

<<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/84efa52e-3879-4d09-b0b9-36362e915866>>

```
    rdf:type              dct:PeriodOfTime ;
    schema:endDate        "2022-12-05T00:00:00Z"^^xsd:dateTime ;
    schema:startDate       "2022-12-06T00:00:00Z"^^xsd:dateTime .
```

<<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/b0bb325d-f408-4dac-80e0-2389d8efb97f>>

```
rdf:type schema:PostalAddress ;
schema:addressCountry "IT" ;
schema:addressLocality "Povo (TN)" ;
schema:postalCode "I-38123" ;
schema:streetAddress "Via Sommarive 9" .

<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/ca07b831-220f-4c49-bc3b-6e8215eb3238>
rdf:type schema:ContactPoint ;
schema:availableLanguage "en-US" , "it-IT" ;
schema:contactType "Member" ;
schema:email "erich.robbi@studenti.unitn.it" .

<https://www.epos-eu.org/epos-dcat-ap#Concept/520bf2b3-6fcc-45ad-971a-b17df5fe98f0>
rdf:type skos:Concept ;
skos:definition "Formal Modeling is the third iTelos phase:\r\nInputs:\r\n * ER Model\r\n * Selected Datasets\r\n *\r\n Reference ontologies\r\nOutputs:\r\n * ETG\r\n *\r\n Dataset syntactically aligned\r\nThe ETG Generation\r\nactivity is internally defined by three sub activities:\r\n*\r\n * Ontology Selection: selection of those ontologies\r\nwhich includes appropriate concepts which can be reused to\r\nmodel the ER\r\n * Language Alignment: identification,\r\nand import in Knowledge Base of concepts and terms to be\r\nused to build the ETG." ;
skos:inScheme <https://www.epos-eu.org/epos-dcat-ap#
ConceptScheme/a21146bd-a38c-49fe-a02f-f8575d1d73e2> ;
skos:prefLabel "Formal Modeling Phase" .

<https://www.epos-eu.org/epos-dcat-ap#Distribution/2d8e5529-0e68-4e0d-af09-49a96f1c6c25>
rdf:type dcat:Distribution ;
dct:conformsTo "https://www.w3.org/OWL/" ;
dct:description "Entity Type Model regarding Education in Trentino." ;
dct:format "https://www.w3.org/OWL/^xsd:anyURI" ;
dct:identifier "https://github.com/samuelebortolotti/Education-Trentino/tree/main/Teleologies/Formal%20Modeling" ;
dct:issued "2022-12-05T00:00:00Z"^^xsd:dateTime ;
dct:language "English" ;
dct:license "https://opensource.org/licenses/MIT"^^xsd:URI ;
```

```
dct:modified      "2022-12-06T00:00:00Z"^^xsd:dateTime ;
dct:rights        "https://opensource.org/licenses/MIT" ;
dct:title         "Entity Type Graph" , "Education in
Trentino" ;
dct:type          "Collection"^^xsd:anyURI ;
dcat:accessURL   "https://github.com/samuelebortolotti/
Education-Trentino"^^xsd:anyURI ;
dcat:downloadURL "https://github.com/samuelebortolotti/
Education-Trentino/tree/main/Teleologies/Formal%20Modeling
"^^xsd:anyURI ;
dcat:mediaType    "application/rdf+xml" ;
foaf:page         "https://github.com/samuelebortolotti/
Education-Trentino" .
```

<<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/cd879ddb-55b6-4620-a0e4-c85a2df9f483>>

```
  rdf:type           schema:PostalAddress ;
  schema:addressCountry "IT" ;
  schema:addressLocality "Trento" ;
  schema:postalCode    "38122" ;
  schema:streetAddress "Via Calepina 14" .
```

<<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/3a8a3044-3e7a-463d-94d9-15c4f4822938>>

```
  rdf:type           schema>ContactPoint ;
  schema:availableLanguage "it-IT" , "en-US" ;
  schema:contactType "Member" ;
  schema:email        "samuele.bortolotti@studenti.unitn.
it" .
```

<<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/a21146bd-a38c-49fe-a02f-f8575d1d73e2>>

```
  rdf:type           skos:ConceptScheme ;
  dct:description   "Education in Trentino" ;
  dct:title         "Status" .
```

<<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/038c1555-1840-4b3c-a08a-a7dfd0378614>>

```
  rdf:type           skos:ConceptScheme ;
  dct:description   "iTilos formal modeling" ;
  dct:title         "Formal Modeling" .
```

<<https://www.epos-eu.org/epos-dcat-ap#Agent/828bf6b9-9f4b-4355-bcd0-d68b188811cb>>

```
rdf:type foaf:Agent ;
dct:type <https://www.epos-eu.org/epos-dcat-ap#Concept/
c7958238-a526-46b6-8a22-7b8b90f4b917> ;
foaf:name "Samuele Bortolotti" .

<https://www.epos-eu.org/epos-dcat-ap#Dataset/3a2e0c02-aebe-44d1-b688
-a5e21d81d088>
rdf:type dcat:Dataset ;
dct:accessRights "Public" ;
dct:accrualPeriodicity "Once"^^xsd:anyURI ;
dct:conformsTo "iTilos" ;
dct:created "2022-12-05T00:00:00Z"^^xsd:dateTime
;
dct:description "Entity Type Graph Metadata regarding
Education in Trentino.", "The project was made for the
Knowledge Graph Engineering (KGE) course held in the
academic year 2022/2023 at the University of Trento." ;
dct:identifier "https://github.com/samuelebortolotti
/Education-Trentino/tree/main/Teleologies/Formal%20
Modeling"^^xsd:anyURI ;
dct:issued "2022-12-06T00:00:00Z"^^xsd:dateTime
;
dct:language "English" ;
dct:modified "2022-12-06T00:00:00Z"^^xsd:dateTime
;
dct:publisher <https://www.epos-eu.org/epos-dcat-ap
#Agent/f9cfb89e-3bb9-41ba-b263-e0386fce30ed> , <https://
www.epos-eu.org/epos-dcat-ap#Organization/2dad489f-4ce3-4
e0d-9122-b70df55b8ef2> , <https://www.epos-eu.org/epos-
dcat-ap#Agent/828bf6b9-9f4b-4355-bcd0-d68b188811cb> ;
dct:temporal <https://www.epos-eu.org/epos-dcat-ap
#PeriodOfTime/84efa52e-3879-4d09-b0b9-36362e915866> ;
dct:title "Entity Type Graph Metadata" ;
dct:type "Collection"^^xsd:anyURI ;
dcat:contactPoint <https://www.epos-eu.org/epos-dcat-ap
#ContactPoint/3a8a3044-3e7a-463d-94d9-15c4f4822938> , <
https://www.epos-eu.org/epos-dcat-ap#ContactPoint/ca07b831
-220f-4c49-bc3b-6e8215eb3238> ;
dcat:distribution <https://www.epos-eu.org/epos-dcat-ap
#Distribution/2d8e5529-0e68-4e0d-af09-49a96f1c6c25> ;
dcat:keyword "KGE" , "Knowledge Graph Engineering"
, "Erich Robbi" , "UnitN" , "University of Trento" , "
Trentino" , "Trento" , "Fausto Giunchiglia" , "Samuele
Bortolotti" , "iTilos" ;
```

```
    dcat:landingPage      "https://github.com/samuelebortolotti
                           /Education-Trentino/" ;
    dcat:theme            <https://www.eops-eu.org/eops-dcat-ap
                           #Concept/520bf2b3-6fcc-45ad-971a-b17df5fe98f0> ;
    foaf:page             "https://github.com/samuelebortolotti
                           /Education-Trentino/" .

<https://www.eops-eu.org/eops-dcat-ap#Agent/f9cfb89e-3bb9-41ba-b263-
e0386fce30ed>
    rdf:type   foaf:Agent ;
    dct:type   <https://www.eops-eu.org/eops-dcat-ap#Concept/
               c7958238-a526-46b6-8a22-7b8b90f4b917> ;
    foaf:name  "Erich Robbi" .
```

6 Data Integration

This section is dedicated to the description of the Data Integration phase. Similar to the previous section, this one tries to describe the many sub-activities carried out by every team member as well as the results of the phase. This section gives specifics about:

- Entity management;
- Data integration phase evaluation.

The data integration phase tries to create the final EG by populating the previously produced ETG. This phase expects the syntactically aligned ETG and datasets as input, while the associated EG, metadata, and project documentation are the Data Integration phase output.

6.1 Data Management

At first, we must acknowledge that there are various ways to represent the same real-world objects; this phenomenon is known as **semantic heterogeneity**.

To accurately align the data to ETG, it is compulsory to address the problem of semantic heterogeneity. In order to fulfill this activity, we aim to combine two independent processes, which are *entity alignment* and *entity matching*.

6.1.1 Entity Alignment

Entity alignment entails converting various entity representations (from many datasets) to a single goal-specific template (ETG). In other words, entity alignment correctly conducts the mapping across different datasets from entities and entity properties to Etypes and Etype properties. In this section, the report focuses on *Entity Alignment*.

During the *Formal Modeling* phase, the datasets were converted to CSV format and syntactically aligned. At this point, we have correctly built 22 different datasets.

As a reminder of the previous phase, here we list a table showing the specific datasets we have decided to employ during the final Knowledge Graph Construction.

Data source	Description
dropouts.csv	contains information about how many students have dropped out of all schools collected
institutes-list.csv	contains information about all collected institutes
invalsi.csv	contains all results of the invalsi test of a given year grouped by municipality
ripetenti.csv	contains information about how many students in a class have failed a year for all schools collected
school-list.csv	contains information about all collected schools
school-responsibles.csv	contains information about all persons in charge of all institutes collected (and therefore their respective schools)
scores.csv	contains, for all schools collected, their respective averaged score.
scrutini.csv	contains informations about how many students in a specific school year have failed their class.
school_reviews-values-aligned.csv	for all schools collected, this data source contains all reviews associated to them
Trentino Commune List.csv	list of all municipality in Trentino
prof_course-values-aligned.csv	every row in this data source represents a study course that is teached by a professor.
unitn_professors_with_infos.csv	contains information about professors of the University of Trento
unitn_responsibles.csv	contains information about people who have a role in the University
tasso_diplomati.csv	for each municipality collected, this resource contains how many citizens have completed a secondary school.
tasso_compimento_obbligo_scolastico.csv	for each municipality collected, this resource contains how many citizens have gone to school until they were 16.
iscritti_centri_formazione.csv	for each municipality collected, this resource contains how many students there are in a professional school
iscritti_infanzia.csv	for each municipality collected, this resource contains how many students there are in a kindergarten
iscritti_scuole_elementari.csv	for each municipality collected, this resource contains how many students there are in a primary school
iscritti_scuole_medie.csv	for each municipality collected, this resource contains how many students there are in a middle school
iscritti_scuole_superiori.csv	for each municipality collected, this resource contains how many students there are in a high school
unitn-courses_values-aligned.csv	contains informations about study courses offered by the University of Trento

As a reference, the following repository hosts all the scripts in Python and Ruby language for the alignment phase: <https://github.com/samuelebortolotti/kge>.

6.2 Entity Matching

Entity matching is the last method of the *iTelos* methodology. The aim is to determine whether several entities in the datasets can reflect the same real-world entity and, as a result, should be integrated into the final EG.



The key point of entity matching is to define, for each entity, an **Identifying Set**. An Identifying Set is a set of Etypes properties that uniquely identify an entity throughout the different datasets. The idea is that identifiers are not always included in datasets, thus, we should find those properties which allow us to uniquely identify an entity.

This procedure was conducted on KarmaLinker, which is a tool compliant with the *iTelos* methodology used for data integration. We can effectively match the datasets to the specified ontology using KarmaLinker. Hence, by defining each Etype URIs, we completely address the semantic heterogeneity problem.

Once the mapping has been performed, KarmaLinker produces the RDF file for each EType, which can then be used in order to load data in GraphDB²².

During this phase we have encountered several issues, the first one regards the already present identifiers. Indeed, during the previous phase, we handcrafted several ids in order to link the institute to the associated schools, and the latter with the study courses. Those naive identifiers were only integer values starting from 1. We have realized that the mapping in the present condition would be ineffective since different Etypes will have the same URI. Moreover, some of the existing data sources come with pre-defined identifiers that are not made of a composition of Etype properties but rather from previous data management operations.

Therefore, we have decided to employ the already available unique identifiers when provided by the data source and to define new ones when this data is missing. As a matter of fact, here we present a table explaining the identifying property we have defined for the different data sources. In detail, when the primary identifier cannot be computed, then we have employed a secondary one.

Etype	Primary	Secondary
Geographic Region	name of the region	
Subregional Academic Division	name of the municipality	
Institute	MIUR code	integer id
School	an integer id	
Institute Contact Information	integer id	
Invalidi	municipality code, school stage, school year and subject	
Professor	UID assigned by Università Digitale	
Study course	UID assigned by Università Digitale	integer identifier
Review	integer id	
School Statistics	integer id	
Education Quality	integer id	

Even when we are sure that the data refer to the same entities, there may be some conflicts when merging dataset values or various representations of the same object. The final decisions in this situation are usually influenced by a particular type of metadata connected to the datasets or directly to the entities within the datasets. One useful tool is provenance, which is metadata that identifies the source of the dataset, entity, or value. As a result, choosing the right values to connect with the Etype attribute is not difficult.

To conclude this section, we list some screenshots showing the entity-matching activity on KarmaLinker:

²²<https://graphdb.ontotext.com/>



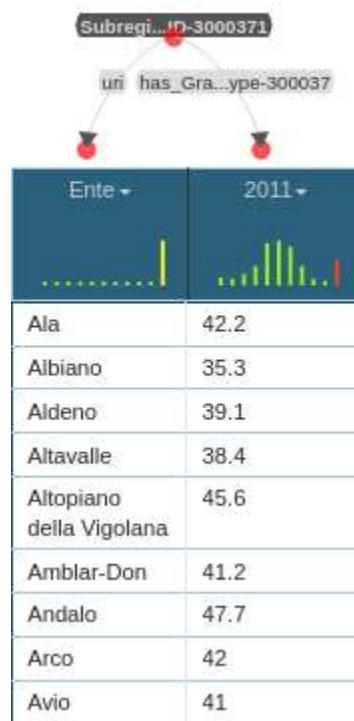
Diagram illustrating the relationships between Institute, School, and Student entities:

```

    graph TD
        Institute["Institute ID-454251"] -- "has_PAT...-Type-45425" --> pat_code["pat_code"]
        Institute -- "include...-Type-45425" --> School["School ID-447781"]
        School -- "produce...-Type-44778" --> Student["Student ID-3000431"]
        Student -- "uri" --> valore["valore"]
        Student -- "uri" --> valore_tot["valore_tot"]
        Student -- "uri" --> valore_d["valore_d"]
        Student -- "uri" --> anno_scol["anno_scol"]
        Student -- "uri" --> institute_miurcode["institute_miurcode"]
        Student -- "uri" --> anno_corso["anno_corso"]
        Student -- "uri" --> id_scuola["id_scuola"]
        Student -- "uri" --> id_stat_3["id stat 3"]
        Student -- "uri" --> id_2020_21_school["id 2020_21_school"]
        valore -- "has_Dro...-ype-300043" --> valore
        valore_tot -- "has_Dro...-ype-300043" --> valore_tot
        valore_d -- "has_Dro...-ype-300043" --> valore_d
        anno_scol -- "has_Yea...-ype-300043" --> anno_scol
        institute_miurcode -- "has_Yea...-ype-300043" --> institute_miurcode
        anno_corso -- "has_Yea...-ype-300043" --> anno_corso
        id_scuola -- "has_Yea...-ype-300043" --> id_scuola
        id_stat_3 -- "has_Yea...-ype-300043" --> id_stat_3
        id_2020_21_school -- "has_Yea...-ype-300043" --> id_2020_21_school
    
```

	pat_code	timetable_type	valore	valore_tot	valore_d	anno_scol	institute_miurcode	anno_corso	id_scuola	id stat 3	id 2020_21_school
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	1		school_id_29284	1_2020_21_schoo...	
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	2		school_id_29284	2_2020_21_schoo...	
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	3		school_id_29284	3_2020_21_schoo...	
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	4		school_id_29284	4_2020_21_schoo...	
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	5		school_id_29284	5_2020_21_schoo...	
221179501	Morning	0.0012048192771...	830.0	1.0	2020-21	TNIS00300A	1		school_id_29803	1_2020_21_schoo...	

Students dropout



Subregional ID-3000371

uri has_Scholarship-type-300037

Ente	2011
Ala	33.2
Albiano	37.3
Aldeno	39.7
Altavalle	33.2
Altopiano della Vigolana	27.1
Amblar-Don	32.4
Andalo	27.2
Arco	33.7
Avio	37.5
Baselga di Pinè	28.5
Bedollo	29.8
Besenello	28.3
Bieno	29.8
Bieggio Superiore	29.9
Bocenago	36.6
Bondone	34.5

Subreqi...ID-3000371

uri has_Nume...ype-300037

Ente ▾	2021 ▾
Rovereto	1413
Ruffrè-Mendola	0
Rumo	0
Sagron Mis	0
Samone	0
San Giovanni di Fassa	0
San Lorenzo Dorsino	0
San Michele all'Adige	325
Sant'Orsola Terme	0
Sanzeno	0
Sarnonico	0
Scurelle	0
Segonzano	0
Sella Giudicarie	0

Subregional ID-3000371

uri has_Number-type-300037

Ente	2021
Ala	0
Albiano	0
Aldeno	0
Altavalle	0
Altopiano della Vigolana	0
Amblar-Don	0
Andalo	0
Arco	173
Avio	0
Baselga di Pinè	0
Bedollo	0
Besenello	0
Bieno	0
Bieggio Superiore	0

Subregionali ID-3000371

uri has_Number-type-300037

Ente	2021
Ala	414
Albiano	104
Aldeno	156
Altavalle	59
Altopiano della Vigolana	262
Amblar-Don	0
Andalo	69
Arco	972
Avio	198
Baselga di Pinè	250
Bedollo	61
Besenello	139
Bieno	0
Bleggio Superiore	80
Bocenago	0
Bondone	0

Subregi...ID-3000371

uri has_Nume...ype-300037

Ente	2021
Ala	234
Albiano	47
Aldeno	77
Altavalle	36
Altopiano della Vigolana	135
Amblar-Don	0
Andalo	36
Arco	408
Avio	116
Baselga di Pinè	130
Bedollo	33
Besenello	73
Bieno	22
Bleggio Superiore	36
Bocenago	0
Bondone	12

Students dropout

Subregi.../D-3000371

uri has_Nume...ype-300037

Iscritti alle scuole medie inferiori	valore
Ala	262
Albiano	41
Aldeno	106
Altavalle	0
Altopiano della Vigolana	144
Amblar-Don	0
Andalo	95
Arco	621
Avio	146
Baselga di Pinè	209
Bedollo	0
Besenello	0
Bieno	0
Bleggio Superiore	0
Bocenago	0
Bondone	0

Diagram illustrating a linked data query result for Institute and School entities:

```

    graph TD
        Institute["Institute ID-454251"] -- "has_PAT...-Type-45425" --> pat_code["pat_code"]
        Institute -- "include...-Type-45425" --> School1["School ID-447781"]
        School1 -- "uri" --> School2["School ID-3000431"]
        School2 -- "produce...-Type-44778" --> School_id["school_id"]
        School2 -- "uri" --> id_stat_3["id stat 3"]
        School2 -- "uri" --> id_stat_4["id stat 4"]
    
```

Institute

pat_code	timetable_type	valore	valore_tot	valore_d	anno_scol	institute_miurcode	anno_corso	id_scuola	id stat 3	id stat 4
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	1	school_id_29284	1_2020_21_schoo...	
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	2	school_id_29284	2_2020_21_schoo...	
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	3	school_id_29284	3_2020_21_schoo...	
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	4	school_id_29284	4_2020_21_schoo...	
222059599	Morning	0.0497237569060...	362.0	18.0	2020-21	TNCF003001	5	school_id_29284	5_2020_21_schoo...	
221179501	Morning	0.0012048192771...	830.0	1.0	2020-21	TNIS00300A	1	school_id_29803	1_2020_21_schoo...	

