UNIVERSITY
OF TRENTO - Italy

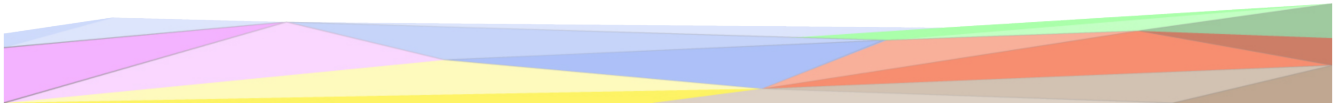
Dipartimento di Ingegneria e Scienza dell'Informazione

– KnowDive Group –

Knowdive Research Group

Document Data:	Reference Persons:
February 8, 2025	Aloisi Deborah, Giaretta Leonardo

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

Data has become an essential part of our lives, but without proper organization and management, it loses its value.

This project has the objective of organizing data about members, alumni, papers and other products of the Knowdive research group in a Knowledge Graph that makes it simpler to query and find information about it.

This will be accomplished using the iTelos methodology, a structured approach designed to streamline the Knowledge Graph Engineering (KGE) process and minimize development effort. The methodology is intended to assist users in addressing the challenges that arise when building purpose-specific knowledge graphs (KGs). It also facilitates the creation of reusable resources, promoting circular data reuse and thus reducing the effort required to generate new resources.

In the iTelos methodology, re-usability is a central principle in KGE, with project documentation playing a critical role in enhancing the re-usability of the resources developed throughout the project. This report aims to document both the process and methodology used, explaining the rationale behind key decisions and how resources were leveraged to achieve the project's objectives. By doing so, the documentation will serve as a valuable resource for external readers, potentially enabling them to reuse the project's outputs for other purposes.

The report is structured as follows:

- Section 2: Definition of the project's purpose and its domain of interest.
- Section 3: Definition of the Data Source.
- Sections 4 and 5: The description of the iTelos process phases.
- Section 6: The description of the evaluation criteria and metrics applied to the project final outcome.
- Section 7: The description of the metadata produced for all (and all kind of) the resources handled and generated by the iTelos process, while executing the project.
- Section 8: Conclusions and open issues summary.

2 Purpose Definition

2.1 Informal Purpose

The final Knowledge Graph (KG) can be utilized as a general-purpose service to assist users in discovering research products (such as papers, reports, and other published outcomes) produced by the Knowdive research group at the University of Trento (DISI). This functionality can be articulated as a user request in the following form:

”A service which helps the users to query and know about all the products and the people that are working or were working in the Knowdive research group at the University of Trento”

2.2 Domain of Interest (DoI)

- Personal context : We expressed our Personal context Purpose from the point of view of a member of the Knowdive research group, as they are the category of users which have a strictly personal interest in this project. ”I want a KG which collects the results of my research and available data from the projects I’m taking part into”
- Reference context: We expressed our Reference Context purpose in a dual way, in order to emphatise two different aspects of this project: the geographical aspect and the research aspect. ”I want a KG which represents the distribution across the globe of Knowdive’s members and projects” ”I want a KG which represents the research products of the Knowdive research products”
- Personal-Reference context Through the Personal-Reference context purpose we were able to express the point of view of a non-member user. ”I want a KG which allows me to access the data regarding the Knowdive group members, their research products and their projects by filtering through various criteria.”
- Reference-Personal context Through the Reference-Personal context purpose we expressed the broader point of view that takes into account both member and non-member users. ”I want a KG which integrates the data about members, research product and projects of the Knowdive Research Group, allowing users to consult data on various scientific topics and members of the group to organize their work results.”

In term of context we define:

- Geographical boundaries: We express the spatial scope with the coordinate of the University of Trento, more precisely with the Department of Information and Computer Science

(DISI) where the Knowdive research group is officially located. The coordinate are:

Department	Latitude	Longitude	Altitude (m)
Dipartimento di Ingegneria e Scienza dell'Informazione	46.0668994	11.1481093	370

- Temporal boundaries: The scope is from the starting of the Knowdive research group in the 2006 to the current year, since a potential user could be both interested in the archived data and current and updated information.
- Domain boundaries: The project concerns the Knowdive Research Group and in particular its member, alumni, papers and other works produced.

2.3 Purpose Formalization

2.3.1 Scenarios definition

1. A Master's student is searching inspiration for the topic of their thesis. They want to see the publications of the Knowdive group so that they can write to a professor that works on their interested topic.
2. A company wants to employ a person with reference to the Knowdive research group, so they want to check if the credentials are correct and if he/she aligns with the position they are offering.
3. A University wants to invite member of the Knowdive research group for a seminar and are searching which member would be the best to contact.
4. A student that will move to another university wants to know if in his/her new university there is a member of the Knowdive group and if the research interests of such a member align with his/hers.
5. An Institution (such as a ranking Institution) wants to measure the growth of the Knowdive group through its research output and/or the current amount of projects its undertaking.
6. A Researcher of the Knowdive group wants to have a neatly organized aggregation of the research products made by him/her and his/hers colleagues.

2.3.2 Personas

Name	Age	Profession	Description
Carlotta	24	Student	Carlotta is approaching the end of her Master's degree, but without a thesis she can't graduate. She doesn't have a precise idea of what she want to study so she is searching inspiration on between other thesis. In doing so she ended on the Knowdive research group page and wants to see what they are studying
Gianpaolo	47	HR personnel	Gianpaolo is one of the HR personnel at company X in Trentino. His company wants to develop a project correlated to a Knowledge graph but they don't have any employee that can take care of it. So his company instructed him to search for a possible candidate to hire for the project and the first idea that has come to his mind is to search for figures connected to the local University.
Mabel	35	Secretary of University	Mabel was tasked by a professor at her university to contact the Knowdive Research group to invite a member of the group to host a seminar. She wants to find who is the best person to contact for the task
Franco	23	Student	Franco is a Master's student with a personal interest in the Human Machine Symbiosis topic. He was planning to try and propose a Project course on this topic, but he doesn't know which professor may be interested. By talking with other students he discovers the existence of the Knowdive group and sets out to find a member that matches his interest.
Anna	22	Student	Anna will be soon finishing her Bachelor's degree and will be moving to another university. Having just developed a passion for Artificial Human Cognition and knowing that the Knowdive group does research on that topic, she searches for connections between the group and her future University.
Valerio	41	Institution X employee	As a part of a future analysis on the research environment in Trentino, Valerio has been tasked by the Institution X to collect data regarding the growth of the various research groups and organization of the Trentino territory. He is currently collecting data on Knowdive and needs the most precise data possible about the amount of members of the group through time, their projects and the amount of research published.
Arianna	37	Knowdive member	Arianna always liked tidy and clear aggregation of data so she want to see all the products done by her research group divided by colleague, categories and so on.

Table 1: Personas

2.3.3 Competency Question (CQs)

Persona	ID	Question
Carlotta	1.1	I'd like to know all the thesis written in collaboration with the Knowdive Research group

Carlotta	1.2	I'd want to see all Master thesis written in collaboration with the Knowdive group
Carlotta	1.3	I'm curious if the people that wrote their thesis in collaboration with the Knowdive group have any postdoctorate work.
Carlotta	1.4	Give me the topics a member of the KRG is studying
Carlotta	1.5	I've finally decided on a topic for my thesis, but I still don't know which member of the KRG to contact. I would need to have a list of all members that work on my chosen topic with the possibility to filter through role and location.
Gianpaolo	2.1	Give me all the current member of the KRG
Gianpaolo	2.2	Give me all the ex-member and their position when they left
Gianpaolo	2.3	Give me only the members that are researchers or professors
Gianpaolo	2.4	Let me search if a specific person has been part of the KRG and get their position when they left and a list of their research products and of the projects they participated in.
Mabel	3.1	Give me a list of the research topics of the KRG.
Mabel	3.2	I want to be able to find all the members of the KRG that are either a professor or a researcher that works or has worked on a given topic.
Mabel	3.3	I want to know who is the coordinator for the seminar activities
Mabel	3.4	I need to get the contact information for a specific member of the KRG.
Mabel	3.5	The person I thought of is unavailable, but I could contact one of his collaborators. I need to find all the members that have worked with him/her on the chosen topic.
Franco	4.1	I want to see all research products produced by the KRG on a specific topic.
Franco	4.2	I want to see each KRG member that has worked on a research product.
Franco	4.3	I know a professor that is part of the KRG, I want to see on which topics he/she worked on during the last years.
Franco	4.4	Are there any members that have worked with the professor I know that have also worked on the topic I prefer?
Anna	5.1	I'd like to know which KRG members are located in my future university.
Anna	5.2	I'd like a list of all non-members of the KRG that have collaborated with the group on a given topic.
Anna	5.3	I'd like a list of all ex-members of the KRG that have worked on a given topic.
Anna	5.4	I would need the location of an ex-member of KRG before he left the group.
Anna	5.5	I would like to know which courses the members of the KRG teach
Anna	5.6	I would like to know who teach the Studies in Human Behavior corse
Valerio	6.1	I need to track how many members the KRG had through the years and who these members are.
Valerio	6.2	I need to track the amount of research products produced by the KRG through the years and fetch the appropriate identifying data about each of those products.

Valerio	6.3	I need to be able to track in which years a current or past member of the KRG has been part of the group, how many and which research products he/she produced during that time.
Valerio	6.4	I need to track the amount of research products on a given topic that has been produced on a given topic and fetch the appropriate identifying data about each of those products.
Valerio	6.5	I need to track the status and duration of each project in which the KRG has partecipated.
Valerio	6.6	For each project, i need to extract all the Universities and Institutions that have partnered with the KRG.
Arianna	7.1	I'd like to have a list of all the research products of the KRG, ordered by date.
Arianna	7.2	I'd like to know, for each project in which the KRG has taken a part in, which members have worked on it.
Arianna	7.3	For each member of the KRG, I would like to have a list of his/her research products.
Arianna	7.4	I want to know in which years a member of the KRG was active the most
Arianna	7.5	I want to know which topic is more researched based on related works created

Table 2: CQs

CQ ID	Common entities	Core Entities	Contextual entities
1.1		Research Product	
1.2		Research Product	
1.3	Person	Research Product	
1.4	Person	Member, Research Topic	
1.5	University	Member, Research Topic	
2.1		Member	
2.2	Alumni		
2.3	Member		
2.4	Member, Research Product, Project		
3.1		Research Topic	
3.2	Member, Research Topic		
3.3	Member		
3.4	Member		
3.5		Member, Research Topic	
4.1		Research Product, Research Topic	
4.2		Member, Research Product	
4.3		Member, Research Topic	
4.4		Member, Research Topic	
5.1	University	Member	
5.2	Person	Research Topic, Research Product	
5.3		Alumni, Research Topic	

5.4	University	Alumni	
5.5	University	Member	Course
5.6	University	Member	Course
6.1		Member, Alumni	
6.2		Research Product	
6.3		Member, Alumni, Research Product	
6.4		Research Product, Research Topic	
6.5		Project	
6.6	University	Project	
7.1		Research Product	
7.2		Project, Member	
7.3		Member, Research Product	
7.4		Member, Project, Research Product	Course
7.5		Research Topic, Research Product	

Table 3: Entities related to CQs

2.3.4 Concepts identification

In a knowledge graph, entities represent real-world objects and are instances of Entity Types (ETypes), which define the types of real-world objects. This section categorizes ETypes based on their popularity categories: Common, Core, and Contextual.

Common entities

Common entities consist of resources that contain information relevant across multiple contexts or domains of interest. While not directly tied to the user's Purpose, they are essential for supporting it within the knowledge graph.

For this project, the identified common entities are:

- University
- Person

Core entities

Core entities are those that contain essential information about the key aspects central to the Purpose, forming the foundational data necessary to construct the knowledge graph. These entities are harder to locate and less reusable than common entities.

For this project, the identified core entities are:

- Member
- Alumni
- Research Product (Thesis, Papers, Reports, etc...)

- Research Topic
- Project

Contextual entities

Contextual entities are those that encompass resources containing specific, often unique information directly related to the user's Purpose. These resources are designed to add distinctive value, setting the application apart from competitors. While core resources are essential for creating a functional application, contextual resources provide a competitive edge. Typically, they are not reusable and often need to be created from scratch.

For this project, the identified contextual entities are:

- Course

2.3.5 ER model definition

The entities previously defined are now used to construct the draft of a possible ER model. The model is going to be improved in the next phase during the information gathering.

