UNIVERSITY OF TRENTO

Trentino Urban Transportation KGE 2022

Diego Barquero Morera, 229577, diego.barqueromorera@studenti.unitn.it
Vahan Petrosyan, 229737, vahan.petrosyan@studenti.unitn.it

Inception

Purpose formalization

"A person currently in an urban area of Trentino region wants to easily move from one place to another by means of public transportation."

Domain of Interest

Public transportation of urban areas of the region of Trentino (Italy) over a period of time of 10 months (between September 2022 and June 2023).

3 Scenarios

Urban areas of Trentino during:

- Working days
- Weekend / national holidays
- Nighttime

5 Personas

With different age, occupation, place of residence / work, and special conditions

13 Competency Questions

For example:

- Rodrigo works at the restaurant during the weekend and needs to find the best time for reaching his workplace at 11 am

Knowledge collection

- schema.org
- GTFS standard

Name	Reference
Agency	Organization
Stops	TrainStation BusStop Trip
Trips and Routes	TravelAction TrainTrip BusTrip
Calendar and Timetable	Schedule Event
Locations	Place

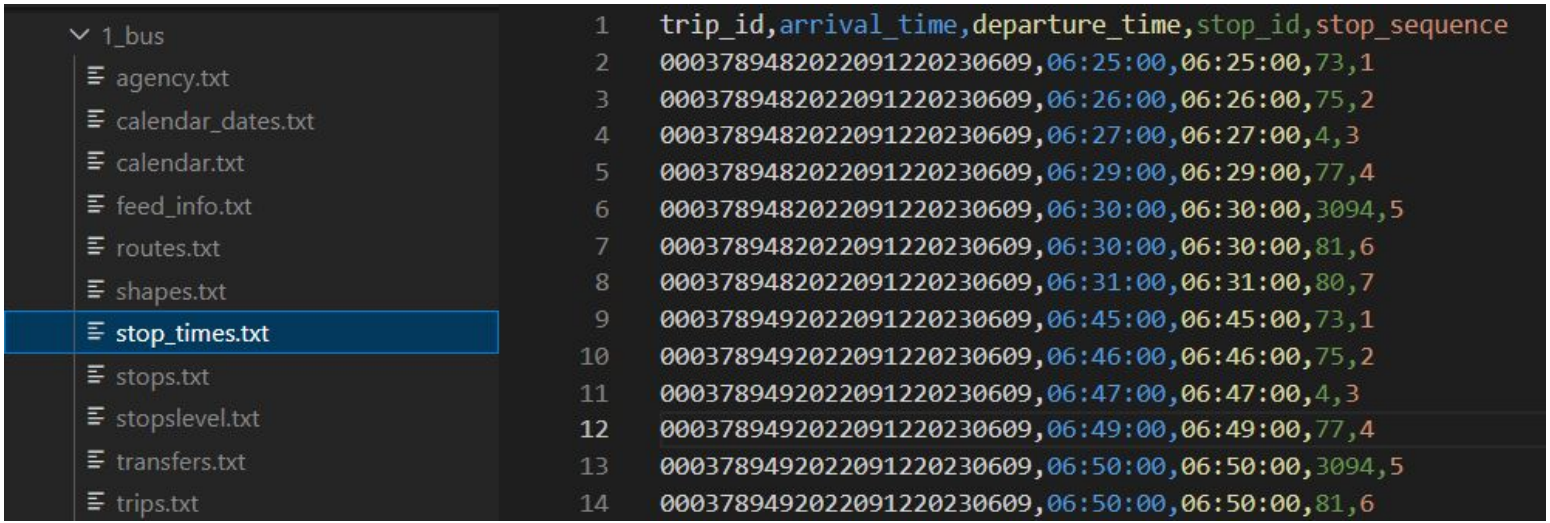
Table 1: Schemas used for the knowledge integration

Data collection

1. Trentino Trasporti - Open Data
2. Trentino Trasporti - Ferrovia
3. Google Maps

Data collection - “1. Open Data”

- Information about bus lines
- CSV files compliant with the GTFS format



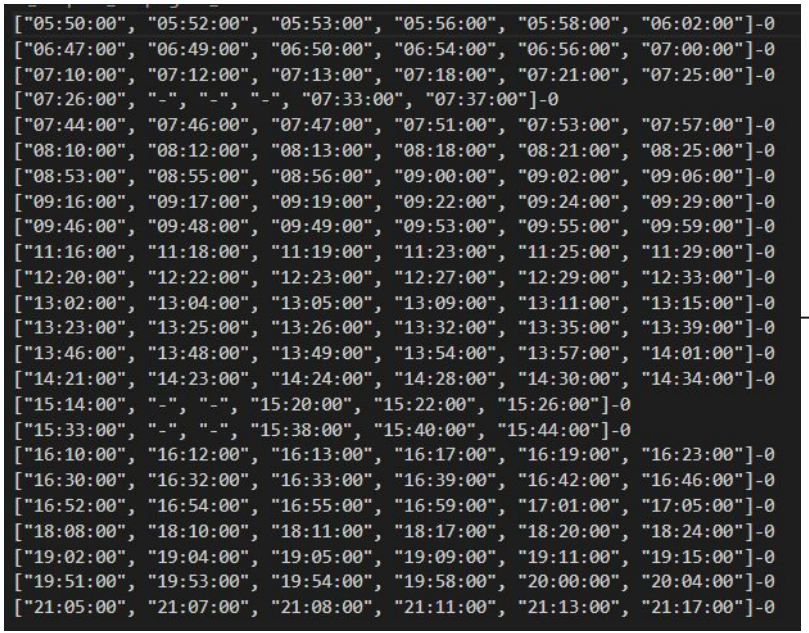
Data collection - “2. Ferrovia”

