UNIVERSITY
OF TRENTO - Italy

KGE 2022 - Project on CNR-Rome data

Submitted to:

Prof. Fausto Giunchiglia
Simone Bocca
(Department of Computer Engineering and Computer Science)

Presented by:

Surya Hembrom (229578)
Surbhi Malhotra (229579)
(Masters- Quantitative and Computational Biology, CIBIO)

PURPOSE

- Integrate CNR-Rome patient's data with synthetic patient data and medicinal data from drug databases.
- Facilitate a second opinion for physician in drug prescription to a patient with comorbidities.
- Make the physician aware of drugs being reassessed, or disapproved in Europe (or Italy).
- Help patients keep track of their medical history.
- Let hospital management handle patient's data efficiently, increase the personalised care of the patients based on their diseases' gravity.

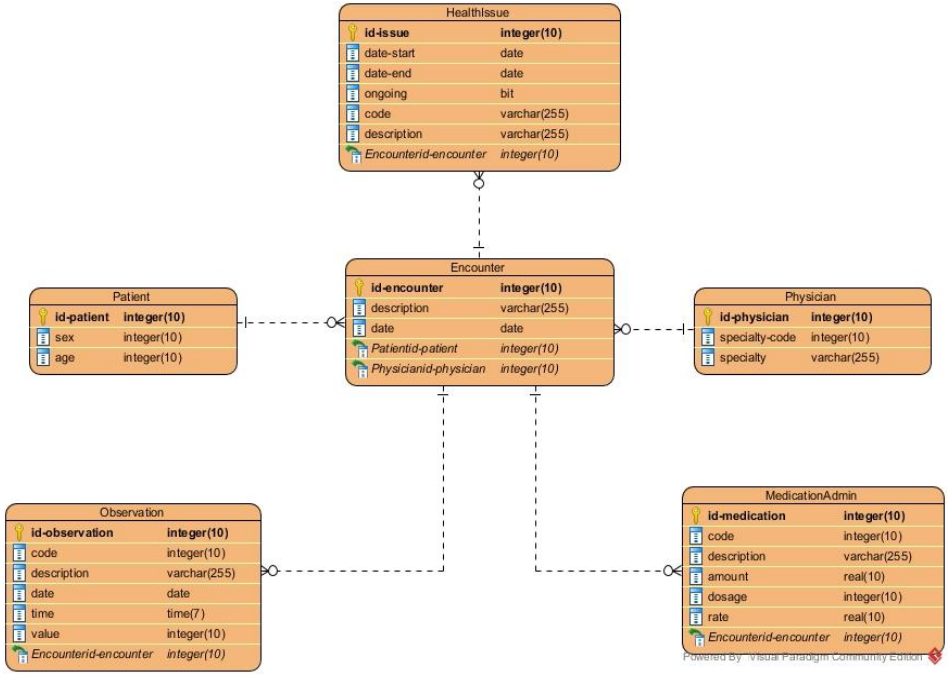
DOMAIN OF INTEREST

- Medical healthcare domain.
- This study is limited only to
    - patients, their diseases, their visits to physician (encounter) and drug administration to the patients in CNR-Rome hospital in the city Roma .

RESOURCES

Knowledge Resources:

- Classes: CNR-Rome data + synthetic datasets.
  - Patient, Encounter, Health Issue, Medication Admin and Physician.
- Add ons : Drug database , Pathology, Pharmacy, Observations and Diagnostics
  - www.schema.org (Medical domain)
  - FHIR standards for the health and medicine related terms or names.



Data Resources:

- CNR-Rome data
- synthetic data of the patients from the website https://synthea.mitre.org/downloads.( )
  - generated with https://www.bestrandoms.com/random-address-in-it.
- The drug data of PharmGKB ( ) https://www.pharmgkb.org/labelAnnotations
  - dataset downloaded from https://www.ema.europa.eu/en/medicines/download- medicine-data under the download link of ”Download table of all EPARs for human and veterinary medicines”.
    - Filtered for Human

PURPOSE FORMALISATION

Scenarios:

1. Hospital management who needs to manage a patient.
2. Physician who needs to provide patients with principal care and consultancy.
3. Pharmacist who needs to manage prescriptions and drugs.
4. Patient who needs to process his own medical information.

Personas:

1. Patient
2. Hospital management
3. Physician/Clinician
4. Test Examiner
5. Pharmacist

CQs:

Scenarios	Personas	Competency Questions
Hospital management who needs to manage a patient	Hospital Management	Hospital staff wants to look for an appointment slot for Matteo with a physician next week.
	Hospital Management	Hospital staff retrieves Andre email address to send him the diagnostic test results.
	Hospital Management	Hsiu is diagnosed with obesity and connects with receptionist to schedule an appointment with the clinician for routine follow up.
Physician who needs to provide patients with principal care and consultancy	Physician	Physician wants to refer to the last medication prescribed to Jayson who is diagnosed with Acute bronchitis.
	Physician	Clinician searches through the list of approved drugs that are available in the local pharmacy for hypertension.
	Physician	Physician searches through database for a drug to prescribe to Hsiu who is diagnosed with Coronary Heart Disease and has active Prediabetes comorbidity.
	Physician	Physician checks if Carvedilol, a drug for Hypertension is available in the Local pharmacy or not.
	Physician	Physician wants to search through drug database for Saxenda prescribed in which comorbidity conditions to avoid side effects of the drug.
	Test Examiner	Test Examiner needs to look for the type of test to conduct for Andre on 2017-02-04.
	Physician	Maria who is a Dutch tourist in Rome and is diagnosed with Vertigo. He asks the physician appointed in Rome CNR to prescribe a drug that is also authorized in Netherlands.
	Test Examiner	Vennie has active Preeclampsia and is referred to the test examiner to conduct Evaluation of uterine fundal height.
	Physician	Clinician looks into the recorded observations of Patient 1 on his visit to MMG on 2019-01-02.
Pharmacist who needs to manage prescriptions and drugs	Pharmacist	Pharmacist is looking for an active component of the prescribed drug for the treatment of hypertension.
	Pharmacist	Pharmacist searches through database if Glidipion (previously Pioglitazone Actavis Group), a drug for Diabetes Mellitus Type 2 still available in the market or withdrawn.
	Pharmacist	Pharmacist wants to check the dosage amount of blopress prescribed to patient ID-1.
	Pharmacist	Pharmacist in Paris city is looking for an alternative of Tranquillante prescribed to an Italian citizen who is diagnosed with hypertension.
Patient who needs to process his own medical information	Patient	Clyde visited Rome CNR on 2009-07-18 to undergo a test and wants to know what health issue is she diagnosed with.
	Patient	Matteo wants to schedule an appointment with a physician next week.
	Patient	Tia who diagnosed with Idiopathic atrophic hypothyroidism records her BP level observations.
	Patient	Agueda has to follow the SoPs/guidelines to take the test for general examination.
	Patient	Mathew who is a Dutch tourist in Rome and is diagnosed with Vertigo. He asks the physician appointed in Rome CNR to prescribe a drug that is authorized in Netherlands.
	Patient	Patient 1 records routine health observations and wants to track the whole record.