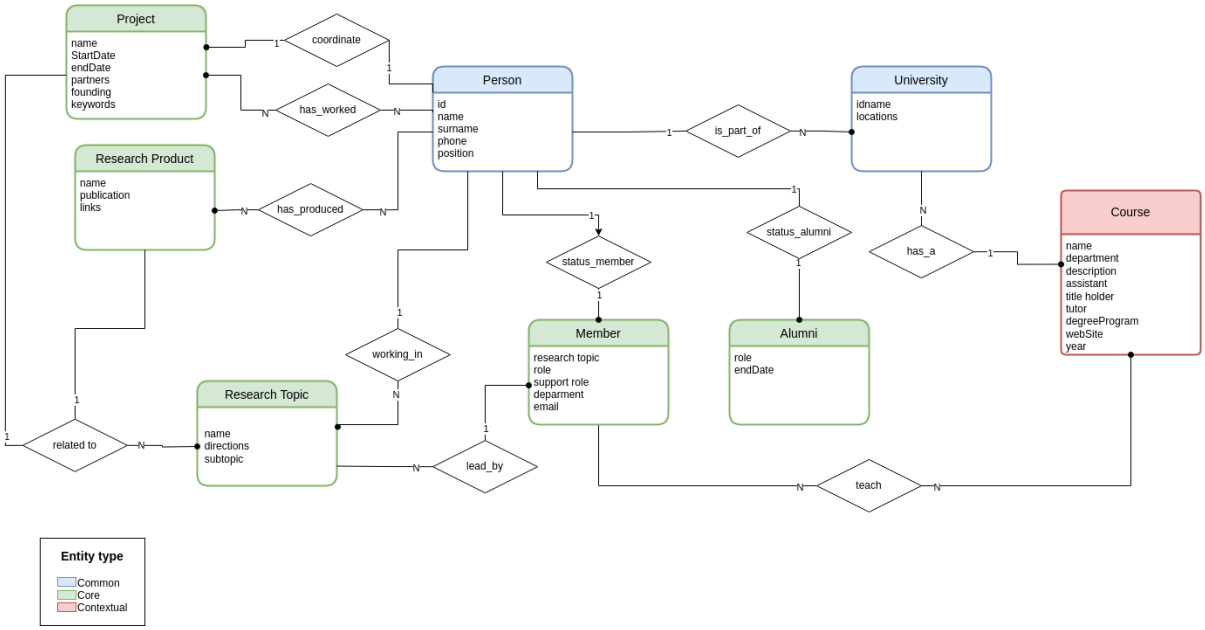


Figure 1: Entity-Relation diagram

3 Information Gathering

3.1 Sources identification

In this project there are three considered types of resources:

3.1.1 Data values datasets

For the purpose of the project we have considered numerous data sources, where some of these have to be scrapped.

The one that were suitable are:

- LiveDataUNITN catalog
  - Courses
  - Research Product
  - Staff
- Knowdive site
  The Knowdive site has been scraped through various means to obtain crucial information, such as the Research Topics of the Knowdive group, which wouldn't have been available otherwise.
  - Research Topic
  - Courses
  - Research Product
  - Project
  - Staff
- OpenStreetMap
  The usage of OpenStreetMap as a data source allows us to put together information about the various universities, especially geographical information.
  The size of the complete dataset was the first difficulty we encountered and the solution has been to download data from this source only when needed.
  The second problem we came across has been the highly variance between the entries of the various universities, with certain universities having entries in OSM for their campuses, other for their structures and others having a single unified entry. This has been resolved during the cleaning of the dataset.
  A third problem has been the lack of data on the coordinates of the universities in most of

the dataset as it is downloadable from OSM. This kind of data has been manually integrated in the dataset during cleaning.

  - University

3.1.2 Knowledge datasets(ontologies)

The knowledge sources chosen for the project were selected based on which ones would be useful in our domain and for our purpose, particularly those containing information regarding the research activities and the organizational structure of an institution. The knowledge sources identified for this project are the following:

- Digital University Ontology, which models and represents the information regarding the research activities of the University of Trento (UNITN), as well as its staff.
- VIVO Core Ontology (VIVO), which is suitable for the academic and research domain.
- Academic Institution Internal Structure Ontology (AIISO) which provides classes and properties to describe the internal organizational structure of an academic institution.
- Semantic Web for Research Communities (SWRC) which is useful for modeling entities of research communities such as persons, organizations, publications (bibliographic metadata) and their relationship.

3.1.3 Language datasets

The language dataset we use for this project are:

- DU-UNITN Language: the dataset collect and describe the terms used into the Digital University data of the University of Trento.
- KRG-UNITN Language: the dataset is created for collecting and describing the term used into the data of the Knowdive Research Group.

3.2 Datasets collection

A part of the data were easily collected, in particular, all the data from the LiveDataNum catalog:

- Courses:LiveDataUNITN-courses
- Research paper: LiveDataUNITN-papers
- Staff: LiveDataUNITN-people

The Courses dataset consist of the following field for each course:

Field name	Description
nome	Name of the course
dipartimento	The academic structure that offers the course;(nome, id)
descrizione	Textual description of the course
docenti	List of professors who teach the course; (nome, cognome, id)
assistenti	List of assistants for the course; (nome, cognome, id)
titolari	List of title holders of the course; (nome, cognome, id)
tutor	List of tutors for the course; (nome, cognome, id)
corsoStudi	Degree program of the course
sitoWeb	Official (Esse3) webpage of the course

The Staff dataset consist of the following field for each person:

Field name	Description
id	Unique identifier of the person
nome	First name
cognome	Last name
telefono	Phone number (formatter as a list)
posizioni	A list of positions;(ruolo, nomeStruttura, idStruttura)

The Research paper dataset consist of the following field for each paper:

Field name	Description
titolo	Title of the paper
tipo	Type of the paper
anno	Year of publication
lingua	Language in which the paper is written (formatter as a list)
autori	A list of people;(nome, cognome, id)
citazioni	Text used for citation in other paper
file	File of the paper;(nome, link, licenza, formato, versione, openAccess)

llThe data format of these sources is semi-structured and in particular JSON format, so we decided to use the JSON format to all the datasets for maintaining homogeneity.

Then we scrapped the Knowdive Research Group Site for information on:

- Member: KRG-UNITN-Member
- Alumni: KRG-UNITN-Alumni
- Research topic: KRG-UNITN-Research-Topic
- Published papers and other products: KRG-UNITN-AHC-Products+KRG-UNITN-LD-Products+KRG-UNITN-KD-Products+KRG-UNITN-HMS-Products=KRG-UNITN-Products
- Projects: KRG-UNITN-Projects
- Courses: KRG-UNITN-Courses

The information obtained was saved and integrated into JSON format compatible with the previous datasets.

The KRG member dataset consist of the following fields for each person:

Field name	Description
id	Unique identifier of the person
name	First name
surname	Last name
role	Role related to the university
researchTopic	Topic in which they work
supportRole	Role related to the Knowdive research group
university	University they are affiliate with
department	Which department they are part of
email	Person's email

The KRG alumni dataset consist of the following fields for each person:

Field name	Description
id	Unique identifier of the person
name	First name
surname	Last name
role	Role related to the university they had
university	University they are affiliate with
endDate	Date in which they left the Knowdive Research Group

The KRG research topics dataset consists of the following fields for each topic:

Field name	Description
name	The name of the research topic
Lead by	Name of the leading member on the topic
Members	list of members that work on the topic
Direction	listed directions on the Knowdive site
sub-topics	listed sub-topics on the Knowdive site
projects	listed projects on the Knowdive site

The KRG products dataset consist of the following fields for each product:

Field name	Description
name	The name of the research product
publication	Where the product is available
authors	The authors of the product
links	A link to the product

The KRG project dataset consist of the following fields for each project:

Field name	Description
name	The name of the project
start Date	When the project started
end Date	When the project ended
coordinator	The Person or University coordinating the project
partners	The list of Universities or Organizations partecipating in the project
members Partecipating	The list of members of the KRG taking part in the project
related Topics	List of the KRG Research Topics related to the project
funding	Organization or Program funding the project
keywords	List of keywords regarding the project

The KRG courses dataset consist of the following fields for each course:

Field name	Description
name	The name of the course
teachers	The listed teachers for the course
year	The year the course
university	The university that hosted the course
webpage	The webpage of the course

Finally, we also collected the OSM data in the XML format. Due to the various problems with data from this source previously mentioned, this data has been processed on a case by case basis. We provide a representation of the final fields in the next section.

The Data dataset collected are enough to answer every CQs we have described in Table 2.

3.3 Datasets cleaning

Data cleaning focuses on removing irrelevant or unnecessary ”noise” from collected datasets, ensuring that the data is accurate, relevant, and useful. We doing so by identifying and removing entities or types within a dataset that have no use and no relevance to the analysis or purpose at hand. Eventually the quality of the dataset is improved, making it more focused and valuable for its intended use.

The data scrapped from the Knowdive website didn’t undergo a substantial cleaning procedure, as most of the work was already done during the scrapping process. Only a couple of the Research Topic dataset’s fields were removed after some consideration, namely the Direction and sub-topics fields as they were, more often than not, repetitive or unnecessary.

Instead the data that come from the LiveDataUNITN catalog and OpenStreetMap need to be removed of the surplus data.

For the LiveDataUNITN catalog we removed for each datasets:

- LiveDataUNITN-Courses: all the courses that are not teached by a member of the KRG and the attributes 'dipartimento', 'tutor' and 'corsoStudi'.
- LiveDataUNITN-Papers: all the paper not written by a member or alumni of the KRG and the attribute 'citazioni'.
- LiveDataUNITN-People: all the people that are not part of the KRG and the attributes 'ssn' and 'cun'.

For the OpenStreetMap data, we cleaned up all the fields used by OSM to connect and reference its various object, leaving only the actual data about the universities. We then added manually the geographical data where it was missing and restructured the data to allow for smoother use.
The dataset now contains a list of Universities identified by their names. For each university a location field has been created, containing a list of structures connected to the specific university. Each element of the locations is structured as follows:

Field name	Description
tag	a collection of data about the location, as available through OSM
_id	ID used by OSM
_timestamp	The timestamp of the last update. Used by OSM.
latitude/longitude	The latitude and longitude values

3.4 Datasets standardization

Data standardization is crucial for ensuring consistency and usability across various datasets, particularly when they come from diverse sources. Aligning different data formats and types, permit to reuse of the data making them more accessible and easier to manage. We have decided to aligning our data to the formats of common standards such as CSV, XML, TSV, JSON, RDF, TTL and OWL to facilitate smoother data integration.

The datasets from the LiveDataNum catalog once downloaded where named with the .txt extension despite being in the JSON format. Removing the extension was enough to align them to the standard.

The dataset obtained from OpenStreetMap has been converted from its original format (XML) to json. This operation has been performed before the actual cleaning.

The dataset created from scrapping the Knowdive website were created in JSON format from the start to align with the previous datasets.

The Knowledge datasets are in the common used format of OWL and TTL.

The Language dataset downloaded is in the CSV format so the dataset we created for the Knowdive Research Group is also in the CSV format for a easy integration.

4 Language Definition

4.1 Concept identification

The objective of this phase is to define purpose-specific language resources: this process allows the transformation of the informal concepts into formal terms.

4.1.1 Concepts Selection

We select all the concepts that will be used to represent the information in the final KG from the ETypes and Data properties of our ER model.

The concepts we are going to formalize for the scope of this project are in the following table divided by each EType. We are going to formalize both the Etypes and their attributes.

Person	**Member**	**Alumni**	**Project**	**Research Products**	**Research Topic**	**University**	**Course**
id	research Topic	role	name	name	name	name	name
name	role	university	startDate	type	lead	location	description
surname	support role	endDate	endDate	year	member		university
phone	department		coordinator	language	projects		teacher
position	email		partners	publication			webpage
	university		member	author			year
			funding	file			
			keyword				
			research Topic				

4.1.2 UKC alignment

We want to define the formal definition for all the concept identified previously so first we exploited the UKC resources to formalize them.