- Information about train lines
- Unstructured files (PDF with timetables)

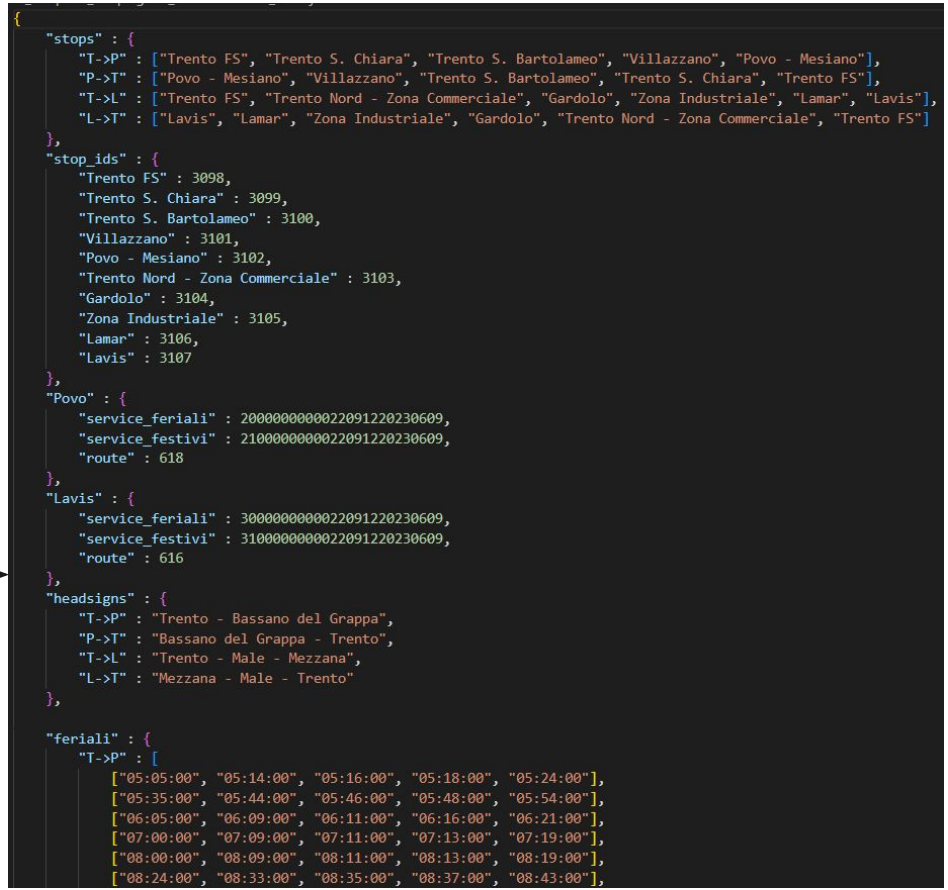
LINEA TRENTO - BASSANO del GRAPPA VALIDO DAL 28 OTTOBRE AL 21 NOVEMBRE 2022

