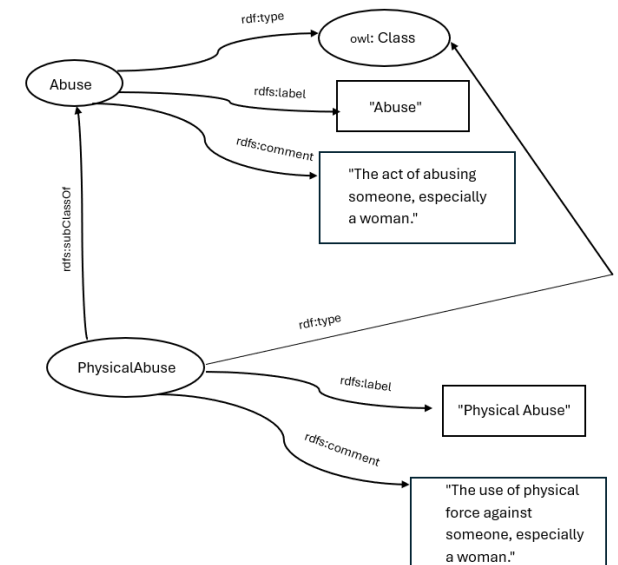
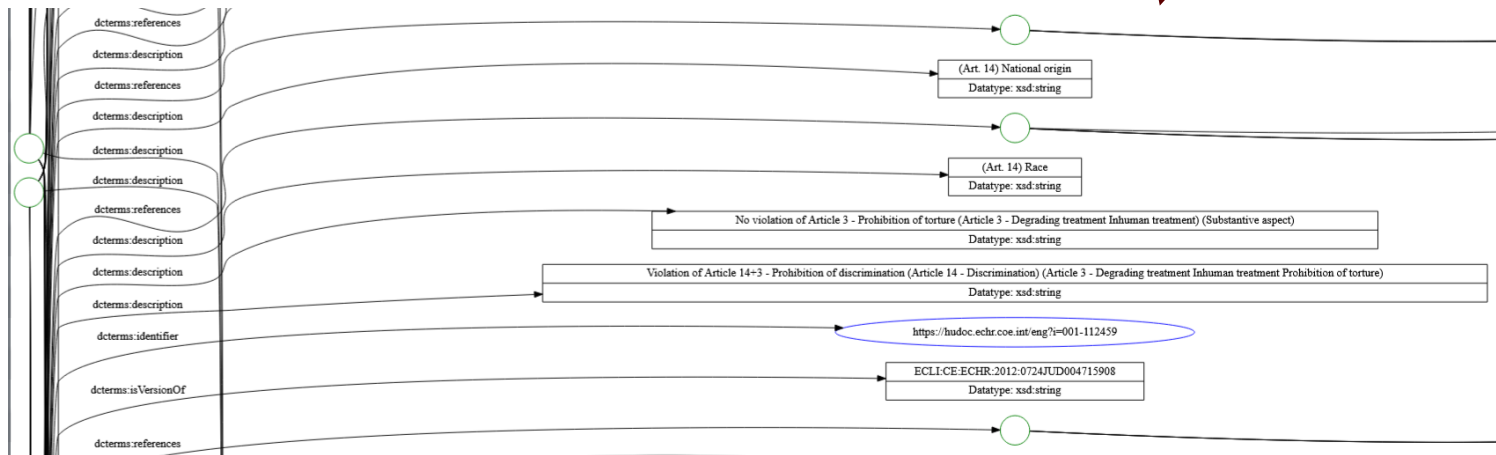

INTEGRAZIONE DI KGS IN AMBITO GIURIDICO PER I CASI DI VIOLENZA SULLE DONNE

di Sara Padovano

CONTESTO E OBIETTIVO

Contesto: Costruzione di un Legal Knowledge Graph per i casi di violenza sulle donne