The major of the concepts formalization were available on UKC. The replicated concept, such as name or year, is formalized only once in the language resource build file. For the missing concept, we have tried to search in other resources like WordNet, but unfortunately we were not able to integrate them into our work. The outcome was that some of them had to be manually declared as new concepts. We defined the following new concept:

- Research topic

- support role
- endDate
- startDate
- keywords
- Research products

4.1.3 Language resource building

The final language resource for our project is:

ConcetID	Word-en	Gloss-en
UKC-36	Person	a human being
UKC-36255	id	a card or badge used to identify the bearer
UKC-2	name	a language unit by which a person or thing is known
UKC-33528	surname	the name used to identify the members of a family (as distinguished from each member's given name)
UKC-34007	phone	the number is used in calling a particular telephone
UKC-2929	position	a job in an organization
UKC-14	Member	one of the persons who compose a social group (especially individuals who have joined and participates in a group organization)
UKC-3646	role	the actions and activities assigned to or required or expected of a person or group
KGE24-13-0002	support role	secondary actions and activities assigned to or required or expected of a person or group distinctive from a role
UKC-43201	department	a specialized division of a large organization
UKC-33270	email	(computer science) a system of world-wide electronic communication in which a computer user can compose a message at one terminal that is generated at the recipient's terminal when he logs in
UKC-47887	Alumnus	a person who has received a degree from a school (high school or college or university)
KGE24-13-0004	endDate	the date in which something ended
UKC-4011	Project, ResearchProject	any piece of work that is undertaken or attempted
KGE24-13-0005	startDate	the date in which something started
UKC-48923	coordinator	someone whose task is to see that work goes harmoniously
UKC-48778	partners	an associate who works with others toward a common goal
UKC-63194	funding	a reserve of money set aside for some purpose
KGE24-13-0006	keyword	a word or a list of words that describe the content of the thing they are associated to
KGE24-13-0003	Research Products	an artifact that has been created by someone or some process for the research purpose
UKC-31362	type	a subdivision of a particular kind of thing

UKC-73248	year	a period of time containing 365 (or 366) days
UKC-33289	language	a systematic means of communicating by the use of sounds or conventional symbols
UKC-34686	publication	a copy of a printed work offered for distribution
UKC-53828	author	writes (books or stories or articles or the like) professionally (for pay)
UKC-34276	file	a set of related records (either written or electronic) kept together
KGE24-13-0001	Research Topic	the subject matter of a research
UKC-46947	leader	a person who rules or guides or inspires others
UKC-44043	University	an educational institution of higher education and research which grants academic degrees in a variety of subjects and provides both undergraduate education and postgraduate education
UKC-50	location	a point or extent in space
UKC-4463	Course	education imparted in a series of lessons or meetings
UKC-3	description	a statement that represents something in words
UKC-53192	teacher	a person whose occupation is teaching
UKC-33648	webpage	a document connected to the World Wide Web and viewable by anyone connected to the internet who has a web browser

This phase definition remains open until the next phases since the data might not fully align with the teleology. The compatibility will be better explored during the knowledge definition and entity definition phase and, if necessary, more data will be added during the process.

4.2 Dataset filtering

Some of our datasets have entities and attributes explicated in the Italian language. We have decided to translate them in the corresponding english version instead of adding the language Italian to the project since the Knowdive Research Group is focused on an international public and not only the Italian one.

Specifically, the translation has been performed on the following files: Cleaned-DU-UNITN-alumni, Cleaned-DU-UNITN-courses, and Cleaned-DU-UNITN-member. The set of files regarding the Research Products has also undergone a translation operation, which has been performed together with an union of multiple files. The results are 4 files, each containing the Research Products of a specific topic. Year-based distinctions are now only possible through the fields of the entries.

Some scraped data also needed some language realignment to guarantee language uniformity across all instances of the same concept. Such files were KRG-UNITN-courses and KRG-UNITN-projects.

5 Knowledge Definition

5.1 KTelos

In this section we will describe the process of Knowledge definition, which will result in the production of a functioning Ontology. Such an Ontology will be fundamental in the next phase to integrate our data and construct the Knowledge Graph.

We model the Teleology first and the Teleontology later ensures that the Knowledge Graph aligns with its intended purpose and domain requirements. By defining the Teleology first, the focus is placed on understanding the primary objectives, goals, and competency questions that the KG must address. Once the purpose is clear, the Teleontology is modeled to structure the knowledge representation accordingly, defining entity types, properties, and relationships that accurately reflect the required domain knowledge. This approach keeps the KG both goal-oriented and structurally sound, minimizing unnecessary complexity while ensuring that the final representation effectively serves its intended use cases.

5.1.1 Teleology definition

Our work on Knowledge Definition began by constructing an Teleology based on our previous work. Therefore, we collected all our EType, object and data proprieties, shown for convenience in the following table, and plugged them into Protégé.

EType	**Object property**	**Data property**		
Person	working_in	author	name	language
Alumnus	has_produced	coordinator	partners	leader
Member	has_a	department	phone	locations
ResearchProduct	lead_by	description	position	members
ResearchProject	is_part_of	email	project	university
ResearchTopic	related_to	endDate	startDate	role
University	coordinates	file	support_role	webpage
Course	status_alumni	funding	surname	year
	status_member	publication	teacher	
	teach	id	keywords	
	has_worked	topic	type	

5.1.2 Teleontology definition

After the construction of our initial Teleology, we used the previously collected Knowledge Resources and other available Ontologies to match our definitions with. This process has made our work interconnected with pre-existing Knowledge resources and, therefore, more reusable

in the future. The following tables show the substitutions that we were able into our own project using pre-existing work, where there is not matched Etype or property from a known ontology it means that we have created them specifically for our purpose.

Old ETypes	Matched Etype	Ontology References
Person	Person	http://num.edu.mn/ontology#Person
Alumnus	Alumni	https://schema.org/alumni
Member	member	https://schema.org/member
ResearchProduct	—	http://knowdive.disi.unitn.it/etype#ResearchProduct
ResearchProject	ResearchProject	http://swrc.ontoware.org/ontology#ResearchProject
ResearchTopic	ResearchTopic	http://swrc.ontoware.org/ontology#ResearchTopic
University	University	http://vivoweb.org/ontology/core#University
Course	Course	http://purl.org/vocab/aiiso/schema#Course

Old Object properties	Matched Object Properties	Ontology References
working_in	expertise_in	http://num.edu.mn/ontology#expertise_in
has_produced	—	http://knowdive.disi.unitn.it/etype#has_produced
has_a	has_subject	http://num.edu.mn/ontology#has_subject
lead_by	headOf	http://swrc.ontoware.org/ontology#headOf
is_part_of	—	http://knowdive.disi.unitn.it/etype#is_part_of
related_to	—	http://knowdive.disi.unitn.it/etype#related_to
coordinates	responsibleFor	http://purl.org/vocab/aiiso/schema#responsibleFor
status_alumni	—	http://knowdive.disi.unitn.it/etype#status_alumni
status_member	—	http://knowdive.disi.unitn.it/etype#status_member
teach	teaches	http://purl.org/vocab/aiiso/schema#teaches
has_worked	worksAtProject	http://swrc.ontoware.org/ontology#worksAtProject

Old Data Properties	Matched Data Properties	Ontology References
author	author	https://schema.org/author
coordinator	projectCoordinator	https://dbpedia.org/ontology/projectCoordinator
department	department	https://schema.org/department
description	description	http://num.edu.mn/ontology#Description
email	email	http://num.edu.mn/ontology#Email
endDate	endDate	http://swrc.ontoware.org/ontology#endDate
file	file	https://trdf.sourceforge.net/provenance/ns.html#File
funding	fund	http://num.edu.mn/ontology#Fund
publication	howpublished	http://swrc.ontoware.org/ontology#howpublished
id	identifier	http://num.edu.mn/ontology#Identifier
keywords	keyword	http://vivoweb.org/ontology/core#Keyword
language	language	http://www.knowdive.disi.unitn.it/etype#Language
leader	leader	https://dbpedia.org/ontology/leader
locations	location	http://swrc.ontoware.org/ontology#location

members	member	https://schema.org/member
name	name	http://swrc.ontoware.org/ontology#name
partners	partner	https://dbpedia.org/ontology/partner
phone	phone	http://swrc.ontoware.org/ontology#phone
position	position	http://www.knowdive.disi.unitn.it/etype#Position
project	project	https://schema.org/Project
startDate	startDate	http://swrc.ontoware.org/ontology#startDate
support_role	—	http://knowdive.disi.unitn.it/etype#support_role
surname	surname	http://num.edu.mn/ontology#Surname
teacher	teacher	http://linkedscience.org/teach/ns#Teacher
topic	topic	http://www.ontotext.com/proton/protontop.html#Topic
type	type	http://swrc.ontoware.org/ontology#type
university	university	http://vivoweb.org/ontology/core#University
role	vcard:role	http://www.w3.org/2006/vcard/ns#role
webpage	website	http://www.knowdive.disi.unitn.it/etype#Website
year	year	http://swrc.ontoware.org/ontology#year

We were able to reuse pre-existing work for almost all of the Entity and the attributes. Instead we needed to define about half of the Object property because they are specific for our project and difficult to match with an existing work.

Following are the entities visualized in Protégé:

Each element in Classes and Data properties contains annotations about the definition and its UKC concept ID. Meanwhile the Object Property are mapped in the description field to the Domains (EType from which the relation starts) and Ranges(EType in which the relation ends).

Finally, though the use of WebVOWL, we were able to visualize our Teleontology.

5.2 Dataset cleaning and formatting

Once our work on the Teleontology was finished, we performed some additional cleaning and formatting on our dataset, in order to create a correspondence between the Knowledge and Data layers and guarantee smoothness of process in the following phases of the project.

Specifically, we performed re-alignment on the all the data scraped in Phase 2 (KRG-UNITN-alumni, KRG-UNITN-members, KRG-UNITN-courses, KRG-UNITN-projects, KRG-UNITN-research-products, KRG-UNITN-research-topics) and the data originally from LiveDataUNITN (DU-UNITN-alumni, DU-UNITN-members, DU-UNITN-courses). Surprisingly, the data from LiveDataUNITN regarding the research products (DU-UNITN-research-products) didn't need any realignment as its fields already conformed to our ontology.

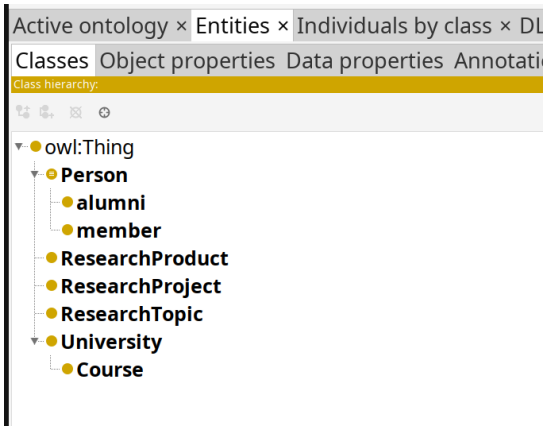


Figure 2: Classes

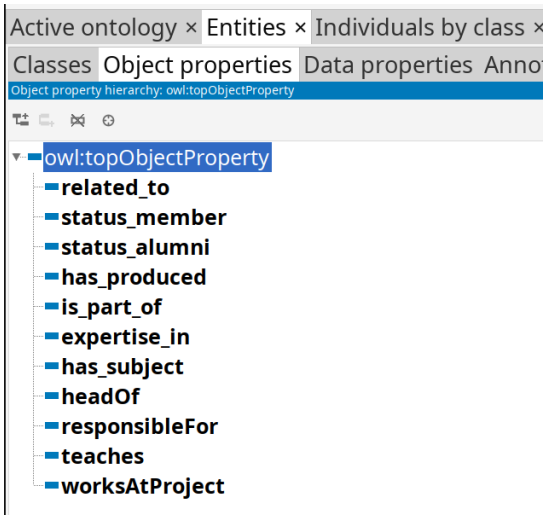


Figure 3: Object property

Figure 4: Data Property

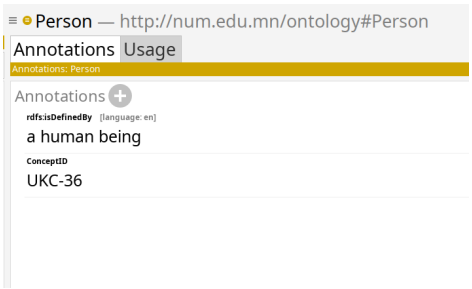


Figure 5: Annotations

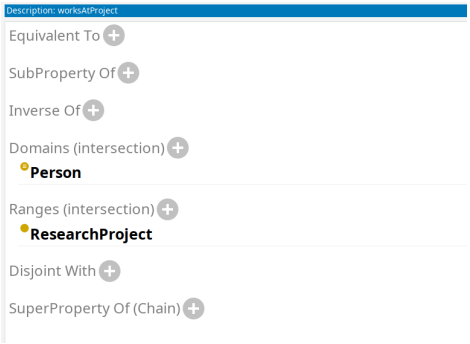


Figure 6: Descriptions

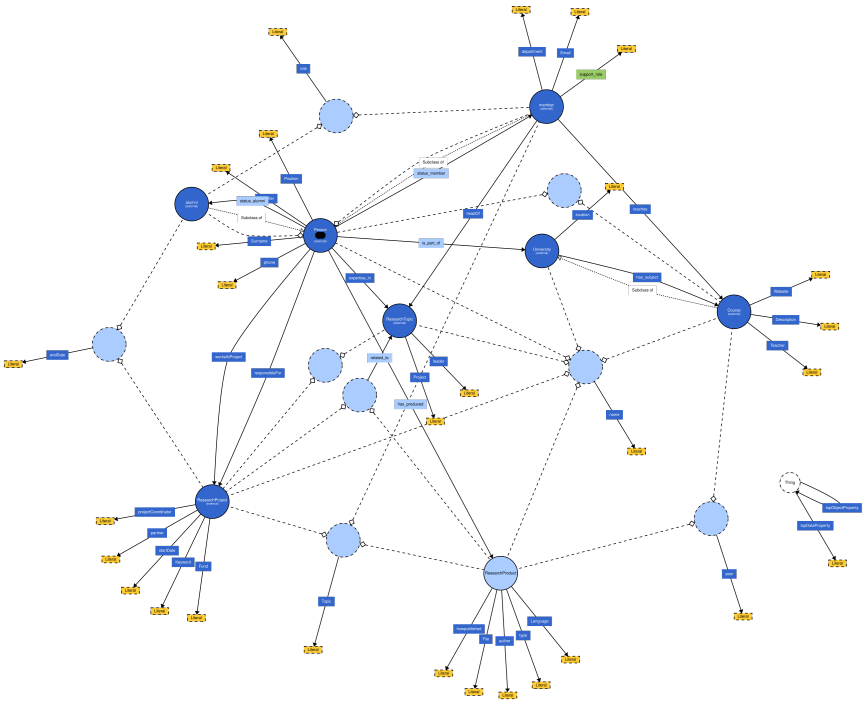


Figure 7: Visualization of our ontology created with WebVOWL

6 Entity Definition

In this phase we aim at merging the knowledge and the data layers into a single structure.