Inception

- Resources scraping and data filtering
- Resource classification
- Ontology creation for separate entities with defined data properties
  - Protégé
- Data layer manipulation and handling
  - Karma tool

Iterative phase: ontology building in concordance to data layer.

Decision making phase: limit the scope of Knowledge Graph (KG) as per the data availability.

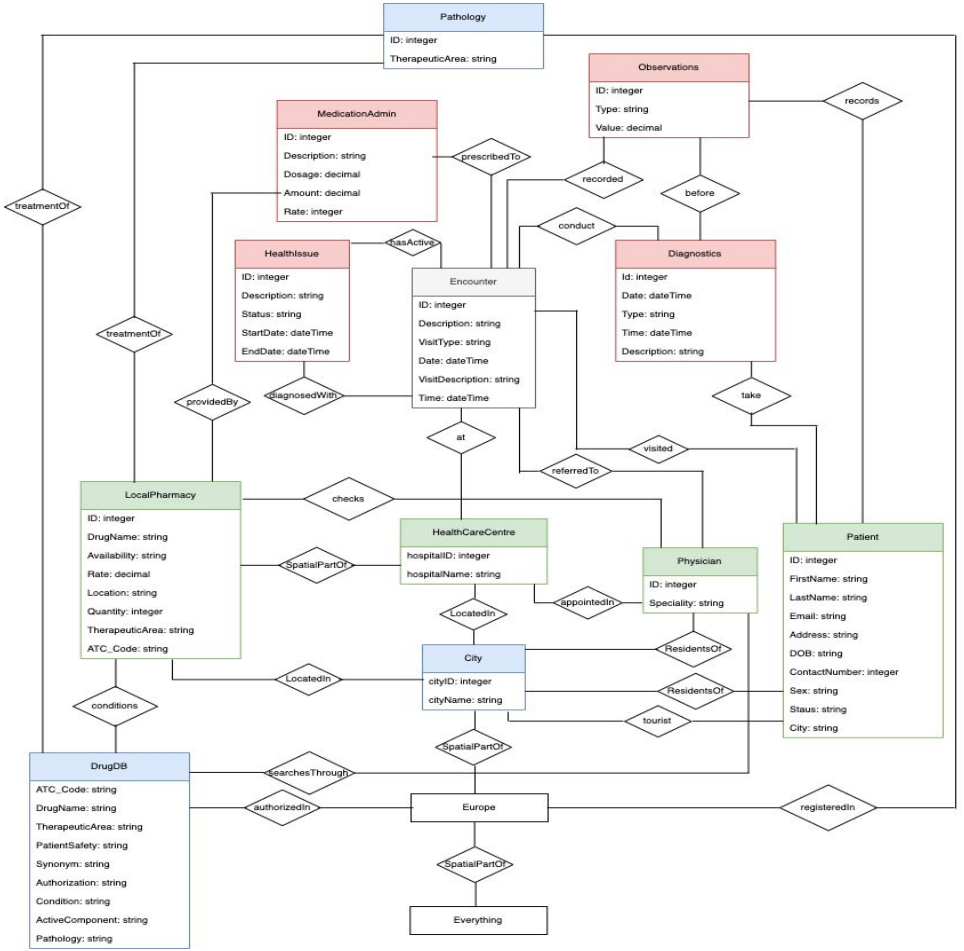
Entities Categorization

- **The common entities :** Drug_database, Pathology and City.
- **The core entities:** Health care centre, Local pharmacy, Patient and Physician
- **The contextual resources collected :** Observations, Health issues, Diagnostics, Encounter and Medication administration.

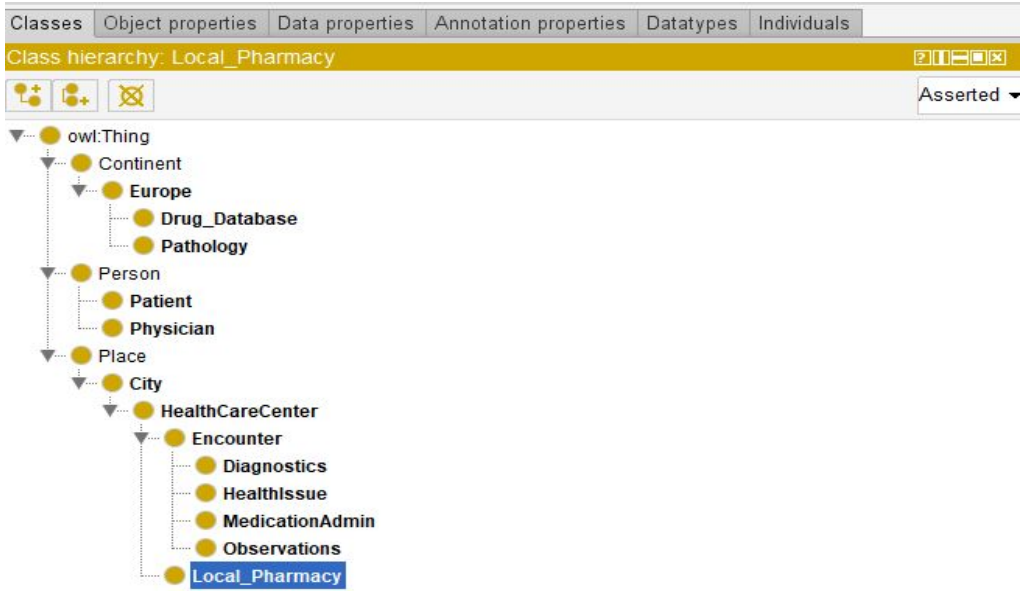
Patient	**Pathology**	**Physician**	**Observations**	**Encounter**	**Medication Admin**
Id	Id	Id	Id	Id	Id
First Name	Therapeutic area	Speciality	Encounter Id	Patient Id	Patient Id
Last name		HCC name	Patient Id	Physician Id	Pharmacy Id
Date of birth			Type of test	Description	Encounter Id
Date of death			Value	Date	Description
Gender				Time	amount
Email				Type of Visit	dosage
Contact number				Visit Description	rate
Address					
City					
Visiting status					
City	**HealthCareCentre**	**Diagnostics**	**Local Pharmacy**	**Drug database**	**Health Issue**
Id	Id	Id	Id	Id	Id
Name	Name	encounter Id	ATC code	ATC Code	Patient Id
Country	City Id	patient Id	pathology Id	Medicine name	Encounter Id
Continent	City name	Diagnostics date	Therapeutic area	Therapeutic area	Description
		Diagnostics time	Drug name	Common name	Status
		Diagnostics description	Dosage	Active ingredient	Start Date
			quantity	Patient safety	End Date
			rate	Authorization	
			shelf location	Marketing	
			Availability Status	Human	
				Conditions	