Institute list

Diagram illustrating a linked data query result for Invalsi and Education entities:

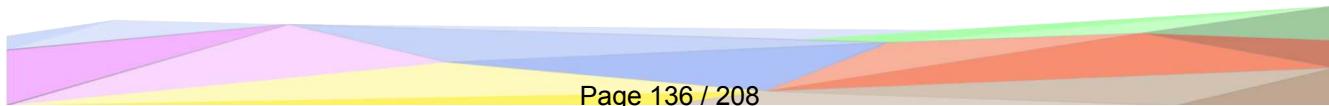
```

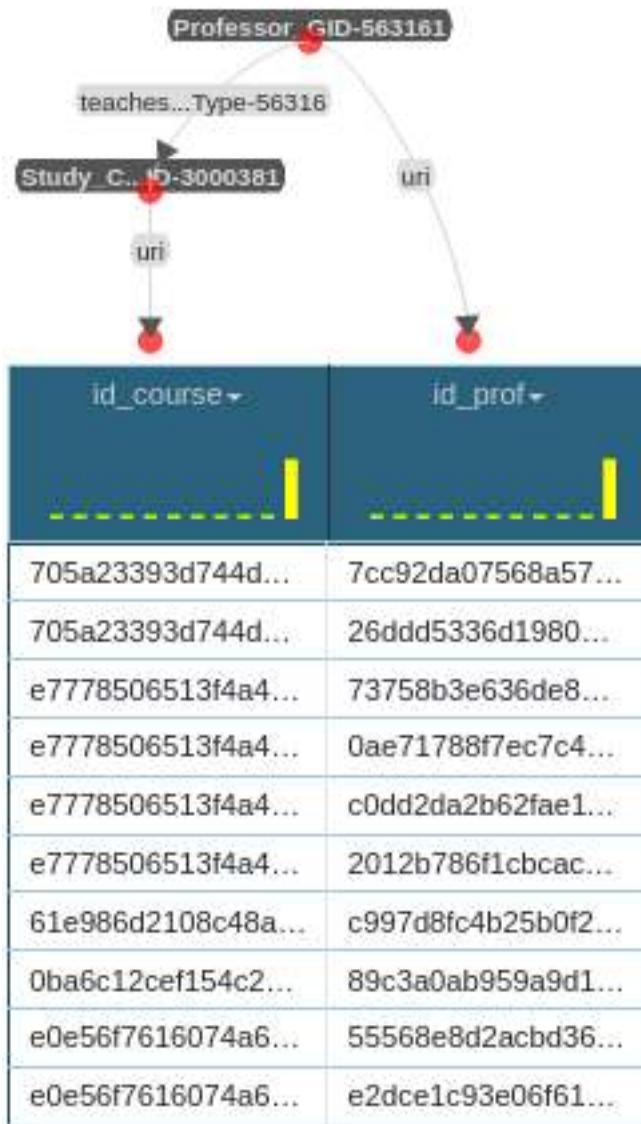
    graph TD
        Educati["Educati ID-3000391"] -- "influen...-ype-300039" --> Invalsi1["Invalsi ID-3000421"]
        Invalsi1 -- "uri" --> Invalsi2["Invalsi ID-3000421"]
        Invalsi2 -- "has_Sub...-ype-300042" --> new_ID["new ID"]
        Invalsi2 -- "has_Ave...-ype-300042" --> comune_name["comune_name"]
        Invalsi2 -- "has_Ave...-ype-300042" --> grado["grado"]
        Invalsi2 -- "has_Ave...-ype-300042" --> subject["subject"]
        Invalsi2 -- "has_Ave...-ype-300042" --> anno_scol["anno_scol"]
        Invalsi2 -- "has_Cov...-ype-300042" --> Punteggio_percentuale_medio["Punteggio_percentuale_medio"]
        Invalsi2 -- "has_Cov...-ype-300042" --> Dev_Std_Punteggio_Percentuale["Dev_Std_Punteggio_Percentuale"]
        Invalsi2 -- "has_Cov...-ype-300042" --> Punteggio_wife_medium["Punteggio_wife_medium"]
        Invalsi2 -- "has_Cov...-ype-300042" --> Dev_Std_Punteggio_wife["Dev_Std_Punteggio_wife"]
        Invalsi2 -- "has_Cov...-ype-300042" --> perc_copertura_stu["perc_copertura_stu"]
        Invalsi2 -- "uri" --> uri["uri"]
    
```

Invalsi

new ID	comune_name	grado	subject	anno_scol	Punteggio_percentuale_medio	Dev_Std_Punteggio_Percentuale	Punteggio_wife_medium	Dev_Std_Punteggio_wife	perc_copertura_stu	uri
del Garda_10_2017...										
invalsi_id_22153_...	Riva del Garda	10	Italian	2017-18	999,00	999,00	219,37	30,30	97,78	education_quality...
invalsi_id_22153_...	Riva del Garda	10	Italian	2017-18	999,00	999,00	219,37	30,30	97,78	education_quality...
invalsi_id_22161_...	Rovereto	10	Italian	2017-18	999,00	999,00	209,80	36,85	95,07	education_quality...
invalsi_id_22161_...	Rovereto	10	Italian	2017-18	999,00	999,00	209,80	36,85	95,07	education_quality...
invalsi_id_22161_...	Rovereto	10	Italian	2017-18	999,00	999,00	209,80	36,85	95,07	education_quality...
invalsi_id_22161_...	Rovereto	10	Italian	2017-18	999,00	999,00	209,80	36,85	95,07	education_quality...

Invalsi





Students dropout

Subregional Subregional ID-3000371 GID-3000371

located_in type-300037

has_NamedEntity type-300037 url

uri Lat_GID-46263 n_GID- StartDate_GID-300006_GID-300012

Nome	lat	lon	start	end	Comune	ID
Fiemme	46.000000	11.110007	2020	2022	Fiemme	Fiemme
Trento	46.066666	11.116667	2020	2022	Tre Ville	Tre Ville
Trento	46.066666	11.116667	2020	2022	Trento	Trento
Trento	46.066666	11.116667	2020	2022	Valdaone	Valdaone
Trento	46.066666	11.116667	2020	2022	Valfloriana	Valfloriana
Trento	46.066666	11.116667	2020	2022	Vallarsa	Vallarsa
Trento	46.066666	11.116667	2020	2022	Vallelaghi	Vallelaghi
Trento	46.066666	11.116667	2020	2022	Varena	Varena
Trento	46.066666	11.116667	2020	2022	Vermiglio	Vermiglio
Trento	46.066666	11.116667	2020	2022	Vignola-Falesina	Vignola-Falesina
Trento	46.066666	11.116667	2020	2022	Villa Lagarina	Villa Lagarina
Trento	46.066666	11.116667	2020	2022	Ville d'Anaunia	Ville d'Anaunia
Trento	46.066666	11.116667	2020	2022	Volano	Volano
Trento	46.066666	11.116667	2020	2022	Zambana	Zambana
Trento	46.066666	11.116667	2020	2022	Ziano di Fiemme	Ziano di Fiemme
Trento	46.066666	11.116667	2020	2022	Novella	Novella
Trento	46.066666	11.116667	2020	2022	Borgo d'Anaunia	Borgo d'Anaunia
Trento	46.066666	11.116667	2020	2022	Ville di Fiemme	Ville di Fiemme
Trento	46.066666	11.116667	2020	2022	Terre d'Adige	Terre d'Adige

Region

Review_GID-3000441

describ...ype-300044

uri School_GID-447781

uri

reviewBody_GID-34405

has_Sen...ype-300044

Column_1	school_id	miur_code	text_review	school_name	sentiment
C	C		all'avanguardia.		
review_uri_7	school_id_65579	TNPS001015	Bellissimo ben organizzato con laboratori all'avanguardia.	LICEO DELLE SCIENZE UMANE	Positive
review_uri_8	school_id_29225	TN1E00900D	Valutazione 0	ASSOCIAZI... R. STEINER - PRIMARIA	Neutral
review_uri_9	school_id_29225	TN1E00900D	Ottimo personale docente ... scuola interessante ... insieme alla	ASSOCIAZI... R. STEINER - PRIMARIA	Positive

Reviews

	valore	valore_d	valore_n	anno_scol	anno_corso	miur_code	school_id	statistics_id
0.0416666666666666...	24.0	1.0	2020-21	5	TNCF005001	school_id_29271	5_2020_21_schoo...	
0.0	58.0	0.0	2020-21	1	TNCF004001	school_id_29273	1_2020_21_schoo...	
0.04285714285714285...	70.0	3.0	2020-21	2	TNCF004001	school_id_29273	2_2020_21_schoo...	
0.19047619047619047...	84.0	16.0	2020-21	3	TNCF004001	school_id_29273	3_2020_21_schoo...	

Ripetenti

School GID-447781

```

graph TD
    SGID[School GID-447781] -- "offers_...Type-44778" --> SCourse[Study Course ID-3000381]
    SGID -- "ufi" --> C1[Column_1]
    SGID -- "ufi" --> C2[teaching_unit]
    SGID -- "ufi" --> C3[date_for_find_type_of_courses_of_study]
    SGID -- "ufi" --> C4[title]
    SGID -- "ufi" --> C5[indirizzo]
    SGID -- "ufi" --> C6[school_id]
    SCourse -- "un" --> C1
    SCourse -- "has_Tit...ype-300038" --> C4
    SCourse -- "has_Art...ype-300038" --> C5
  
```

study_course_uri	id	date	qualification	operator	school_id
study_course_uri_89	2776435	2022-11-12	QUALIFICA PROFESSIONALE	Operatore ai servizi di impresa	school_id_29277
study_course_uri_90	2776435	2022-11-12	QUALIFICA PROFESSIONALE	Operatore ai servizi di vendita	school_id_29277
study_course_uri_91	2776435	2022-11-12	DIPLOMA PROFESSIONALE	Tecnico commerciale delle vendite	school_id_29277
study_course_uri_92	2776435	2022-11-12	DIPLOMA PROFESSIONALE	Tecnico dei servizi amministrativi e contabili	school_id_29277
study_course_uri_93	4254581	2022-11-12	QUALIFICA PROFESSIONALE	Operatore ai servizi di impresa	school_id_29279
study_course_uri_94	4254581	2022-11-12	QUALIFICA PROFESSIONALE	Operatore ai servizi di vendita	school_id_29279

School courses

Institute GID-15425

```

graph TD
    IGID[Institute GID-15425] -- "institute_...Type-15425" --> SGID[School GID-447781]
    SGID -- "ufi" --> C1[school_type]
    SGID -- "ufi" --> C2[address]
    SGID -- "ufi" --> C3[phone_number]
    SGID -- "ufi" --> C4[fax_number]
    SGID -- "ufi" --> C5[school_email]
    SGID -- "ufi" --> C6[management_email]
    SGID -- "ufi" --> C7[office_email]
    SGID -- "ufi" --> C8[website]
    SGID -- "ufi" --> C9[mail_code]
    SGID -- "ufi" --> C10[school_name]
    SGID -- "ufi" --> C11[school_id]
    SGID -- "ufi" --> C12[institute_id]
    SGID -- "ufi" --> C13[municipality]
    SGID -- "ufi" --> C14[institute_name]
    SGID -- "ufi" --> C15[stage]
  
```

Autonomous Management	Riviera Campus, 3	+39 0464 921300	+39 0464 921303	dip.vialedelcampo@...	dip.vialedelcampo@...	https://vialecampo...	TNCF007001	SETTORE SERVIZI	school_id	CNAF-Tono-d-16	Riva del Garda	TNCF007001	Professional School
Autonomous Management	Via Asmara, 14	+39 0461 923306	+39 0461 923305	dip.viaasmara@pm...	dip.viaasmara@pm...	https://viaasmara.e...	TNCF008001	SETTORE INDUSTRIA E ARTIGIANATO	school_id_29266	Risarcimento-Formazione-Professionale-CHAP-Viazzana2	Trento	TNCF008001	Professional School
Autonomous Management	Via Asmara, 14	+39 0461 923306	+39 0461 923305	dip.viaasmara@pm...	dip.viaasmara@pm...	https://viaasmara.e...	TNCF008001	SETTORE INDUSTRIA E ARTIGIANATO ECRAL	school_id_29267	Risarcimento-Formazione-Professionale-CHAP-Viazzana2	Trento	TNCF008001	Professional School
Autonomous Management	Viale Roma, 93	+39 0461 765146	+39 0461 765146	dipviterbo@pmcp...	dipviterbo@pmcp...	http://www.open...	TNCP772STA	SETTORE SERVIZI	school_id_29268	Risarcimento-Formazione-Professionale-Osp-Amica-	Lavis-Terme	TNCI_NUR_C	Professional School
Autonomous Management	Via Selenite, 5	+39 0464 437171	+39 0464 437171	dipviterbo@pmcp...	dipviterbo@pmcp...	http://www.open...	TNCI05001	SETTORE SERVIZI	school_id_29271	Risarcimento-Formazione-Professionale-Osp-Amica-	Rovereto	TNCI05001	Professional School

School list

Nome...	Cognome...	Titolo...	id_istituto...	institute_uri...	School responsible
Moreno	Murano	President of Managing Body	128	TN_NO_MIUR_C...	
Silvano	Medves	Pedagogican Coordinator	129	TN_NO_MIUR_C...	
Giorgio	Rella	President of Managing Body	130	TN_NO_MIUR_C...	
Roberta	Medeghini	Pedagogican Coordinator	131	TN_NO_MIUR_C...	

Diagram illustrating the relationship between School_GID-447781 and Education_ID-3000391:

```

graph TD
    A[School_GID-447781] -- "has_met...Type-44778" --> B[Educational_ID-3000391]
    A -- "uri" --> C[school_id=]
    A -- "uri" --> D[miur_code=]
    A -- "uri" --> E[score=]
    A -- "uri" --> F[school_name=]
    A -- "uri" --> G[uri=]
    B -- "has_Use...ype-300039" --> C
    B -- "uri" --> D
    B -- "uri" --> E
    B -- "uri" --> F
    B -- "uri" --> G

```

school_id=	miur_code=	score=	school_name=	uri=
C	SCUOLA SECONDARIA DI PRIMO GRADO "M. POLA"	C
school_id_29510	TNMM8130...	5,0	ISTITUTO TECNICO PER IL SETTORE ECONOMICO -	education_quality_2
school_id_29783	TNTD001011	5,0	ISTITUTO TECNICO PER IL SETTORE ECONOMICO	education_quality_3
school_id_29784	TNTD001509	5,0	ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO	education_quality_4
school_id_29785	TNTL00101B	4,7	ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO	education_quality_5
school_id_29786	TNTD001509	5,0	ISTITUTO	education_quality_6

Reviews Score

has_Tim...Type-44778	has_Num...Type-44778			produce...Type-44778			
uri				uri			
tip...orario	num_anni	etichetta	percentuale_ammes... si	anno	miur_code	school_id	statistics_id
Morning		2021-22	82.75862068965517	4	TNCF008001	school_id_29266	4_2021_22_schoo...
Morning	5	2021-22	94.11764705882352	5	TNCF008001	school_id_29266	5_2021_22_schoo...
Evening	5	2021-22	0.0	1	TNCF008001	school_id_29267	1_2021_22_schoo...
Evening	5	2021-22	0.0	2	TNCF008001	school_id_29267	2_2021_22_schoo...

Scrutini

Diagram illustrating the relationship between School (GID-447781) and Study_C (ID-3000381). The School offers the Study_C, which has an articulation and a title.

```

graph TD
    School((School GID-447781)) -- "offers_..._Type-44778" --> StudyC((Study_C ID-3000381))
    StudyC -- "uri" --> Articulation[Articulation]
    StudyC -- "uri" --> Title[Title]
    
```

<code>id_course</code>	<code>articulation</code>	<code>title</code>	<code>school_id</code>
705a23393d744d...	Databases	Informatica (L2)	school_id_000000...
e7778506513f4a4...	Analisi matematica 1	Ingegneria Civile (L2)	school_id_000000...
61e986d2108c48a...	Econometrics	Behavioural and Applied Economics - Economia Comportamentale e	school_id_000000...
Oba6c12cef154c2...	Agiografia II	Beni culturali (L2)	school_id_000000...
e0e56f7616074a6...	Signal, Image and Video	Biologia Quantitativa e Computazionale (LM)	school_id_000000...
e7efb421b7d14c6...	Immagini di futuro	Previsione Sociale (M2)	school_id_000000...
5141f1110c9c4f79...	Filologia latina - LM	Filologia e critica letteraria (LM)	school_id_000000...

Unitn courses

Professor GID-563161

reachab...Type-45425 works_a...Type-56316

uri has_Nam...ype-3000371as_Sur... GID-34003 contactPoint has_CUN...Type-56316 has_SSD...Type-56316

telephone_GID-34494

uni

id → Bruno Callegher Teaching fellow Università degli studi di Trento

name → Moreno Ferrarese Teaching assistant Università degli studi di Trento

surname → Barbieri Full professor 0461 281326 Scienze politiche e sociali (14) SOCIOLOGIA DEI PROCESSI ECONOMICI E Università degli studi di Trento

role →

phone_number →

cun →

ssd →

institute_id → C

lc0c137493b2ec9...	Bruno	Callegher	Teaching fellow				Università degli studi di Trento
ff561286a936502...	Moreno	Ferrarese	Teaching assistant				Università degli studi di Trento
069872e46afdf11...	Paolo	Barbieri	Full professor	0461 281326	Scienze politiche e sociali (14)	SOCIOLOGIA DEI PROCESSI ECONOMICI E	Università degli studi di Trento

Unitn professors

Unitn responsible

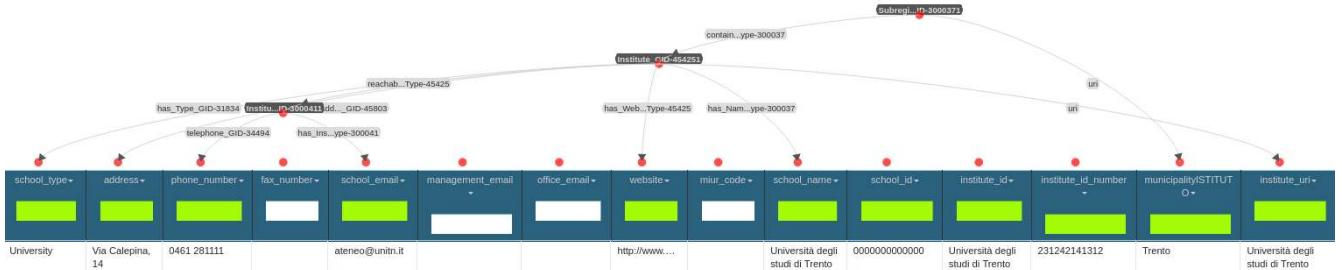
	id	name	surname	role	phone_number	institute_id
7b2760d7ee3b64...	Jonathan Brys	Bibee	Staff	0461 282980	Università degli studi di Trento	
a30ddcbf731ab31...	Marco	Buiatti	Staff	0464 808178	Università degli studi di Trento	
080044aed21e27...	Marianna	Menestrina	Staff	0461 283257	Università degli studi di Trento	
73669baa79d6b7...	Francesca	Tomasi	Head	0461 283653	Università degli studi di Trento	
95b0f4bb58c5bbb...	Mirian Soledad	Ayala De Chenu Abente	Staff	0461 282094	Università degli studi di Trento	
54f690ca02a2f965...	Massimo	Eccel	Staff	0461 281108	Università degli studi di Trento	

Unitn responsible

Unitn school

school_type	address	phone_number	fax_number	school_email	management_email	office_email	website	mail_code	school_name	school_id	university_id	institute_id	institute_id_number	institute_type	institute_uri	page
University	Via Calepina, 14	0461 281111		ateneo@unitn.it			http://www....		Università degli studi di Trento	231242141312	Trento			ISTITUTO	Università degli studi di Trento	University

Unitn school



Unitn institute

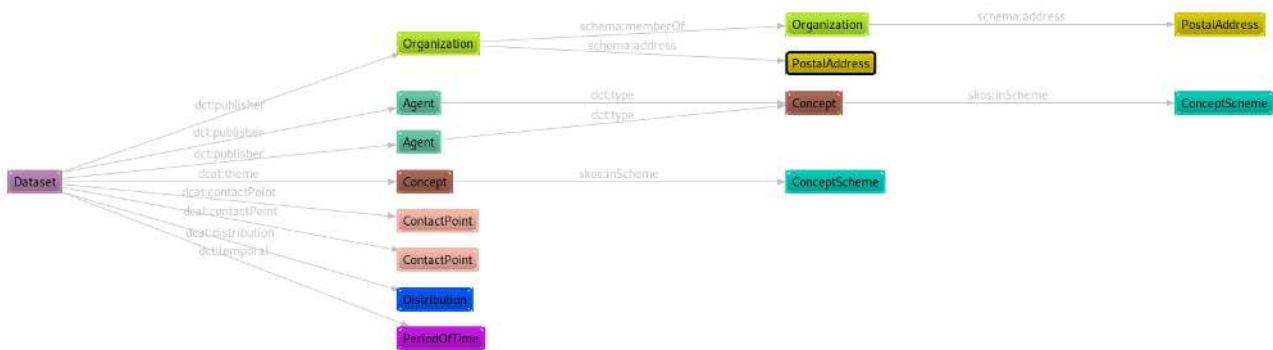
6.3 Evaluation

During the last phase of the *iTelos* methodology, the evaluation is crucial since directly provides feedback concerning the effectiveness and applicability of the final model.

Concerning the final evaluation, the key point is to employ the Knowledge Graph in real-world applications and employ query tools such as SQL or SPARQL to resolve CQs we have defined during the inception phase. To this end, we provide the final assessment of our work in the section concerning the Knowledge Graph outcome exploitation.

6.4 Metadata

Along with the final alignment, we have produced the associated Metadata compliant with the DCAT [2] vocabulary, thanks to the SHAPENess [9] metadata editor.



Data Integration Metadata

```

@prefix : <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix schema: <http://schema.org/> .
@prefix spdx: <http://spdx.org/rdf/terms#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix gsp: <http://www.opengis.net/ont/geosparql#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dqv: <http://www.w3.org/ns/dqv#> .