6.1 Entity matching

Real-world entities, identified by their values, can be represented through various properties and property values across different datasets. This challenge, known as the 'entity matching problem', has two primary implications: the need to identify the appropriate set of properties across datasets where the same entity may have multiple representations, and the need to resolve discrepancies in property values when multiple representations share the same properties but have differing values.

In our specific case, most of our datasets contained the only instances of a specific entity, with one important exception. The Alumni and Member data consists of four intersecting sets representing real people through various combinations of attributes. To solve this problem, we firstly observed our data. Most fields were overlapping and their values were matching. This was consistent with the modeling we did on our ontology, which identifies Member and Alumni as sub-entities of the Person Etype. We decided to separate the redundant data into a single dataset, while mantaining non-overlapping fields into separate datasets. To be able to correctly identify the correspondance between these two different kinds of data, this operation has been performed together with the Entity Identification process.

6.2 Entity identification

The problem of Entity identification is that of finding or creating appropriate identifiers for the entities in our datasets, which is a fundamental step when constructing a functional database.

Our first step in this process was to try to find appropriate attributes to use as URIs for each Etype. For most of our data, this seemed like the optimal solution, with two exceptions:

- The Person, Member and Alumni data, which, as mentioned before, now consists of non-overlapping data. Our first option to identify this data was to use the name and surname of each individual as an identifier, but after careful consideration, we opted to construct unique identifiers in order to avoid the homonimity problem.
- The Course dataset, which in real life scenarios is often referred not by its full name, but by an acronym. We have then enriched our data with the appropriate acronyms and chose them as our URN.

6.3 Data mapping

Our work in the previous phases and sub-phases has provided with all the pieces that we needed to finally merge the Knowledge and Data layers of our Knowledge Graph. To achieve this result, we used the Karma tool to manually match our datasets with our teleontology.

This operation, which initially sounded trivial, came with its own problem to solve. Although Karma has proven to be a rather intuitive tool, we found ourselves in need of extending some of the work we did in previous phases to accommodate the mapping operation.

We have to add a column in the Karma tool in the alumni and member data that we have used as a duplicate of the identifier that matched with the dataset Person. This was done to be able to assign a uri to the class Person and the class Member and Alumni during the data mapping of the file member.json and alumni.json and avoid a blank class during the visualization of the Knowledge Graph in GraphDB.

Moreover, our previous modeling assumed some degree of freedom when modeling the data regarding Locations, but all the column merging functions provided by Karma were insufficient to reach a satisfactory result and the complete elimination of such data from our Knowledge Graph seemed too drastic as a measure.

We decided to backtrack to the Language Definition and the Knowledge Definition phases to extend our work to better model such data. Specifically, the following concepts had to be added to the work we did on Language Definition:

- Location (ConceptID:132*), which technically already existed as a concept, but is now defined as an Etype.
- Info (ConceptID:KGE24-13-0011), which is a new Etype. Its function is to mantain all the additional data we have gathered through OSM about our locations.
- Latitude (ConceptID:46263*)
- Longitude (ConceptID:46270*)
- Timestamp (ConceptID:KGE24-13-0012)
- Value ((ConceptID:KGE24-13-0013)

* ConceptIDs marked with this smbol are part of http://num.edu.mn/ontology .

This new concepts have then been integrated into our teleontology using the Protégé tool. We also defined two new Object properties to model the relations between our new Etypes.

- has_info, for which we were unable to find a correspective in existing ontologies.
- has_location, which was provided by http://num.edu.mn/ontology

Using this extended ontology, we performed the mapping between our Knowledge and Data layers. Next we provide some images of some of the mappings performed using Karma