Informal Modeling

- ER diagram: conceptualize our purpose considering the proposed CQs, built upon the defined scenarios.
- Contextual teleology building.
- Class hierarchy.

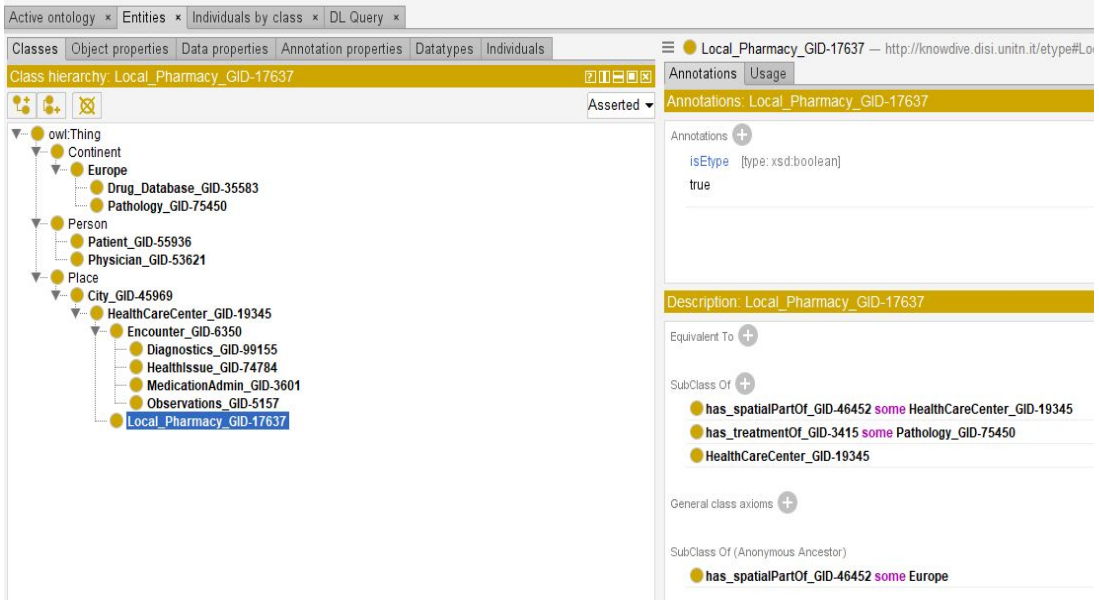


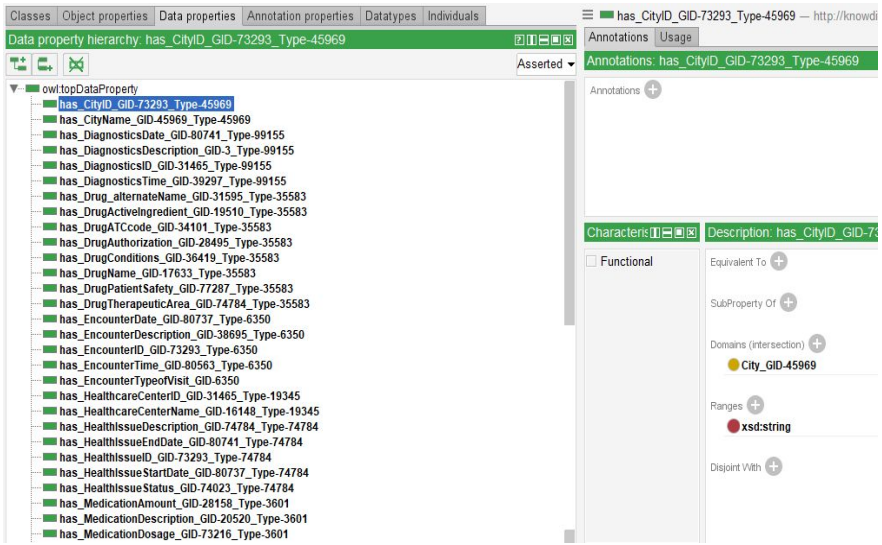
Teleology

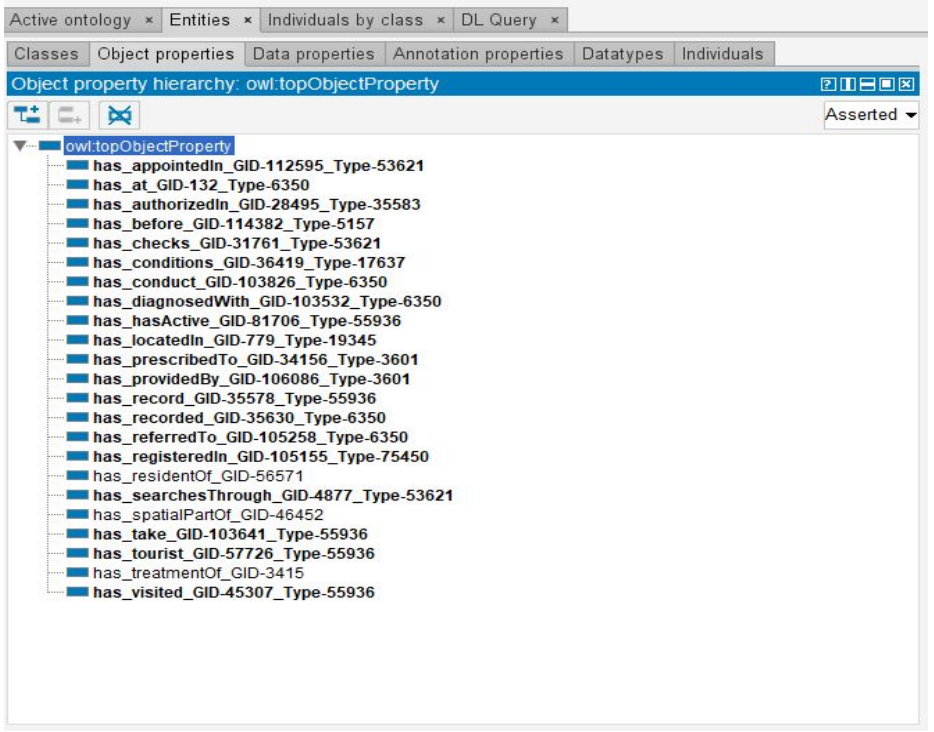

Classes
Object properties
Data properties
Annotation properties
Datatypes
Individuals
Class hierarchy: Local_Pharmacy
Asserted
owl:Thing
Continent
Europe
Drug_Database
Pathology
Person
Patient
Physician
Place
City
HealthCareCenter
Encounter
Diagnostics
HealthIssue
MedicationAdmin
Observations
Local_Pharmacy

Formal Modeling

- ETG (Entity Type Graph) generation
  - Protégé tool
- Language alignment
  - KOS web tool

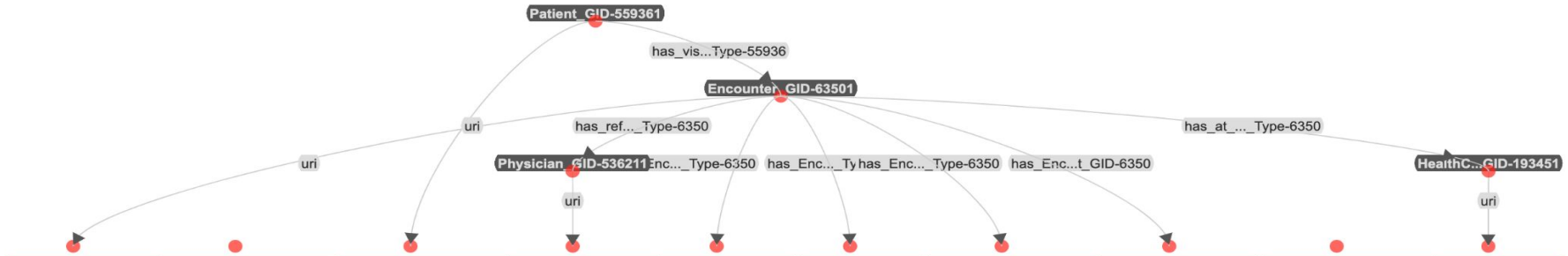



Classes
Object properties
Data properties
Annotation properties
Datatypes
Individuals
Data property hierarchy: has_CityID_GID-73293_Type-45969
Asserted
owl:topDataProperty
has_CityID_GID-73293_Type-45969
has_CityName_GID-45969_Type-45969
has_DiagnosticsDate_GID-80741_Type-99155
has_DiagnosticsDescription_GID-3_Type-99155
has_DiagnosticsID_GID-31465_Type-99155
has_DiagnosticsTime_GID-39297_Type-99155
has_Drug_alternateName_GID-31595_Type-35583
