

HETOR

Data Integration Project Presentation

Marco Di Panfilo, Marco Zenere

Table of Contents

1. Introduction
 - 1.1. Motivation Scenario
 - 1.2. Domain Specification
2. System Architecture
3. Data Cleaning
4. Database
5. Ontology
6. Mapping
7. Application / Demo

Motivation Scenario

The Italian cultural heritage is one of the largest around the world

Cultural heritage is a strategic value for the social and economic development of a country

HETOR is a project for promoting the cultural heritage in Campania



Domain Specification

HETOR comprises several domains related to the cultural heritage in Campania

Itineraries domain:

Itineraries - Caserta [CSV]

Itineraries - Nocera Inferiore [CSV]

Itineraries - Nocera [CSV]

Intineraties - Costiera Amalfitana [CSV]

Categoria itinerario	Nome	Comune	Località	Via	Cap	Provincia	Regione	Tipologia	Secolo di costruzione
culturale	Monumento ai caduti Caserta	Caserta	Monumento	Corso Trieste	81100	CE	Campania	Monumento Commemorativo	XX
culturale	Monumento Luigi Vanvitelli	Caserta	Statua	Piazza Vanvitelli	81100	CE	Campania	Statua Commemorativa	XX
culturale	Cattedrale San Michele Arcangelo	Caserta	Cattedrale	Piazza Duomo	81100	CE	Campania	Edificio Religioso	XIX
culturale	Parrocchia Buon Pastore	Caserta	Chiesa	Piazza Pilesti	81100	CE	Campania	Edificio Religioso	XXI
culturale	Monumento ai Garibaldini	Santa Maria Capua Vetere	Monumento	Villa Comunale	81055	CE	Campania	Monumento Commemorativo	XIX
culturale	Le Madonelle	San Nicola la Strada	Monumento	Via Appia	81020	CE	Campania	Edificio Religioso	XX
culturale	Chiesa di Sant'Antonio di Padova	Caserta	Chiesa	Corso Giannone	81100	CE	Campania	Edificio Religioso	NULL
culturale	Complesso monumentale di San Leucio	Caserta	Monumento	Belvedere Di San Leucio	81100	CE	Campania	Complesso Monumentale	XIX
culturale	Reale Tenuta Carditello	San Tammaro	Monumento	Sp229	81050	CE	Campania	Monumento	XVII
culturale	Chiesa di San Ferdinando Re	Caserta	Chiesa	Via Atrio Superiore	81100	CE	Campania	Edificio Religioso	XVIII
culturale	Real Convitto Borbonico	San Nicola la Strada	Orfanotrofio / Convento	Villa Comunale	81020	CE	Campania	Edificio Religioso	XVIII
culturale	Casa del fascio	Caserta	Sala Conferenze	Piazza Matteotti	81100	CE	Campania	Edificio Cittadino	XX
culturale	Statua dell'Agricoltura	Caserta	Statua	Via Raffaele Gasparri	81100	CE	Campania	Statua	XIX
culturale	Basilica Del Corpus Domini	Maddaloni	Basilica	Piazza De Sivo	81024	CE	Campania	Edificio Religioso	XVIII
culturale	Statua della Carità	Marcianise	Statua	Piazza Carità	81025	CE	Campania	Statua	XIX
culturale	Monumento al milite Ignoto	Maddaloni	Monumento	Piazza Della Vittoria	81024	CE	Campania	Statua Commemorativa	XX
culturale	Teatro Ricciardi	Capua	Teatro	Largo Porta Napoli	81043	CE	Campania	Edificio Cittadino	XVIII
culturale	Monumento ai Caduti di Capua	Capua	Monumento	Via Duomo	81043	CE	Campania	Statua Commemorativa	XX
culturale	Teatro Garibaldi	Santa Maria Capua Vetere	Teatro	Corso Garibaldi	81055	CE	Campania	Edificio Cittadino	XIX
culturale	Monumento ai fratelli De Simone	Santa Maria Capua Vetere	Monumento	Incrocio Tra Via Giuseppe Sirtori E Via Uccella	81055	CE	Campania	Statua Commemorativa	XX
culturale	Monumento a Domenico Cimarosa	Aversa	Monumento	Piazza Mazzini	81031	CE	Campania	Statua Commemorativa	XX
culturale	Monumento a Pietro Rosano	Aversa	Monumento	Villa Comunale	81031	CE	Campania	Statua Commemorativa	XX
culturale	Monumento ai Caduti In Guerra	Aversa	Monumento	Piazza Municipio	81031	CE	Campania	Statua Commemorativa	XX
culturale	Chiesa di San Luca Evangelista	Casapulla	Chiesa	Via Appia	81020	CE	Campania	Edificio Religioso	XX
culturale	Statua Della Carità	Marcianise	Monumento	Piazza Carità	81025	CE	Campania	Statua	XIX
culturale	Torre Dell'Orologio	San Prisco	Torre	Via Michele Monaco	81054	CE	Campania	Edificio Cittadino	XVIII
culturale	Convitto Nazione Giordano Bruno	Maddaloni	Scuola	Via San Francesco D'Assisi	81024	CE	Campania	Ex Edificio Cittadino	XIX
culturale	Santuario di San Salvatore	Orta di Atella	Santuario	Via San Salvatore	81030	CE	Campania	Edificio Religioso	XXI