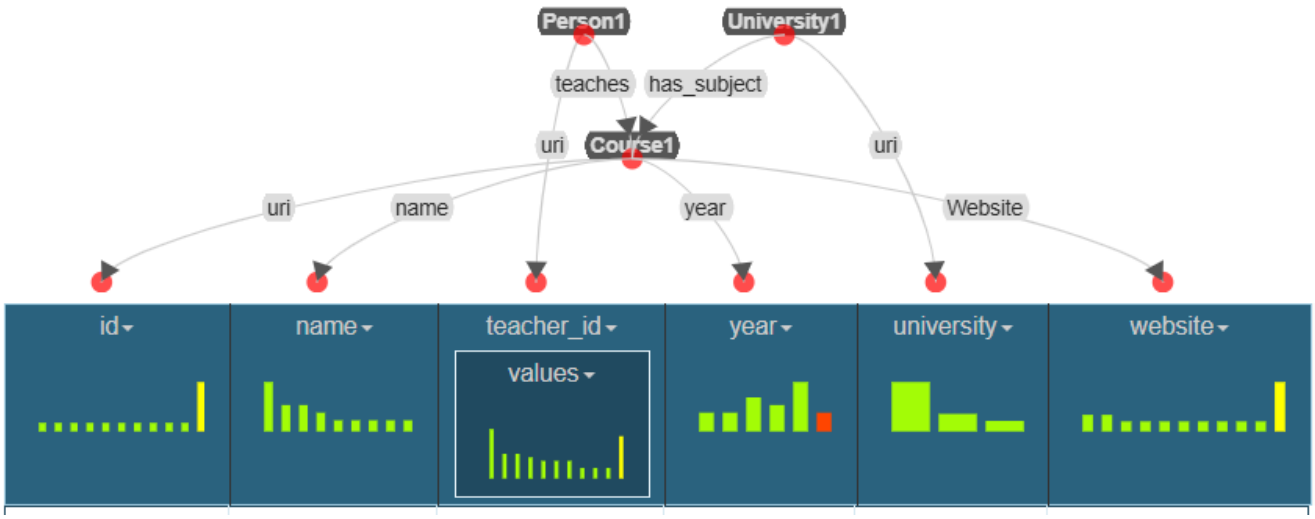


Figure 8: Mapping of Course using Karma

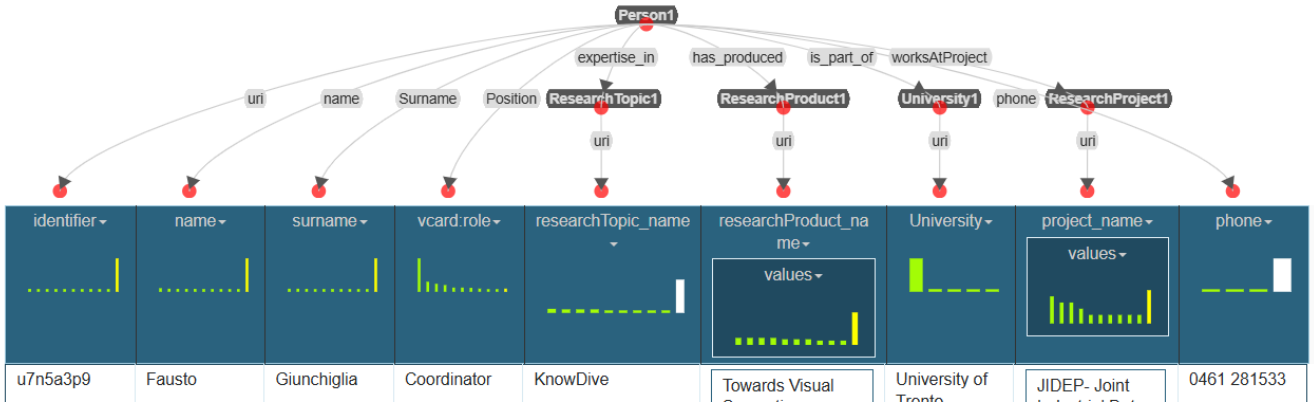


Figure 9: Mapping of Person using Karma

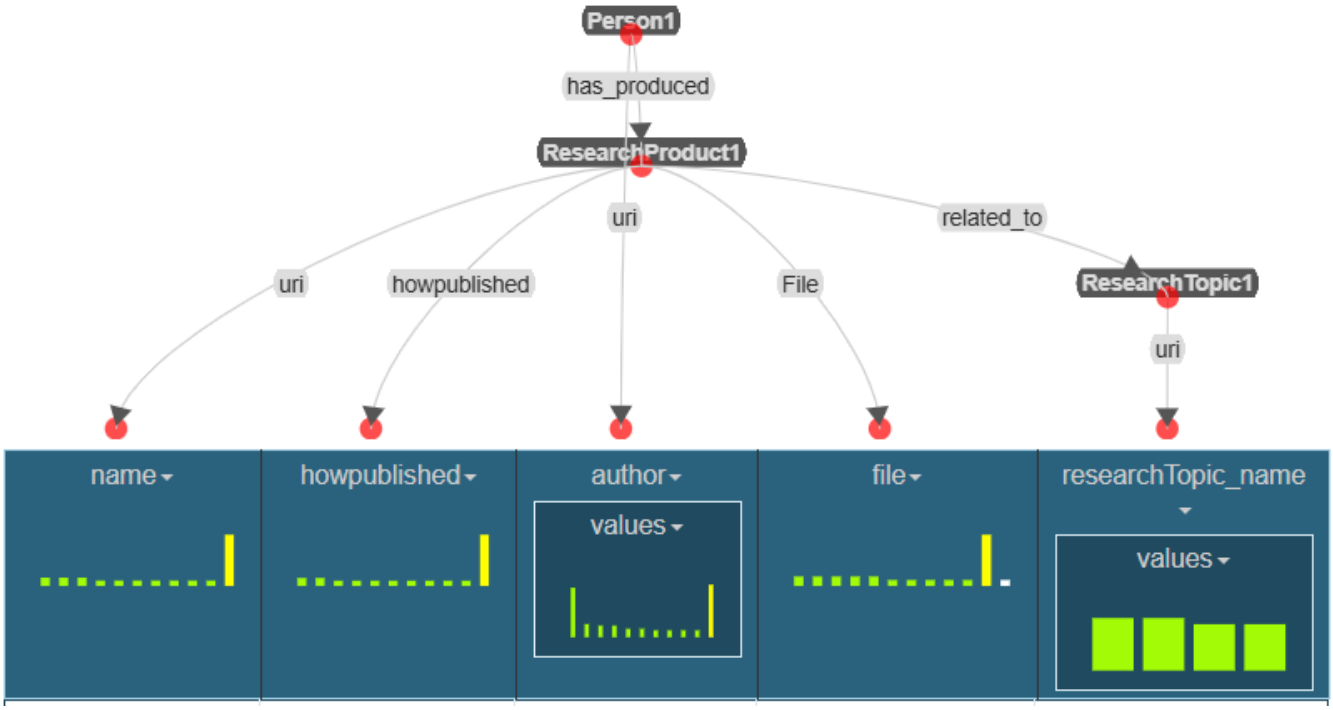


Figure 10: Mapping of Product using Karma

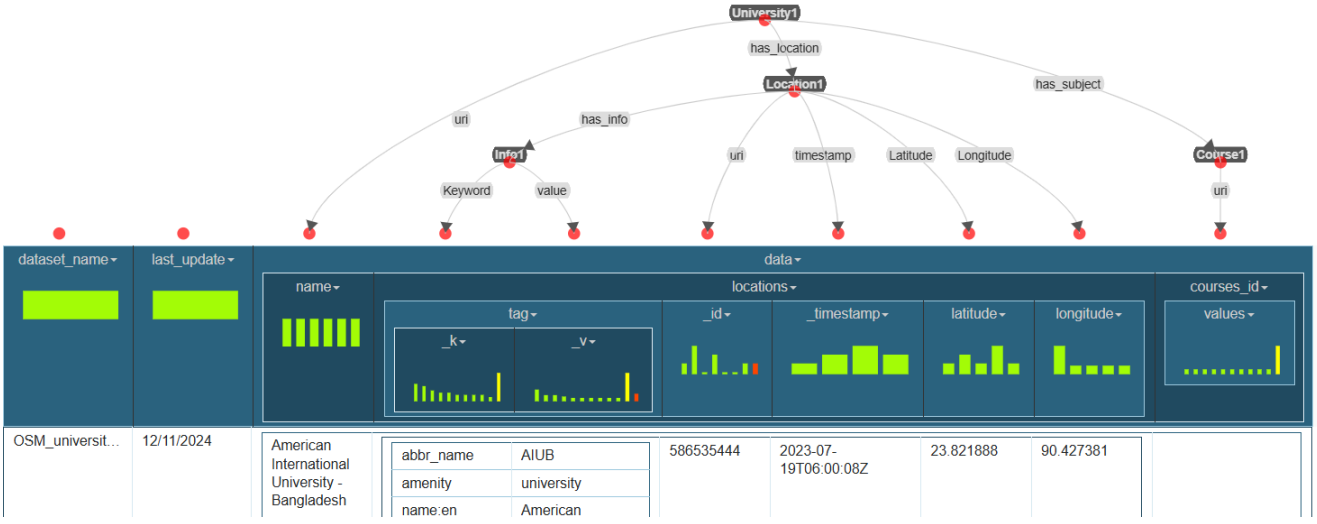


Figure 11: Mapping of University using Karma

7 Evaluation

This section aims to provide some information statistics about the final Knowledge Graph, evaluating the layer and the execution fo the query .

Evaluating a Knowledge Graph (KG) is essential to ensure its quality, accuracy, and effectiveness in representing structured knowledge. The evaluation process not only assesses the final KG but also verifies the entire construction process. The iTelos methodology incorporates evaluation activities at the end of each phase to determine whether the intermediate outputs are suitable for the next step or require revision. It defines criteria for assessing both primary and secondary objectives. The primary objective is purpose satisfaction, which measures how well the KG answers Competency Questions (CQs) by evaluating the teleontology and data connectivity. The secondary objective is reusability, which assesses how effectively the KG aligns with reference ontologies. By systematically evaluating these aspects, KG construction ensures reliability, completeness, and long-term usability.

7.1 Knowledge Graph information statistics

The EType in the graph are 10:

EType	Instances
Person	248
Alumni	181
Info	44
Member	38
Research Product	27
Course	23
Research Project	15
Research Topic	11
Location	7
University	6

Figure 12: Class hierarchy in GraphDB

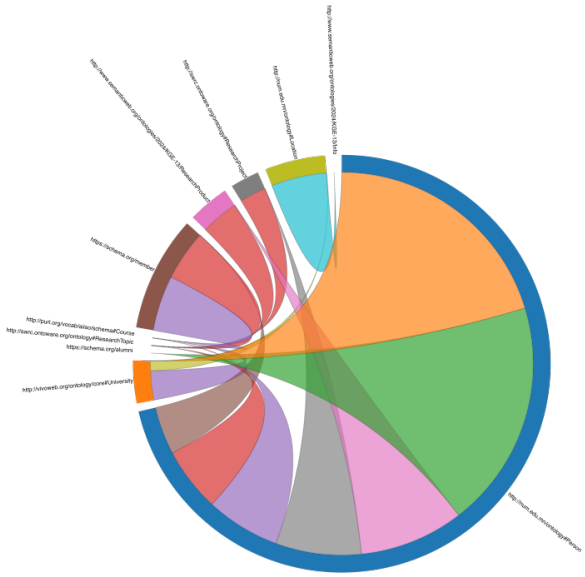


Figure 13: Class relationship in GraphDB

7.2 Knowledge layer evaluation

The final phase of the data integration process is the evaluation of the KG. In particular, four metrics are considered:

- EType coverage with respect to the Competency Questions
- Property coverage with respect to the Competency Questions
- EType coverage with respect to the Reference Ontologies
- Property coverage with respect to the Reference Ontologies

These four coverage metrics assess the extent to which a knowledge graph (KG) represents a specific portion of knowledge, structured as ETypes and properties. Specifically, the first two metrics evaluate how well the ETG covers the entities and properties derived from the Competency Questions (CQs), while the last two assess the ETG's coverage of ETypes and properties extracted from the reference ontologies. Additionally, a less formal but still meaningful metric is included: the number of Competency Questions successfully answered out of the total.

7.2.1 EType coverage with respect to the Competency Questions

EType coverage (CovE) measures how well the teleontology (T) represents the entity types (ETypes) extracted from a set of Competency Questions (CQs). It is calculated as the ratio of the intersection between the ETypes from CQs (CQE) and the ETypes in the teleontology (TE) to the total number of CQE. A higher CovE value indicates better alignment between the teleontology and the knowledge required by the CQs. The formula ensures that the teleontology effectively covers the expected entity types for accurate knowledge representation.

$\text{Cov}_E\left(CQ_E\right) = \frac{|CQ_E \cap T_E|}{CQ_E} = \frac{|8 \cap 10|}{8} = 1$

7.2.2 Property coverage with respect to the Competency Questions

Property coverage (CovP) evaluates how well the teleontology (T) represents the properties extracted from a set of Competency Questions (CQs). It is calculated as the ratio of the intersection between the properties from CQs (CQp) and the properties in the teleontology (Tp) to the total number of CQp.

$\text{Cov}_p\left(CQ_p\right) = \frac{|CQ_p \cap T_p|}{CQ_p}$

Since no properties were extracted from the Competency Questions, it was not possible to evaluate this metric.

7.2.3 EType coverage with respect to the Reference Ontologies

EType coverage (CovE) at the reference ontology (RO) level measures how well the teleontology (T) aligns with the entity types (ETypes) extracted from reference ontologies (ROs).

$\text{Cov}_E\left(RO_E\right) = \frac{|RO_E \cap T_E|}{RO_E} == \frac{|8 \cap 10|}{8} = 1$

7.2.4 Property coverage with respect to the Reference Ontologies

Property coverage (CovP) at the reference ontology (RO) level evaluates how well the teleontology (T) aligns with the properties extracted from reference ontologies (ROs).

$\text{Cov}_E\left(RO_E\right) = \frac{|RO_E \cap T_E|}{RO_E} = \frac{|28 \cap 34|}{28} = 1$

7.2.5 CQ coverage

$\text{Cov}(CQ) = \frac{|CQ_A| - |CQ_D|}{|CQ|} = \frac{26-0}{35} = 0,74$

In this equation, |CQA| is the number of CQ answered, |CQD| is the number of CQs which produced duplicated results and |CQ| is the total number of CQs. . There are no applicable duplicated queries.

The queries that are not doable with the current data are 1.1, 1.2, 1.3, 5.2, 5.3, 6.1, 6.2, 6.3, 7.4.

For the 1.* CQ the main problem is the missing information about the ResearchProduct of the type of product; we can't know if it a thesis or something else.

For the 5.* CQ the missing information is about the research topic the alumni were studying.

For the 6.* and 7.* CQ the setback was the missing of the year when a person began to be part of the KnowDive group.

The query 4.4 is theoretical executable but we were not able to implement it due to limited knowledge regarding the SPARQL language.

7.3 Data layer evaluation

Evaluating the data layer of a Knowledge Graph is essential to understanding its connectivity and structural coherence. The iTelos methodology assesses how "dense" or "connected" the KG is both at the end of the construction process and during its development. Connectivity is evaluated at two stages: first, on the final KG to determine its overall interconnectedness, and second, throughout the construction process to measure how each dataset contributes to improving connectivity. The impact of a dataset on connectivity can vary depending on whether it is evaluated in isolation or within the evolving KG, as entity matching conflicts and their resolutions affect the final structure. The evaluation focuses on two key dimensions: entity connectivity, which measures how well entities are linked to each other, and property connectivity, which assesses the relationships between entities and their associated properties. These aspects ensures that the KG is both well-structured and effectively represents the intended knowledge domain.

The Entity and Property Connectivity can be calculated by using the following connectivity matrix.

	Person	Alumni	Info	Member	Research Product	Course	Research Project	Research Topic	Location	University
Person	9	1		1	1		2	1		1
Alumni	1	2								
Info										
Member	1			6		1		1		
Research Product	1				5			1		
Course				1		6				1
Research Project	2						9	1		
Research Topic	1			1	1		1	5		
Location			1							1
University	1					1			1	6

Table 9: Connectivity matrix

7.3.1 Entity connectivity

(X, Y) is a cell in the connectivity matrix, and OP(X) is the number of object properties of the ETypes X.

$$EC(X) = \frac{\sum_{Y=1}^{N}(X,Y)}{OP(X)}$$

$$EC(KG) = \sum_{X=1}^{N} EC(X) = 9$$

7.3.2 Property connectivity

(X, Y) is a cell in the connectivity matrix, and DP(X) is the number of data properties of the ETypes X.

$$PC(X) = \frac{(X,X)}{DP(X)}$$

$$PC(KG) = \sum_{X=1}^{N} PC(X) = 8$$

7.4 Query execution

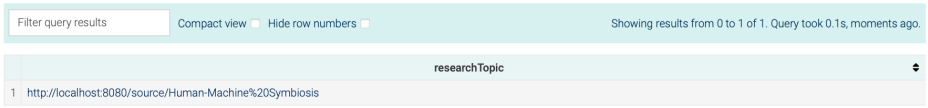
Finally, we carried out the testing phase of the project to verify that our final knowledge graph fully met its intended purpose. To achieve this, we executed a set of SPARQL queries, each corresponding to a Competency Question defined during the purpose formalization phase. The knowledge graph was stored and queried using GraphDB, a graph database that leverages graph structures—comprising nodes, edges, and properties—to facilitate semantic queries and efficiently represent and store data.

7.4.1 Query 1.4

Execution time: 0.1s

```
PREFIX general: <http://swrc.ontoware.org/ontology#>
PREFIX person: <http://num.edu.mn/ontology#>


SELECT ?researchTopic
WHERE {
  ?person person:Surname "Angeli" ;
          general:name "Alessia" ;
          person:expertise_in ?researchTopic .
}
```

In this query the person chosen is Alessia Angeli:



7.4.2 Query 1.5

Execution time: 0.1s

```
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX ds: <http://www.knowdive.disi.unitn.it/etype#>
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT ?name ?surname
WHERE {
  ?x person:expertise_in <http://localhost:8080/source/Artificial%20Human%20Cognition> ;
    general:name ?name;
    person:Surname ?surname.

  ?x ds:Position "External Member".
  ?x sw:is_part_of <http://localhost:8080/source/University%20of%20Minho>
}
```

In this query, the research topic chosen is Artificial Human Cognition, the position is External member and is part of the University of Minho:

7.4.3 Query 2.1

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT ?name ?surname
WHERE {

   ?member sw:status_member ?Course;
         general:name ?name ;
         person:Surname ?surname .
}
```

This query returns:

	name	surname
1	"Fausto"	"Giunchiglia"
2	"Adriano"	"Tavares"
3	"Ivano"	"Bison"
4	"Hao"	"Xu"
5	"Khandaker Tabin"	"Hasan"
6	"Gabor"	"Bella"
7	"Alessio"	"Zamboni"
8	"Yamini"	"Chandrashekar"
9	"Rui"	"Zhang"
10	"Andrea"	"Bontempelli"
11	"Mayukh"	"Bagchi"
12	"Simone"	"Bocca"
13	"Marco"	"Barbieri"
14	"Matteo"	"Busso"
15	"Xiaoyue"	"Li"
16	"Kayongo"	"Ivan"
17	"Xiaolei"	"Diao"
18	"Shandy"	"Darma"
19	"Haonan"	"Zhao"
20	"Ali"	"Hamza"

7.4.4 Query 2.2

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>
PREFIX ds: <http://www.knowdive.disi.unitn.it/etype#>

SELECT ?name ?surname ?position
WHERE {
  ?alumni sw:status_alumni ?value;
          general:name ?name ;
          person:Surname ?surname ;
          ds:Position ?position .
}
```