```

```

@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix sh: <http://www.w3.org/ns/shacl#> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix locn: <http://www.w3.org/ns/locn#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix adms: <http://www.w3.org/ns/adms#> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix http: <http://www.w3.org/2006/http#> .
@prefix dash: <http://datashapes.org/dash#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

<https://www.epos-eu.org/epos-dcat-ap#Organization/2dad489f-4ce3-4e0d-9122-b70df55b8ef2>
    rdf:type schema:Organization ;
    schema:address <https://www.epos-eu.org/epos-dcat-ap#PostalAddress/b0bb325d-f408-4dac-80e0-2389d8efb97f> ;
    schema:email "knowdive@disi.unitn.it" ;
    schema:identifier "http://knowdive.disi.unitn.it"^^xsd:anyURI ;
    schema:legalName "Knowdive" ;
    schema:leiCode "http://knowdive.disi.unitn.it" ;
    schema:logo "http://knowdive.disi.unitn.it/wp-content/uploads/knowdive-new-logo.png"^^xsd:anyURI ;
    schema:memberOf <https://www.epos-eu.org/epos-dcat-ap#Organization/78574deb-e40e-4c05-a5c6-40beeff01aa0> ;
    schema:url "http://knowdive.disi.unitn.it"^^xsd:anyURI .

<https://www.epos-eu.org/epos-dcat-ap#Concept/c7958238-a526-46b6-8a22-7b8b90f4b917>
    rdf:type skos:Concept ;
    skos:definition "Data Integration" ;
    skos:inScheme <https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/038c1555-1840-4b3c-a08a-a7dfd0378614> ;
    skos:prefLabel "KGE" .

```

<<https://www.epos-eu.org/epos-dcat-ap#Organization/78574deb-e40e-4c05-a5c6-40beef01aa0>>

 rdf:type schema:Organization ;
 schema:address <https://www.epos-eu.org/epos-dcat-ap#PostalAddress/cd879ddb-55b6-4620-a0e4-c85a2df9f483> ;
 schema:email "ateneo@unitn.it" , "ateneo@pec.unitn.it"
 ;
 schema:identifier "www.unitn.it"^^xsd:anyURI ;
 schema:legalName "Università degli Studi di Trento" ;
 schema:leiCode "00340520220" ;
 schema:logo https://static-cdn.unitn.it/sites/www.unitn.it/themes/unitn_theme/images/newlogo_unitn_en.png^^xsd:anyURI ;
 schema:telephone "0461281111" ;
 schema:url "www.unitn.it"^^xsd:anyURI .

<<https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/84efa52e-3879-4d09-b0b9-36362e915866>>

 rdf:type dct:PeriodOfTime ;
 schema:endDate "2022-12-11T00:00:00Z"^^xsd:dateTime ;
 schema:startDate "2022-12-09T00:00:00Z"^^xsd:dateTime .

<<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/b0bb325d-f408-4dac-80e0-2389d8efb97f>>

 rdf:type schema:PostalAddress ;
 schema:addressCountry "IT" ;
 schema:addressLocality "Povo (TN)" ;
 schema:postalCode "I-38123" ;
 schema:streetAddress "Via Sommarive 9" .

<<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/ca07b831-220f-4c49-bc3b-6e8215eb3238>>

 rdf:type schema>ContactPoint ;
 schema:availableLanguage "en-US" , "it-IT" ;
 schema:contactType "Member" ;
 schema:email "erich.robbi@studenti.unitn.it" .

<<https://www.epos-eu.org/epos-dcat-ap#Concept/520bf2b3-6fcc-45ad-971a-b17df5fe98f0>>

 rdf:type skos:Concept ;
 skos:definition "Data Integration is the third iTelos phase
 :\r\nInputs:\n * ETG\r\n * Dataset syntactically aligned\r\n\nOutputs:\n * Final Knowledge Graph" ;

```
skos:inScheme      <https://www.epos-eu.org/epos-dcat-ap#
ConceptScheme/a21146bd-a38c-49fe-a02f-f8575d1d73e2> ;
skos:prefLabel     "Data Integration Phase" .
```

<<https://www.epos-eu.org/epos-dcat-ap#Distribution/2d8e5529-0e68-4e0d-af09-49a96f1c6c25>>

```
    rdf:type          dcat:Distribution ;
    dct:conformsTo   "https://www.w3.org/OWL/" ;
    dct:description   "Data Integration regarding Education in Trentino." ;
    dct:format        "https://www.w3.org/OWL/^xsd:anyURI" ;
    dct:identifier    "https://github.com/samuelebortolotti/Education-Trentino/tree/main/Teleologies/Data%20Integration"^^xsd:anyURI ;
    dct:issued        "2022-12-09T00:00:00Z"^^xsd:dateTime ;
    dct:language      "English" ;
    dct:license       "https://opensource.org/licenses/MIT"^^xsd:anyURI ;
    dct:modified      "2022-12-11T00:00:00Z"^^xsd:dateTime ;
    dct:rights        "https://opensource.org/licenses/MIT" ;
    dct:title         "Education in Trentino", "Data Integration" ;
    dct:type          "Information"^^xsd:anyURI ;
    dcat:accessURL   "https://github.com/samuelebortolotti/Education-Trentino"^^xsd:anyURI ;
    dcat:downloadURL  "https://github.com/samuelebortolotti/Education-Trentino/tree/main/Teleologies/Data%20Integration"^^xsd:anyURI ;
    dcat:mediaType    "application/rdf+xml" ;
    foaf:page         "https://github.com/samuelebortolotti/Education-Trentino" .
```

<<https://www.epos-eu.org/epos-dcat-ap#PostalAddress/cd879ddb-55b6-4620-a0e4-c85a2df9f483>>

```
    rdf:type          schema:PostalAddress ;
    schema:addressCountry "IT" ;
    schema:addressLocality "Trento" ;
    schema:postalCode   "38122" ;
    schema:streetAddress "Via Calepina 14" .
```

<<https://www.epos-eu.org/epos-dcat-ap#ContactPoint/3a8a3044-3e7a-463d-94d9-15c4f4822938>>

```
    rdf:type          schema>ContactPoint ;
    schema:availableLanguage "it-IT", "en-US" ;
```

```
    schema:contactType          "Member" ;
    schema:email                "samuele.bortolotti@studenti.unitn.
                                it" .

<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/a21146bd-a38c-49
fe-a02f-f8575d1d73e2>
    rdf:type          skos:ConceptScheme ;
    dct:description  "Education in Trentino" ;
    dct:title        "Status" .

<https://www.epos-eu.org/epos-dcat-ap#ConceptScheme/038c1555-1840-4
b3c-a08a-a7dfd0378614>
    rdf:type          skos:ConceptScheme ;
    dct:description  "iTilos formal modeling" ;
    dct:title        "Formal Modeling" .

<https://www.epos-eu.org/epos-dcat-ap#Agent/828bf6b9-9f4b-4355-bcd0-
d68b188811cb>
    rdf:type      foaf:Agent ;
    dct:type     <https://www.epos-eu.org/epos-dcat-ap#Concept/
                  c7958238-a526-46b6-8a22-7b8b90f4b917> ;
    foaf:name   "Samuele Bortolotti" .

<https://www.epos-eu.org/epos-dcat-ap#Dataset/3a2e0c02-aebe-44d1-b688
-a5e21d81d088>
    rdf:type          dcat:Dataset ;
    dct:accessRights "Public" ;
    dct:accrualPeriodicity "Once"^^xsd:anyURI ;
    dct:conformsTo   "iTilos" ;
    dct:created      "2022-12-09T00:00:00Z"^^xsd:dateTime
                      ;
    dct:description   "Data Integration Metadata regarding
                      Education in Trentino." , "The project was made for the
                      Knowledge Graph Engineering (KGE) course held in the
                      academic year 2022/2023 at the University of Trento." ;
    dct:identifier    "https://github.com/samuelebortolotti
                      /Education-Trentino/tree/main/Teleologies/Data%20
                      Integration"^^xsd:anyURI ;
    dct:issued        "2022-12-09T00:00:00Z"^^xsd:dateTime
                      ;
    dct:language      "English" ;
    dct:modified      "2022-12-11T00:00:00Z"^^xsd:dateTime
                      ;
```

```
dct:publisher <https://www.epos-eu.org/epos-dcat-ap#Agent/f9cfb89e-3bb9-41ba-b263-e0386fce30ed> , <https://www.epos-eu.org/epos-dcat-ap#Organization/2dad489f-4ce3-4e0d-9122-b70df55b8ef2> , <https://www.epos-eu.org/epos-dcat-ap#Agent/828bf6b9-9f4b-4355-bcd0-d68b188811cb> ;  
dct:temporal <https://www.epos-eu.org/epos-dcat-ap#PeriodOfTime/84efa52e-3879-4d09-b0b9-36362e915866> ;  
dct:title "Entity Type Graph Metadata" ;  
dct:type "Information"^^xsd:anyURI ;  
dcat:contactPoint <https://www.epos-eu.org/epos-dcat-ap#ContactPoint/3a8a3044-3e7a-463d-94d9-15c4f4822938> , <https://www.epos-eu.org/epos-dcat-ap#ContactPoint/ca07b831-220f-4c49-bc3b-6e8215eb3238> ;  
dcat:distribution <https://www.epos-eu.org/epos-dcat-ap#Distribution/2d8e5529-0e68-4e0d-af09-49a96f1c6c25> ;  
dcat:keyword "KGE" , "Knowledge Graph Engineering" , "Erich Robbi" , "UnitN" , "University of Trento" , "Trentino" , "Trento" , "Fausto Giunchiglia" , "Samuele Bortolotti" , "iTilos" ;  
dcat:landingPage "https://github.com/samuelebortolotti/Education-Trentino/" ;  
dcat:theme <https://www.epos-eu.org/epos-dcat-ap#Concept/520bf2b3-6fcc-45ad-971a-b17df5fe98f0> ;  
foaf:page "https://github.com/samuelebortolotti/Education-Trentino/" .  
  
<https://www.epos-eu.org/epos-dcat-ap#Agent/f9cfb89e-3bb9-41ba-b263-e0386fce30ed>  
rdf:type foaf:Agent ;  
dct:type <https://www.epos-eu.org/epos-dcat-ap#Concept/c7958238-a526-46b6-8a22-7b8b90f4b917> ;  
foaf:name "Erich Robbi" .
```

7 Outcome Exploitation

The current document's concluding section tries to describe the results of the data integration procedure.

The final Knowledge Graph (KB) data, including the number of Etypes and characteristics, the number of entities for each Etype, and so on, are reported first. Additionally, using SPARQL, we demonstrate additional application examples based on our model.

A few basic conclusions are presented at the end of this section, covering the project's consideration and ideas for future development.

7.1 Knowledge Graph information

After completing this project, we successfully built a Knowledge Graph for Education in Trentino which included all relevant information about educational facilities and courses offered in the Trentino region.

We estimated the total numbers of our final Knowledge Graph components, including Etypes and attributes. The table below displays some information concerning the Knowledge Graph which was built thanks to the *Explore* functionality of GraphDB.

We have a total of 33 classes, where 11 are from the original teleology and the remaining ones are from reference ontologies.

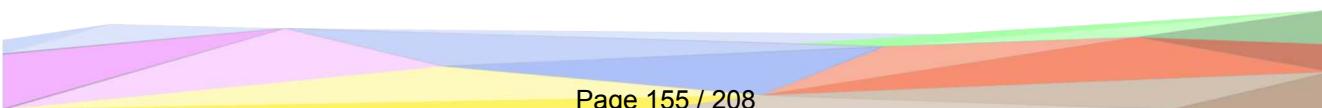
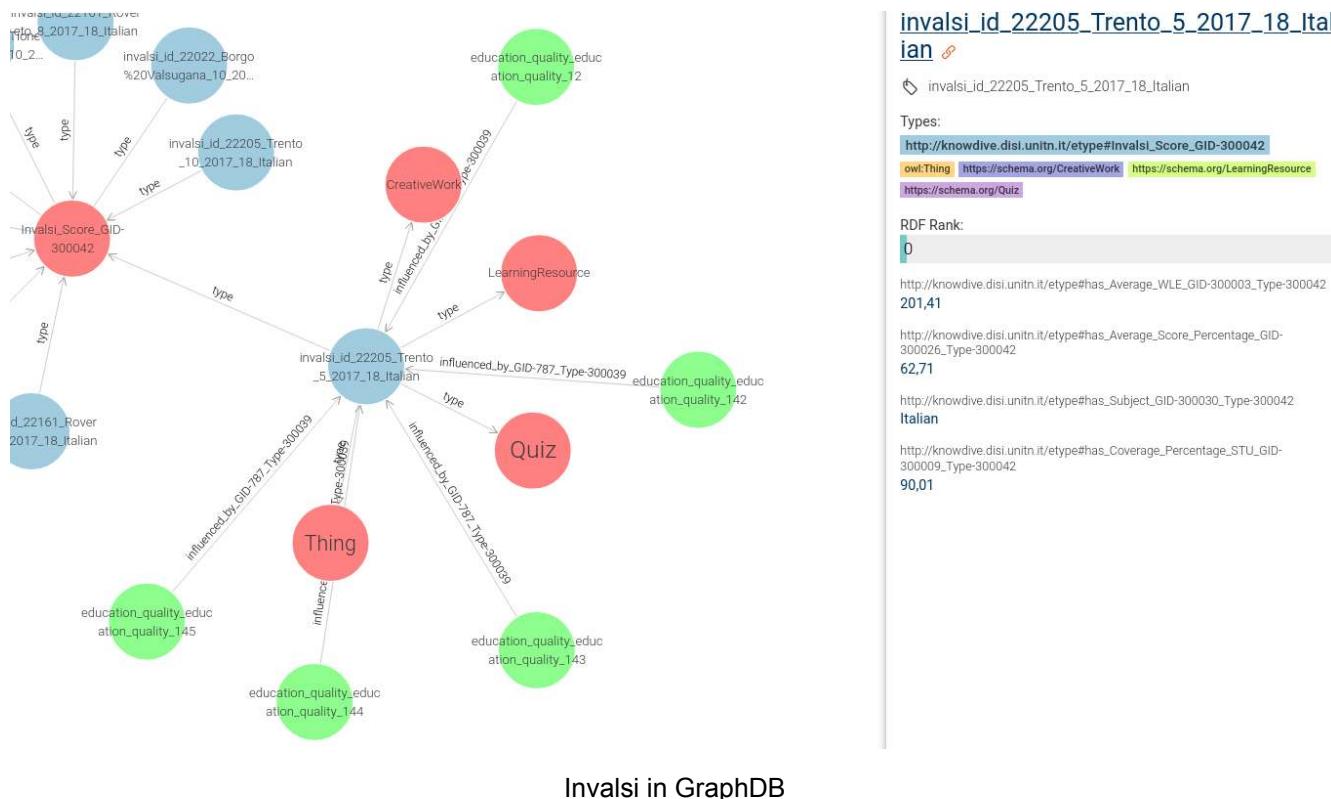
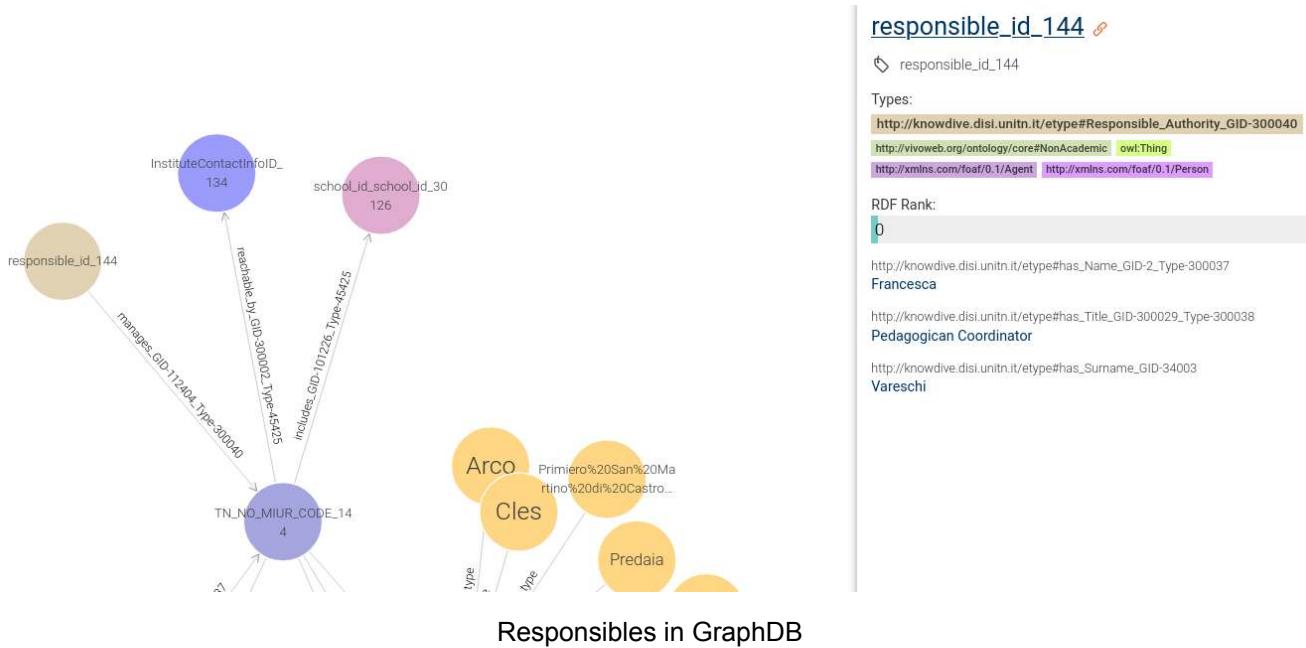
Etype	Instances
Subregional Academic Division	183
Institute	273
School	724
Institute Contact Information	273
Responsible Authority	1239
Study Course	3326
Professor	1804
School Statistics	3846
Review	944
Education Quality	344
Invalsi score	36

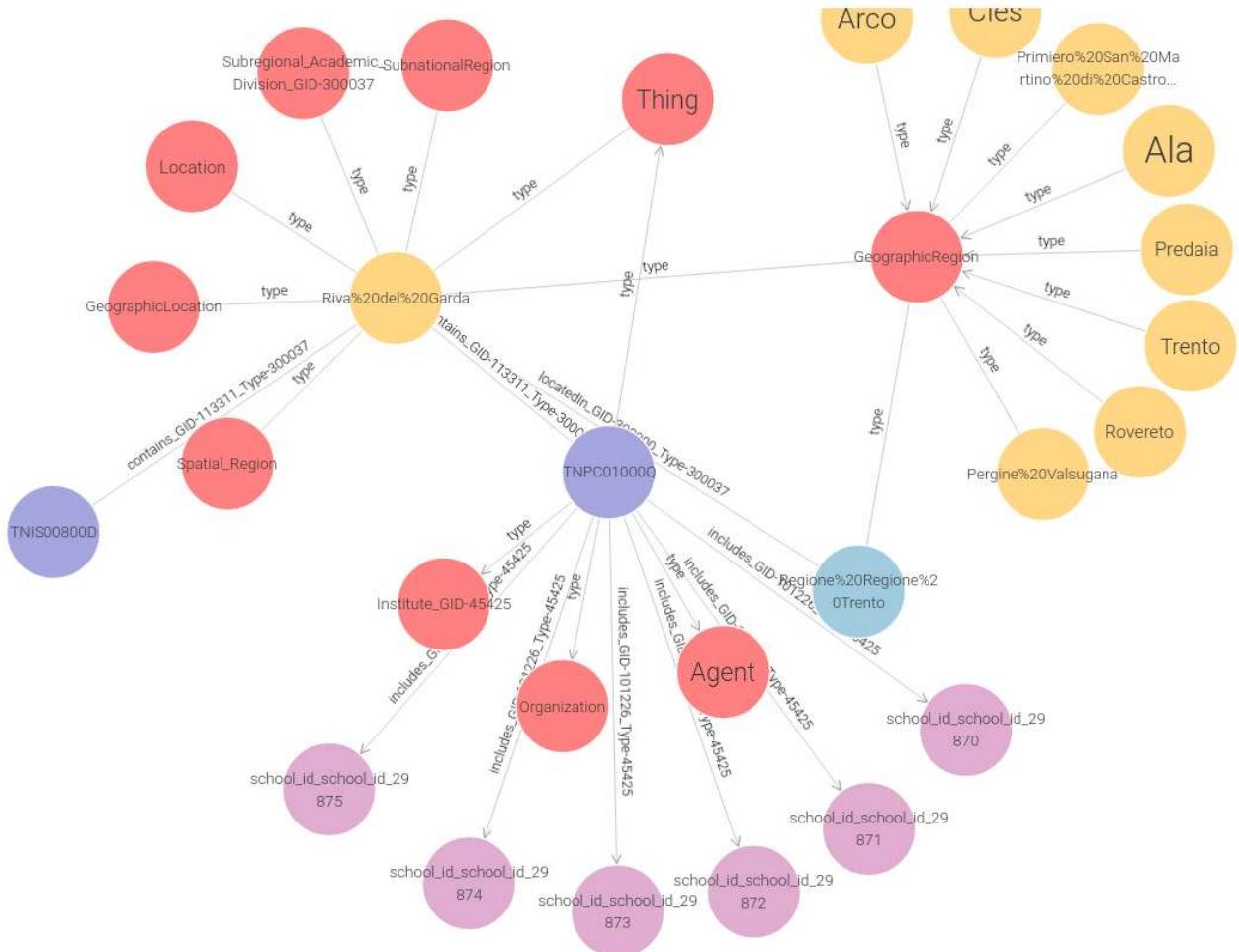
Interestingly, we have found out that the number of subregional academic divisions is higher (183) with respect to the real amount of municipalities that Trentino has (166). The reason behind that is related to the data we have collected. Indeed, some institutes or schools were connected to former municipalities. As an effective representation *Brez*²³ and *Cagnò*²⁴ which are instances of our Subregional Academic Division Etype are not municipalities anymore, since from January 2020 they were merged with other institutions. However, a subregional academic division is more specific than a standard municipality, thus the semantics of the data is still preserved.