		BUS	BUS	BUS	R	BUS	BUS	BUS	BUS	BUS	BUS	BUS	BUS	BUS	BUS	BUS	BUS	BUS	R	BUS	R	BUS	R	BUS	BUS	BUS	BUS	BUS
		VE151	TN311	TN305	16109	TN327	TT03	TT05	TT915	TT919	TT921	TN301	TT923	TT925	TN313	TT927	TN307	TT929	16119	TN329	16931	TT933	16935	TT835	TT935	TT937	TN325	TN309
					G	G	G	G		G	G																	
	IF	TI veneto	TI	TI	TI	TI	TT	TT	TT	TT	TT	TI	TT	TT	TI	TT	TI	TT	TI		TT	TT	TT	TT	TT	TT	TI	TI
	NOTE	(1)	(1)	(1)			(1)	(1)	(3)	(2)	(4)		(1)	(1)	(1)	(3)	(1)	(1)			(1)	(1)	(1)	(1)	(3)	(1)	(1)	(1)
Trento	p.	-	5:05	5:35	6:05	-	7:00	8:00	8:05	8:24	8:24	9:05	9:35	10:05	11:05	11:05	12:05	12:35	13:05	-	13:35	14:05	15:05	-	15:05	15:35	16:05	16:35
Trento S. Chiara	p.	-	5:14	5:44	6:09	-	7:09	8:09	8:14	8:33	8:33	9:14	9:44	10:14	11:14	11:14	12:14	12:44	13:09	-	13:39	14:14	15:09	-	15:14	15:44	16:14	16:44
Trento S. Bartolameo	p.	-	5:16	5:46	6:11	-	7:11	8:11	8:16	8:35	8:35	9:16	9:46	10:16	11:16	11:16	12:16	12:46	13:11	-	13:41	14:16	15:11	-	15:16	15:46	16:16	16:46
Villazzano	p.	-	5:18	5:48	6:16	-	7:13	8:13	8:18	8:37	8:37	9:18	9:48	10:18	11:18	11:18	12:18	12:48	13:16	-	13:46	14:18	15:16	-	15:18	15:48	16:18	16:48
Povo – Mesiano	p.	-	5:24	5:54	6:21	-	7:19	8:19	8:24	8:43	8:43	9:24	9:54	10:24	11:24	11:24	12:24	12:54	13:21	-	13:51	14:24	15:21	-	15:24	15:54	16:24	16:54
Pergine Valsugana	a.	-	5:34	6:04	6:31	-	7:29	8:29	8:34	8:53	8:53	9:34	10:04	10:34	11:34	11:34	12:34	13:04	13:31	-	14:01	14:34	15:31	-	15:34	16:04	16:34	17:04
	p.	-	5:35	6:05	6:32	-	7:30	8:30	8:35	9:02	9:02	9:35	10:05	10:35	11:35	11:35	12:35	13:05	13:32	-	14:02	14:35	15:32	-	15:35	16:05	16:35	17:05
S. Cristoforo al L. I.	p.	-	5:40	6:10	6:36	-	7:35	8:35	8:40	9:07	9:07	9:40	10:10	10:40	11:40	11:40	12:40	13:10	13:36	-	14:06	14:40	15:36	-	15:40	16:10	16:40	17:10
Calceranica	p.	-	5:45	6:15	6:41	-	7:40	8:40	8:45	9:12	9:12	9:45	10:15	10:45	11:45	11:45	12:45	13:15	13:41	-	14:11	14:45	15:41	-	15:45	16:15	16:45	17:15

Data extraction - “2. Ferrovia”



```
["05:50:00", "05:52:00", "05:53:00", "05:56:00", "05:58:00", "06:02:00"]-0
["06:47:00", "06:49:00", "06:50:00", "06:54:00", "06:56:00", "07:00:00"]-0
["07:10:00", "07:12:00", "07:13:00", "07:18:00", "07:21:00", "07:25:00"]-0
["07:26:00", "-", "-", "-", "07:33:00", "07:37:00"]-0
["07:44:00", "07:46:00", "07:47:00", "07:51:00", "07:53:00", "07:57:00"]-0
["08:10:00", "08:12:00", "08:13:00", "08:18:00", "08:21:00", "08:25:00"]-0
["08:53:00", "08:55:00", "08:56:00", "09:00:00", "09:02:00", "09:06:00"]-0
["09:16:00", "09:17:00", "09:19:00", "09:22:00", "09:24:00", "09:29:00"]-0
["09:46:00", "09:48:00", "09:49:00", "09:53:00", "09:55:00", "09:59:00"]-0
["11:16:00", "11:18:00", "11:19:00", "11:23:00", "11:25:00", "11:29:00"]-0
["12:20:00", "12:22:00", "12:23:00", "12:27:00", "12:29:00", "12:33:00"]-0
["13:02:00", "13:04:00", "13:05:00", "13:09:00", "13:11:00", "13:15:00"]-0
["13:23:00", "13:25:00", "13:26:00", "13:32:00", "13:35:00", "13:39:00"]-0
["13:46:00", "13:48:00", "13:49:00", "13:54:00", "13:57:00", "14:01:00"]-0
["14:21:00", "14:23:00", "14:24:00", "14:28:00", "14:30:00", "14:34:00"]-0
["15:14:00", "-", "-", "15:20:00", "15:22:00", "15:26:00"]-0
["15:33:00", "-", "-", "15:38:00", "15:40:00", "15:44:00"]-0
["16:10:00", "16:12:00", "16:13:00", "16:17:00", "16:19:00", "16:23:00"]-0
["16:30:00", "16:32:00", "16:33:00", "16:39:00", "16:42:00", "16:46:00"]-0
["16:52:00", "16:54:00", "16:55:00", "16:59:00", "17:01:00", "17:05:00"]-0
["18:08:00", "18:10:00", "18:11:00", "18:17:00", "18:20:00", "18:24:00"]-0
["19:02:00", "19:04:00", "19:05:00", "19:09:00", "19:11:00", "19:15:00"]-0
["19:51:00", "19:53:00", "19:54:00", "19:58:00", "20:00:00", "20:04:00"]-0
["21:05:00", "21:07:00", "21:08:00", "21:11:00", "21:13:00", "21:17:00"]-0
```