This query returns:

	name	surname	position
1	"Hao"	"Xu"	"PhD Student"
2	"Gabor"	"Bella"	"Research Fellow"
3	"Rui"	"Duro"	"Master Student"
4	"JoÃƒÂ£o Manuel"	"GonÃƒÂ§alves Silva"	"Master Student"
5	"Pavel"	"Shvaiko"	"Research Fellow"
6	"Fabio"	"Corubolo"	"Research Fellow"
7	"Biswanath"	"Dutta"	"Research Fellow"
8	"Elena"	"Pavan"	"Research Fellow"
9	"Gaia"	"Trecarchi"	"Research Fellow"
10	"Maria"	"Menendez"	"Research Fellow"
11	"Juan"	"Pane"	"Research Fellow"
12	"Aliaksandr"	"Autayeu"	"Research Fellow"
13	"Uladzimir"	"Kharkevich"	"Research Fellow"
14	"Hui"	"Yang"	"Research Fellow"
15	"Pil Ho"	"Kim"	"Research Fellow"
16	"Mladjan"	"Jovanovic"	"Research Fellow"
17	"Stefano"	"Teso"	"Research Fellow"
18	"Mattia"	"Fumagalli"	"Research Fellow"
19	"Mattia"	"Zeni"	"Research Fellow"
20	"Kyriakos"	"Kyriakou"	"Research Fellow"
21	"Alexander"	"Ivanyukovich"	"PhD Student"
22	"Denys"	"Babenko"	"PhD Student"
23	"Md Ahsan-ul"	"Morshed"	"PhD Student"

7.4.5 Query 2.3

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX ds: <http://www.knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>

SELECT ?name ?surname
WHERE {
   ?member sw:status_member ?Course;
         general:name ?name ;
         person:Surname ?surname ;
         ds:Position ?position .
  FILTER (CONTAINS(LCASE(?position), "researcher") || CONTAINS(LCASE(?position), "professor"))
}
```

This query returns:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 3 of 3. Query took 0.1s, moments ago.

	name	surname
1	"Ali"	"Hamza"
2	"Roy Alia"	"Asiku"
3	"Dragos"	"Gruia"

7.4.6 Query 2.4

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX general: <http://swrc.ontoware.org/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>

SELECT ?position ?researchProduct ?project
WHERE {
  ?person general:name "Mattia";
          person:Surname "Fumagalli" ;
          sw:Position ?position .

    OPTIONAL{?person general:worksAtProject ?project;}
    OPTIONAL{?person sw:has_produced ?researchProduct;}
}
```

In this query, the person chosen is Mattia Fumagalli:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 3 of 3. Query took 0.1s, moments ago.

	position	researchProduct	project
1	"Research Fellow"	http://localhost:8080/source/On%20Knowledge%20Diversity	
2	"Research Fellow"	http://localhost:8080/source/Teleologies-%20Objects%2C%20actions%20and%20functions.	
3	"Research Fellow"	http://localhost:8080/source/Concepts%20as%20(Recognition)%20Abilities	

7.4.7 Query 3.1

Execution time: 0.1s

```
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT DISTINCT ?topic
WHERE {
  ?topic a general:ResearchTopic ;
}
```

This query returns:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 11 of 11. Query took 0.1s, moments ago.

	topic
1	http://localhost:8080/source/Language%20Diversity
2	http://localhost:8080/source/KnowDive
3	http://localhost:8080/source/Artificial%20Human%20Cognition
4	http://localhost:8080/source/Human-Machine%20Symbiosis

7.4.8 Query 3.2

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT ?name ?surname
WHERE {

   ?member sw:status_member ?Course;
         general:name ?name ;
         person:Surname ?surname ;
         sw:Position ?position ;
        person:expertise_in <http://localhost:8080/source/Artificial%20Human%20Cognition> .


  # Filter only Researchers or Professors
  FILTER (CONTAINS(LCASE(?position), "researcher") || CONTAINS(LCASE(?position), "professor"))
}
```

In this query, the research topic chosen is Artificial Human Cognition:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 1 of 1. Query took 0.2s, moments ago.

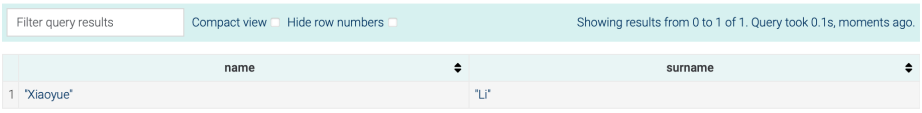
	name	surname
1	"Dragos"	"Gruia"

7.4.9 Query 3.3

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX so: <https://schema.org/>
PREFIX general: <http://swrc.ontoware.org/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>

SELECT ?name ?surname
WHERE {
    ?member a so:member;
        sw:support_role "Seminar Coordinator".

    ?member sw:status_member ?value;
        general:name ?name ;
        person:Surname ?surname .

}
```

This query returns:



Filter query results | Compact view | Hide row numbers | Showing results from 0 to 1 of 1. Query took 0.1s, moments ago.

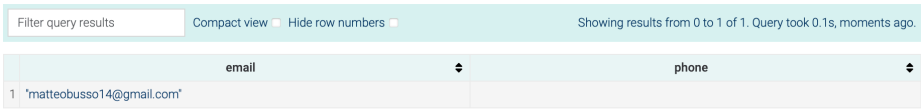
	name	surname
1	"Xiaoyue"	"Li"

7.4.10 Query 3.4

Execution time: 0.1s

```
PREFIX so: <https://schema.org/>
PREFIX general: <http://swrc.ontoware.org/ontology#>
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>

SELECT ?email ?phone
WHERE {
    ?member a so:member;
    OPTIONAL{?member person:Email ?email .}

    ?member sw:status_member ?value;
        general:name "Matteo" ;
        person:Surname "Busso" .
     OPTIONAL{?member general:phone ?phone .}
}
```

In this query, the person chosen is Matteo Busso:



7.4.11 Query 3.5

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT ?name ?surname
WHERE {
  ?person general:name "Gabor" ;
            person:Surname "Bella";
            person:expertise_in ?topic.
   ?person2 a person:Person;
        person:expertise_in ?topic;
        general:name ?name;
        person:Surname ?surname.


}
```

In this query, the person chosen is Gabor Bella:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 18 of 18. Query took 0.1s, moments ago.

	name	surname
1	"Hao"	"Xu"
2	"Gabor"	"Bella"
3	"Yamini"	"Chandrashekar"
4	"Shandy"	"Darma"
5	"Hadi"	"Khalilia"
6	"Kubra"	"Korkmaz"
7	"Virendra"	"Mehta"
8	"Patrik Calin"	"Maniu"
9	"Gabor"	"Bella"
10	"Hao"	"Xu"
11	"Gabor"	"Bella"
12	"Yamini"	"Chandrashekar"
13	"Shandy"	"Darma"
14	"Hadi"	"Khalilia"
15	"Kubra"	"Korkmaz"
16	"Virendra"	"Mehta"
17	"Patrik Calin"	"Maniu"
18	"Gabor"	"Bella"

7.4.12 Query 4.1

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT DISTINCT ?products
WHERE {

   ?products sw:related_to <http://localhost:8080/source/Language%20Diversity> .

}
```

In this query, the research topic chosen is Language Diversity:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 13 of 13. Query took 0.1s, moments ago.

	products
1	http://localhost:8080/source/Cyprus%20Center%20for%20Algorithmic%20Transparency%20(CyCAT)
2	http://localhost:8080/source/Language%20and%20Domain%20Aware%20Lightweight%20Ontology%20Matching
3	http://localhost:8080/source/Cross-Border%20Medical%20Research%20using%20Multi-Layered%20and%20Distributed%20Knowledge.
4	http://localhost:8080/source/One%20World%20-%20Seven%20Thousand%20Languages
5	http://localhost:8080/source/Lexical%20diversity%20in%20kinship%20across%20languages%20and%20dialects
6	http://localhost:8080/source/Understanding%20and%20Exploiting%20Language%20Diversity
7	http://localhost:8080/source/CogNet-%20a%20Large-Scale%20Cognate%20Database
8	http://localhost:8080/source/Domain-Based%20Sense%20Disambiguation%20on%20Multilingual%20Structured%20Data
9	http://localhost:8080/source/InteropEHRate-%20Health%20Records%20in%20People%27s%20Hands%20Across%20Europe
10	http://localhost:8080/source/SPRINT%20-%20Graph-Based%20Data%20Federation%20for%20Healthcare%20Data%20Science
11	http://localhost:8080/source/Safe%20Haven%201%20-%20Healthcare%20Data%20Safe%20Havens
12	http://localhost:8080/source/SHIB%20-%20Safe%20Haven%20in%20a%20Box
13	http://localhost:8080/source/Safe%20Haven%202%20-%20Digital%20Health%20Data%20Federation

7.4.13 Query 4.2

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT DISTINCT ?name ?surname
WHERE {
  ?person sw:has_produced
<http://localhost:8080/source/Improving%20time%20use%20measurement%20with%20personal%20big%20data%2
0collection%20-%20the%20experience%20of%20the%20European%20Big%20Data%20Hackathon%202019.>;
    general:name ?name;
    person:Surname ?surname.

}
```

In this query, the product chosen is ”Improving time measurement with personal big data collection”:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 3 of 3. Query took 0.3s, moments ago.

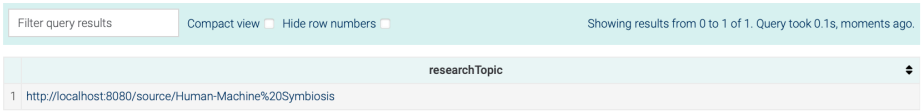
	name	surname
1	"Fausto"	"Giunchiglia"
2	"Ivano"	"Bison"
3	"Mattia"	"Zeni"

7.4.14 Query 4.3

Execution time: 0.1s

```
PREFIX kge: <http://knowdive.disi.unitn.it/etype#>
PREFIX project: <http://swrc.ontoware.org/ontology#>
PREFIX person: <http://num.edu.mn/ontology#>


SELECT ?researchTopic
WHERE {
  ?person project:name "Alessia" ;
          person:Surname "Angeli" ;
          person:expertise_in ?researchTopic .
}
```

In this query, the person chosen is Alessia Angeli:



7.4.15 Query 5.1

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT DISTINCT ?name ?surname
WHERE {
  ?member sw:status_member ?Course;
          general:name ?name ;
          person:Surname ?surname ;
          sw:is_part_of <http://localhost:8080/source/University%20of%20Jilin>.

}
```

In this query the university chosen is the University of Jilin:

7.4.16 Query 5.4

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT DISTINCT ?name ?surname ?university ?where
WHERE {
  ?alumni sw:status_alumni ?value ;
          general:name ?name ;
          person:Surname ?surname ;
          sw:is_part_of ?university .
  ?university person:has_location ?where
}
```

This query returns:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 156 of 156. Query took 0.1s, moments ago.

	name	surname	university	where
1	"Hao"	"Xu"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
2	"Gabor"	"Bella"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
3	"Rui"	"Duro"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
4	"JoÃƒÂ£o Manuel"	"GonÃƒÂ§alves Silva"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
5	"Pavel"	"Shvaiko"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
6	"Fabio"	"Corubolo"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
7	"Biswanath"	"Dutta"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
8	"Elena"	"Pavan"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
9	"Gaia"	"Trecarchi"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
10	"Maria"	"Menendez"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
11	"Juan"	"Pane"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
12	"Aliaksandr"	"Autayeu"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948
13	"Uladzimir"	"Kharkevich"	http://localhost:8080/source/University%20of%20Trento	http://localhost:8080/source/2776349948

7.4.17 Query 5.5

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX teach: <http://purl.org/vocab/aiiso/schema#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT DISTINCT ?name ?surname ?courseName ?year
WHERE {
  ?person a person:Person ;
          general:name ?name ;
          person:Surname ?surname ;
          teach:teaches ?course .

  ?course general:name ?courseName ;
          general:year ?year .
}
ORDER BY ?surname ?name ?year
```

This query returns:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 46 of 46. Query took 0.1s, moments ago.