²³<https://it.wikipedia.org/wiki/Brez>

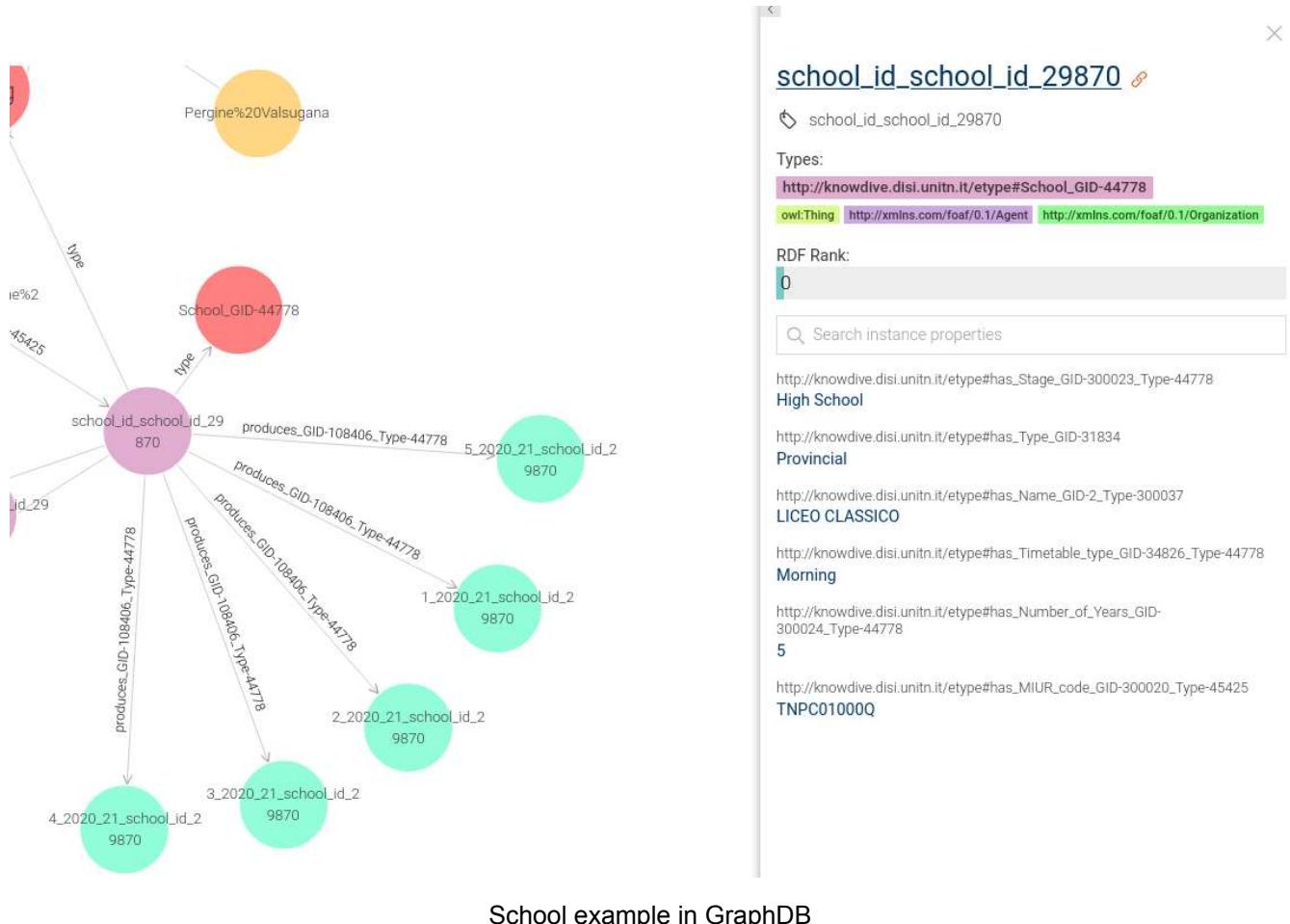
²⁴<https://it.wikipedia.org/wiki/Cagn%C3%B2>

Here we present some screenshots concerning the final Knowledge Graph thanks to the Visual Graph functionality of GraphDB.





General School and Institute overview in GraphDB



School example in GraphDB

7.2 Model application

The objective of this project is to integrate data and knowledge related to Trentino educational facilities into a knowledge graph so that users can use it to get the answers to their queries.

In compliance with the final objective, we used SPARQL²⁵ via GraphDB to see whether our model can be applied to real-world scenarios. As evidenced by the outcomes of certain deployed instances, our model answers the queries in a satisfactory way.

The following sections provide information regarding the questions, SPARQLs outcomes, and visualization.

7.2.1 SPARQL

SPARQL is an RDF query language that can access and operate with data that is stored in Resource Description Framework (RDF) format. It is regarded as one of the fundamental technologies of the semantic web and was turned into a standard by the RDF Data Access Working Group (DAWG) of the World Wide Web Consortium [15]. In the following section, we list most of the SPARQL queries employed to answer the competency questions. The missing competency

²⁵<https://www.w3.org/TR/rdf-sparql-query/>

questions can be trivially answered by looking at the result of the previous ones. Moreover, since the knowledge graph comprises a huge quantity of data, we have decided to integrate bonus queries which may provide additional and useful information for the project's purpose.

7.2.1.1 Competency Question 1

Give me how many schools have courses on Software Engineering

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select (count(?name) as ?count) where {
    ?school rdf:type ds:School_GID-44778 .
    ?course rdf:type ds:Study_Course_GID-300038 .
    ?school ds:has_Name_GID-2_Type-300037 ?name .
    ?course ds:has_Articulation_GID-300007_Type-300038 ?articulation .
    ?school ds:offers_GID-111716_Type-44778 ?course .
    FILTER(contains(lcase(?articulation), "informatica"))
}
```

Result

The screenshot shows a SPARQL query interface with the following details:

- Query Editor:** Displays the SPARQL query code.
- Run Button:** A red button labeled "Run".
- Download as:** A button to download the results in various formats.
- Table:** The results are presented in a table format.
- Raw Response:** An option to view the raw SPARQL response.
- Pivot Table:** An option to view the results as a pivot table.
- Google Chart:** An option to view the results as a chart.
- Filter query results:** A text input field for filtering the results.
- Showing results from 1 to 1 of 1. Query took 0.3s, moments ago.** A message indicating the query execution status.
- Data Table:**

	count
1	*23*^^xsd:integer

7.2.1.2 Competency Question 2

Give me schools with courses on Software Engineering

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select distinct ?institute_name ?name
where {
  ?institute rdf:type
    ds:Institute_GID-45425 .
  ?school rdf:type
    ds:School_GID-44778 .
  ?course rdf:type
    ds:Study_Course_GID-300038 .

  ?school
    ds:has_Name_GID-2_Type-300037
    ?name .
  ?course ds:has_Title_GID-300029_Type-300038 ?title
    .
  ?course ds:has_Articulation_GID-30007_Type-300038 ?articulation
    .
  ?institute
    ds:has_Name_GID-2_Type-300037
    ?institute_name .
  ?school ds:offers_GID-111716_Type-44778 ?course
    .
  ?institute ds:includes_GID-101226_Type-45425
    ?school.
  FILTER(contains(lcase(?articulation), "informatica"))
}
  
```

Result

The screenshot shows a SPARQL query interface with the following details:

- Query Editor:** The top half displays the SPARQL query with syntax highlighting for prefixes, variables, and predicates.
- Run Button:** A red "Run" button is located on the right side of the editor.
- Results Table:** The bottom half shows a table with two columns: "institute_name" and "name". The table contains 6 rows of data, each with a numerical index (1-6) and corresponding names.

	institute_name	name
1	"Università degli studi di Trento"	"Università degli studi di Trento"
2	"ISTITUTO DI ISTRUZIONE 'LA ROSA BIANCA' - CAVALESE"	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE - PREDAZZO"
3	"ISTITUTO TECNICO ECONOMICO E TECNOLOGICO 'C.A. PILATTI' - CLES"	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO"
4	"ISTITUTO TECNICO TECNOLOGICO 'G. MARCONI' - ROVERETO"	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO"
5	"ISTITUTO TECNICO TECNOLOGICO 'G. MARCONI' - ROVERETO"	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE"
6	"ISTITUTO TECNICO TECNOLOGICO 'M. BUONARROTI' - TRENTO"	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO"

7.2.1.3 Competency Question 3 and 11

*Give me a list of courses at the University of Trento I can consult when I finish High School
 Get me all study programs the University of Trento is proposing*

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?name ?title ?articulation
where {
    ?school rdf:type
        ds:School_GID-44778 .
    ?course rdf:type
        ds:Study_Course_GID-300038 .
    ?school
        ds:has_Name_GID-2_Type-300037
        ?name .
    ?course ds:has_Title_GID-300029_Ty
        pe-300038 ?title
        .
    ?course ds:has_Articulation_GID-30
        0007_Type-300038 ?articulation
        .
    ?school ds:offers_GID-111716_Type-
        44778 ?course
        .
FILTER(?school =
    <http://localhost:8080/source/
    school_id_000000000000>)
```

Result

The screenshot shows a SPARQL query interface with the following components:

- Code Editor:** Displays the SPARQL query above.
- Results Table:** Shows a table with three columns: name, title, and articulation. The data is as follows:

	name	title	articulation
1	"Università degli studi di Trento"	"Informatica (L2)"	"Databases"
2	"Università degli studi di Trento"	"Ingegneria Civile (L2)"	"Analisi matematica 1"
3	"Università degli studi di Trento"	"Behavioural and Applied Economics - Economia Comportamentale e Applicata (LM)"	"Econometrics"
4	"Università degli studi di Trento"	"Beni culturali (L2)"	"Agiografia II"
5	"Università degli studi di Trento"	"Biologia Quantitativa e Computazionale (LM)"	"Signal, Image and Video"

- Preview Area:** Shows a snippet of the query results with some rows highlighted in orange.

7.2.1.4 Competency Question 4

Give me schools that prepare me to become a Software Engineer

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?name ?title ?articulation
where {
  ?school rdf:type ds:School_GID-44778 .
  ?course rdf:type
    ds:Study_Course_GID-300038 .
  ?school
    ds:has_Name_GID-2_Type-300037
    ?name .
  ?course ds:has_Title_GID-300029_Ty
    pe-300038 ?title
    .
  ?course ds:has_Articulation_GID-30
    0007_Type-300038 ?articulation
    .
  ?school ds:offers_GID-111716_Type-
    44778 ?course
    .
  FILTER(contains(lcase(?articulatio
    n),
    "informatica"))
}
```

Result

The screenshot shows a SPARQL query results interface. At the top, there are two tabs: 'Subregional_Academic_Division...' and 'School_GID-44778'. Below the tabs is a code editor containing the SPARQL query. To the right of the code editor are several icons: a save icon, a copy icon, a link icon, a refresh icon, and a magnifying glass icon. A red 'Run' button is located at the bottom right of the code editor. Below the code editor, there are four buttons: 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. The 'Table' button is highlighted. At the bottom of the interface, there is a 'Filter query results' input field and a message stating 'Showing results from 1 to 23 of 23. Query took 0.2s, moments ago.' The main area displays an 8x3 table with the following data:

	name	title	articulation
1	"Università degli studi di Trento"	"Scienze e Tecnologie Biomolecolari (L2)"	"Informatica"
2	"Università degli studi di Trento"	"Ingegneria Industriale (L2)"	"Informatica"
3	"Università degli studi di Trento"	"Medicina e Chirurgia (LM6)"	"Fisica e informatica"
4	"Università degli studi di Trento"	"Viticoltura ed Enologia (L2)"	"Informatica di base (abilità informatiche)"
5	"Università degli studi di Trento"	"Giurisprudenza (LM5)"	"Diritto penale dell'informatica"
6	"Università degli studi di Trento"	"Interfacce e Tecnologie della Comunicazione (L2)"	"Informatica ed elementi di programmazione II"
7	"Università degli studi di Trento"	"Interfacce e Tecnologie della Comunicazione (L2)"	"Informatica ed elementi di programmazione I"
8	"Università degli studi di Trento"	"Matematica (L2)"	"Informatica"

7.2.1.5 Competency Question 5

Give me popular schools

This is the sole Competency Question that has remained unanswered. The reason is that the latter was thought of in the early stages of the project when we did not yet have a clear idea of which data to collect and from where. During the evaluation phase, we realized we had not collected data that could be useful for quantifying a school's popularity factor.

Popularity may be measured in a variety of ways, depending on the goals of the individual. In our scenario, we considered rating a school's popularity based on the number of students who attend it. Unfortunately, this is not feasible since we have not been able to determine how many pupils attend each institution but rather how many students are located within each subregional academic division.

However, it is still possible to query the schools that got the highest number of reviews, which indeed, can be thought of as a form of "popularity", which unfortunately differs from the one we formulated during the inception phase:

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/type#>
select ?institute_name ?school_name (count(?review) as ?review_cont) where {
    ?school rdf:type ds:School_GID-44778 .
    ?institute rdf:type ds:Institute_GID-45425 .
    ?review rdf:type ds:Review_GID-300044 .

    ?school ds:offers_GID-111716_Type-44778 ?course .
    ?institute ds:includes_GID-101226_Type-45425 ?school .

    ?school ds:has_Name_GID-2_Type-300037 ?school_name .
    ?institute ds:has_Name_GID-2_Type-300037 ?institute_name .
    ?review ds:describes_GID-104985_Type-300044 ?school .

} group by ?institute_name ?school_name
order by desc(?review_cont)
LIMIT 5
```

Result

institute_name	school_name	review_cont
1 'FONDAZIONE EDMUND MACH - ISTITUTO AG RARIO SAN MICHELE ALLADIGE'	'SETTORE AGRICOLTURA E AMBIENTE'	'247'^xsd:integer
2 'ISTITUTO TECNICO TECNOLOGICO 'M. BUON ARROTT' - TRENTO'	'ISTITUTO TECNICO PER IL SETTORE TECNOL OGICO'	'120'^xsd:integer
3 'ISTITUTO TECNICO TECNOLOGICO 'G. MARCONI' - ROVERETO'	'ISTITUTO TECNICO PER IL SETTORE TECNOL OGICO'	'119'^xsd:integer
4 'GIUSEPPE VERONESI - CENTRO DI ISTRUZIONE SCOLASTICA E DI FORMAZIONE PROFESSIONALE'	'SETTORE INDUSTRIA E ARTIGIANATO'	'78'^xsd:integer
5 'ISTITUTO TECNICO TECNOLOGICO 'M. BUON ARROTT' - TRENTO'	'ISTITUTO TECNICO PER IL SETTORE TECNOL OGICO SERALE'	'60'^xsd:integer

7.2.1.6 Competency Question 6

Give me schools that have courses in the multidisciplinary area

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?nome_scuola ?nome_istituto (count(?corso)
as ?number_of_courses) where {
?corso rdf:type ds:Study_Course_GID-300038 .
?scuola rdf:type ds:School_GID-44778 .
?istituto rdf:type ds:Institute_GID-45425 .

?istituto ds:includes_GID-101226_Type-45425
?scuola .
?scuola ds:offers_GID-111716_Type-44778 ?corso
.

?scuola ds:has_Name_GID-2_Type-300037
?nome_scuola .
?istituto ds:has_Name_GID-2_Type-300037
?nome_istituto .
} group by ?nome_scuola ?nome_istituto
having(?number_of_courses > 2) order by
desc(?number_of_courses)
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query, various configuration buttons (e.g., Run, Download as), and a table below showing the results. The table has columns: nome_scuola, nome_istituto, and number_of_courses. The results are as follows:

	nome_scuola	nome_istituto	number_of_courses
1	"Università degli studi di Trento"	"Università degli studi di Trento"	"2041"**xsd:integer
2	"SETTORE INDUSTRIA E ARTIGIANATO"	"GIUSEPPE VERONESI - CENTRO DI ISTRUZIONE SCOLASTICA E DI FORMAZIONE PROFESSIONALE"	"13"**xsd:integer
3	"SETTORE AGRICOLTURA E AMBIENTE"	"FONDAZIONE EDMUND MACH - ISTITUTO AGRAZIO SAN MICHELE ALL'ADIGE"	"13"**xsd:integer
4	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO"	"ISTITUTO TECNICO TECNOLOGICO 'M. BUONARROTI' - TRENTO"	"12"**xsd:integer
5	"SETTORE INDUSTRIA E ARTIGIANATO"	"CENTRO FORMAZIONE PROFESSIONALE ENAIP - VILLAZZANO"	"11"**xsd:integer
6	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO"	"ISTITUTO TECNICO ECONOMICO E TECNOLOGICO 'C. PILATI' - CLES"	"9"**xsd:integer

7.2.1.7 Competency Question 7

Give me the schools' contact information

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX sc: <https://schema.org/>
select ?name ?office_mail
    ?institute_mail ?management_mail
    ?telephone ?fax where {
        ?school rdf:type
            → ds:School_GID-44778 .
        ?institute rdf:type
            → ds:Institute_GID-45425 .
        ?contact rdf:type ds:Institute_Con
            → tact_Information_GID-300041
            → .
        ?institute ds:includes_GID-101226_
            → Type-45425 ?school
            → .
        ?institute ds:reachable_by_GID-300
            → 002_Type-45425 ?contact
            → .
        ?school
            → ds:has_Name_GID-2_Type-300037
            → ?name .

        ?contact ds:has_Office_Mail_GID-30
            → 0017_Type-300041 ?office_mail
            → .
        ?contact ds:has_Institute_Mail_GID_
            → -300035_Type-300041
            → ?institute_mail .

        ?contact ds:has_Management_Mail_GI
            → D-300016_Type-300041
            → ?management_mail .

        ?contact sc:telephone_GID-34494
            → ?telephone .

        ?contact sc:faxNumber_GID-300008
            → ?fax .

    }

```

Result

	name	office_mail	institute_mail	management_mail	telephone	fax
1	"SCUOLA PRIMARIA 'C. CORRADI' NOVALEDO"	"segr.ic.centrovalsgana@scuole.provincia.it"	"ic.centrovalsgana@pec.provincia.it"	"dir.ic.centrovalsgana@scuole.provincia.it"	"+39 0461 764581"	"+39 0461 771046"
2	"SCUOLA PRIMARIA 'P. MARTINELLI' RONCEGO TERMENZO"	"segr.ic.centrovalsgana@scuole.provincia.it"	"ic.centrovalsgana@pec.provincia.it"	"dir.ic.centrovalsgana@scuole.provincia.it"	+39 0461 764581"	+39 0461 771046"
3	"SCUOLA SECONDARIA DI PRIMO GRADO 'M. POLA' RONCEGO TERME"	"segr.ic.centrovalsgana@scuole.provincia.it"	"ic.centrovalsgana@pec.provincia.it"	"dir.ic.centrovalsgana@scuole.provincia.it"	+39 0461 764581"	+39 0461 771046"
4	"SCUOLA PRIMARIA 'T. V. GOZZER' CASTELNUOVO"	"segr.ic.centrovalsgana@scuole.provincia.it"	"ic.centrovalsgana@pec.provincia.it"	"dir.ic.centrovalsgana@scuole.provincia.it"	+39 0461 764581"	+39 0461 771046"
5	"SCUOLA PRIMARIA MARTER"	"segr.ic.centrovalsgana@scuole.provincia.it"	"ic.centrovalsgana@pec.provincia.it"	"dir.ic.centrovalsgana@scuole.provincia.it"	+39 0461 764581"	+39 0461 771046"

7.2.1.8 Competency Question 8

Give me study programs of the schools in Trentino

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?institute_school ?name_school
?articulation ?title where {
  ?course rdf:type
    ds:Study_Course_GID-300038 .
  ?institute rdf:type
    ds:Institute_GID-45425 .
  ?school rdf:type
    ds:School_GID-44778 .