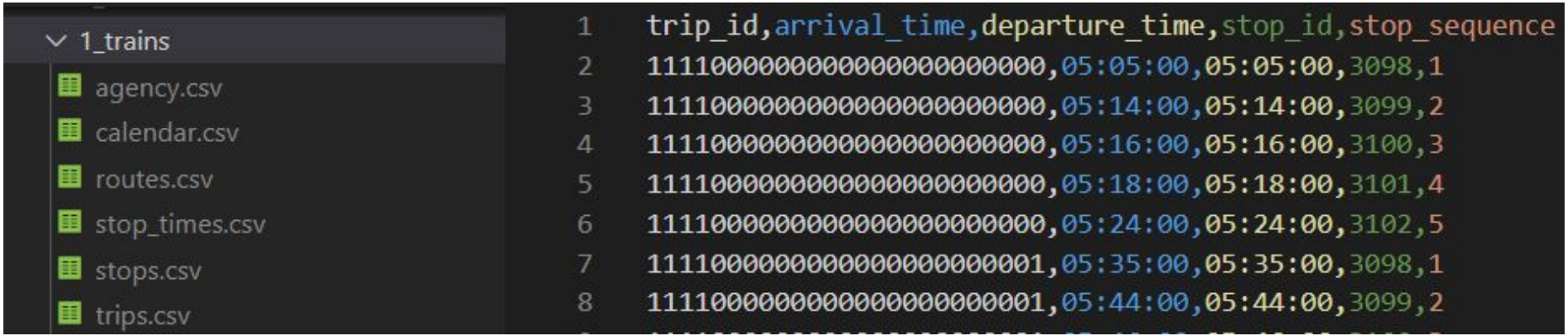


```
{
    "stops" : {
        "T->P" : ["Trento FS", "Trento S. Chiara", "Trento S. Bartolameo", "Villazzano", "Povo - Mesiano"],
        "P->T" : ["Povo - Mesiano", "Villazzano", "Trento S. Bartolameo", "Trento S. Chiara", "Trento FS"],
        "T->L" : ["Trento FS", "Trento Nord - Zona Commerciale", "Gardolo", "Zona Industriale", "Lamar", "Lavis"],
        "L->T" : ["Lavis", "Lamar", "Zona Industriale", "Gardolo", "Trento Nord - Zona Commerciale", "Trento FS"]
    },
    "stop_ids" : {
        "Trento FS" : 3098,
        "Trento S. Chiara" : 3099,
        "Trento S. Bartolameo" : 3100,
        "Villazzano" : 3101,
        "Povo - Mesiano" : 3102,
        "Trento Nord - Zona Commerciale" : 3103,
        "Gardolo" : 3104,
        "Zona Industriale" : 3105,
        "Lamar" : 3106,
        "Lavis" : 3107
    },
    "Povo" : {
        "service_feriali" : 2000000000022091220230609,
        "service_festivi" : 2100000000022091220230609,
        "route" : 618
    },
    "Lavis" : {
        "service_feriali" : 3000000000022091220230609,
        "service_festivi" : 3100000000022091220230609,
        "route" : 616
    },
    "headsigns" : {
        "T->P" : "Trento - Bassano del Grappa",
        "P->T" : "Bassano del Grappa - Trento",
        "T->L" : "Trento - Male - Mezzana",
        "L->T" : "Mezzana - Male - Trento"
    },

    "feriali" : {
        "T->P" : [
            ["05:05:00", "05:14:00", "05:16:00", "05:18:00", "05:24:00"],
            ["05:35:00", "05:44:00", "05:46:00", "05:48:00", "05:54:00"],
            ["06:05:00", "06:09:00", "06:11:00", "06:16:00", "06:21:00"],
            ["07:00:00", "07:09:00", "07:11:00", "07:13:00", "07:19:00"],
            ["08:00:00", "08:09:00", "08:11:00", "08:13:00", "08:19:00"],
            ["08:24:00", "08:33:00", "08:35:00", "08:37:00", "08:43:00"],
```

Data formatting - “2. Ferrovia”

The intermediate files were programmatically formatted into CSV files compliant with the GTFS format



Informal Modeling

Teleology Foundations

E-types chosen according to the inception phase

- Transportation_agency
- Line
- Trip
- Schedule
- Calendar
- Stop

Components	Common	Core	Contextual
Object	location date time person	stop line schedule	bus train wheelchair workday weekend morning afternoon night holiday
Action	decision making	-	maintainance raining
Function	student worker	trip	departure arrival cover available seats

Table 5: Teleology components classified as object, action or function, according to the information extracted from the CQs.

Teleology Foundations

Property name	Description	Domain(s)	Range(s)
has_location_in	Urban area where a stop is located	Stop	Region
has_stops	Sequence of stops which are included in a trip	Schedule	Stop
has_lines	Lines operated by a transportation organization	Agency	Line
has_trips	Trips included in the route	Line	Trip
has_calendar	Weekdays when the current trip runs	Trip	Calendar
has_schedule	Timetable of transportation means	Stop, Trip	Schedule

Table 7: Object Properties extracted from the CQs.