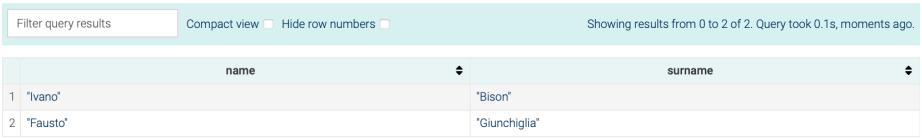
	name	surname	courseName	year
1	"Mayukh"	"Bagchi"	"Knowledge and Data Integration (KDI)"	"2020"
2	"Mayukh"	"Bagchi"	"Knowledge and Data Integration (KDI)"	"2021"
3	"Mayukh"	"Bagchi"	"Knowledge Graph Engineering (KGE)"	"2022"
4	"Mayukh"	"Bagchi"	"Knowledge Graph Engineering (KGE)"	"2023"
5	"Mayukh"	"Bagchi"	"Knowledge Graph Engineering (KGE)"	"2024"
6	"Ivano"	"Bison"	"Studies on Human Behaviour (SHB)"	"2020"
7	"Ivano"	"Bison"	"Studies on Human Behaviour (SHB)"	"2021"
8	"Ivano"	"Bison"	"Studies on Human Behaviour (SHB)"	"2022"
9	"Ivano"	"Bison"	"Studies on Human Behaviour: procuring the data (SHB)"	"2023"
10	"Ivano"	"Bison"	"Studies on Human Behaviour: procuring the data (SHB)"	"2024"
11	"Ivano"	"Bison"	"Data engineering for the quantified self (SHB)"	"2024"
12	"Simone"	"Bocca"	"Knowledge and Data Integration (KDI)"	"2020"
13	"Simone"	"Bocca"	"Knowledge and Data Integration (KDI)"	"2021"
14	"Simone"	"Bocca"	"Knowledge Graph Engineering (KGE)"	"2022"
15	"Simone"	"Bocca"	"Knowledge Graph Engineering (KGE)"	"2023"
16	"Simone"	"Bocca"	"Knowledge Graph Engineering (KGE)"	"2024"
17	"Simone"	"Bocca"	"Knowledge Graph Engineering (ICSI500)"	"2024"

7.4.18 Query 5.6

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX teach: <http://purl.org/vocab/aiiso/schema#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT DISTINCT ?name ?surname
WHERE {
  ?person a person:Person ;
          general:name ?name ;
          person:Surname ?surname ;
          teach:teaches <http://localhost:8080/source/SHB_2021> .
}
ORDER BY ?surname ?name
```

In this query, the course chosen is the course of Studies on Human Behaviour in 2021:



	name	surname
1	"Ivano"	"Bison"
2	"Fausto"	"Giunchiglia"

7.4.19 Query 6.4

Execution time: 0.1s

```
PREFIX file: <https://trdf.sourceforge.net/provenance/ns.html#>
PREFIX sw: <http://www.knowdive.disi.unitn.it/etype#>
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT ?product ?pub ?person ?name ?surname ?file
WHERE {
   ?product sw:related_to <http://localhost:8080/source/Artificial%20Human%20Cognition>;
        general:howpublished ?pub;
        file:File ?file.

   ?person sw:has_produced ?product;
    OPTIONAL{?person general:name ?name;}
    OPTIONAL{?person person:Surname ?surname.}

}ORDER BY ?product
```

In this query, the research topic chosen is Artificial Human Cognition:

	product	pub	person	name	surname	file
1	http://localhost:8080/source/Artificial%20Intelligence%20in%202027	"IJCAI, AI Matters, vol. 4, 2017"	http://localhost:8080/source/u7n5a3p9	"Fausto"	"Giunchiglia"	"https://drive.google.com/file/d/1X9RTD2uK1uWC6Ub0Y14YXIIdnD-GGs3c/view"
2	http://localhost:8080/source/Artificial%20Intelligence%20in%202027	"IJCAI, AI Matters, vol. 4, 2017"	http://localhost:8080/source/Maria%20Gini			"https://drive.google.com/file/d/1X9RTD2uK1uWC6Ub0Y14YXIIdnD-GGs3c/view"
3	http://localhost:8080/source/Artificial%20Intelligence%20in%202027	"IJCAI, AI Matters, vol. 4, 2017"	http://localhost:8080/source/Noa%20Agmon			"https://drive.google.com/file/d/1X9RTD2uK1uWC6Ub0Y14YXIIdnD-GGs3c/view"
4	http://localhost:8080/source/Artificial%20Intelligence%20in%202027	"IJCAI, AI Matters, vol. 4, 2017"	http://localhost:8080/source/Kevin%20Leyton%27Brown			"https://drive.google.com/file/d/1X9RTD2uK1uWC6Ub0Y14YXIIdnD-GGs3c/view"
5	http://localhost:8080/source/Artificial%20Intelligence%20in%202027	"IJCAI, AI Matters, vol. 4, 2017"	http://localhost:8080/source/Sven%20Koenig			"https://drive.google.com/file/d/1X9RTD2uK1uWC6Ub0Y14YXIIdnD-GGs3c/view"
6	http://localhost:8080/source/Computational%20Humanism	"On Line presentation at DISI, 2017. Trento, Italy"	http://localhost:8080/source/u7n5a3p9	"Fausto"	"Giunchiglia"	"https://drive.google.com/file/d/1rVUC6CSTVq9KowmrqfPxUAVg9h9AP9yq/view"
7	http://localhost:8080/source/Concepts%20as%20(Recognition)%20Abilities	"In Formal Ontology in Information System: Proceedings of the 9th International Conference (FOIS 2016), vol. 283, p. 153. IOS Press,2016"	http://localhost:8080/source/u7n5a3p9	"Fausto"	"Giunchiglia"	"https://drive.google.com/file/d/1XCUXEUVG_m0lY2hevCKOdHYIvDnkBpf8/view"
8	http://localhost:8080/source/Concepts%20as%20(Recognition)%20Abilities	"In Formal Ontology in Information System: Proceedings of the 9th International Conference (FOIS 2016), vol. 283, p. 153. IOS Press,2016"	http://localhost:8080/source/ID3n4o5p6	"Mattia"	"Fumagalli"	"https://drive.google.com/file/d/1XCUXEUVG_m0lY2hevCKOdHYIvDnkBpf8/view"

7.4.20 Query 6.5

Execution time: 0.1s

```
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT ?project ?startDate ?endDate
WHERE {
    ?project a general:ResearchProject;
         general:startDate ?startDate;
         general:endDate ?endDate.


}
```

This query returns:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 15 of 15. Query took 0.1s, moments ago.

http://localhost:8080/source/JIDEP-%20Joint%20Industrial%20Data%20Exchange%20Platform

		startDate	endDate
1	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"January 2013"	"December 2016"
2	http://localhost:8080/source/Wenet%20-%20The%20Internet%20of%20Us	"January 2019"	"December 2022"
3	http://localhost:8080/source/Cyprus%20Center%20for%20Algorithmic%20Transparency%20(CyCAT)	"November 2018"	"(Ongoing)"
4	http://localhost:8080/source/WhiteRabbit%20-%20Private%20Data%20Manager	"April 2017"	"July 2018"
5	http://localhost:8080/source/QROWD%20-%20Because%20Big%20Data%20Integration%20is%20Humanly%20Possible	"December 2016"	"November 2019"
6	http://localhost:8080/source/JIDEP-%20Joint%20Industrial%20Data%20Exchange%20Platform	"1 June 2022"	"31 May 2025"
7	http://localhost:8080/source/Digital%20University	"January 2014"	"(Ongoing)"
8	http://localhost:8080/source/InteropEHRate-%20Health%20Records%20in%20People%27s%20Hands%20Across%20Europe	"1 January 2019"	"30 June 2022"
9	http://localhost:8080/source/SPRINT%20-%20Graph-Based%20Data%20Federation%20for%20Healthcare%20Data%20Science	"April 2019"	"December 2019"
10	http://localhost:8080/source/Delphi%20â€"%20Discovering%20Life%20Patterns	"01/09/2021"	"31/08/2023"
11	http://localhost:8080/source/ISTAT%20-%20SmartSurveys	"June 2021"	"June 2026"
12	http://localhost:8080/source/Eurostat	"December 2018"	"April 2019"

7.4.21 Query 6.6

Execution time: 0.1s

```
PREFIX general: <http://swrc.ontoware.org/ontology#>
prefix db: <https://dbpedia.org/ontology/>

SELECT ?project ?partner
WHERE {
    ?project a general:ResearchProject;
         db:partner ?partner.

}
```

This query returns:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 95 of 95. Query took 0.1s, moments ago.

	project	partner
1	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"BEN-GURION UNIVERSITY OF THE NEGEV"
2	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"Austria"
3	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"KARLSTADS UNIVERSITET"
4	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"U-HOPPER SRL"
5	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"United Kingdom"
6	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"THE UNIVERSITY OF EDINBURGH"
7	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"Sweden"
8	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"MASTERS AND SCHOLARS OF THE UNIVERSITY OF OXFORD"
9	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	" TECHNISCHE UNIVERSITAET WIEN"
10	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"IMAGINARY SRL"
11	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"Israel"
12	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"Italy"
13	http://localhost:8080/source/SmartSociety%20-%20Hybrid%20and%20Diversity-Aware%20Collective%20Adaptive%20Systems	"UNIVERSITY OF SOUTHAMPTON"

7.4.22 Query 7.1

Execution time: 0.1s

```
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>

SELECT DISTINCT ?products
WHERE {
   ?products a sw:ResearchProduct .

}
```

This query returns:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 27 of 27. Query took 0.1s, moments ago.

http://localhost:8080/source/One%20World%20-%20Seven%20Thousand%20Languages

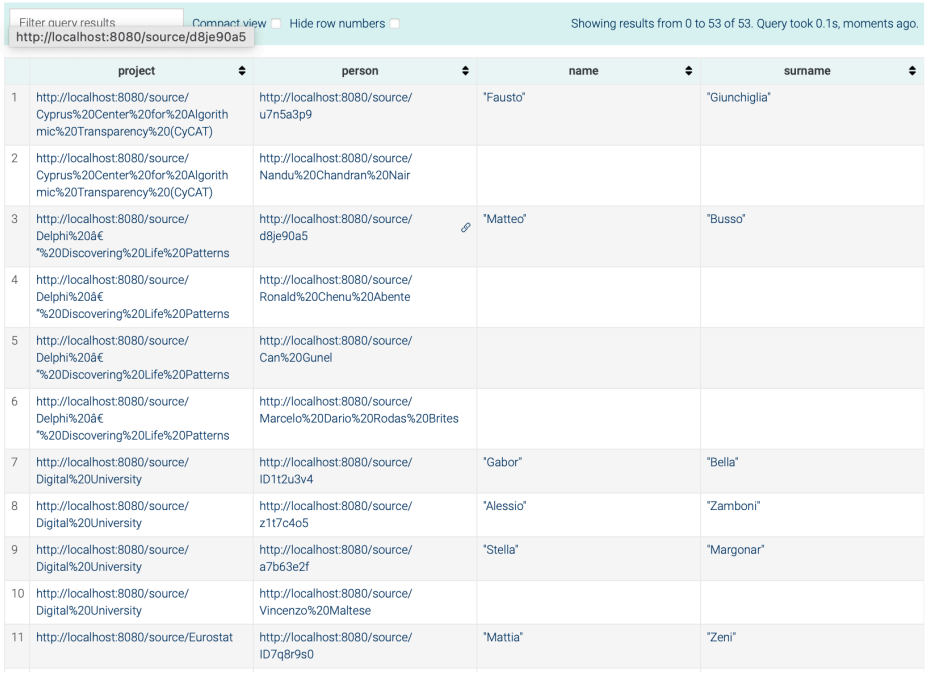
	products
1	http://localhost:8080/source/Language%20and%20Domain%20Aware%20Lightweight%20Ontology%20Matching
2	http://localhost:8080/source/%20Managing%20diversity%20in%20Knowledge
3	http://localhost:8080/source/Cross-Border%20Medical%20Research%20using%20Multi-Layered%20and%20Distributed%20Knowledge.
4	http://localhost:8080/source/One%20World%20-%20Seven%20Thousand%20Languages
5	http://localhost:8080/source/The%20future%20of%20AI
6	http://localhost:8080/source/Computational%20Humanism
7	http://localhost:8080/source/Towards%20Visual%20Semantics
8	http://localhost:8080/source/Lexical%20diversity%20in%20kinship%20across%20languages%20and%20dialects
9	http://localhost:8080/source/On%20Knowledge%20Diversity
10	http://localhost:8080/source/Understanding%20and%20Exploiting%20Language%20Diversity
11	http://localhost:8080/source/Teleologies-%20Objects%2C%20actions%20and%20functions.
12	http://localhost:8080/source/Mobile%20social%20media%20usage%20and%20academic%20performance
13	http://localhost:8080/source/CogNet-%20a%20Large-Scale%20Cognate%20Database
14	http://localhost:8080/source/Domains%20and%20context-%20first%20steps%20towards%20managing%20diversity%20in%20knowledge
15	http://localhost:8080/source/Personal%20context%20modelling%20and%20annotation
16	http://localhost:8080/source/Multi-device%20activity%20logging
17	http://localhost:8080/source/The%20open%20platform%20for%20personal%20lifelogging-%20the%20elifelog%20architecture
18	http://localhost:8080/source/Improving%20time%20use%20measurement%20with%20personal%20big%20data%20collection%20-%20the%20experience%20of%20the%20European%20Big%20Data%20Hackathon%202019.
19	http://localhost:8080/source/Continual%20egocentric%20object%20recognition
20	http://localhost:8080/source/Concepts%20as%20(Recognition)%20Abilities

7.4.23 Query 7.2

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT ?project  ?person ?name ?surname
WHERE {
    ?person general:worksAtProject ?project;
    OPTIONAL{?person general:name ?name ;}
    OPTIONAL{?person person:Surname ?surname ;}

} ORDER BY ?project
```