  ?institute ds:includes_GID-101226_
    Type-45425 ?school
  .
  ?school ds:offers_GID-111716_Type-
    44778 ?course
  .
  ?institute
    ds:has_Name_GID-2_Type-300037
    ?institute_school .
  ?school
    ds:has_Name_GID-2_Type-300037
    ?name_school .
  ?course ds:has_Articulation_GID-30
    0007_Type-300038 ?articulation
  .
  ?course ds:has_Title_GID-300029_Ty
    pe-300038 ?title
  .
}
```

Result

The screenshot shows the Subregional_Academic RDF interface. At the top, there is a search bar with the hint "Hint: 'abc' matches 'abC', 'ab c*' and 'ab-c*'". Below the search bar is the SPARQL query code. To the right of the query, there are several icons for interacting with the data, including a red 'Run' button. Below the query, there are tabs for 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. A 'Download as' button is also present. The main area displays the results of the query, which are listed in a table with columns: institute_school, nome_school, articulation, and title. The results show various study programs offered by the University of Trento.

	institute_school	nome_school	articulation	title
1	"Università degli studi di Trento"	"Università degli studi di Trento"	"Databases"	"Ingegneria Civile (L2)"
2	"Università degli studi di Trento"	"Università degli studi di Trento"	"Analisi matematica 1"	"Ingegneria Civile (L2)"
3	"Università degli studi di Trento"	"Università degli studi di Trento"	"Econometrics"	"Behavioural and Applied Economics - Economia Comportamentale e Applicata (LM)"
4	"Università degli studi di Trento"	"Università degli studi di Trento"	"Agiografia II"	"Beni culturali (L2)"
5	"Università degli studi di Trento"	"Università degli studi di Trento"	"Signal, Image and Video"	"Biologia Quantitativa e Computazionale (LM)"
6	"Università degli studi di Trento"	"Università degli studi di Trento"	"Immagini di futuro"	"Previsione Sociale (M2)"

7.2.1.9 Competency Question 9

Give me the University of Trento's study programs similar to mine

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?institute_name ?school_name
?articulation ?title where {
    ?institute rdf:type
        → ds:Institute_GID-45425 .
    ?school rdf:type
        → ds:School_GID-44778 .
    ?course rdf:type
        → ds:Study_Course_GID-300038 .

    ?institute ds:includes_GID-101226_
        → Type-45425 ?school
        → .
    ?school ds:offers_GID-111716_Type-
        → 44778 ?course
        → .

    ?course ds:has_Articulation_GID-30
        → 0007_Type-300038 ?articulation
        → .
    ?course ds:has_Title_GID-300029_Ty
        → pe-300038 ?title
        → .

    ?institute
        → ds:has_Name_GID-2_Type-300037
        → ?institute_name .
    ?school
        → ds:has_Name_GID-2_Type-300037
        → ?school_name .
    FILTER(contains(lcase(?articulatio
        → n), "chimica") && ?school =
        → <http://localhost:8080/source/
        → school_id_000000000000>
}
}
```

Result

The screenshot shows a SPARQL query results interface. The query has been run successfully, returning 20 results. The results are displayed in a table with columns: Institute_name, school_name, articulation, and title. The table contains 6 rows of data.

Institute_name	school_name	articulation	title
1 "Università degli studi di Trento"	"Università degli studi di Trento"	"Chimica con esercitazioni di laboratorio"	"Fisica (L2)"
2 "Università degli studi di Trento"	"Università degli studi di Trento"	"Chimica e biochimica"	"Medicina e Chirurgia (LM6)"
3 "Università degli studi di Trento"	"Università degli studi di Trento"	"Chimica"	"Ingegneria per il territorio (L2)"
4 "Università degli studi di Trento"	"Università degli studi di Trento"	"Laboratorio di chimica"	"Ingegneria Industriale (L2)"
5 "Università degli studi di Trento"	"Università degli studi di Trento"	"Chimica per il restauro"	"Beni culturali (L2)"
6 "Università degli studi di Trento"	"Università degli studi di Trento"	"Chimica e microbiologia enologica"	"Viticoltura ed Enologia (L2)"

7.2.1.10 Competency Question 10

Give me information about the university's biology department

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?institute_name ?website
?address ?type where {
    ?institute rdf:type
        ds:Institute_GID-45425 .
    ?school rdf:type
        ds:School_GID-44778 .

?institute ds:includes_GID-101226_
    Type-45425 ?school
    .
    ?institute
        ds:has_Name_GID-2_Type-300037
        ?institute_name .
?institute ds:has_Website_GID-3412_
    6_Type-45425 ?website
    .
    ?institute
        ds:has_Address_GID-45803
        ?address .
?institute ds:has_Type_GID-31834
    ?type .
FILTER(?institute = <http://localhost:8080/source/Università%20degli%20studi%20di%20Trento>)
}
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the SPARQL query. Below it is a toolbar with various icons for copy, paste, refresh, and download. A status bar at the bottom right indicates "Showing results from 1 to 1 of 1. Query took 0.1s, today at 18:27." The main area displays a table with four columns: Institute_name, website, address, and type. There is one row of data: "Università degli studi di Trento", "http://www.unitn.it", "Via Calepina, 14", and "University".

Institute_name	website	address	type
"Università degli studi di Trento"	"http://www.unitn.it"	"Via Calepina, 14"	"University"

In the original version of our knowledge graph, we intended to model also university department. However, in the latter refinement, we have collapsed everything within the Study Course entity which was meant to represent a study course regardless of whether it is from a university or another school. This is why, we are returning general information about the University of Trento as an institution, rather than of the specific department.

7.2.1.11 Competency Question 12

Give me Kindergartens near Pergine Valsugana

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name where {
    ?institute rdf:type
        → ds:Institute_GID-45425 .
    ?school rdf:type
        → ds:School_GID-44778 .
    ?municipality rdf:type
        → ds:Subregional_Academic_Division_GID-300037
        → .
    ?municipality ds:contains_GID-1133
        → 11_Type-300037
        → ?institute.
    ?institute ds:includes_GID-101226_
        → Type-45425 ?school
        → .

    ?school
        → ds:has_Name_GID-2_Type-300037
        → ?school_name .
    ?institute ds:has_Type_GID-31834
        → ?type .
    ?municipality
        → ds:has_Name_GID-2_Type-300037
        → ?municipality_name .
    FILTER(?type = "Kindergarten" &&
        → ?municipality_name = "Pergine
        → Valsugana")
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query and some metadata like 'Subregional_Academic_Division_GID-300037' and 'School_GID-44778'. Below the code editor are several icons for file operations (Save, Open, Copy, etc.). A 'Run' button is also present. Below the code editor, there are tabs for 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. The main area displays a table with one column labeled 'school_name'. The table contains five rows of data:

school_name
"SCUOLA DELL'INFANZIA DI ISCHIA DI PERGINE"
"SCUOLA DELL'INFANZIA DI MADRANO"
"SCUOLA DELL'INFANZIA DI SUSÀ"
"SCUOLA DELL'INFANZIA DI PERGINE "G.B. CHIMELLI"
"SCUOLA DELL'INFANZIA DI RONCOGNONE DI PERGINE"

A message at the bottom of the table area says 'Showing results from 1 to 5 of 5. Query took 0.2s, moments ago.'

7.2.1.12 Competency Question 13

Give me Kindergartens in Trentino

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/eype#>
PREFIX vivo: <http://vivoweb.org/ontology/core#>
select ?school_name where {
    ?institute rdf:type
        → ds:Institute_GID-45425 .
    ?school rdf:type
        → ds:School_GID-44778 .
    ?municipality rdf:type
        → ds:Subregional_Academic_Division_GID-300037
        → on_GID-300037
        → .
    ?region rdf:type
        → vivo:GeographicRegion .

    ?municipality ds:locatedIn_GID-300_000_Type-300037 ?region
    → .
    ?municipality ds:contains_GID-1133_11_Type-300037 ?institute
    → .
    ?institute ds:includes_GID-101226_Type-45425 ?school
    → .

    ?school
        → ds:has_Name_GID-2_Type-300037
        → ?school_name .
    ?institute ds:has_Type_GID-31834
        → ?type
    FILTER(?type = "Kindergarten")
}
```

Result

The screenshot shows a SPARQL query results interface. The query has been run, and the results are displayed in a table. The table has one column labeled 'school_name' containing the names of eight kindergartens. The names listed are: "ASILO COLLE FIORITO", "SCUOLA DELL'INFANZIA DI VIGO DI FASSA", "SCUOLA DELL'INFANZIA DI VIGO DI TON", "SCUOLA DELL'INFANZIA DI VILLAMONTAGNA", "SCUOLA DELL'INFANZIA DI VILLAZZANO", "SCUOLA DELL'INFANZIA DI VOLANO", "SCUOLA DELL'INFANZIA DI ALBIANO", and "SCUOLA DELL'INFANZIA DI ALDENO". The interface includes various buttons for saving, printing, and sharing the results.

school_name
"ASILO COLLE FIORITO"
"SCUOLA DELL'INFANZIA DI VIGO DI FASSA"
"SCUOLA DELL'INFANZIA DI VIGO DI TON"
"SCUOLA DELL'INFANZIA DI VILLAMONTAGNA"
"SCUOLA DELL'INFANZIA DI VILLAZZANO"
"SCUOLA DELL'INFANZIA DI VOLANO"
"SCUOLA DELL'INFANZIA DI ALBIANO"
"SCUOLA DELL'INFANZIA DI ALDENO"

7.2.1.13 Competency Question 14

Give me contact informations about this particular Kindergarten

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX sc: <https://schema.org/>
select ?type ?school_name ?office_mail
    ?institute_mail ?management_mail ?telephone
    ?fax where {
        ?institute rdf:type ds:Institute_GID-45425 .
        ?school rdf:type ds:School_GID-44778 .
        ?contact rdf:type ds:Institute_Contact_Informa_
        tion_GID-300041
        ?institute ds:includes_GID-101226_Type-45425
        ?school .
        ?institute
        → ds:reachable_by_GID-300002_Type-45425
        → ?contact .

        ?school ds:has_Name_GID-2_Type-300037
        → ?school_name .
        optional { ?institute ds:has_Type_GID-31834
        → ?type . }
        optional { ?contact
        → ds:has_Office_Mail_GID-300017_Type-300041
        → ?office_mail . }
        optional { ?contact ds:has_Institute_Mail_GID-
        → 300035_Type-300041 ?institute_mail .
        }
        optional { ?contact ds:has_Management_Mail_GID_
        → -300016_Type-300041 ?management_mail .
        }
        optional { ?contact sc:telephone_GID-34494
        → ?telephone . }
        optional { ?contact sc:faxNumber_GID-300008
        → ?fax . }
        FILTER(?type = "Kindergarten")
    }

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the SPARQL query. Below it is a table with 5 rows of data. The columns are labeled: type, school_name, office_mail, institute_mail, management_mail, telephone, and fax. The data is as follows:

	type	school_name	office_mail	institute_mail	management_mail	telephone	fax
1	"Kindergarten"	"ASILO COLLE FIORITO"					
2	"Kindergarten"	"SCUOLA DELLA INFANZIA DI VIGO DI FASSA"	"vigo.fassa.segretario@fpmtn.it"	"vigo.fassa.presidente@fpmtn.it"		"0462-763194"	
3	"Kindergarten"	"SCUOLA DELLA INFANZIA DI VIGO DI TONI"	"vigo.ton.segretario@fpmtn.it"	"vigo.ton.segretario@fpmtn.it"		"0461-657897"	
4	"Kindergarten"	"SCUOLA DELLA INFANZIA DI VILLAGGIAMONTAGNA"	"villamontagna.segretario@fpmtn.it"	"villamontagna.segretario@fpmtn.it"		"0461-209045"	
5	"Kindergarten"	"SCUOLA DELLA INFANZIA DI VILLAZZANO"	"villazzano.segretario@fpmtn.it"	"villazzano.segretario@fpmtn.it"		"0461-920382"	

7.2.1.14 Competency Question 15

Give me the list of professors of University of Trento

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/.it/etype#>
select ?professor_fullname where {
    ?professor rdf:type ds:Professor_GID-56316 .
    ?school rdf:type ds:School_GID-44778 .
    ?institute rdf:type ds:Institute_GID-45425 .

    ?professor ds:works_at_GID-3007_Type-56316 ?institute
    .
    ?institute ds:includes_GID-101226_
    Type-45425 ?school
    .

    ?professor
    <ds:has_Name_GID-2_Type-300037>
    ?professor_name .
    ?professor
    <ds:has_Surname_GID-34003>
    ?professor_surname .
    BIND(CONCAT(?professor_name, " ", ?professor_surname) AS
    ?professor_fullname)
    FILTER(?school =
    <http://localhost:8080/source/
    school_id_0000000000000000>)
}
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query. Below it is a table with the results. The table has one column labeled "professor_fullname". The results are as follows:

professor_fullname
1 "Yannis Velegakis"
2 "Atefah Zareh Chahoki"
3 "Francesco Menegale"
4 "Alberto Valli"
5 "Mattia Sensi"
6 "Alessio Marinelli"
7 "Carlo Fezzi"
8 "Francesco De Natale"

7.2.1.15 Competency Question 16

Give me the courses a professor teaches

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?professor_fullname
    ?course_title ?course_articulation
    where {
        ?professor rdf:type
            ds:Professor_GID-56316 .
        ?course rdf:type
            ds:Study_Course_GID-300038 .

        ?professor ds:teaches_GID-104366_Type-56316 ?course
        .
        ?professor
            ds:has_Name_GID-2_Type-300037
            ?professor_name .
        ?professor
            ds:has_Surname_GID-34003
            ?professor_surname .
        ?course ds:has_Title_GID-300029_Type-300038 ?course_title
        .
        ?course ds:has_Articulation_GID-30007_Type-300038
            ?course_articulation .
        BIND(CONCAT(?professor_name, " ", ?professor_surname) AS
            ?professor_fullname)
    }
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, there are tabs for 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. Below the tabs is a search bar labeled 'Filter query results' and a note indicating 'Showing results from 1 to 1,000 of 4,670. Query took 0.1s, moments ago'. The main area displays a table with three columns: 'professor_fullname', 'course_title', and 'course_articulation'. The table contains 7 rows of data, all belonging to the same professor ('Yannis Velegrakis').

	professor_fullname	course_title	course_articulation
1	"Yannis Velegrakis"	"Informatica (L2)"	"Databases"
2	"Yannis Velegrakis"	"Informatica (L2)"	"Databases"
3	"Yannis Velegrakis"	"Ingegneria Informatica, delle Comunicazioni ed Elettronica (L2)"	"Databases"
4	"Yannis Velegrakis"	"Ingegneria dell'Informazione e delle Comunicazioni (LM)"	"Data mining"
5	"Yannis Velegrakis"	"Data Science (LM)"	"Data Mining"
6	"Yannis Velegrakis"	"INFORMATICA (LM)"	"Data mining"
7	"Yannis Velegrakis"	"MATEMATICA (LM)"	"Data mining"

7.2.1.16 Competency Question 17

Give me contact information about a particular professor of the University of Trento

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX sc: <https://schema.org/>
select ?professor_name
    ?professor_surname ?fax ?telephone
    where {
        ?professor rdf:type
            ds:Professor_GID-56316 .
        ?school rdf:type
            ds:School_GID-44778 .
        ?contact rdf:type sc>ContactPoint .
        ?institute rdf:type
            ds:Institute_GID-45425 .

        ?professor ds:works_at_GID-3007_Type-56316 ?institute
            pe-56316 ?institute
            .
        ?institute ds:includes_GID-101226_Type-45425 ?school
            .
        ?professor ds:reachable_by_GID-30002_Type-45425 ?contact
            .
        optional{?professor
            ds:has_Name_GID-2_Type-300037
            ?professor_name .}
        optional{?professor
            ds:has_Surname_GID-34003
            ?professor_surname .}
        optional{?contact
            sc:faxNumber_GID-300008 ?fax .}
        optional{?contact
            sc:telephone_GID-34494
            ?telephone .}
    }
}

```

Result

The screenshot shows a SPARQL query editor with the following details:

- Query Editor:** The top part displays the SPARQL query with numbered lines. Lines 4-18 show the main query structure, while lines 14-18 show optional clauses for professor names and surnames.
- Result Table:** Below the editor is a table showing the results of the query. The columns are labeled: professor_name, professor_surname, fax, and telephone.
- Data:** The table contains 8 rows of data, each corresponding to a professor. The data is as follows:

	professor_name	professor_surname	fax	telephone
1	"Yannis"	"Velegrakis"		"0461 283986"
2	"Alberto"	"Valli"		"0461 281580"
3	"Carlo"	"Fezzi"		"0461 282215"
4	"Francesco"	"De Natale"		"0461 282058"
5	"Francesca"	"Odella"		"0461 281313"
6	"Erica"	"Santini"		"0461 282308"
7	"Andrea"	"Carla"		"0464 808465"
8	"Alessandro"	"Carlotto"		"0461 281612"

7.2.1.17 Competency Question 18

Give me ratings about Institute M. Buonarroti in Trento

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX sc: <https://schema.org/>
select ?institute_name ?school_name
?body ?sentiment where {
?institute rdf:type
    ds:Institute_GID-45425 .
?school rdf:type
    ds:School_GID-44778 .
?review rdf:type
    ds:Review_GID-300044 .

?institute ds:includes_GID-101226_
    Type-45425 ?school
    .
?review ds:describes_GID-104985_Ty
    pe-300044 ?school
    .

?institute
    ds:has_Name_GID-2_Type-300037
    ?institute_name .
?school
    ds:has_Name_GID-2_Type-300037
    ?school_name .
?review sc:reviewBody_GID-34405
    ?body .
?review ds:has_Sentiment_GID-32333_
    _Type-300044
    ?sentiment

FILTER(contains(lcase(?institute_n
    ame), "istituto tecnico
    tecnologico \"m. buonarroti\""
    - trento"))
}
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the SPARQL query. Below it is a table with the following data:

	Institute_name	school_name	body	sentiment
1	'ISTITUTO TECNICO TECNOLOGICO "M. BUONARROTI" - TRENTO'	'ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO'	"Ben rifornito per quanto riguarda i prodotti enogastronomici ma quasi privo di prodotti culinari"	"Positive"
2	'ISTITUTO TECNICO TECNOLOGICO "M. BUONARROTI" - TRENTO'	'ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO'	"Personale che non guarda cosa gli scrive no comment neanche mezza stella"	"Neutral"
3	'ISTITUTO TECNICO TECNOLOGICO "M. BUONARROTI" - TRENTO'	'ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO'	"È l'Istituto tecnico tecnologico di Trento. Dispone di tanti laboratori attrezzati adatti a realizzare numerose esperienze. Gli indirizzi di studio disponibili sono elettronica, chimica, meccanica, informatica ed edilizia. È facilmente raggiungibile dal centro storico e dispone di un ampio parcheggio nel cortile per gli insegnanti."	"Neutral"

7.2.1.18 Competency Question 19

Give me the top-rated kindergartens in Trentino

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name ?municipality_name ?score
where {
  ?municipality rdf:type ds:Subregional_Academic_Division . 
  ?municipality rdf:type ds:Subregional_Academic_Division_GID-300037 .
  ?institute rdf:type ds:Institute_GID-45425 .
  ?school rdf:type ds:School_GID-44778 .
  ?quality rdf:type ds:Education_Quality_GID-300039 .

  ?municipality
  ds:contains_GID-113311_Type-300037
  ?institute .
  ?institute ds:includes_GID-101226_Type-45425
  ?school .
  ?school ds:has_metric_GID-300001_Type-44778
  ?quality .

  ?municipality ds:has_Name_GID-2_Type-300037
  ?municipality_name .
  ?institute ds:has_Type_GID-31834 ?type .
  ?school ds:has_Name_GID-2_Type-300037
  ?school_name .

  ?quality
  ds:has_User_Score_GID-31336_Type-300039
  ?score .
  FILTER(?type = "Kindergarten")
} order by desc(?score) LIMIT 10
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the SPARQL query. Below it is a table with the following data:

	school_name	municipality_name	score
1	"ASILO COLLE FIORITO"	"Rovereto"	'5,0'
2	"SCUOLA DELL'INFANZIA DI ANDALO"	"Andalo"	'5,0'
3	"SCUOLA DELL'INFANZIA DI BORGO VAL SUGANA"	"Borgo Valsugana"	'5,0'
4	"SCUOLA DELL'INFANZIA DI PIEVE TESINO"	"Pieve Tesino"	'5,0'
5	"SCUOLA DELL'INFANZIA DI POVO"	"Trento"	'5,0'
6	"SCUOLA DELL'INFANZIA DI TRENTO 'P EDROTTI"	"Trento"	'5,0'
7	"SCUOLA DELL'INFANZIA DI TRENTO CO GNOLA 'B. KOFLER"	"Trento"	'5,0'
8	"SCUOLA DELL'INFANZIA DI MADRANO"	"Pergine Valsugana"	'5,0'

7.2.1.19 Competency Question 20

Give me the top rated kindergartens near Pergine Valsugana

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name ?municipality_name ?score
where {
  ?municipality rdf:type ds:Subregional_Academic_Division_GID-300037 .
  ?institute rdf:type ds:Institute_GID-45425 .
  ?school rdf:type ds:School_GID-44778 .
  ?quality rdf:type ds:Education_Quality_GID-300039 .

  ?municipality
    ds:contains_GID-113311_Type-300037
    ?institute .
  ?institute ds:includes_GID-101226_Type-45425
    ?school .
  ?school ds:has_metric_GID-300001_Type-44778
    ?quality .