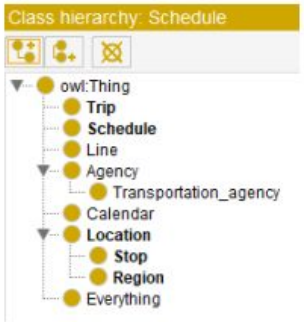
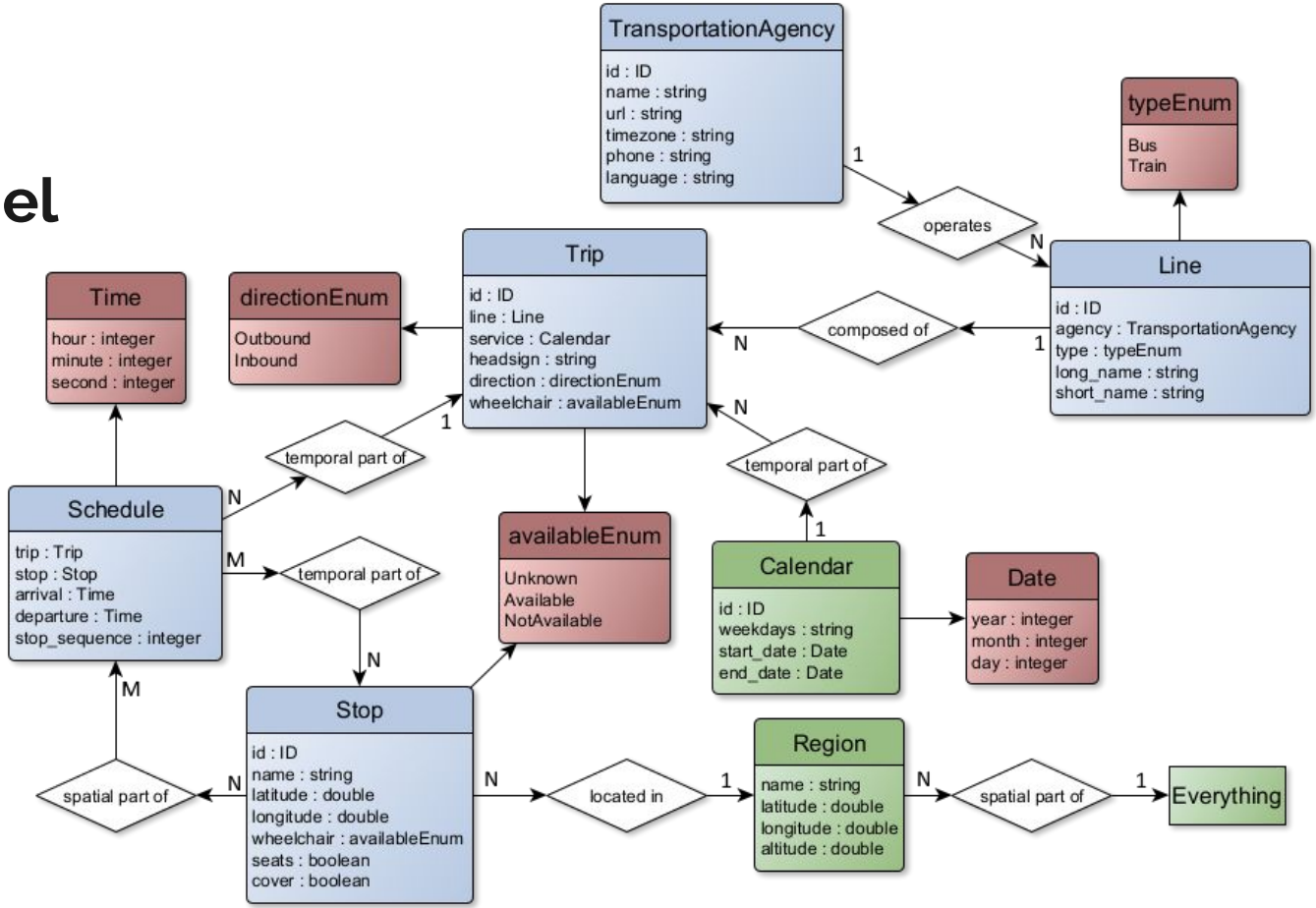


Figure 4: Updated ontology after the informal modelling

Property name	Description	Domain(s)	Range(s)
has_id	Unique identifier of an instance/entity	Agency Stop Line Trip Calendar	xsd:int
has_name	Name of an instance/entity	Agency Line Location	xsd:string
has_phone	Phone number of an organization	Agency	xsd:string
has_language	Official language of an organization	Agency	xsd:string
has_timezone	Time zone of an organization	Agency	xsd:string
has_url	Website of an organization	Agency	xsd:string
has_headsign	Short description of the trip	Trip	xsd:string
has_direction	Direction of the trip/line	Trip	xsd:boolean
has_wheelchair_accessibility	Availability of facilities for people with reduced mobility	Trip Stop	xsd:boolean
has_covering	Availability of a covering in a stop	Stop	xsd:boolean
has_seats	Availability of seats in a stop	Stop	xsd:boolean
has_latitude	Latitude of the location	Location	xsd:double
has_longitude	Longitude of the location	Location	xsd:double
has_altitude	Altitude of the location	Region	xsd:double
has_arrival_time	Arrival time of the transportation mean in a stop	Schedule	xsd:dateTime
has_departure_time	Departure time of the transportation mean in a stop	Schedule	xsd:dateTime
has_stops_sequence	Sequence of stops which are included in the current trip	Schedule	xsd:int
has_weekdays	Weekdays when the current trip runs	Calendar	xsd:string
has_end_date	End date of a specific trip	Calendar	xsd:dateTime
has_start_date	Start date of a specific trip	Calendar	xsd:dateTime
has_long_name	Full name of the line	Line	xsd:string
has_type	Type of a transportation mean, i.e., train or bus	Line	xsd:int