has_DrugActiveIngredient_GID-19510_Type-35583
has_DrugATCcode_GID-34101_Type-35583
has_DrugAuthorization_GID-28495_Type-35583
has_DrugConditions_GID-36419_Type-35583
has_DrugName_GID-17633_Type-35583
has_DrugPatientSafety_GID-77287_Type-35583
has_DrugTherapeuticArea_GID-74784_Type-35583
has_EncounterDate_GID-80737_Type-6350
has_EncounterDescription_GID-38695_Type-6350
has_EncounterID_GID-73293_Type-6350
has_EncounterTime_GID-80563_Type-6350
has_EncounterTypeofVisit_GID-6350
has_HealthcareCenterID_GID-31465_Type-19345
has_HealthcareCenterName_GID-16148_Type-19345
has_HealthIssueDescription_GID-74784_Type-74784
has_HealthIssueEndDate_GID-80741_Type-74784
has_HealthIssueID_GID-73293_Type-74784
has_HealthIssueStartDate_GID-80737_Type-74784
has_HealthIssueStatus_GID-74023_Type-74784
has_MedicationAmount_GID-28158_Type-3601
has_MedicationDescription_GID-20520_Type-3601
has_MedicationDosage_GID-73216_Type-3601
has_CityID_GID-73293_Type-45969 — http://knowdi
Annotations
Usage
Annotations: has_CityID_GID-73293_Type-45969
Annotations
Characteris
Description: has_CityID_GID-7
Functional
Equivalent To
SubProperty Of
Domains (intersection)
City_GID-45969
Ranges
xsd:string
Disjoint With