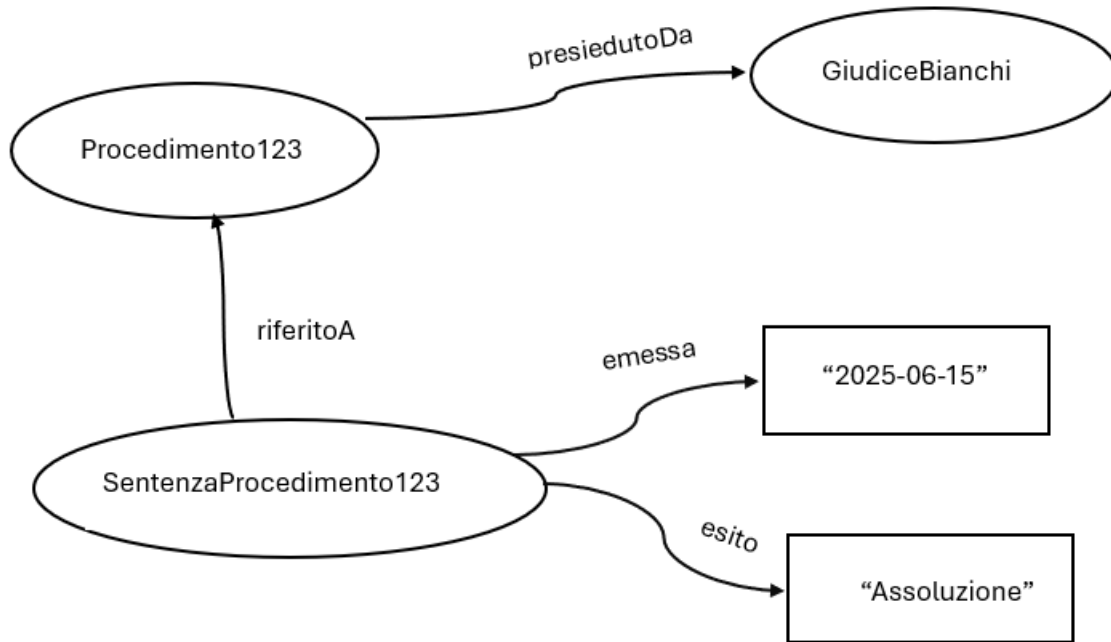
Obiettivo: Integrazione dei due Knowledge Graphs (KGs) già creati in questo contesto, al fine di avere un unico KG interrogabile e onnicomprensivo



DEFINIZIONI BASE

Un Knowledge Graph è un grafo di dati inteso a rappresentare la conoscenza del mondo reale:

- conforme a un modello di dati basato su grafi
- i nodi rappresentano entità del dominio di appartenenza
- gli archi (o collegamenti) rappresentano diverse relazioni tra queste entità
- grafo di dati potenzialmente arricchito con uno schema:
 - ontologie impiegate per definire e ragionare sulla semantica dei nodi e degli archi



INTEGRAZIONE DI KGS: STATO DELL'ARTE

Utilizzo LLM:

- Vasta conoscenza pregressa;
- Domini specializzati.

1. Integrazione modulare¹

CONTRO:

- Ontologie non modulari;
- Utilizzo non automatizzato che richiede l'aiuto di esperti;

2. Integrazione per intero²

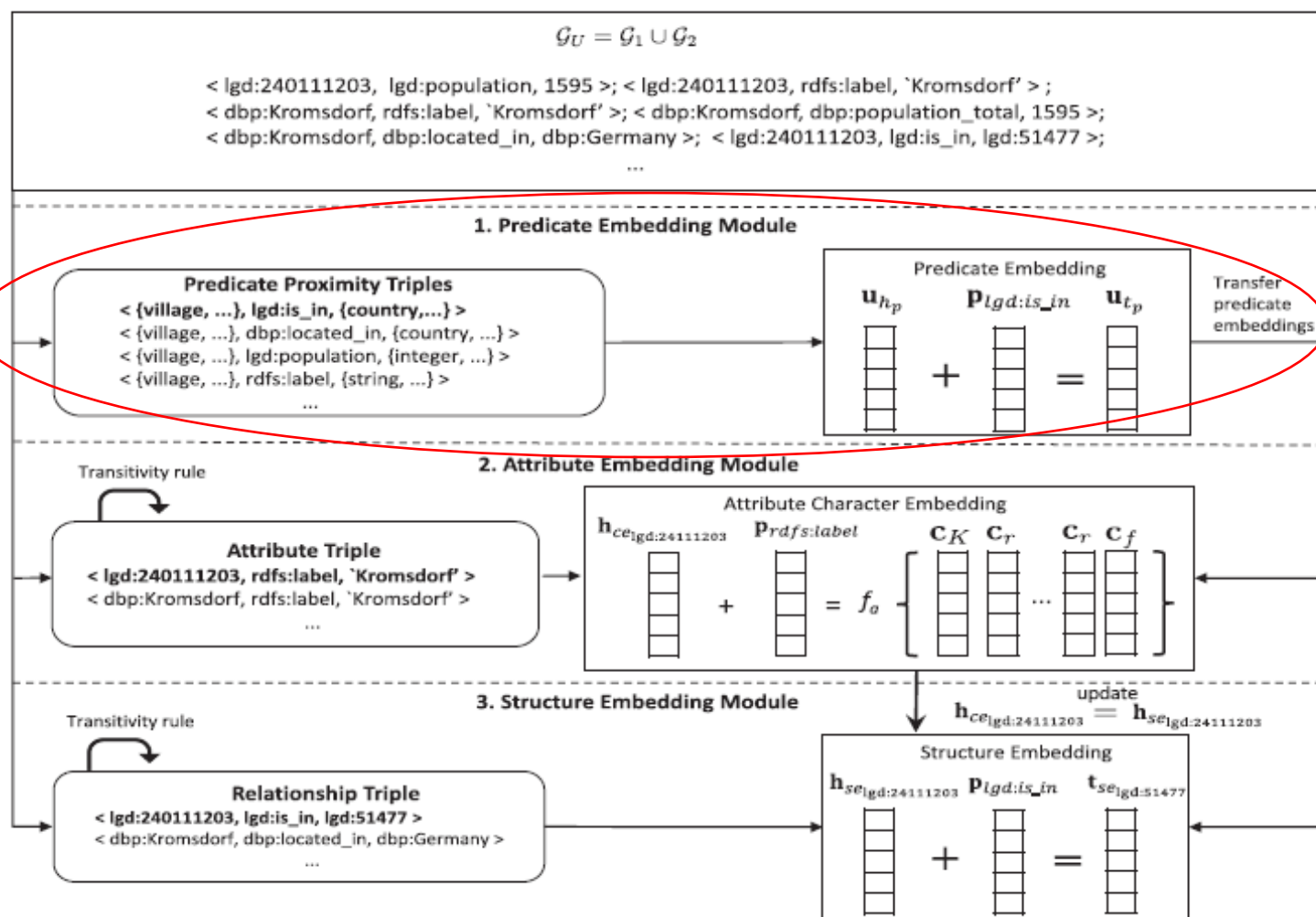
Si sviluppano sistemi per l'integrazione per intero dei KGs, senza moduli. Tra questi sistemi sviluppati vi è **AutoAlign**³