  ?municipality ds:has_Name_GID-2_Type-300037
    ?municipality_name .
  ?institute ds:has_Type_GID-31834 ?type .
  ?school ds:has_Name_GID-2_Type-300037
    ?school_name .
  ?quality
    ds:has_User_Score_GID-31336_Type-300039
    ?score .
  FILTER(?type = "Kindergarten" &&
        ?municipality_name = "Pergine Valsugana")
} order by desc(?score) LIMIT 10

```

Result

	school_name	municipality_name	score
1	'SCUOLA DELL'INFANZIA DI MADRANO'	'Pergine Valsugana'	'5,0'
2	'SCUOLA DELL'INFANZIA DI ISCHIA DI PERGINE'	'Pergine Valsugana'	'5,0'

7.2.1.20 Competency Question 21

Give me the top-rated schools in Trentino near my home

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name ?municipality_name ?score
where {
  ?municipality rdf:type ds:Subregional_Academic_Division_GID-300037 .
  ?institute rdf:type ds:Institute_GID-45425 .
  ?school rdf:type ds:School_GID-44778 .
  ?quality rdf:type ds:Education_Quality_GID-300039 .

  ?municipality
    ds:contains_GID-113311_Type-300037
    ?institute .
  ?institute ds:includes_GID-101226_Type-45425
    ?school .
  ?school ds:has_metric_GID-300001_Type-44778
    ?quality .

  ?municipality ds:has_Name_GID-2_Type-300037
    ?municipality_name .
  ?institute ds:has_Type_GID-31834 ?type .
  ?school ds:has_Name_GID-2_Type-300037
    ?school_name .
  ?quality
    ds:has_User_Score_GID-31336_Type-300039
    ?score .
  FILTER(?municipality_name = "Vallelaghi")
} order by desc(?score)

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor containing the SPARQL query. Below the code editor is a toolbar with various icons for file operations like save, copy, and run. Below the toolbar is a table with the following data:

	school_name	municipality_name	score
1	"SCUOLA PRIMARIA PIETRAMURATA"	"Vallelaghi"	"4,7"
2	"SCUOLA PRIMARIA TERLAGO"	"Vallelaghi"	"4,5"
3	"SCUOLA PRIMARIA VEZZANO"	"Vallelaghi"	"4,0"
4	"SCUOLA SECONDARIA DI PRIMO GRADO 'NUOVA EUROPA' DRO"	"Vallelaghi"	"2,3"

7.2.1.21 Competency Question 22

Give me % of students admitted to the next class for a particular school

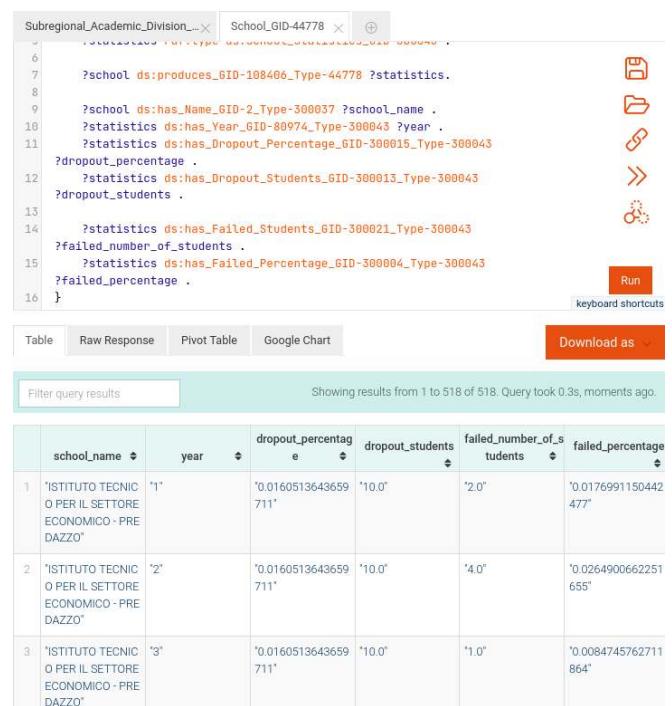
SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name ?year
    ?dropout_percentage
    ?dropout_students
    ?failed_number_of_students
    ?failed_percentage where {
        ?school rdf:type
            <ds:School_GID-44778> .
        ?statistics rdf:type ds:School_Statistics_GID-300043
        .
        ?school ds:produces_GID-108406_Type-44778
        <ds:has_Percentage> ?statistics.
```



```
?school
    <ds:has_Name_GID-2_Type-300037>
    <?school_name> .
?statistics <ds:has_Year_GID-80974_Type-300043>
<?year> .
?statistics <ds:has_Dropout_Percentage_GID-300015_Type-300043>
<?dropout_percentage> .
?statistics <ds:has_Dropout_Students_GID-300013_Type-300043>
<?dropout_students> .
?statistics <ds:has_Failed_Students_GID-300021_Type-300043>
<?failed_number_of_students> .
?statistics <ds:has_Failed_Percentage_GID-300004_Type-300043>
<?failed_percentage> .
}
```

Result



The screenshot shows a SPARQL query results interface. At the top, there are two tabs: "Subregional_Academic_Division..." and "School_GID-44778". Below the tabs is a code editor with the SPARQL query. To the right of the code editor are several icons for file operations (Save, Open, Copy, Print, etc.) and a "Run" button. Below the code editor are four buttons: "Table", "Raw Response", "Pivot Table", and "Google Chart". A "Filter query results" input field is at the bottom left, and a message "Showing results from 1 to 518 of 518. Query took 0.3s, moments ago." is at the bottom right. The main area displays a table with six columns: school_name, year, dropout_percentage, dropout_students, failed_number_of_students, and failed_percentage. There are three rows of data, each corresponding to a different school with its respective statistics.

	school_name	year	dropout_percentage	dropout_students	failed_number_of_students	failed_percentage
1	"ISTITUTO TECNICO PER IL SETTORE ECONOMICO - PREDAZZO"	"1"	"0.0160513643659711"	"10.0"	"2.0"	"0.0176991150442477"
2	"ISTITUTO TECNICO PER IL SETTORE ECONOMICO - PREDAZZO"	"2"	"0.0160513643659711"	"10.0"	"4.0"	"0.0264900662251655"
3	"ISTITUTO TECNICO PER IL SETTORE ECONOMICO - PREDAZZO"	"3"	"0.0160513643659711"	"10.0"	"1.0"	"0.0084745762711864"

7.2.1.22 Competency Question 23

Give me % of students that failed the year for an institute

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?institute_name ?year
    ?failed_number_of_students
    ?failed_percentage where {
        ?institute rdf:type
            ds:Institute_GID-45425 .
        ?school rdf:type
            ds:School_GID-44778 .
        ?statistics rdf:type ds:School_Statistics_GID-300043
        .
        ?institute ds:includes_GID-101226_
            Type-45425 ?school
        .
        ?school ds:produces_GID-108406_Typ
            e-44778
        ?statistics.

        ?institute
        ds:has_Name_GID-2_Type-300037
        ?institute_name .

        ?school
        ds:has_Name_GID-2_Type-300037
        ?school_name .

        ?statistics ds:has_Year_GID-80974_
            Type-300043 ?year
        .
        ?statistics ds:has_Failed_Students_
            _GID-300021_Type-300043
        ?failed_number_of_students .

        ?statistics ds:has_Failed_Percentage_
            _GID-300004_Type-300043
        ?failed_percentage .

    } order by ?year
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a query editor window with the following SPARQL code:

```
Subregional_Academic_Division_--> School_GID-44778 × ⊕
SELECT ?institute_name ?year ?failed_number_of_students ?failed_percentage
WHERE {
  ?institute rdf:type ds:Institute_GID-45425 .
  ?school rdf:type ds:School_GID-44778 .
  ?statistics rdf:type ds:School_Statistics_GID-300043 .
  ?institute ds:includes_GID-101226_Type-45425 ?school .
  ?school ds:produces_GID-108406_Type-44778 ?statistics .
  ?institute ds:has_Name_GID-2_Type-300037 ?institute_name .
  ?school ds:has_Name_GID-2_Type-300037 ?school_name .
  ?statistics ds:has_Year_GID-80974_Type-300043 ?year .
  ?statistics ds:has_Failed_Students_GID-300021_Type-300043
  ?failed_number_of_students .
  ?statistics ds:has_Failed_Percentage_GID-300004_Type-300043
  ?failed_percentage .
}
```

Below the query editor is a table with the following data:

	institute_name	year	failed_number_of_students	failed_percentage
1	'ISTITUTO DI ISTRUZIONE "LA ROSA BIANCA" - CAVALESE'	"1"	"2.0"	"0.0176991150442477"
2	'ISTITUTO DI ISTRUZIONE "LA ROSA BIANCA" - CAVALESE'	"1"	"2.0"	"0.0176991150442477"
3	'ISTITUTO DI ISTRUZIONE "LA ROSA BIANCA" - CAVALESE'	"1"	"2.0"	"0.0176991150442477"
4	'ISTITUTO DI ISTRUZIONE "LA ROSA BIANCA" - CAVALESE'	"1"	"2.0"	"0.0176991150442477"

7.2.1.23 Competency Question 24

Give me the invalsi score of a particular school

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name ?subject
    ?average_score_percentage
    ?average_wle ?coverage_percentage
    where {
        ?school rdf:type
            → ds:School_GID-44778 .
        ?quality rdf:type ds:Education_Quality_GID-300039
            → .
        ?invalsi rdf:type
            → ds:Invalsi_Score_GID-300042 .

        ?school ds:has_metric_GID-300001_Type-44778 ?quality
            → .
        ?quality ds:influenced_by_GID-787_Type-300039 ?invalsi
            → .

        ?school
            → ds:has_Name_GID-2_Type-300037
            → ?school_name .
        ?invalsi ds:has_Subject_GID-300030_Type-300042 ?subject
            → .
        ?invalsi ds:has_Average_Score_Percentage_GID-300026_Type-300042
            → ?average_score_percentage .
        ?invalsi ds:has_Average_WLE_GID-300030_Type-300042 ?average_wle
            → .
        ?invalsi ds:has_Coverage_Percentage_GID-300009_Type-300042
            → ?coverage_percentage .
    }
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, there are tabs for 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. A 'Run' button is located on the right. Below the tabs is a 'Filter query results' input field and a message stating 'Showing results from 1 to 185 of 185. Query took 0.1s, moments ago.' The main area displays a table with the following columns: school_name, subject, average_score_percent, average_wle, and coverage_percentag. The data is as follows:

	school_name	subject	average_score_percent	average_wle	coverage_percentag
1	'SCUOLA PRIMARIA 'A. BETTÀ' ALA'	"Italian"	"60,40"	"198,67"	"95,37"
2	'SCUOLA PRIMARIA G ARDASCUOLA ARCO'	"Italian"	"58,19"	"191,78"	"92,81"
3	'SCUOLA PRIMARIA " G. SEGANTINI' ARCO'	"Italian"	"58,19"	"191,78"	"92,81"
4	'SCUOLA PRIMARIA B OLOGNANO'	"Italian"	"58,19"	"191,78"	"92,81"
5	'SCUOLA PRIMARIA R OMARZOLLO'	"Italian"	"58,19"	"191,78"	"92,81"

7.2.1.24 Competency Question 25

Give me % of student that abandoned a particular institute

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?institute_name ?year
    ?dropout_number_of_students
    ?dropout_percentage where {
        ?institute rdf:type
            ds:Institute_GID-45425 .
        ?school rdf:type
            ds:School_GID-44778 .
        ?statistics rdf:type ds:School_Statistics_GID-300043
            .
        ?institute ds:includes_GID-101226_
            Type-45425 ?school
            .
        ?school ds:produces_GID-108406_Typ
            e-44778
            .
        ?statistics.

        ?institute
            ds:has_Name_GID-2_Type-300037
            .
        ?institute_name .

        ?school
            ds:has_Name_GID-2_Type-300037
            .
        ?school_name .

        ?statistics ds:has_Year_GID-80974_
            Type-300043 ?year
            .
        ?statistics ds:has_Dropout_Student_
            s_GID-300013_Type-300043
            .
        ?dropout_number_of_students .

        ?statistics ds:has_Dropout_Percent_
            age_GID-300015_Type-300043
            .
        ?dropout_percentage .

    } order by ?year
    desc(?dropout_percentage)

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query. Below it is a table with three rows of data. The table has four columns: Institute_name, year, dropout_number_of_students, and dropout_percentage. The data is as follows:

Institute_name	year	dropout_number_of_students	dropout_percentage
'CENTRO FORMAZIONE PROFESSIONALE ENAIP - VARONE DI RIVA DEL GARDA'	'1"	'22.0"	'0.0589812332439678"
'ISTITUTO FORMAZIONE PROFESSIONALE SERVIZI ALLA PERSONA E LEGNO 'S. PERTINI' - TRENTO'	'1"	'20.0"	'0.0511508951406649"
'ISTITUTO FORMAZIONE PROFESSIONALE SERVIZI ALLA PERSONA E LEGNO 'S. PERTINI' - TRENTO'	'1"	'20.0"	'0.0511508951406649"

Below the table, a message says "Showing results from 1 to 1,000 of 1,985. Query took 0.1s, minutes ago."

7.2.1.25 Competency Question 26

Give me schools that offer evening programs in Trentino

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name ?title
?articulation where {
  ?school rdf:type
    ↳ ds:School_GID-44778 .
  ?course rdf:type
    ↳ ds:Study_Course_GID-300038 .

  ?school ds:offers_GID-111716_Type-
    ↳ 44778 ?course
    ↳ .

  ?school ds:has_Timetable_type_GID-
    ↳ 34826_Type-44778 ?time_table
    ↳ .

  ?school
    ↳ ds:has_Name_GID-2_Type-300037
    ↳ ?school_name .

  ?course ds:has_Title_GID-300029_Type-
    ↳ -300038 ?title
    ↳ .

  ?course ds:has_Articulation_GID-300007_Type-300038 ?articulation
    ↳ .

  FILTER(?time_table = "Evening")
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query. Below it is a table with four columns: school_name, title, and articulation. The table contains four rows of data. At the bottom of the interface, there is a navigation bar with buttons for Table, Raw Response, Pivot Table, Google Chart, Filter query results, and Download as.

school_name	title	articulation
1 'ISTITUTO TECNICO PER IL SETTORE ECONOMICO SERALE - PREDAZZO'	'DIPLOMA ESAME DI STATO CONCLUSIVO DEL SECONDO CICLO'	"Diploma di istituto tecnico settore economico indirizzo "amministrazione, finanza e marketing"
2 'ISTITUTO TECNICO PER IL SETTORE ECONOLOGICO SERALE - PREDAZZO'	'DIPLOMA ESAME DI STATO CONCLUSIVO DEL SECONDO CICLO'	"Diploma di istituto tecnico settore tecnologico indirizzo "costruzioni, ambiente e territorio"
3 'ISTITUTO TECNICO PER IL SETTORE ECONOLOGICO SERALE - PREDAZZO'	'DIPLOMA ESAME DI STATO CONCLUSIVO DEL SECONDO CICLO'	"Diploma di istituto tecnico settore tecnologico indirizzo "informatica e telecomunicazioni" articulazione "informatica"
4 'SETTORE INDUSTRIA E ARTIGIANATO SERALE'	'QUALIFICA PROFESSIONALE'	"Operatore edile ad indirizzo completamento e finitura"

7.2.1.26 Competency Question 27

Give me courses provided by evening schools

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name where {
    ?school rdf:type
        → ds:School_GID-44778 .
    ?course rdf:type
        → ds:Study_Course_GID-300038 .

    ?school ds:offers_GID-111716_Type-
        → 44778 ?course
        → .

    ?school ds:has_Timetable_type_GID-
        → 34826_Type-44778 ?time_table
        → .

    ?school
        → ds:has_Name_GID-2_Type-300037
        → ?school_name .
    ?course ds:has_Title_GID-300029_Ty
        → pe-300038 ?title
        → .

    ?course ds:has_Articulation_GID-30
        → 0007_Type-300038 ?articulation
        → .

FILTER(?time_table = "Evening")
}
```

Result

The screenshot shows a SPARQL query results interface. At the top, there are PREFIX declarations for rdf and ds, and a WHERE clause selecting ?school_name where ?school offers a course with a specific timetable type (44778) and has a name. The results are filtered to show only entries where the timetable type is "Evening". The results table has one column labeled "school_name" with 7 rows of data.

school_name
'ISTITUTO TECNICO PER IL SETTORE ECONOMICO SERALE - PREDAZZO'
'ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE - PREDAZZO'
'ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE - PREDAZZO'
'SETTORE INDUSTRIA E ARTIGIANATO SERALE'
'SETTORE INDUSTRIA E ARTIGIANATO SERALE'
'SETTORE INDUSTRIA E ARTIGIANATO SERALE'
'SETTORE SERVIZI SERALE'

7.2.1.27 Competency Question 28

Give me contact information about a particular evening school

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX sc: <https://schema.org/>
select ?school_name ?office_mail ?institute_mail
    ?management_mail ?telephone ?fax where {
        ?institute rdf:type ds:Institute_GID-45425
        .
        ?school rdf:type ds:School_GID-44778 .
        ?contact rdf:type ds:Institute_Contact_Information_GID-300041
        .
        ?institute ds:includes_GID-101226_Type-45425
        .
        ?school .
        ?institute
        .
        ?school ds:reachable_by_GID-300002_Type-45425
        .
        ?contact .

        ?school
        .
        ?ds:has_Timetable_type_GID-34826_Type-44778
        .
        ?time_table .
        ?school ds:has_Name_GID-2_Type-300037
        .
        ?school_name .

        ?contact
        .
        ?ds:has_Office_Mail_GID-300017_Type-300041
        .
        ?office_mail .
        ?contact ds:has_Institute_Mail_GID-300035_Type-300041 ?institute_mail
        .
        ?contact ds:has_Management_Mail_GID-300016_Type-300041 ?management_mail
        .
        ?contact sc:telephone_GID-34494 ?telephone .
        ?contact sc:faxNumber_GID-300008 ?fax .

        FILTER(?time_table = "Evening")
    }

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query. Below it is a table with the results. The table has columns: school_name, office_mail, institute_mail, management_mail, telephone, and fax. There are three rows of data.

	school_name	office_mail	institute_mail	management_mail	telephone	fax
1	'ISTITUTO TECNICO PER IL SETTORE ECONOMICO SERALE - PREDAZZO'	'segr.iicavalese@scuole.provincia.tn.it'	'weisserose@pec.scuole.provincia.tn.it'	'dir.iicavalese@scuole.provincia.tn.it'	'+39 0462 341449"	+'39 0462 248071"
2	'ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE - PREDAZZO'	'segr.iicavalese@scuole.provincia.tn.it'	'weisserose@pec.scuole.provincia.tn.it'	'dir.iicavalese@scuole.provincia.tn.it'	'+39 0462 341449"	+'39 0462 248071"
3	'ISTITUTO TECNICO PER IL SETTORE ECONOMICO SERALE'	'segr.martini.mezzolombardo@scuole.provincia.tn.it"	'martini@pec.provincia.tn.it'	'dir.martini.mezzolombardo@scuole.provincia.tn.it'	'+39 0461 601122"	+'39 0461 601470"

7.2.1.28 Competency Question 29 and 36

Give me % of students admitted to the next class for a particular night school

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?institute_name ?year
?admitted_percentage where {
    ?institute rdf:type
        ds:Institute_GID-45425 .
    ?school rdf:type
        ds:School_GID-44778 .
    ?statistics rdf:type ds:School_Statistics_GID-300043
    .
    ?institute ds:includes_GID-101226_Type-45425 ?school
    .
    ?school ds:produces_GID-108406_Type-44778
    ?statistics .
    ?institute
        ds:has_Name_GID-2_Type-300037
        ?institute_name .
    ?school
        ds:has_Name_GID-2_Type-300037
        ?school_name .
    ?school ds:has_Timetable_type_GID-34826_Type-44778 ?time_table
    .
    ?statistics ds:has_Year_GID-80974_Type-300043 ?year
    .
    ?statistics ds:has_Admitted_Percentage_GID-300036_Type-300043
    ?admitted_percentage .
}
FILTER(?time_table = "Evening")
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the SPARQL query. Below it is a toolbar with various icons for file operations. The main area displays a table with three columns: Institute_name, year, and admitted_percentage. The table contains five rows of data, each representing a different institute with its name, year, and admitted percentage.

	Institute_name	year	admitted_percentage
1	'ISTITUTO DI ISTRUZIONE "LA ROSA BI ANCA" - CAVALESE'	'1'	'0.0'
2	'ISTITUTO DI ISTRUZIONE "LA ROSA BI ANCA" - CAVALESE'	'2'	'0.0'
3	'ISTITUTO DI ISTRUZIONE "LA ROSA BI ANCA" - CAVALESE'	'3'	'18.181818181818183'
4	'ISTITUTO DI ISTRUZIONE "LA ROSA BI ANCA" - CAVALESE'	'4'	'40.909090909090914'
5	'ISTITUTO DI ISTRUZIONE "LA ROSA BI ANCA" - CAVALESE'	'5'	'70.0'

7.2.1.29 Competency Question 30

Give me courses provided by evening schools

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name ?title
?articulation where {
    ?school rdf:type
        → ds:School_GID-44778 .
    ?course rdf:type
        → ds:Study_Course_GID-300038 .

    ?school ds:offers_GID-111716_Type-
        → 44778 ?course
        → .

    ?school ds:has_Timetable_type_GID-
        → 34826_Type-44778 ?time_table
        → .

?school
    → ds:has_Name_GID-2_Type-300037
    → ?school_name .
?course ds:has_Title_GID-300029_Ty
    → pe-300038 ?title
    → .

?course ds:has_Articulation_GID-30
    → 0007_Type-300038 ?articulation
    → .

FILTER(?time_table = "Evening")
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, there are PREFIX declarations for rdf and ds, followed by a SELECT statement. The query retrieves school names, titles, and articulations for courses offered by schools that provide evening classes. The results are displayed in a table with columns: school_name, title, and articulation. There are 56 rows in the table, showing three entries. Each entry includes the name of the school, the course title, and a brief description of the articulation. The interface includes various buttons for saving, printing, and navigating through the results.

school_name	title	articulation
1 "ISTITUTO TECNICO PER IL SETTORE ECONOMICO SERALE - PREDAZZO"	"DIPLOMA ESAME DI STATO CONCLUSIVO DEL SECONDO CICLO"	"Diploma di istituto tecnico settore economico indirizzo "amministrazione, finanza e marketing"
2 "ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE - PREDAZZO"	"DIPLOMA ESAME DI STATO CONCLUSIVO DEL SECONDO CICLO"	"Diploma di istituto tecnico settore tecnologico indirizzo "costruzioni, ambiente e territorio"
3 "ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE - PREDAZZO"	"DIPLOMA ESAME DI STATO CONCLUSIVO DEL SECONDO CICLO"	"Diploma di istituto tecnico settore tecnologico indirizzo "informatica e telecomunicazioni" articolazione "informatica"

7.2.1.30 Competency Question 37

Give me the number of enrolled students of a particular school

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?municipality_name ?n_k ?n_ps
?n_ms ?n_hs ?n_cf where {
  ?municipality rdf:type
    ds:Subregional_Academic_Division
  on_GID-300037
  .
  ?municipality
    ds:has_School_Completion_Rate_
  GID-300011_Type-300037
    ?school_completion_rate .
  ?municipality
    ds:has_Name_GID-2_Type-300037
    ?municipality_name .
  ?municipality ds:has_Number_of_Stu_
    dents_Enrolled_in_Kindergarten_
  GID-300025_Type-300037 ?n_k
  .
  ?municipality ds:has_Number_of_Stu_
    dents_Enrolled_in_Primary_Scho_
  ol_GID-300022_Type-300037
    ?n_ps .
  ?municipality ds:has_Number_of_Stu_
    dents_Enrolled_in_Middle_Schoo_
  l_GID-300032_Type-300037 ?n_ms
  .
  ?municipality ds:has_Number_of_Stu_
    dents_Enrolled_in_Highschool_G_
  ID-300018_Type-300037 ?n_hs
  .
  ?municipality ds:has_Number_of_Stu_
    dents_Enrolled_in_CF_GID-30002_
  7_Type-300037 ?n_cf
  .
}
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query and a run button. Below it is a table with 7 rows of data. The columns are labeled: municipality_name, n_k, n_ps, n_ms, n_hs, and n_cf. The data is as follows:

	municipality_name	n_k	n_ps	n_ms	n_hs	n_cf
1	"Trento"	'2915'	'5529'	'3882'	'8631'	'1923'
2	"Roncogno Terme"	'66'	'142'	'146'	'0'	'0'
3	"Cavalese"	'70'	'180'	'307'	'366'	'0'
4	"Brentonico"	'99'	'177'	'125'	'0'	'0'
5	"Rovereto"	'1041'	'2124'	'1812'	'4582'	'1413'
6	"Arco"	'408'	'972'	'621'	'173'	'352'
7	"Pergine Valsugana"	'559'	'1058'	'801'	'671'	'0'

As mentioned above, the previous query does not address completely the query, this is due to the fact that we did not find information related to the number of students enrolled in a specific school. Hence, this question answers how many students there are within schools of a specific municipality.

7.2.1.31 Competency Question 38

Give me the graduation rate of a particular municipality in Trentino

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?municipality_name ?graduation_rate where {
    ?municipality rdf:type ds:Subregional_Academic_Division_GID-300037 .
    ?municipality ds:has_Graduate_Rate_GID-300010_Type-300037 ?graduation_rate .
    ?municipality ds:has_Name_GID-2_Type-300037 ?municipality_name .
}
```