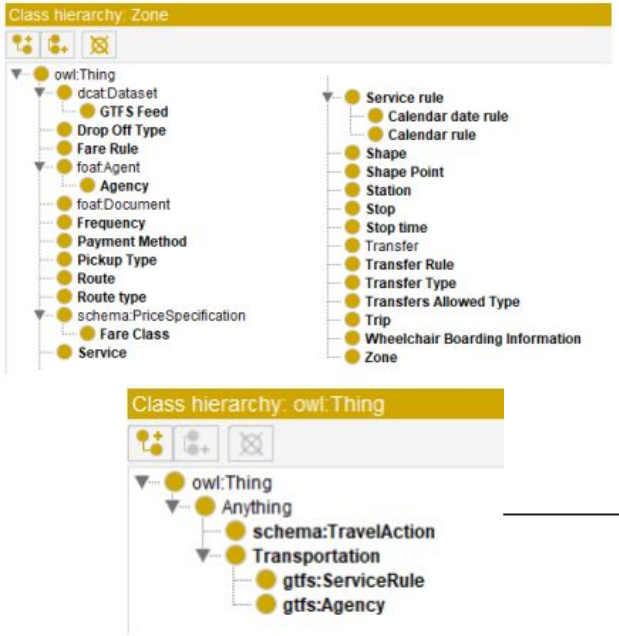
Table 6: Data Properties extracted from the CQs.

ER model


TransportationAgency
id : ID
name : string
url : string
timezone : string
phone : string
language : string
1
operates
N
typeEnum
Bus
Train
Line
id : ID
agency : TransportationAgency
type : typeEnum
long_name : string
short_name : string
1
composed of
N
Trip
id : ID
line : Line
service : Calendar
headsign : string
direction : directionEnum
wheelchair : availableEnum
directionEnum
Outbound
Inbound
Time
hour : integer
minute : integer
second : integer
N
temporal part of
1
N
temporal part of
1
Schedule
trip : Trip
stop : Stop
arrival : Time
departure : Time
stop_sequence : integer
M
temporal part of
N
availableEnum
Unknown
Available
NotAvailable
Calendar
id : ID
weekdays : string
start_date : Date
end_date : Date
Date
year : integer
month : integer
day : integer
M
spatial part of
N
Stop
id : ID
name : string
latitude : double
longitude : double
wheelchair : availableEnum
seats : boolean
cover : boolean
N
located in
1
Region
name : string
latitude : double
longitude : double
altitude : double
N
spatial part of
1
Everything

Formal Modeling

ETG generation



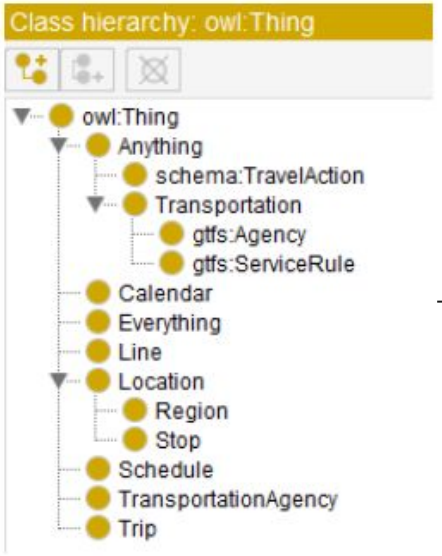


Figure 8: Intermediate teleontology.

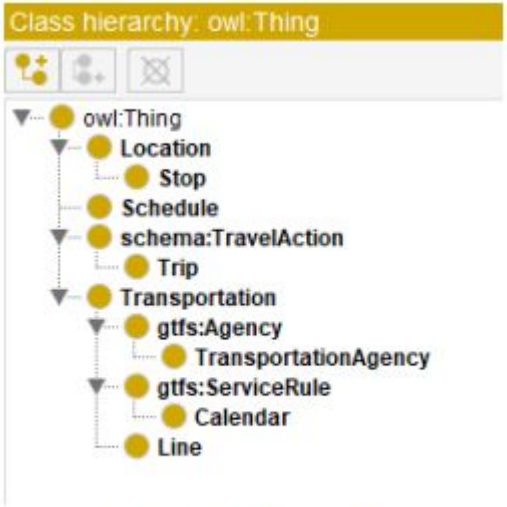


Figure 9: Formalized schema.

Language Alignment

- Etypes were standardized following the UKC standard

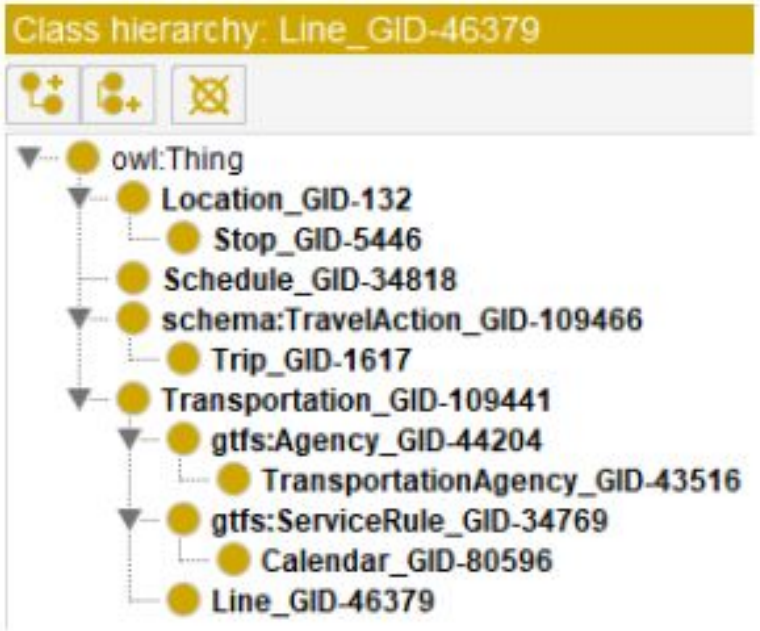


Figure 10: Language aligned ETG.