Active ontology
Entities
Individuals by class
DL Query
Classes
Object properties
Data properties
Annotation properties
Datatypes
Individuals
Object property hierarchy: owl:topObjectProperty
Asserted
owl:topObjectProperty
has_appointedIn_GID-112595_Type-53621
has_at_GID-132_Type-6350
has_authorizedIn_GID-28495_Type-35583
has_before_GID-114382_Type-5157
has_checks_GID-31761_Type-53621
has_conditions_GID-36419_Type-17637
has_conduct_GID-103826_Type-6350
has_diagnosedWith_GID-103532_Type-6350
has_hasActive_GID-81706_Type-55936
has_locatedIn_GID-779_Type-19345
has_prescribedTo_GID-34156_Type-3601
has_providedBy_GID-106086_Type-3601
has_record_GID-35578_Type-55936
has_recorded_GID-35630_Type-6350
has_referredTo_GID-105258_Type-6350
has_registeredIn_GID-105155_Type-75450
has_residentOf_GID-56571
has_searchesThrough_GID-4877_Type-53621
has_spatialPartOf_GID-46452
has_take_GID-103641_Type-55936
has_tourist_GID-57726_Type-55936
has_treatmentOf_GID-3415
has_visited_GID-45307_Type-55936

Data Integration

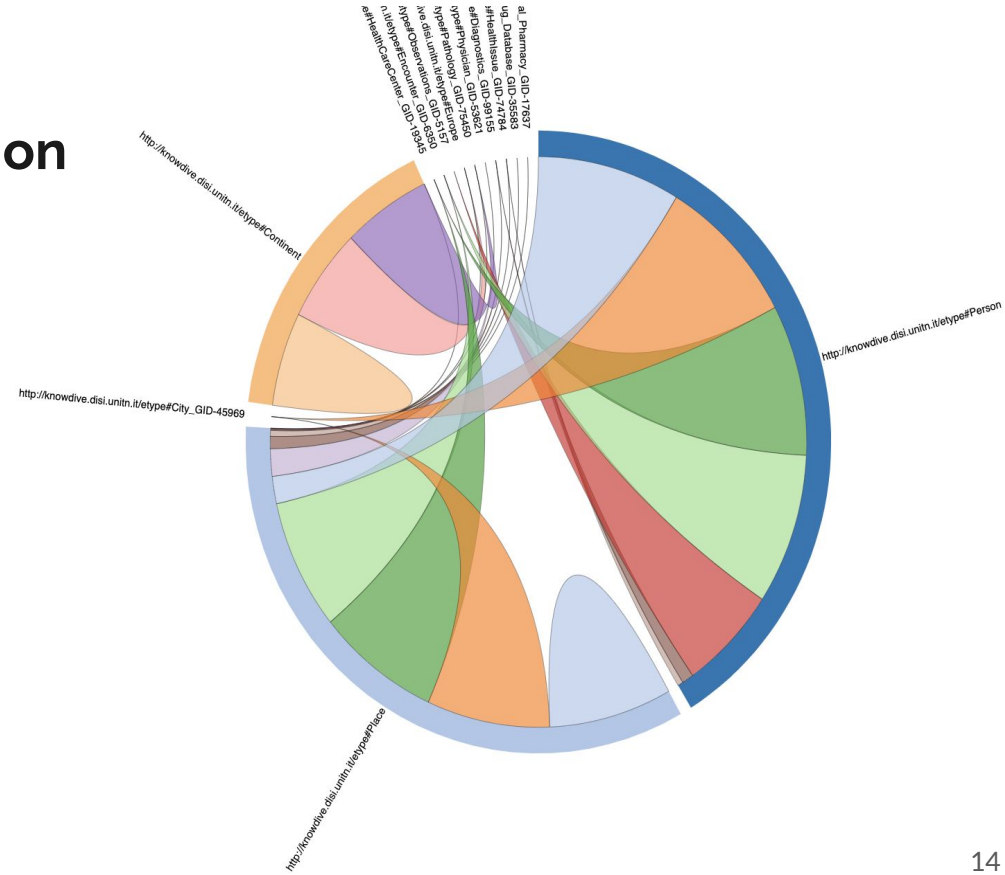
- Formal teleontology-Data Integration
  - .ttl files produced for 12 classes (.owl file) and its corresponding dataset (.csv file)
- URIs - for each class
- Relationship links to URIs of another class with defined object properties