System Architecture



Data Cleaning

Cleaning of the dataset through the extract, transform, and load process (ETL)

Alignment of municipalities to ISTAT names

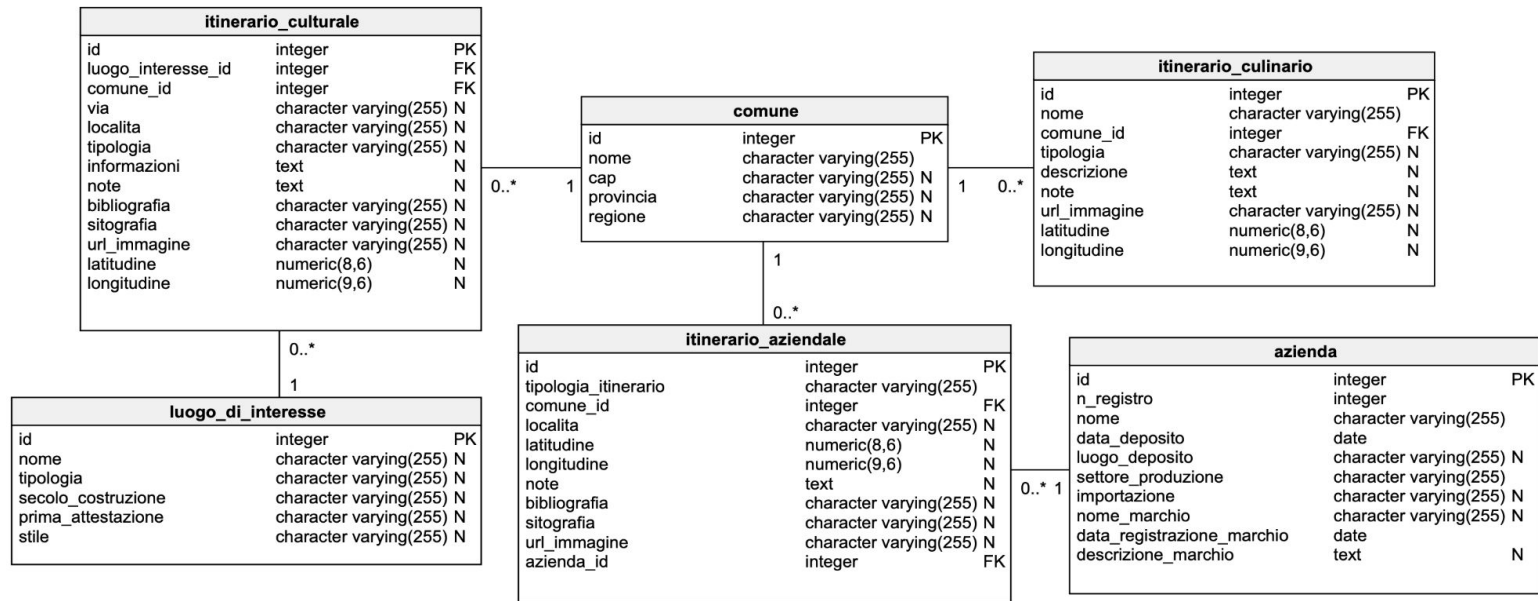
- String matching technique (Perfect match or Levensthein-distance)