Result

The screenshot shows a SPARQL query results interface. At the top, there are several tabs: 'unnamed', 'unnamed', 'unnamed', 'unnamed', 'unnamed', 'http://enlris.com/foaf/0.1...', 'Subregional_Academic_Div...', and 'School_GID-44778'. To the right of these tabs are icons for copy, download, and refresh. Below the tabs is a red 'Run' button and a note: 'Press Alt+Enter to keyboard shortcuts'. Underneath the tabs, there are buttons for 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. A 'Download as' button is located on the far right. A 'Filter query results' input field is present. The main area displays a table with two columns: 'municipality_name' and 'graduation_rate'. The table contains 166 rows, with the first few rows listed below:

	municipality_name	graduation_rate
1	'Trento'	'40.1'
2	'Roncoiglio Terme'	'41'
3	'Cavalese'	'38.8'
4	'Brentonico'	'41.7'
5	'Rovereto'	'41.7'
6	'Aldo'	'42'
7	'Perge Valtagana'	'41.7'

7.2.1.32 Competency Question 39

Give me the number of people who didn't finish school in a certain municipality in Trentino

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name ((100 - xsd:float(?completion_rate)) as
→ ?not_have_a_degree) where {
?municipality rdf:type ds:Subregional_Academic_Division_GID-300037 .

?municipality ds:has_School_Completion_Rate_GID-300011_Type-300037
→ ?completion_rate .
?municipality ds:has_Name_GID-2_Type-300037 ?municipality_name .
}
```

Result

	municipality_name	not_have_a_degree
1	"Trento"	71.2**not fin
2	"Rhoenigo Terme"	71.2**not fin
3	"Cavalese"	62.7**not fin
4	"Trentonino"	71.1**not fin
5	"Tiroveneto"	71.5**not fin
6	"Arosa"	76.3**not fin

Since we have not been able to retrieve the exact number of people with a degree, we have decided to use the completion rate, which is a piece of information we were able to retrieve. Therefore, the query is not completely solved.

7.2.1.33 Competency Question 40

Give me a list of the number of schools per municipality

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name
  (sum(xsd:integer(?schools_number)))
  as ?total_number) where {
  ?municipality rdf:type
    ds:Subregional_Academic_Division_GID-300037
  .
  ?institute rdf:type
    ds:Institute_GID-45425 .

  ?municipality ds:contains_GID-11331_11_Type-300037 ?institute
  .
  ?institute ds:includes_GID-101226_Type-45425 ?school
  .
  ?institute ds:has_Number_of_School_s_GID-300019_Type-45425
  ?schools_number .
  ?municipality
    ds:has_Name_GID-2_Type-300037
    ?municipality_name .
} group by (?municipality_name)

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query, and below it, a table displaying the results. The table has two columns: 'municipality_name' and 'total_number'. The results are as follows:

	municipality_name	total_number
1	"Trento"	*603**xsd:integer
2	"Roncegno Terme"	*83**xsd:integer
3	"Cavalese"	*164**xsd:integer
4	"Brentonico"	*6**xsd:integer
5	"Rovereto"	*309**xsd:integer
6	"Arco"	*55**xsd:integer
7	"Pergine Valsugana"	*335**xsd:integer

7.2.1.34 Bonus Query 1

Schools with the highest number of positive reviews

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name (count(?sentiment) as ?positive_count) where {
    ?school rdf:type
        → ds:School_GID-44778 .
    ?review rdf:type
        → ds:Review_GID-300044 .

    ?review ds:describes_GID-104985_Ty
        → pe-300044 ?school
        → .

    ?school
        → ds:has_Name_GID-2_Type-300037
        → ?school_name .
    ?review ds:has_Sentiment_GID-32333
        → _Type-300044 ?sentiment
        → .
    FILTER(?sentiment = "Positive")
} group by (?school_name)
order by desc(?positive_count)
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query and some metadata like 'Subregional_Academic_Division...' and 'School_GID-44778'. Below the code editor are several icons for file operations. A red 'Run' button is visible. Below the run button are tabs for 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. A 'Download as' button is also present. The main area displays a table with two columns: 'school_name' and 'positive_count'. The table contains five rows of data. A search bar 'Filter query results' and a note 'Showing results from 1 to 126 of 126. Query took 0.1s, moments ago.' are at the bottom.

	school_name	positive_count
1	'ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO'	*57***xsd:integer
2	'ISTITUTO TECNICO PER IL SETTORE ECONOMICO'	*48***xsd:integer
3	'LICEO SCIENTIFICO OPZIONE SCIENZE APPLICATE'	*41***xsd:integer
4	'SETTORE INDUSTRIA E ARTIGIANATO'	*32***xsd:integer
5	'ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE'	*31***xsd:integer

7.2.1.35 Bonus Query 2

Schools with the highest number of negative reviews

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name (count(?sentiment)
as ?negative_count) where {
?school rdf:type
  → ds:School_GID-44778 .
?review rdf:type
  → ds:Review_GID-300044 .

?review ds:describes_GID-104985_Ty
  → pe-300044 ?school
  → .

?school
  → ds:has_Name_GID-2_Type-300037
  → ?school_name .
?review ds:has_Sentiment_GID-32333
  → _Type-300044 ?sentiment
  → .
FILTER(?sentiment = "Negative")
} group by (?school_name)
order by desc(?negative_count)

```

Result

The screenshot shows a SPARQL query results interface. At the top, there are two tabs: 'Subregional_Academic_Division...' and 'School_GID-44778'. Below the tabs is a code editor with the SPARQL query. To the right of the code editor are several icons for file operations: copy, paste, save, and print. A red 'Run' button is located at the bottom right of the code editor. Below the code editor are four buttons: 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. A red 'Download as' button is also at the bottom right. A search bar labeled 'Filter query results' is at the bottom left, and a message 'Showing results from 1 to 39 of 39. Query took 0.1s, moments ago.' is at the bottom right. The main area displays a table with six rows, each containing a school name and its negative count.

	school_name	negative_count
1	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO"	*12***xsd:integer
2	"LICEO SCIENTIFICO OPZIONE SCIENZE APPLICATE"	*8***xsd:integer
3	"ISTITUTO TECNICO PER IL SETTORE TECNOLOGICO SERALE"	*8***xsd:integer
4	"ISTITUTO TECNICO PER IL SETTORE ECONOMICO"	*7***xsd:integer
5	"LICEO SCIENTIFICO"	*6***xsd:integer
6	"LICEO DELLE SCIENZE UMANE"	*6***xsd:integer

7.2.1.36 Bonus Query 3

Get some schools quality metrics

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etyp#>
select distinct ?institute_name ?school_name ?user_score ?positive_reviews
<- ?average_score_percentage ?average_wle ?coverage_percentage where {
  ?school rdf:type ds:School_GID-44778 .
  ?review rdf:type ds:Review_GID-300044 .
  ?quality rdf:type ds:Education_Quality_GID-300039 .
  ?invalsi rdf:type ds:Invalsi_Score_GID-300042 .
  ?institute rdf:type ds:Institute_GID-45425 .

  ?school ds:has_metric_GID-300001_Type-44778 ?quality .
  ?quality ds:influenced_by_GID-87_Type-300039 ?invalsi .
  ?review ds:describes_GID-104985_Type-300044 ?school .
  ?institute ds:includes_GID-101226_Type-45425 ?school .

?institute ds:has_Name_GID-2_Type-300037 ?institute_name .
?school ds:has_Name_GID-2_Type-300037 ?school_name .
  ?invalsi ds:has_Subject_GID-300030_Type-300042 ?subject .
  ?invalsi ds:has_Average_Score_Percentage_GID-300026 ?average_score_percentage .
  ?invalsi ds:has_Average_WLE_GID-300003_Type-300042 ?average_wle .
  ?invalsi ds:has_Coverage_Percentage_STU_GID-300009_Type-300042 ?coverage_percentage .
  ?quality ds:has_User_Score_GID-31336_Type-300039 ?user_score .
{
  select ?local_school_name (count(?sentiment) as ?positive_reviews) where {
    ?local_school rdf:type ds:School_GID-44778 .
    ?local_review rdf:type ds:Review_GID-300044 .
    ?local_review ds:describes_GID-104985_Type-300044 ?local_school .
    ?local_school ds:has_Name_GID-2_Type-300037 ?local_school_name .
    ?local_review ds:has_Sentiment_GID-32333_Type-300044 ?sentiment .
    FILTER(?sentiment = "Positive")
  } group by (?local_school_name)
}
FILTER(?school_name = ?local_school_name)
} order by desc(?user_score) desc(?positive_reviews) desc(?average_score_percentage)
<- desc(?average_wle) desc(?coverage_percentage)

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the SPARQL query. Below the code editor are several buttons: 'Table', 'Raw Response', 'Pivot Table', 'Google Chart', 'Run' (highlighted in red), and 'Download as'. A status bar at the bottom indicates 'Showing results from 1 to 145 of 145. Query took 0.4s, moments ago.' The main area displays a table with the following data:

Institute_name	school_name	user_score	positive_reviews	average_score_percentage	coverage_percentage
1 'ISTITUTO COMPRENSIVO ROVERETO EST'	'SCUOLA PRIMA 'REGINA ELENA' ROVERETO'	'5,0'	'3***sd integer'	'63,83'	'204,52'
2 'SCUOLA PARITARIA LA VELA - FONDAZIONE FAMILIA MATERNA ROVERETO'	'SCUOLA SECONDA 'NDARIA DI PRIMO GRADO 'LA VELA' ROVERETO'	'5,0'	'1***sd integer'	'99,00'	'207,46'
3 'COOPERATIVA SOCIALE SARA A FAMIGLIA ONLUS'	'SCUOLA PRIMA SOCIALE SARA A FAMIGLIA ONLUS'	'5,0'	'1***sd integer'	'62,71'	'201,41'

As mentioned multiple in the sections, the definition of education quality we have provided includes several metrics, which we list below:

- Google Reviews users' score;
- number of positive reviews;
- Score in the invalsi test

7.2.1.37 Bonus Query 4

Get responsible information per institute

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?institute_name
    ?responsible_name
    ?responsible_surname
    ?responsible_title where {
        ?institute rdf:type
            ds:Institute_GID-45425 .
        ?responsible rdf:type ds:Responsible_Authority_GID-300040
        .
        ?responsible ds:manages_GID-112404
            _Type-300040 ?institute
        .
        ?institute
            ds:has_Name_GID-2_Type-300037
            ?institute_name .
        ?responsible
            ds:has_Name_GID-2_Type-300037
            ?responsible_name .
        ?responsible
            ds:has_Surname_GID-34003
            ?responsible_surname .
        ?responsible ds:has_Title_GID-3000
            29_Type-300038
            ?responsible_title
    }
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query and some annotations. Below it is a table with the results. The table has four columns: institute_name, responsible_name, responsible_surname, and responsible_title. The data is as follows:

	institute_name	responsible_name	responsible_surname	responsible_title
1	"Università degli studi di Trento"	"Antonella"	"Degl'Innocenti"	"Deputy dean"
2	"Università degli studi di Trento"	"Andrea"	"Rosani"	"Head"
3	"Università degli studi di Trento"	"Caterina"	"Mordegila"	"Deputy coordinator"
4	"Università degli studi di Trento"	"Roberto"	"Passerone"	"Deputy coordinator"
5	"Università degli studi di Trento"	"Barbara"	"Treccani"	"Deputy coordinator"
6	"Università degli studi di Trento"	"Cinzia"	"Picciocchi"	"Rector's delegate for the Animal Welfare Committee"
7	"Università degli studi di Trento"	"Matteo"	"Cosulich"	"Deputy dean"

7.2.1.38 Bonus Query 5

Get the schools' information

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
select ?school_name ?school_type
    → ?school_number_of_years ?school_stage
    → ?school_miur_code ?school_timetable
    → ?institute_name ?institute_type
    → ?institute_website ?institute_number_of_school
    → ?institute_miur_code ?institute_pat where {
        ?institute rdf:type ds:Institute_GID-45425 .
        ?school rdf:type ds:School_GID-44778 .

        ?institute
            → ds:includes_GID-101226_Type-45425
            → ?school .

?institute ds:has_Name_GID-2_Type-300037
    → ?institute_name .
?institute ds:has_Type_GID-31834
    → ?institute_type .
?institute ds:has_Website_GID-34126_Type-45425
    → ?institute_website .
?institute ds:has_Number_of_Schools_GID-300019_
    → _Type-45425 ?institute_number_of_school
    → .
?institute
    → ds:has_MIUR_code_GID-300020_Type-45425
    → ?institute_miur_code .
?institute
    → ds:has_PAT_code_GID-300005_Type-45425
    → ?institute_pat .

?school ds:has_Name_GID-2_Type-300037
    → ?school_name .
?school ds:has_Type_GID-31834 ?school_type .
?school ds:has_Number_of_Years_GID-300024_Type_]
    → -44778 ?school_number_of_years
    → .
?school ds:has_Stage_GID-300023_Type-44778
    → ?school_stage .
?school ds:has_MIUR_code_GID-300020_Type-45425
    → ?school_miur_code .
?school
    → ds:has_Timetable_type_GID-34826_Type-44778
    → ?school_timetable
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the SPARQL query. Below it is a toolbar with various icons for file operations. The main area displays the results in a 'Pivot Table' format. The table has several columns representing different attributes: school_name, school_type, school_number_of_years, school_stage, school_miur_code, school_timetable, institute_name, institute_type, institute_website, institute_number_of_school, institute_miur_code, institute_pat, institute_name, institute_type, institute_website, institute_number_of_school, institute_miur_code, and institute_pat. There are two rows of data in the table.

school_name	school_type	school_number_of_years	school_stage	school_miur_code	school_timetable	institute_name	institute_type	institute_website	institute_number_of_school	institute_miur_code	institute_pat
'SCUOLA PRIMARIA "C. CORRIADI' NOV ALEDO'	'Provincia'	'5'	'Primary School'	'TNEE813 064'	'Morning'	'ISTITUTO COMPRENSIVO CENTRO VALSUGANA'	'Educational and training institution'	'http://www.wicentrovalsugana.it'	'9'	'TNIC813 00R'	'2218829 02'
'SCUOLA PRIMARIA "P. MARTINI' RO NCEGNO TERME'	'Provincia'	'5'	'Primary School'	'TNEE813 075'	'Morning'	'ISTITUTO COMPRENSIVO CENTRO VALSUGANA'	'Educational and training institution'	'http://www.wicentrovalsugana.it'	'9'	'TNIC813 00R'	'2218829 02'

7.2.1.39 Bonus Query 6

Get the number of students per school type grouped by subregional academic division

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name ?n_k ?n_ps ?n_ms ?n_hs ?n_cf ?total_number_of_schools where {
    ?municipality rdf:type ds:Subregional_Academic_Division_GID-300037 .

    ?municipality ds:has_School_Completion_Rate_GID-300011_Type-300037
    ↳ ?school_completion_rate .
    ?municipality ds:has_Name_GID-2_Type-300037 ?municipality_name .
    ?municipality ds:has_Number_of_Students_Enrolled_in_Kindergarten_GID-300025_Type-300037
    ↳ ?n_k .
    ?municipality
    ↳ ds:has_Number_of_Students_Enrolled_in_Primary_School_GID-300022_Type-300037 ?n_ps .
    ?municipality
    ↳ ds:has_Number_of_Students_Enrolled_in_Middle_School_GID-300032_Type-300037 ?n_ms .
    ?municipality ds:has_Number_of_Students_Enrolled_in_Highschool_GID-300018_Type-300037
    ↳ ?n_hs .
    ?municipality ds:has_Number_of_Students_Enrolled_in_CF_GID-300027_Type-300037 ?n_cf .

{
    select ?local_municipality_name (sum(xsd:integer(?institute_number_of_school)) as
    ↳ ?total_number_of_schools) where {
        ?local_municipality rdf:type
        ↳ ds:Subregional_Academic_Division_GID-300037 .
        ?institute rdf:type ds:Institute_GID-45425 .

        ?local_municipality ds:contains_GID-113311_Type-300037 ?institute .