Data integration

Knowledge Graph Construction

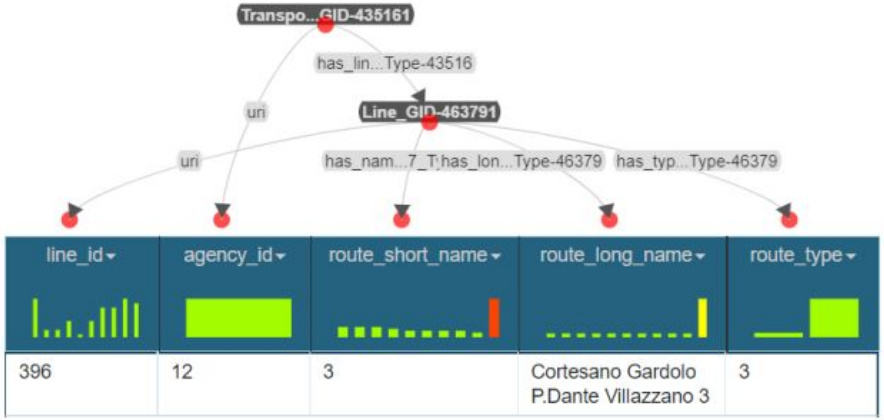


Figure 11: Data Integration in KarmaLinker: Lines

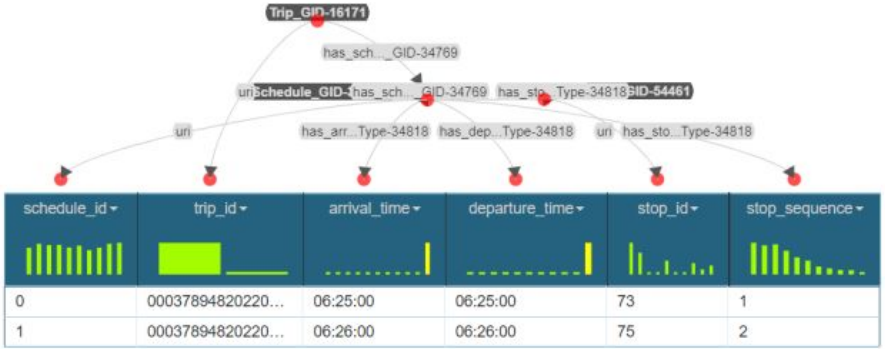


Figure 12: Data Integration in KarmaLinker: Schedule

Exploitation

Knowledge Graph Statistics

Coverage:

- 1 for **e-types**
- 24/26 (0.92) for **properties**

- 6 etypes, namely:
  1. TransportationAgency_GID-43516 with **1 instance** and **387 links** to other etypes.
  2. Calendar_GID-80596 with **43 instances** and approximately **3K links** to other etypes.
  3. Line_GID-46379 with **43 instances** and approximately **7K links** to other etypes.
  4. Stop_GID-5446 with **1112 instances** and approximately **160K links** to other etypes.
  5. Trip_GID-1617 with **3127 instances** and approximately **85K links** to other etypes.
  6. Schedule_GID-34818 with **77923 instances** and approximately **389K links** to other etypes.
- 5 object properties
- 21 data properties

```
############################
#CQ4
############################
# Rodrigo works at the restaurant during the weekend and needs to find the best time for
reaching his workplace at 11 am.
PREFIX rdf: <http://knowdive.disi.unitn.it/etype#>

select distinct ?name_start ?name_end ?departure ?arrival ?linenumber ?linename ?weekdays
where {
    ?line a rdf:Line_GID-46379;
        rdf:has_name_GID-34017_Type-132 ?linenumber;
        rdf:has_long_name_GID-34017_Type-46379 ?linename;
        rdf:has_trips_GID-1501_Type-46379 ?trip.

    ?trip
        rdf:has_calendar_GID-80596_Type-1617 ?calendar;
        rdf:has_schedule_GID-34769 ?schedule_start;
        rdf:has_schedule_GID-34769 ?schedule_end.

    ?calendar rdf:has_weekdays_GID-80597_Type-80596 ?weekdays.

    ?schedule_start
        rdf:has_stops_GID-5446_Type-34818 ?stops_start;
        rdf:has_departure_time_GID-80846_Type-34818 ?departure.

    ?schedule_end
        rdf:has_stops_GID-5446_Type-34818 ?stops_end;
        rdf:has_arrival_time_GID-80845_Type-34818 ?arrival.

    ?stops_start rdf:has_name_GID-34017_Type-132 ?name_start.
    ?stops_end   rdf:has_name_GID-34017_Type-132 ?name_end.

    FILTER (!regex (?weekdays, "00$"))
    FILTER (?arrival < "11:00:00")
    FILTER (?departure < ?arrival)

    FILTER (CONTAINS(?name_start, "Piazza Dante"))
    FILTER (CONTAINS(?name_end, "Mattarello Gotarda"))

}
ORDER BY DESC (?arrival) LIMIT 10
```

Rodrigo works at the restaurant during the weekend and needs to find the best time for reaching his workplace at 11 am.