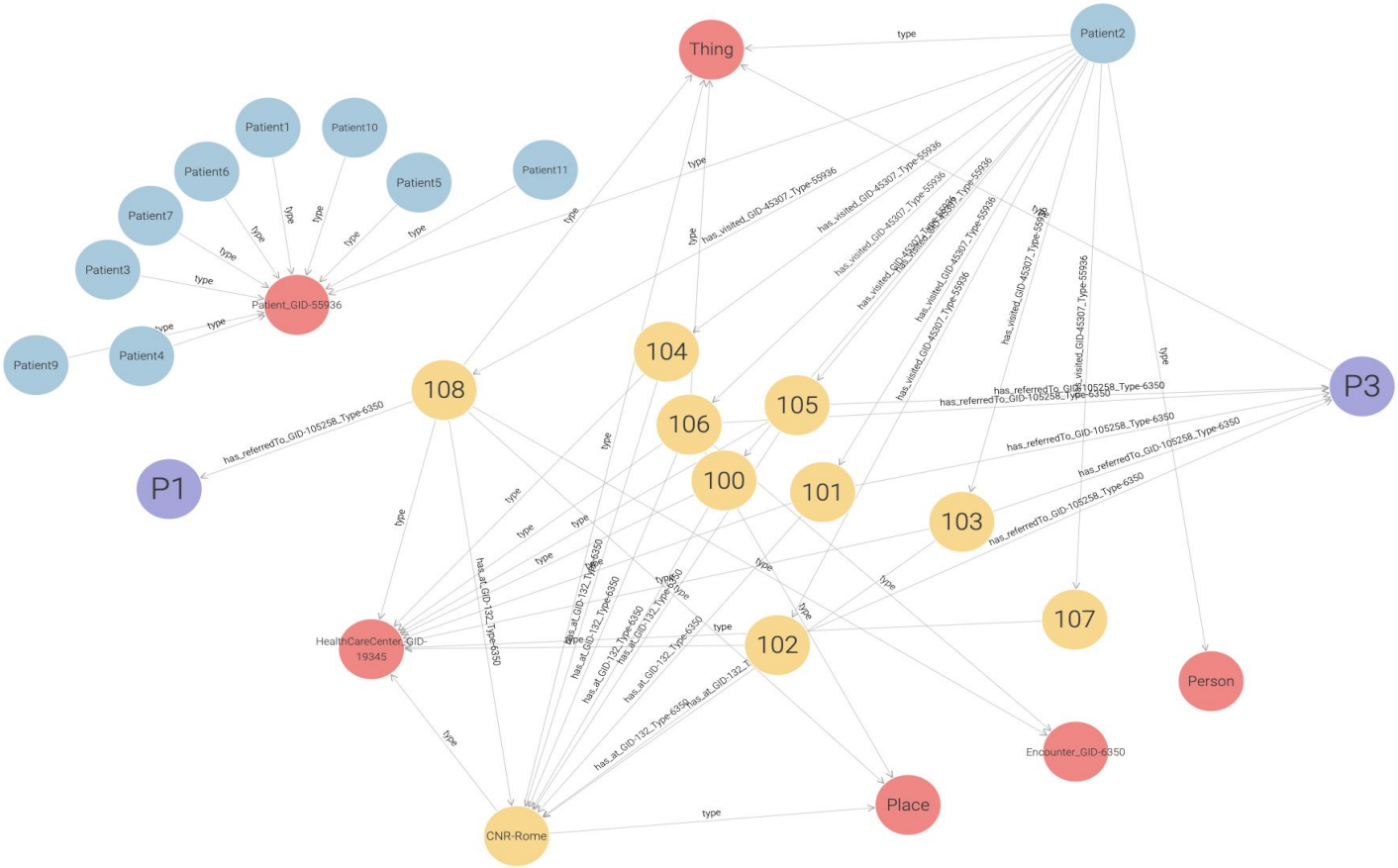


id-encounter	patient-id	id-patient_karma	id-physician	description	date	time	Type_of_visit	Visit_Description	HCCname
1	1.0	Patient1	P1	visita MMG	2010-01-01	12:00			CNR-Rome
2	1.0	Patient1	P2	misurazione pressione	2019-01-01	11:00			CNR-Rome

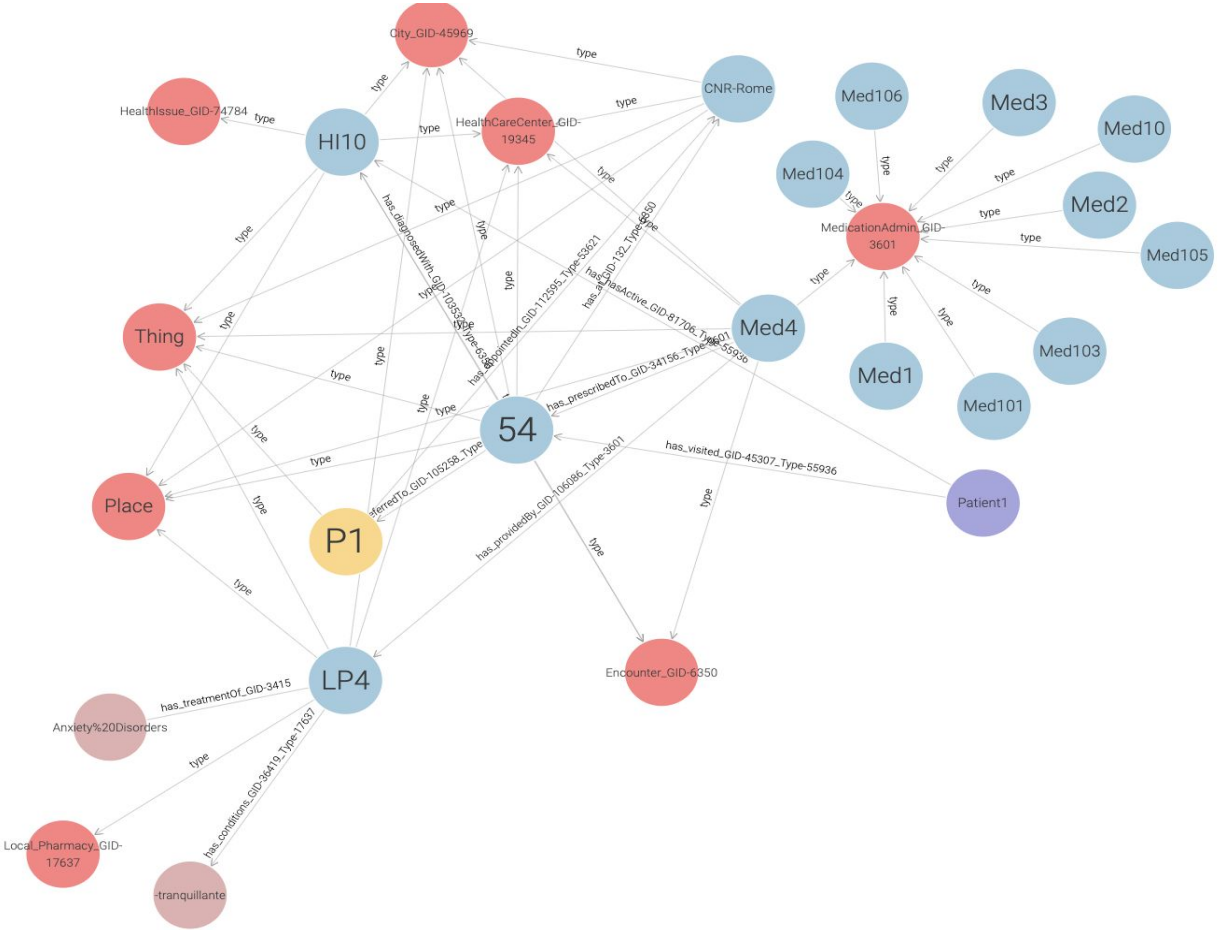
KG Construction & Exploitation

- GraphDB
  - 12 .ttl files imported
  - Linked Relationships
    - KG construction
  - Visualised graph for different classes.
  - SPARQL queries tested for each CQs





Knowledge graph exploitation for Patient: The Patient2 (blue) connects with P1 (purple), i.e., physician or MMG and P3 (purple), i.e. diagnostic tests through Encounters labelled in number 108 (yellow) to P1 and other Encounters (yellow) connected to P3 respectively. The several relationships are denoted on the edges. All encounters are linked to CNR-Rome (yellow): Healthcare Center.