- Manual correction

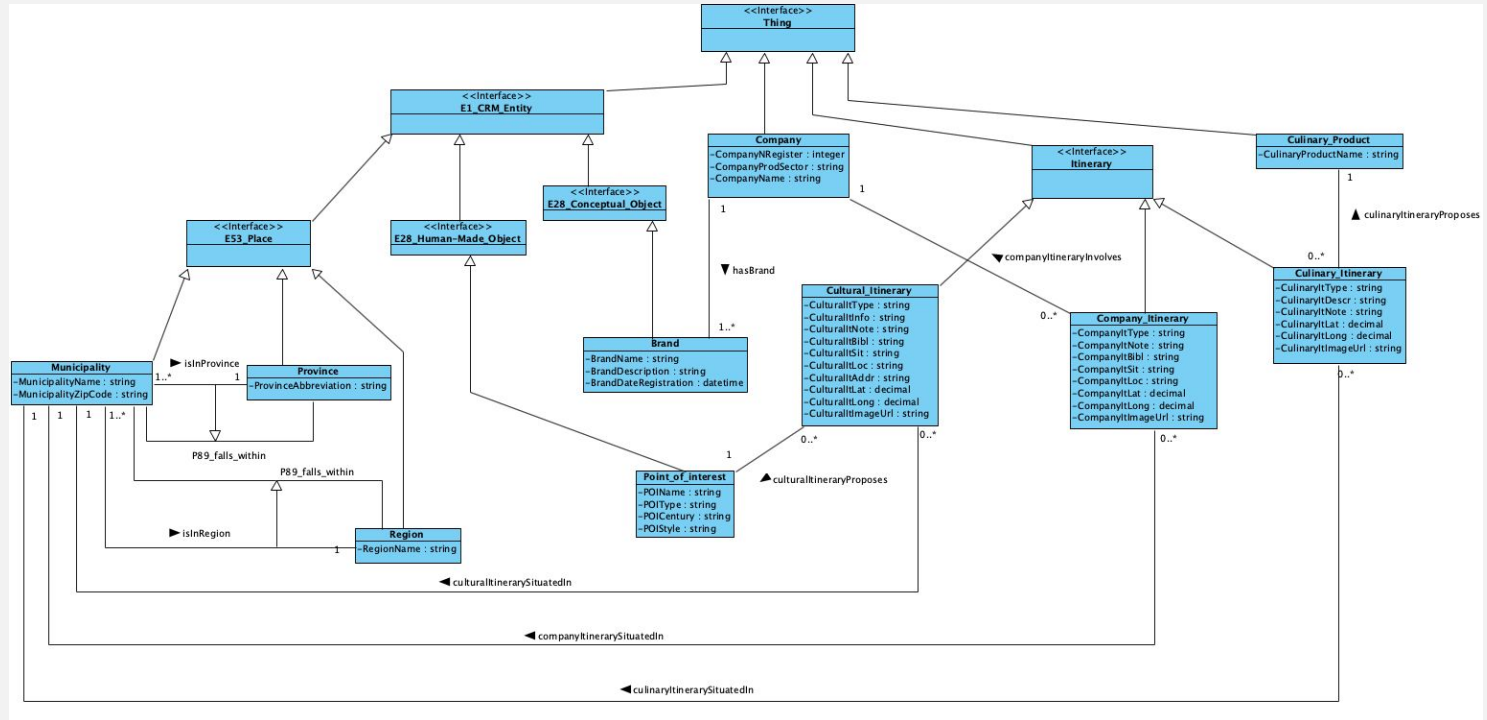
Fixing of latitude and longitude geographical coordinates

Empty values or other form of them to NULL values (Consistency Check)

Database



Ontology



Mapping

Two **Mapping Pattern** has been followed:

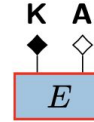
- Entity, like Itineraries

Relational schema and constraints:

$$T_E(\underline{\mathbf{K}}, \mathbf{A})$$

Mapping assertion:

$$\begin{aligned} s &: T_E \\ t &: C_E(\mathbf{E}/\{\mathbf{K}\}), \\ &\quad \{d_A(\mathbf{E}/\{\mathbf{K}\}, \{A\})\}_{A \in \mathbf{K} \cup \mathbf{A}} \end{aligned}$$