KG exploitation through SparQL

Rodrigo works at the restaurant during the weekend and needs to find the best time for reaching his workplace at 11 am.

	name_start	name_end	departure	arrival	linenumber	linename	weekdays
1	"Piazza Dante "Stazione Fs"	"Mattarello Gotarda"	"10:37:00"	"10:58:00"	"8"	"Centochiavi Piazza Dante Mattarello"	"1111110"
2	"Piazza Dante "Stazione Fs"	"Mattarello Gotarda"	"10:37:00"	"10:58:00"	"Rosmini "Cimitero" / Muse"	"Centochiavi Piazza Dante Mattarello"	"1111110"
3	"Piazza Dante "Stazione Fs"	"Mattarello Gotarda"	"10:07:00"	"10:28:00"	"8"	"Centochiavi Piazza Dante Mattarello"	"1111110"
4	"Piazza Dante "Stazione Fs"	"Mattarello Gotarda"	"10:07:00"	"10:28:00"	"Rosmini "Cimitero" / Muse"	"Centochiavi Piazza Dante Mattarello"	"1111110"
5	"Piazza Dante "Stazione Fs"	"Mattarello Gotarda"	"09:50:00"	"10:11:00"	"8"	"Centochiavi Piazza Dante Mattarello"	"0000001"

```
##########################
#CQ5
##########################
# Giulia is going to a birthday party next Saturday in the city centre
(Trento). The party tends to end at midnight, so she is wondering when is
the last bus to return back home.
PREFIX rdf: <http://knowdive.disi.unitn.it/etype#>
PREFIX omgeo: <http://www.ontotext.com/owlim/geo#>

select distinct ?stopname ?departure ?linenumber ?linename
where {
      ?line a rdf:Line_GID-46379;
            rdf:has_name_GID-34017_Type-132 ?linenumber;
            rdf:has_long_name_GID-34017_Type-46379 ?linename;
            rdf:has_trips_GID-1501_Type-46379 ?trip.

      ?trip rdf:has_calendar_GID-80596_Type-1617 ?calendar;
            rdf:has_schedule_GID-34769 ?schedule_start;
            rdf:has_schedule_GID-34769 ?schedule_end.

      ?calendar rdf:has_weekdays_GID-80597_Type-80596 ?weekdays;
            FILTER regex (?weekdays, "1[01]$").

      ?schedule_start rdf:has_stops_GID-5446_Type-34818 ?stops_start;
            rdf:has_departure_time_GID-80846_Type-34818 ?departure.

      ?schedule_end
            rdf:has_stops_GID-5446_Type-34818 ?stops_end;
            rdf:has_arrival_time_GID-80845_Type-34818 ?arrival.

      ?stops_start rdf:has_name_GID-34017_Type-132 ?stopname;
            rdf:has_latitude_GID-46264_Type-132 ?lat;
            rdf:has_longitude_GID-46270_Type-132 ?long.

      ?stops_end rdf:has_name_GID-34017_Type-132 ?destiny.

      FILTER (?departure < ?arrival)

      ### Any stop at less than 500m from Piazza Dante train station
      FILTER (omgeo:distance(46.07209811, 11.11955396, ?lat, ?long) < 0.5)
      FILTER (CONTAINS(?destiny, "Gardolo 4 Nov. \"Piscina"))
}
ORDER BY DESC (?departure) LIMIT 5
```

Giulia is going to a birthday party next Saturday in the city centre (Trento). The party tends to end at midnight, so she is wondering when is the last bus to return back home.

KG exploitation through SparQL

Giulia is going to a birthday party next Saturday in the city centre (Trento). The party tends to end at midnight, so she is wondering when is the last bus to return back home.

	stopname	departure	linenumber	linename
1	"Romagnosi Vannetti"	"23:41:00"	"3"	"Cortesano Gardolo P.Dante Villazzano 3"
2	"Sanzio "Castello"	"23:39:00"	"3"	"Cortesano Gardolo P.Dante Villazzano 3"
3	"Romagnosi Vannetti"	"23:11:00"	"3"	"Cortesano Gardolo P.Dante Villazzano 3"
4	"Sanzio "Castello"	"23:09:00"	"3"	"Cortesano Gardolo P.Dante Villazzano 3"
5	"Romagnosi Vannetti"	"22:41:00"	"3"	"Cortesano Gardolo P.Dante Villazzano 3"

Open Issues & Conclusions

UNIVERSITY OF TRENTO

Trentino Urban Transportation KGE 2022

Diego Barquero Morera, 229577, diego.barqueromorera@studenti.unitn.it
Vahan Petrosyan, 229737, vahan.petrosyan@studenti.unitn.it