This query returns:



Filter query results Compact view Hide row numbers Showing results from 0 to 53 of 53. Query took 0.1s, moments ago.

http://localhost:8080/source/d8je90a5

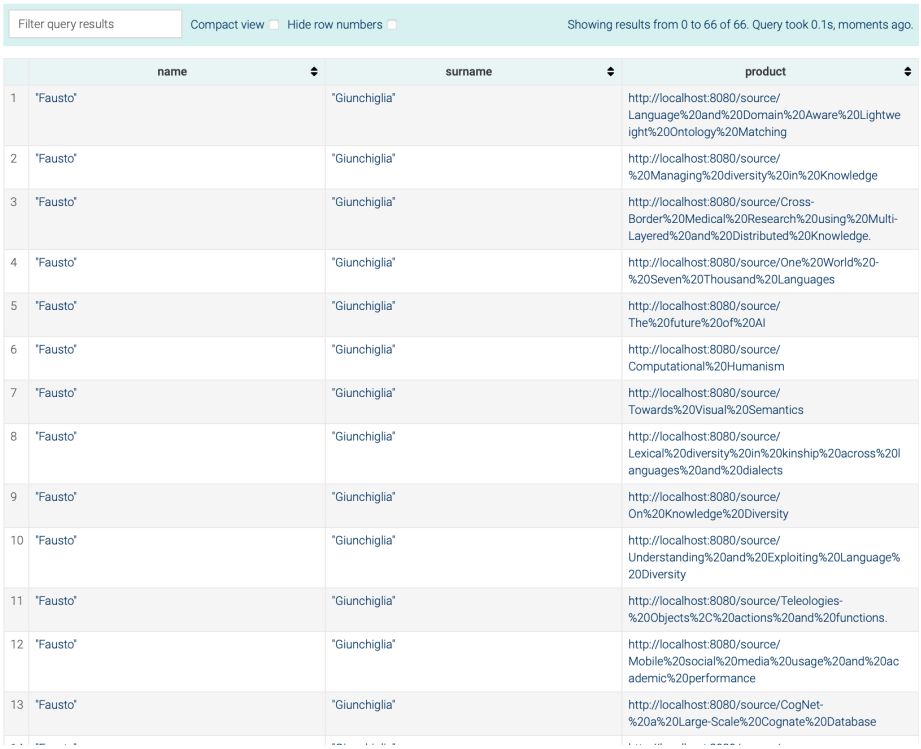
	project	person	name	surname
1	http://localhost:8080/source/Cyprus%20Center%20for%20Algorithmic%20Transparency%20(CyCAT)	http://localhost:8080/source/u7n5a3p9	"Fausto"	"Giunchiglia"
2	http://localhost:8080/source/Cyprus%20Center%20for%20Algorithmic%20Transparency%20(CyCAT)	http://localhost:8080/source/Nandu%20Chandran%20Nair		
3	http://localhost:8080/source/Delphi%20â€"%20Discovering%20Life%20Patterns	http://localhost:8080/source/d8je90a5	"Matteo"	"Busso"
4	http://localhost:8080/source/Delphi%20â€"%20Discovering%20Life%20Patterns	http://localhost:8080/source/Ronald%20Chenu%20Abente		
5	http://localhost:8080/source/Delphi%20â€"%20Discovering%20Life%20Patterns	http://localhost:8080/source/Can%20Gunel		
6	http://localhost:8080/source/Delphi%20â€"%20Discovering%20Life%20Patterns	http://localhost:8080/source/Marcelo%20Dario%20Rodas%20Brites		
7	http://localhost:8080/source/Digital%20University	http://localhost:8080/source/ID1t2u3v4	"Gabor"	"Bella"
8	http://localhost:8080/source/Digital%20University	http://localhost:8080/source/z1t7c4o5	"Alessio"	"Zamboni"
9	http://localhost:8080/source/Digital%20University	http://localhost:8080/source/a7b63e2f	"Stella"	"Margonar"
10	http://localhost:8080/source/Digital%20University	http://localhost:8080/source/Vincenzo%20Maltese		
11	http://localhost:8080/source/Eurostat	http://localhost:8080/source/ID7q8r9s0	"Mattia"	"Zeni"

7.4.24 Query 7.3

Execution time: 0.1s

```
PREFIX person: <http://num.edu.mn/ontology#>
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>
PREFIX general: <http://swrc.ontoware.org/ontology#>

SELECT ?name ?surname ?product
WHERE {

   ?member sw:status_member ?Course;
         general:name ?name ;
         person:Surname ?surname ;

    OPTIONAL{?member sw:has_produced ?product .}

}
```

This query returns:



Filter query results | Compact view | Hide row numbers | Showing results from 0 to 66 of 66. Query took 0.1s, moments ago.

	name	surname	product
1	"Fausto"	"Giunchiglia"	http://localhost:8080/source/Language%20and%20Domain%20Aware%20Lightweight%20Ontology%20Matching
2	"Fausto"	"Giunchiglia"	http://localhost:8080/source/%20Managing%20diversity%20in%20Knowledge
3	"Fausto"	"Giunchiglia"	http://localhost:8080/source/Cross-Border%20Medical%20Research%20using%20Multi-Layered%20and%20Distributed%20Knowledge.
4	"Fausto"	"Giunchiglia"	http://localhost:8080/source/One%20World%20-%20Seven%20Thousand%20Languages
5	"Fausto"	"Giunchiglia"	http://localhost:8080/source/The%20future%20of%20AI
6	"Fausto"	"Giunchiglia"	http://localhost:8080/source/Computational%20Humanism
7	"Fausto"	"Giunchiglia"	http://localhost:8080/source/Towards%20Visual%20Semantics
8	"Fausto"	"Giunchiglia"	http://localhost:8080/source/Lexical%20diversity%20in%20kinship%20across%20languages%20and%20dialects
9	"Fausto"	"Giunchiglia"	http://localhost:8080/source/On%20Knowledge%20Diversity
10	"Fausto"	"Giunchiglia"	http://localhost:8080/source/Understanding%20and%20Exploiting%20Language%20Diversity
11	"Fausto"	"Giunchiglia"	http://localhost:8080/source/Teleologies-%20Objects%2C%20actions%20and%20functions.
12	"Fausto"	"Giunchiglia"	http://localhost:8080/source/Mobile%20social%20media%20usage%20and%20academic%20performance
13	"Fausto"	"Giunchiglia"	http://localhost:8080/source/CogNet-%20a%20Large-Scale%20Cognate%20Database

7.4.25 Query 7.5

Execution time: 0.1s

```
PREFIX sw: <http://knowdive.disi.unitn.it/etype#>

SELECT ?researchTopic (COUNT(?product) AS ?productCount)
WHERE {
  ?product a sw:ResearchProduct ;
          sw:related_to ?researchTopic .
}
GROUP BY ?researchTopic
ORDER BY DESC(?productCount)
```

This query returns:

Filter query results | Compact view | Hide row numbers | Showing results from 0 to 4 of 4. Query took 0.2s, moments ago.

	researchTopic	productCount
1	http://localhost:8080/source/Knowledge%20Diversity	"8"^^xsd:integer
2	http://localhost:8080/source/Artificial%20Human%20Cognition	"8"^^xsd:integer
3	http://localhost:8080/source/Language%20Diversity	"7"^^xsd:integer
4	http://localhost:8080/source/Human-Machine%20Symbiosis	"7"^^xsd:integer

8 Metadata Definition

In this section the report collects the definitions of all the metadata defined for the different resources produced along the whole process. The metadata defined in this phase describes both the final outcome of the project, and the intermediate outcome of each phase, subdivided as appropriate through the various categories of metadata.

The 4 categories of metadata are: Project, Language, Knowledge and Data. We will make separate considerations regarding information contained in our metadata for each category.
For the sake of clarity, our metadata refers both to resources we collected and to resources we have produced during the project. In this second case, to avoid excessive redundancy, the metadata refers to the most recent version, which can be found in the Phase 5 - Data Definition folder of our repository.

8.1 Project Metadata

The project metadata can be found in the Project_metadata.xlsx file and provides relevant information regarding our project as a whole. Specifically, the Metadata describes the project through the following fields: Title, URL, Keywords, Type, Description, Start Date, End Date, Funding Agency, Input, Output, Coordinator, Observations. The URL and Funding Agency fields have been assigned a placeholder value. The Type fields has been set as ”Knowledge Graph Generation”, meanwhile the Input and Output fields describe the premise of the project for the KGE course: starting with an informal purpose and some suggested sources, produce a complete KG. The Observations field has been deemed necessary to convey information regarding the fact that this project is the work of students and is, by consequence, been used as a tool to learn. Other fields are used in an intuitive manner and need no explicit discussion.

8.2 Language Metadata

The language metadata is found in the Datasets_Metadata.xlsx file and conveys information regarding the various language resources that we initially collected and the final language resource produced through our work. Specifically, the metadata refers to the version of our language file that has been modified in the Entity definition phase.

The fields used in our metadata to describe our language resources are : Title, Description, Type, Size, FileFormat, Keywords, Source Info. The Description field contains a short description of the nature of the Language resource and to which data resources it is applicable. The Keywords field allowed us to define certain keywords to enable a faster search of our resources

in larger collections. The keyword Language specifically has been used to mark Language resources. The Source Info field contains information regarding both the source of the original files and the means used to aquire them. For example, both the original KRG-UNITN Language and our KGE24-13-UNITN Language are reported as sourceless and as a product of our project. Other fields are used in an intuitive manner and need no explicit discussion.

8.3 Knowledge Metadata

The knowledge metadata is found in the Datasets_Metadata.xlsx file and conveys information regarding the various knowledge resources that we initially collected and the final knowledge resource produced through our work. Specifically, the metadata refers to the version of our ontology that has been extended in the Entity definition phase. The fields used in our metadata to describe our language resources are : Title, Description, Type, Size, FileFormat, Keywords, Source Info. The Description field contains a short description that conveys in natural language certain critical information, such as the designed domain for the knowledge resource. The Keywords field contains terms that can be used to enable faster search of the resources through larger collection and the keyword Ontology has been used to mark our knowledge resources. The Source Info field contains information regarding both the source of the original files and the means used to aquire them. For example, the UNITN-DU ontology has been reported as having the source LiveData Unitn catalog, but has no information regarding the means of aquiring it, meanwhile our KGE24-13-ontology is reported without a source and as constructed during the project. Other fields are used in an intuitive manner and need no explicit discussion.

8.4 Data Metadata

The Data metadata is found in the Datasets_Metadata.xlsx file and conveys information regarding the various data resources that we initially collected and the final data resources produced through our work. The fields used in our metadata to describe our language resources are : Title, Description, Type, Size, FileFormat, Keywords, Source Info. The description field contains information regarding the nature of contained information and also specifies other properties, such as if the data has been obtained thorugh scraping or if the datasets is the product of processes of our project. The Keywords field consists of terms that can be used to search for our resources in larger collections. The Source Info field contains information regarding both the source of the original files and the means used to aquire them. For example, all KRG-UNITN sources are reported as scraped from the Knowdive Research Group website, meanwhile all the final data resources are reported without a source and as constrcuted during our project. Other fields are used in an intuitive manner and need no explicit discussion.

9 Open Issues

This section aims at explicitly list the remaining problems present in this project and to propose or otherwise discuss potential solutions.

For the sake of relevance, the problems related to conflicting schedules, other project and other typical issues that arise in a project between students will be omitted.

The following are the issues still open within our work:

- Lack of data: by reviewing the CQ coverage of our work, we realized how important data with high density of information is. In our case specifically, the lack of information on certain topics caused a significant loss in applicability in real case scenarios of our project. Trying to retrieve such information also seem unadvisable as it is unclear where to look for it and how much time it would take to construct more complete data resources.
- The late relevance of Metadata: the problem of metadata is problem that has come to our attention only in the later phases of our project and our work to produce them has been mainly one about backtracking our work in order to reconstruct the most complete metadata possible. Methodologically, we believe that intermediate Metadata creation steps at the end of each phase would have been a more functional solution and one that would have produced metadata richer in information.

10 Additional material

In this section we have collected the links and references to all additional materials related to our project.

Github: KGE-Project Repository
Webpage: KGE-Webpage