        ?institute ds:has_Number_of_Schools_GID-300019_Type-45425
        ↳ ?institute_number_of_school .
        ?local_municipality ds:has_Name_GID-2_Type-300037
        ↳ ?local_municipality_name .
    } group by(?local_municipality_name)
}
FILTER(?municipality_name = ?local_municipality_name)
}

```

Result

The screenshot shows a SPARQL query results interface. At the top, the query is displayed in a code editor-like area with syntax highlighting. Below the code, there are several buttons: 'Table' (selected), 'Raw Response', 'Pivot Table', 'Google Chart', 'Run' (highlighted in red), and 'Download as'. A status bar at the bottom indicates 'Showing results from 1 to 95 of 95. Query took 0.1s, moments ago.' The main area displays a table with the following data:

municipality_name	n_k	n_ps	n_ms	n_hs	n_cf	total_number_of_schools
1 'Trento'	'2915'	'5529'	'3882'	'8631'	'1023'	'133'**xsd:integer
2 'Ronciglio Terme'	'60'	'142'	'146'	'0'	'0'	'11'**xsd:integer
3 'Cavalese'	'70'	'180'	'307'	'366'	'0'	'20'**xsd:integer
4 'Brentonico'	'99'	'177'	'125'	'0'	'0'	'4'**xsd:integer
5 'Rovereto'	'1041'	'2124'	'1812'	'4582'	'1413'	'99'**xsd:integer
6 'Arco'	'408'	'972'	'621'	'173'	'352'	'15'**xsd:integer
7 'Pergine Valsugana' a	'559'	'1058'	'801'	'671'	'0'	'39'**xsd:integer
8 'Borgo Valsugana'	'153'	'277'	'197'	'891'	'188'	'27'**xsd:integer

7.2.1.40 Bonus Query 7

Get the number of schools per municipality

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name (sum(xsd:int) as ?total_number_of_schools) where {
?municipality rdf:type ds:Subregional_Academic_Division_GID-300037 .
?institute rdf:type ds:Institute_GID-45425 .
?municipality ds:contains_GID-113311_Type-300037 ?institute .
?municipality ds:has_Name_GID-2_Type-300037 ?municipality_name .
?institute ds:has_Number_of_School_GID-300019_Type-45425 ?institute_number_of_school .
} group by(?municipality_name)

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the query and some metadata about the dataset (e.g., Subregional_Academic_Division_GID-44778). Below the code editor is a table with two columns: 'municipality_name' and 'total_number_of_schools'. The table contains 9 rows of data, each representing a municipality and its total number of schools. The data is as follows:

municipality_name	total_number_of_schools
'Trento'	'133'^xsd:integer
'Roncogno Terme'	'11'^xsd:integer
'Cavalese'	'20'^xsd:integer
'Brentonico'	'4'^xsd:integer
'Rovereto'	'69'^xsd:integer
'Arco'	'15'^xsd:integer
'Pergine Valsugana'	'39'^xsd:integer
'Borgo Valsugana'	'27'^xsd:integer
'Cles'	'19'^xsd:integer

7.2.1.41 Bonus Query 8

Get the highest dropout percentage municipality on average

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name
  (avg(xsd:float(?dropout_percentage)) as
  ?average_dropout_percentage) where {
    ?municipality rdf:type ds:Subregional_Academic_Division_GID-300037
    . ?school rdf:type ds:School_GID-44778 .
    ?institute rdf:type ds:Institute_GID-45425 .
    ?statistics rdf:type
      ds:School_Statistics_GID-300043 .

    ?municipality
      ds:contains_GID-113311_Type-300037
      ?institute .
    ?institute ds:includes_GID-101226_Type-45425
      ?school .
    ?school ds:produces_GID-108406_Type-44778
      ?statistics .

    ?municipality ds:has_Name_GID-2_Type-300037
    ?municipality_name .
    ?statistics ds:has_Dropout_Percentage_GID-300043
      15_Type-300043 ?dropout_percentage
      .

} group by(?municipality_name) order by
  desc(?average_dropout_percentage)
  
```

Result

The screenshot shows a SPARQL query results interface. At the top, the query is displayed in a code editor-like area with syntax highlighting. Below the code, there are tabs for 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. The 'Table' tab is selected, showing a list of 12 municipalities and their average failed percentage. The table has two columns: 'municipality_name' and 'average_failed_percentage'. The data is as follows:

municipality_name	average_failed_percentage
'Borgo Valsugana'	'0.08928867'^xsd:float
'Levico Terme'	'0.0810447'^xsd:float
'Trento'	'0.06473365'^xsd:float
'Riva del Garda'	'0.05326094'^xsd:float
'Cavalese'	'0.040567085'^xsd:float
'Pergine Valsugana'	'0.040280197'^xsd:float
'Rovereto'	'0.03336359'^xsd:float
'Cles'	'0.028623004'^xsd:float
'Tione di Trento'	'0.018560596'^xsd:float
'Mezzolombardo'	'0.0066493847'^xsd:float

At the bottom right of the interface, there is a 'Run' button and a 'Download as' button.

7.2.1.42 Bonus Query 9

Get the highest failed percentage per municipality

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name
  (avg(xsd:float(?failed_percentage)) as
  ?average_failed_percentage) where {
    ?municipality rdf:type ds:Subregional_Academic_Division_GID-300037
    ?school rdf:type ds:School_GID-44778 .
    ?institute rdf:type ds:Institute_GID-45425 .
    ?statistics rdf:type
      ds:School_Statistics_GID-300043 .

    ?municipality
      ds:contains_GID-113311_Type-300037
      ?institute .
    ?institute ds:includes_GID-101226_Type-45425
      ?school .
    ?school ds:produces_GID-108406_Type-44778
      ?statistics .

    ?municipality ds:has_Name_GID-2_Type-300037
    ?municipality_name .
    ?statistics ds:has_Failed_Percentage_GID-3000043
      ?failed_percentage
      .
} group by(?municipality_name) order by
  desc(?average_failed_percentage)
  
```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor with the SPARQL query. Below it is a toolbar with various icons for file operations. The main area displays a table with the following data:

municipality_name	average_admitted_percentage
'Vallelaghi'	'100.0'^xsd:float
'Ledro'	'100.0'^xsd:float
'Altopiano della Vigolana'	'100.0'^xsd:float
'Ossana'	'100.0'^xsd:float
'Avio'	'100.0'^xsd:float
'Predaia'	'100.0'^xsd:float
'Arcò'	'99.97195'^xsd:float
'Volano'	'99.92785'^xsd:float
'Comano Terme'	'99.86897'^xsd:float
'Castel Ivano'	'99.80946'^xsd:float
'Cembra Lisignago'	'99.80769'^xsd:float

7.2.1.43 Bonus Query 10

Get the highest admitted percentage per municipality

SPARQL Query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name
  (avg(xsd:float(?admitted_percentage)) as
  ?average_admitted_percentage) where {
    ?municipality rdf:type ds:Subregional_Academic |
      _Division_GID-300037
    .
    ?school rdf:type ds:School_GID-44778 .
    ?institute rdf:type ds:Institute_GID-45425 .
    ?statistics rdf:type
      ds:School_Statistics_GID-300043 .

    ?municipality
      ds:contains_GID-113311_Type-300037
      ?institute .
    ?institute ds:includes_GID-101226_Type-45425
      ?school .
    ?school ds:produces_GID-108406_Type-44778
      ?statistics .

    ?municipality ds:has_Name_GID-2_Type-300037
    < ?municipality_name .
    ?statistics ds:has_Admitted_Percentage_GID-300 |
      036_Type-300043 ?admitted_percentage
    .
} group by(?municipality_name) order by
  desc(?average_admitted_percentage)

```

Result

The screenshot shows a SPARQL query results interface. At the top, there is a code editor containing the SPARQL query. Below the code editor is a table with the following data:

	municipality_name	average_admitted_percentage
1	'Vallelaghi'	'100.0'^xsd:float
2	'Ledro'	'100.0'^xsd:float
3	'Altopiano della Vigolana'	'100.0'^xsd:float
4	'Ossana'	'100.0'^xsd:float
5	'Avio'	'100.0'^xsd:float
6	'Predaia'	'100.0'^xsd:float
7	'Arco'	'99.97195'^xsd:float
8	'Volano'	'99.92785'^xsd:float

7.2.1.44 Bonus Query 11

Get the highest average dropout students number

SPARQL Query

```

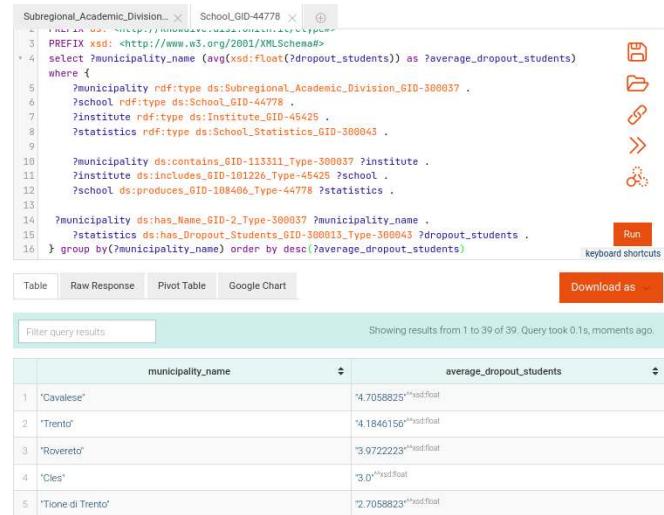
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name
  (avg(xsd:float(?dropout_students)) as
  ?average_dropout_students) where {
    ?municipality rdf:type ds:Subregional_Academic_Division_GID-300037
    _Division_GID-300037
    .
    ?school rdf:type ds:School_GID-44778 .
    ?institute rdf:type ds:Institute_GID-45425 .
    ?statistics rdf:type
      ds:School_Statistics_GID-300043 .

    ?municipality
    ds:contains_GID-113311_Type-300037
    ?institute .
    ?institute ds:includes_GID-101226_Type-45425
    ?school .
    ?school ds:produces_GID-108406_Type-44778
    ?statistics .

    ?municipality ds:has_Name_GID-2_Type-300037
    ?municipality_name .
    ?statistics ds:has_Dropout_Students_GID-300013
    _Type-300043 ?dropout_students
    .

} group by(?municipality_name) order by
desc(?average_dropout_students)
  
```

Result



7.2.1.45 Bonus Query 12

Get the highest metrics in invalsi score per municipality

SPARQL Query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ds: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?municipality_name ?subject
    → (avg(xsd:float(?average_wle)) as
    → ?average_average_wle) where {
    ?municipality rdf:type ds:Subregional_Academic_Division...
    → _Division_GID-300037
    → .
    ?school rdf:type ds:School_GID-44778 .
    ?institute rdf:type ds:Institute_GID-45425 .
    ?quality rdf:type
    → ds:Education_Quality_GID-300039 .
    ?invalsi rdf:type ds:Invalsi_Score_GID-300042 .

    ?municipality
    → ds:contains_GID-113311_Type-300037
    → ?institute .
    ?institute ds:includes_GID-101226_Type-45425
    → ?school .
    ?school ds:has_metric_GID-300001_Type-44778
    → ?quality .
    ?quality ds:influenced_by_GID-787_Type-300039
    → ?invalsi .

    ?municipality ds:has_Name_GID-2_Type-300037
    → ?municipality_name .
    ?invalsi ds:has_Subject_GID-300030_Type-300042
    → ?subject .
    ?invalsi
    → ds:has_Average_WLE_GID-300003_Type-300042
    → ?average_wle .
} group by ?municipality_name ?subject order by
    → desc(?average_average_wle)
```

Result

The screenshot shows a SPARQL query results interface. The query has been run and is displaying the results in a table format. The table has three columns: municipality_name, subject, and average_average_wle. There are three rows of data:

municipality_name	subject	average_average_wle
'Ala'	"Italian"	"0.0"
'Arco'	"Italian"	"0.0"
'Baselga di Piné'	"Italian"	"0.0"

7.2.2 Project conclusion and future work

Regarding the project's results, we think that the objective of creating a Knowledge Graph that contains details and data on educational institutions in Trentino has been attained.

Our project's outcomes are practical and usable in real-world situations. What is more, is that the validation phase has shown a positive detail: the teleology and the KG design were completed in a logical manner. This is evident from the fact that the queries created to address the CQs ended up having a relatively straightforward form.

Although the final Knowledge Graph is straightforwardly applicable as is, there is still room for improvement.

As a future work, we propose to:

- integrate data related to other regions;
- retrieve data for defining the popularity of a school;
- integrates high/middle school professors;
- gather more precise information for the timetables of the courses.

7.2.2.1 Final Evaluation

Regarding KG evaluations, iTelos offers a set of measures.

One of the most beneficial among them is the **coverage** which is the extent to which a KG covers a particular body of knowledge (represented as Etypes and properties).

At first, we used coverage to assess the Knowledge layer of the final KG. In detail, the first evaluation tries to understand:

- how much of the extracted entities and characteristics from the CQs the ETG covers
- the extent to which the reference ontologies' Etypes and characteristics are covered by the ETG.

The Coverage is computed as the ratio between the intersection of α and β and the whole α sets:

$$Cov = (\alpha \cap \beta)/\alpha = C/(A + C)$$

where:

- α is a portion of knowledge to be verified;
- β is the KG's Knowledge Layer;

Concerning the **ETG vs CQs** metrics, the etype coverage was the first we assessed, yielding the following result:

$$Cov_e(CQ_e) = \frac{|CQ_e \cap ETG_e|}{CQ_e} = \frac{11}{12} \approx 0.91$$

Concerning instead the property coverage, we did not define the explicit properties required for answering the competency question during the purpose formalization. However, since our Knowledge Graph successfully answered almost all the competency questions, and the only missing information regards the popularity which in our initial definition relates to the number of students enrolled in that specific school, we claim that we miss one property in those ones required to answer the competency question. Moreover, as we use roughly 40 properties in the answer to the competency question, we can approximately estimate that:

$$Cov_p(CQ_p) = \frac{|CQ_p \cap ETG_p|}{CQ_p} \approx \frac{39}{40} \approx 0.975$$

According to the **ETG vs Reference Ontologies (ROs)**, we have:



- RO_E is the number of Etypes extracted from the ROs, which in our case is 22 coming from both VIVO and schema.org.
- ETG_E is the number of Etypes of the ETG, which is 13 since we have considered the version in which the Etypes of the teleology were not already merged with those of the reference ontologies.
- RO_p is the number of properties extracted from the ROs, which in our case is approximately 24 coming from both VIVO and schema.org. Those properties come from the schema.org entities only, since the VIVO entities have come with no predefined data properties as we have mentioned in the *Formal modeling* section.
- ETG_p is the number of properties of the ETG, which is approximately 60.

Hence, the **Etype coverage** is

$$Cov_e(RO_e) = \frac{|RO_e \cap ETG_e|}{RO_e} = \frac{5}{22} \approx 0.227$$

specifically, the intersection gives rise to {Region, Contact Information, Study Course, Institute, School} while the **property coverage** is:

$$Cov_p(RO_p) = \frac{|RO_p \cap ETG_p|}{RO_p} = \frac{6}{24} = 0.25$$

where the matched properties are: {email, telephone, faxNumber, ratingValue, reviewBody, itemReviewed}.

From this analysis, we can understand that the final knowledge graph is almost perfectly aligned with the purpose we had in mind during the initial phase of *iTelos*. On the other hand, the reusability of our knowledge graph is not exceptional. We believe that the main reason behind this result is the huge amount of contextual information (both related to the purpose we have to pursue).

A second evaluation we can make regards the understanding of how *dense* or *connected* the KG is in order to evaluate the data layer of the KG.

A KG's connectivity can be assessed in two dimensions:

- **Entity connectivity:** The degree to which the entities are interconnected;
- **Property connectivity:** It refers to how closely an entity is linked to its properties.

To comprehend the connection of the KG, we can assess two important factors:

- determine how connected the KG is after the process is complete;
- determine how much the connectivity of the final KG is improved by each dataset that is processed during the process.

To evaluate the connectivity of the final KG, we have measured the following:



- **Entity connectivity:** Therefore the number of entities $E(T)$ for each Etype T in the KG are:

$$\sum_{k=1}^N E(T_k) = 12992$$

While the number of object property values not null $Op(T)$, for each Etype T in the KG:

$$\sum_{k=1}^N Op(T_k) = 7968$$

To compute such information, we have decided to ask the GraphDB using SPARQL:

```
select ?p (count(distinct ?entity) AS ?count_entity) where{ ?entity ?p ?o } group by(?p)
← order by desc(?count_entity)
```

- **Property connectivity:** Concerning the final metric, we have to compute the number of data property values not null $Dp(T)$, for each Etype T in the KG. By employing the following query we know that:

```
select (count(?p) AS ?cnt) where { ?s ?p ?o}
```

$$\sum_{k=1}^N Dp(T_k) = 116332$$

From this evaluation, we can understand that the final knowledge graph is quite large and connected. However, there are several missing data properties which is why we have less property connectivity than entity connectivity.

8 Issues

This section tries to outline any challenges or issues that have not necessarily remained unresolved during the whole process.

Generally speaking about the project, the only significant issue is that, as previously said, it is currently not feasible to produce a popularity index for a school. As a result, the finished product is incomplete; nevertheless, with further research, data may be gathered to create the above-described index.

As for the tools used during the project, there have been issues with their employment, which have caused several delays in our plan.

As far as the KOS system is concerned, a first project was created without immediately adding the collaborator, thinking that the system would let us do it afterward, but this was not the case. To fix this, a new project was created without making the aforementioned mistake.

The biggest mistake we made and the one that cost us most of the time was how we employed Protégé at first. To be more explicit, we directly imported the VIVO teleontology into Protégé in order to attempt to align our teleology with a reference ontology. After the alignment procedure was completed, the file was successfully saved. After that, it was important to do the language alignment of the aforementioned file. This required importing it into KOS, which allowed each specified object and attribute to be given a semantic meaning. However, the system failed to read the custom-made properties and entities for as-yet-unknown reasons. It appears that after numerous hours of research and fruitless attempts, the functionality of importing an already-existing ontology into Protégé has been shown to inadvertently "corrupt" the file that we saved, or alternatively the KOS system was unable to process entity imported from other RDF/OWL files. We reached this conclusion because KOS accurately identified the features that required linguistic alignment after completely reconstructing the aligned teleontology from scratch.

References

- [1] Federico Bianchi, Debora Nozza, and Dirk Hovy. ““FEEL-IT: Emotion and Sentiment Classification for the Italian Language””. In: *Proceedings of the 11th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Association for Computational Linguistics, 2021.
- [2] *Data Catalog Vocabulary (DCAT) - Version 2*. <https://www.w3.org/TR/vocab-dcat-2/>. [Online; accessed 23- November-2022]. 2022.
- [3] Edward L Deci and Richard M Ryan. “The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior”. In: *Psychological inquiry* 11.4 (2000), pp. 227–268.
- [4] Fausto Giunchiglia, Simone Bocca, and Mayukh Bagchi. *Knowledge-Graph-Engineering course, UniTh*. <https://unitn-knowledge-graph-engineering.github.io/KGE2022-website>. Accessed: 2022-10-29.
- [5] Fausto Giunchiglia et al. *iTelos – Purpose Driven Knowledge Graph Generation*. 2021. doi: 10.48550/ARXIV.2105.09418. url: <https://arxiv.org/abs/2105.09418>.
- [6] Michael Grüninger and Mark S. Fox. “The Role of Competency Questions in Enterprise Engineering”. In: *Benchmarking — Theory and Practice*. Ed. by Asbjørn Rolstadås. Boston, MA: Springer US, 1995, pp. 22–31. isbn: 978-0-387-34847-6. doi: 10.1007/978-0-387-34847-6_3. url: https://doi.org/10.1007/978-0-387-34847-6_3.
- [7] *INVALSI - Servizio Statistico*. <https://invalsi-serviziostatistico.cineca.it/>. Accessed: 2022-10-29.
- [8] *Istituto di Statistica della Provincia di Trento*. http://anteprime.provincia.tn.it/pat_statistica_new/. [Online; accessed 29-October-2022]. 2022.
- [9] Rossana Paciello and Valerio Vinciarelli. “SHAPEness Metadata Editor”. In: (Nov. 2021). Download binary release: <https://github.com/epos-eu/SHAPEness-Metadata-Editor/releases/tag/1.2.1>. doi: 10.5281/zenodo.5714433.
- [10] Norman Reider. “A Theory of Human Motivation: AH Maslow, Psychological Rev., L, 1943, pp. 370–396”. In: *Psychoanalytic Quarterly* 14 (1945), pp. 144–145.
- [11] *Trentino Digitale, Open Data project*. <https://www.trentinodigitale.it/Infrastrutture/Piattaforme-abilitanti/Open-Data>. Accessed: 2022-10-29.
- [12] Ashish Vaswani et al. *Attention Is All You Need*. 2017. doi: 10.48550/ARXIV.1706.03762. url: <https://arxiv.org/abs/1706.03762>.
- [13] *Vivoscuola, Il portale della scuola in Trentino*. <https://www.vivoscuola.it/>. Accessed: 2022-11-11.
- [14] Wikipedia contributors. *Google Developers — Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/w/index.php?title=Google_Developers&oldid=1117452416. [Online; accessed 29-October-2022]. 2022.
- [15] Wikipedia contributors. *SPARQL — Wikipedia, The Free Encyclopedia*. [Online; accessed 14-December-2022]. 2022. url: <https://en.wikipedia.org/w/index.php?title=SPARQL&oldid=1101084257>.
- [16] Mark D Wilkinson et al. “The FAIR Guiding Principles for scientific data management and stewardship”. In: *Scientific Data* 3.1 (Mar. 2016), p. 160018.