KG exploitation for Medication administration : Knowledge graph from teleontology integrated with the medication data of Patient1. The Med4 (blue) medication is administered to the patient at Encounter 54 (blue) with local pharmacy drug ID, LP4 (blue) for a given healthissue, HI10 (blue) the encounter number on the prescription by physician or MMG, P1 (yellow).

CQs tested with SPARQl query

Competency Question (CQ) : Pharmacist is looking for an active component/ingredient of the prescribed drug for the treatment of hypertension.

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
PREFIX etype:<http://knowdive.disi.unitn.it/etype#>

select ?drugname ?disease ?activeComponent where
{
  ?drugid a etype:Drug_Database_GID-35583.
  ?drugid etype:has_treatmentOf_GID-3415 ?disease.
    filter(?disease= <http://localhost:8080/source/Hypertension>)
    ?drugid etype:has_DrugName_GID-17633_Type-35583 ?drugname.
  ?drugid etype:has_DrugActiveIngredient_GID-19510_Type-35583 ?activeComponent.
}limit 5
```

	drugname	disease	activeComponent
1	"Exforge"	http://localhost:8080/source/Hypertension	"valsartan, amlodipine (as amlodipine besilate)"
2	"Dafiro"	http://localhost:8080/source/Hypertension	"amlodipine, valsartan"
3	"Copalia"	http://localhost:8080/source/Hypertension	"valsartan, amlodipine (as amlodipine besilate)"
4	"Exforge HCT"	http://localhost:8080/source/Hypertension	"valsartan, hydrochlorothiazide, Amlodipine besilate"

Competency Question: Hospital staff retrieves Andre email address to send him the diagnostic test results.

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
PREFIX etype:<http://knowdive.disi.unitn.it/etype#>

select * where
{
  ?patientId a etype:Patient_GID-55936.
    optional{?patientId etype:has_PatientFirstName_GID-2_Type-55936 ?firstName.}
    filter(?firstName ="Andre").
  ?patientId etype:has_PatientEmail_GID-105296_Type-55936 ?patientEmail.
}
```

Run

keyboard shortcuts

Table | Raw Response | Pivot Table | Google Chart

Download as

Filter query results

Showing results from 1 to 1 of 1. Query took 0.1s, moments ago.

	patientId	firstName	patientEmail
1	http://localhost:8080/source/Patient5	"Andre"	"Andre610.Murray856@gmail.com"

CQ : Physician wants to refer to the last medication prescribed to Jayson who is diagnosed with Acute bronchitis

Unnamed | CQ4_updated

```
select ?firstName ?encDate ?healssId ?heaDesc ?heaStatus ?MedDesc where
{
  ?patientId a etype:Patient_GID-55936.
    optional{?patientId etype:has_PatientFirstName_GID-2_Type-55936 ?firstName.}
    filter(?firstName ="Jayson")
   ?patientId etype:has_visited_GID-45307_Type-55936 ?encId.
   ?encId etype:has_EncounterDate_GID-80737_Type-6350 ?encDate.
   ?encId etype:has_diagnosedWith_GID-103532_Type-6350 ?healssId.
   ?healssId etype:has_HealthIssueDescription_GID-74784_Type-74784 ?heaDesc.
    ?healssId etype:has_HealthIssueStatus_GID-74023_Type-74784 ?heaStatus.
    filter(?heaStatus="active")
    optional{
        ?MedId etype:has_prescribedTo_GID-34156_Type-3601 ?encId.
        ?MedId etype:has_MedicationDescription_GID-20520_Type-3601 ?MedDesc.}
```

Run

Press Alt+Enter to keyboard shortcuts

Table | Raw Response | Pivot Table | Google Chart

Download as

Filter query results

Showing results from 1 to 1 of 1. Query took 0.1s, minutes ago.

	firstName	encDate	healssId	heaDesc	heaStatus	MedDesc
1	"Jayson"	"2019-04-23"	http://localhost:8080/sou	"Acute bronchitis"	"active"	"Acetaminophen 325 MG Oral Tablet"

CQ: Clinician searches through the list of approved drugs that are available in the local pharmacy for hypertension.

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

select * where
{
  ?pharmacyid a etype:Local_Pharmacy_GID-17637.
   ?pharmacyid etype:has_treatmentOf_GID-3415 ?disease.
    filter(?disease= <http://localhost:8080/source/Hypertension>)
        ?pharmacyid etype:has_PharmacyDrugName_GID-17633_Type-17637 ?drug.
        ?pharmacyid etype:has_PharmacyDrugAvailability_GID-26190_Type-17637 ?availability.
    optional{
        ?pharmacyid  etype:has_conditions_GID-36419_Type-17637 ?drugDB.
        ?drugDB etype:has_DrugAuthorization_GID-28495_Type-35583 ?Approval.
    }
}
```

Run

keyboard shortcuts

Table | Raw Response | Pivot Table | Google Chart

Download as

Filter query results

Showing results from 1 to 3 of 3. Query took 0.1s, moments ago.

	pharmacyid	disease	drug	availability	drugDB	Approval
1	http://localhost:8080/sou	http://localhost:8080/sou	"blopress"	"Available"		
2	http://localhost:8080/sou	http://localhost:8080/sou	"blopress"	"Available"		
3	http://localhost:8080/sou	http://localhost:8080/sou	"Carvedilol"	"Unavailable"		

CQ: Patient1 records routine health observations and wants to track the whole record.

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

select ?encDate ?obstype ?obsvalue where
{
    ?patientId a etype:Patient_GID-55936.
     filter(?patientId =<http://localhost:8080/source/Patient1>)
    ?patientId etype:has_record_GID-35578_Type-55936 ?obsId.
    ?obsId etype:has_ObservationType_GID-5157 ?obstype.
    ?obsId etype:has_ObservationValue_GID-31912_Type-5157 ?obsvalue.
    ?obsId etype:has_recorded_GID-35630_Type-6350 ?encId.
    ?encId etype:has_EncounterDate_GID-80737_Type-6350 ?encDate.
}ORDER BY DESC (?encDate) limit 10
```