Ontology axioms:

$$\{\exists d_A \sqsubseteq C_E\}_{A \in \mathbf{K} \cup \mathbf{A}}$$

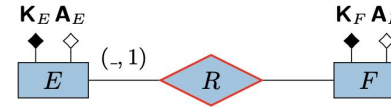
- Relationship with Merging,

Relational schema and constraints:

$$\begin{aligned} T_F(\underline{\mathbf{K}_F}, \mathbf{A}_F) \\ \uparrow \\ T_E(\underline{\mathbf{K}_E}, \mathbf{K}_{EF}, \mathbf{A}_E) \end{aligned}$$

Mapping assertion:

$$\begin{aligned} s &: T_E \\ t &: p_R(\mathbf{E}/\{\mathbf{K}_E\}, \mathbf{F}/\{\mathbf{K}_{EF}\}) \end{aligned}$$



Ontology axioms:

$$\begin{aligned} \exists p_R \sqsubseteq C_E \\ \exists p_R^- \sqsubseteq C_F \end{aligned}$$

Mapping

Two **Mapping Pattern** has been followed:

- Entity, e.g. Itineraries, Municipality

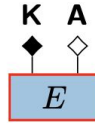
Relational schema and constraints:

$T_E(\underline{K}, A)$

Mapping assertion:

$s : T_E$

$t : C_E(:E/\{K\}),$
 $\{d_A(:E/\{K\}, \{A\})\}_{A \in KUA}$



Ontology axioms:

$\{ \exists d_A \sqsubseteq C_E \}_{A \in KUA}$

Edit Triples Map

Mapping ID: Municipality

Target (Triples Template):
:municipality/{id} a :Municipality ; :MunicipalityName {nome}^^xsd:string ; :MunicipalityZipCode {cap}^^xsd:string .

Source (SQL Query):
select id, nome, cap
from comune;

SQL Query results:

Execute the SQL query (100 rows)

Update Abbrechen

Mapping

Two **Mapping Pattern** has been followed:

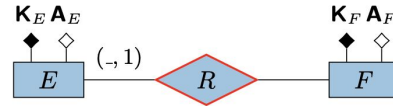
- Relationship with Merging, e.g. `culturalItinerarySituatdIn`

Relational schema and constraints:

$T_F(\underline{K_F}, \underline{A_F})$
 $T_E(\underline{K_E}, \underline{K_{EF}}, \underline{A_E})$

Mapping assertion:

$s : T_E$
 $t : p_R(\cdot E / \{K_E\}, \cdot F / \{K_{EF}\})$



Ontology axioms:

$\exists p_R \sqsubseteq C_E$
 $\exists p_R^- \sqsubseteq C_F$

Edit Triples Map

Mapping ID: `Object property: culturalItinerarySituatdIn`

Target (Triples Template):
`:culturalItinerary/{it_cult_id} :culturalItinerarySituatdIn :municipality/{c_id} .`

Source (SQL Query):

```
select it_cult.id as it_cult_id, it_cult.comune_id as c_id
from itinerario_culturale as it_cult
```

SQL Query results:

Execute the SQL query (100 rows)

Update Abbrechen

Application/Demo



- Low code Python web framework
- Fast prototyping, with predefined methods

Trova il tuo prossimo itinerario!



Filters Compila il modulo e trova i migliori suggerimenti per te. **Map**

Regione:

Provincia:

Comune:

Categoria: ☐ aziendale ☐ culinario ☐ culturale

Cerca nel titolo:

Cerca nella descrizione:

Result short

Itinerari trovati:

Itinerari aziendali: 47
Itinerari culinari: 9
Itinerari culturali: 106

Result long

Itinerari aziendali:

	Immagine	Nome azienda	Settore di produzione	Nome marchio	Località	Comune	Provincia
1		SOCIETÀ ANONIMA NOCERINA CONSERVE ALIMENTARI [...]	Agro alimentare	Pantera	Mercato, Arenula	Nocera Inferiore	SA



- SPARQL Query languages IS used to retrieve informations