Table | Raw Response | Pivot Table | Google Chart

Download as

Filter query results

⚠ Showing results from 1 to 10 of 10. Query took 0.1s, on 2022-12-28 at 18:53.

	encDate	obstype	obsvalue
1	"2019-03-27"	"BPmax"	"135.0"
2	"2019-03-27"	"BPmin"	"82.0"
3	"2019-03-26"	"BPmax"	"136.0"
4	"2019-03-26"	"BPmin"	"85.0"
5	"2019-03-25"	"BPmax"	"137.0"

Scenarios	Personas	Competency Questions
Hospital management who needs to manage a patient	Hospital Management	Hospital staff wants to look for an appointment slot for Matteo with a physician next week.
	Hospital Management	Hospital staff retrieves Andre email address to send him the diagnostic test results.
	Hospital Management	Hsiu is diagnosed with obesity and connects with receptionist to schedule an appointment with the clinician for routine follow up.
Physician who needs to provide patients with principal care and consultancy	Physician	Physician wants to refer to the last medication prescribed to Jayson who is diagnosed with Acute bronchitis.
	Physician	Clinician searches through the list of approved drugs that are available in the local pharmacy for hypertension.
	Physician	Physician searches through database for a drug to prescribe to Hsiu who is diagnosed with Coronary Heart Disease and has active Prediabetes comorbidity.
	Physician	Physician checks if Carvedilol, a drug for Hypertension is available in the Local pharmacy or not.
	Physician	Physician wants to search through drug database for Saxenda prescribed in which comorbidity conditions to avoid side effects of the drug.
	Test Examiner	Test Examiner needs to look for the type of test to conduct for Andre on 2017-02-04.
	Physician	Maria who is a Dutch tourist in Rome and is diagnosed with Vertigo. He asks the physician appointed in Rome CNR to prescribe a drug that is also authorized in Netherlands
	Test Examiner	Vennie has active Preeclampsia and is referred to the test examiner to conduct Evaluation of uterine fundal height.
	Physician	Clinician looks into the recorded observations of Patient 1 on his visit to MMG on 2019-01-02.
Pharmacist who needs to manage prescriptions and drugs	Pharmacist	Pharmacist is looking for an active component of the prescribed drug for the treatment of hypertension.
	Pharmacist	Pharmacist searches through database if Glidipion (previously Pioglitazone Actavis Group), a drug for Diabetes Mellitus Type 2 still available in the market or withdrawn.
	Pharmacist	Pharmacist wants to check the dosage amount of blopress prescribed to patient ID-1.
	Pharmacist	Pharmacist in Paris city is looking for an alternative of Tranquillante prescribed to an Italian citizen who is diagnosed with hypertension.
Patient who needs to process his own medical information	Patient	Clyde visited Rome CNR on 2009-07-18 to undergo a test and wants to know what health issue is she diagnosed with.
	Patient	Matteo wants to schedule an appointment with a physician next week.
	Patient	Tia who diagnosed with Idiopathic atrophic hypothyroidism records her BP level observations.
	Patient	Agueda has to follow the SoPs/guidelines to take the test for general examination.
	Patient	Mathew who is a Dutch tourist in Rome and is diagnosed with Vertigo. He asks the physician appointed in Rome CNR to prescribe a drug that is authorized in Netherlands.
	Patient	Patient 1 records routine health observations and wants to track the whole record.

Figure 13: The competency questions, CQs which were answered with the SPARQL Query. Green represents the whole CQ query was resolved, yellow represents a certain part of the CQ was resolved, red represents the unresolved CQs.

Evaluation

Table 4: Coverage evaluation.

Evaluation Type	Coverage
eType Coverage	0.8
Property Coverage	0.86

Table 5: Connectivity metrics.

Metric Type	Connectivity		
Entities	Entity type	Object Property	Data Property
City	2	2	2
Healthcare Center	2	2	3
Drug Database	2	2	9
Pathology	2	2	2
Local pharmacy	3	3	9
Physician	2	2	3
Encounter	6	6	8
Diagnostics	2	2	6
Health Issue	2	2	5
Medication Admin	2	2	7
Observations	2	2	5
Patient	4	4	11

Open Issues

- Different languages not integrated
  - Eg: a tourist visits CNR-Rome or goes to local pharmacy to buy his medicines.
- Dataset unavailability:
  - SoPs/guidelines for the diagnostics tests.
  - Appointment scheduling
  - Country specific drug database
    - eg: Farmabank in Italy
- Limited data
  - Only one hospital, one patient.

Conclusions and Future Aspects

- Nearly 77% of the CQs (17 out of 22 CQs) fetched.
  - Covered important CQs related to medical and healthcare domain.
- More datasets -> retain more information for all the patients -> help the hospital management share patients' records and health status to other hospitals for diseases whose treatment facility unavailable in CNR-Rome without loss of time or medical information.
- knowledge graphs -> help digitalization of the health care sector and devise methodologies for developing precision medicine for several patients in a fast pace.

References

1] Giunchiglia, F., Bocca, S., Fumagalli, M., Bagchi, M., Zamboni, A. (2021). iTelos-Building reusable knowledge graphs. arXiv preprint arXiv:2105.09418.

[2] Walonoski, J., Kramer, M., Nichols, J., Quina, A., Moesel, C., Hall, D., ... McLachlan, S. (2018). Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record. *Journal of the American Medical Informatics Association, 25*(3), 230-238.

[3] Whirl-Carrillo, M., Huddart, R., Gong, L., Sangkuhl, K., Thorn, C. F., Whaley, R., Klein, T. E. (2021). An evidence-based framework for evaluating pharmacogenomics knowledge for personalized medicine. *Clinical Pharmacology & Therapeutics, 110*(3), 563-572.

[4] Musen, M. A. (2015). The protégé project: a look back and a look forward. *AI matters, 1*(4), 4-12.

[5] Knoblock, C. A., Szekely, P., Ambite, J. L., Goel, A., Gupta, S., Lerman, K., ...Mallick, P. (2012, May). Semi automatically mapping structured sources into the semantic web. In *Extended Semantic Web Conference* (pp. 375-390). Springer, Berlin, Heidelberg.