1- Cogan Shimizu and Pascal Hitzler. *Accelerating knowledge graph and ontology engineering with large language models*. Journal of Web Semantics, 2025

2- Carl Yang, Ran Xu, Linhao Luo, and Shirui Pan. *Knowledge Graph and Large Language Model Co-learning via Structure-oriented Retrieval Augmented Generation*. IEEE Data Eng. Bull., 2024

3- Rui Zhang, Yixin Su, Bayu Distiawan Trisedya, Xiaoyan Zhao, Min Yang, Hong Cheng, and Jianzhong Qi. *AutoAlign: Fully Automatic and Effective Knowledge Graph Alignment enabled by Large Language Models*. 2023.

AUTOALIGN: MAIN ARCHITECTURE



APPROCCIO

Concetto base: predicati simili hanno tipi di entità simili.

Grafo di prossimità dei predicati è un grafo in cui al posto delle entità sono presenti i rispettivi tipi.

dai due insiemi di tipi ritrovati (uno per KG)

Creazione del prompt per ritrovare i tipi simili:

Now you are an expert in linguistics and knowledge graphs. I will give you two sets of words, indicating the entity types from two knowledge graphs. You need to identify all the word pairs from the two sets that are synonyms. For example, if the first set has the word 'people' and the second set has the word 'person', you need to identify the two words being synonyms and return me the pair (people, person) or for example, if you have StrasbourgCaseLaw in the first set and LegalCase in the second (StrasbourgCaseLaw, LegalCase). Pay attention to details because for example decision and LegalJudgment aren't synonyms and don't consider literal types. Now the following are the two sets: Set 1: {list_to_str(typeset1)} Set 2: {list_to_str(typeset2)} Please return all the pairs that are synonyms from the two sets regarding entity types. Do not output the pairs if they are exactly the same. Remember you only need to return the pairs, each pair in one line. Each pair contain two types, one from Set 1 and another from Set 2, in the format (type1, type2). Please pay very attention that they have to be synonyms or have a strong semantic connection. For example (jurist, perpatator) are not synonyms. Remember the synonyms have to be from the two different sets not from the same, the pair must be (type1, type2) not (type1, typw1) neither (type2,type2). Only (type1, type2).

LLM

prompt

GPT ✓

google/flan-t5-small ✗

google/flan-t5-large ✗

ritrovati i sinonimi

Sostituzione dei sinonimi ritrovati

Embedding

✗

VALUTAZIONE QUALITATIVA

- Utilizzo LLM locale con *google/flan-t5-small*:

```
Consequence, Victim
```

- Utilizzo LLM locale con *google/flan-t5-large*:

```
hasConsequence, hasLegalJudgment, hasConsequence, hasLegalJudgment, hasContext
```

- Utilizzo modello di embedding *all-MiniLM-L6-v2* con threshold pari a 0,54:

```
(StrasbourgCaseLaw, LegalCase)  
(jurist, LegalJudgment)  
(jurist, LegalCase)  
(legal case, hasLegalCase)  
(legal case, LegalJudgment)  
(legal case, LegalArticle)  
(legal case, LegalCase)  
(court decision, LegalCase)  
(legal profession, LegalJudgment)  
(legal profession, LegalArticle)  
(legal profession, LegalCase)
```

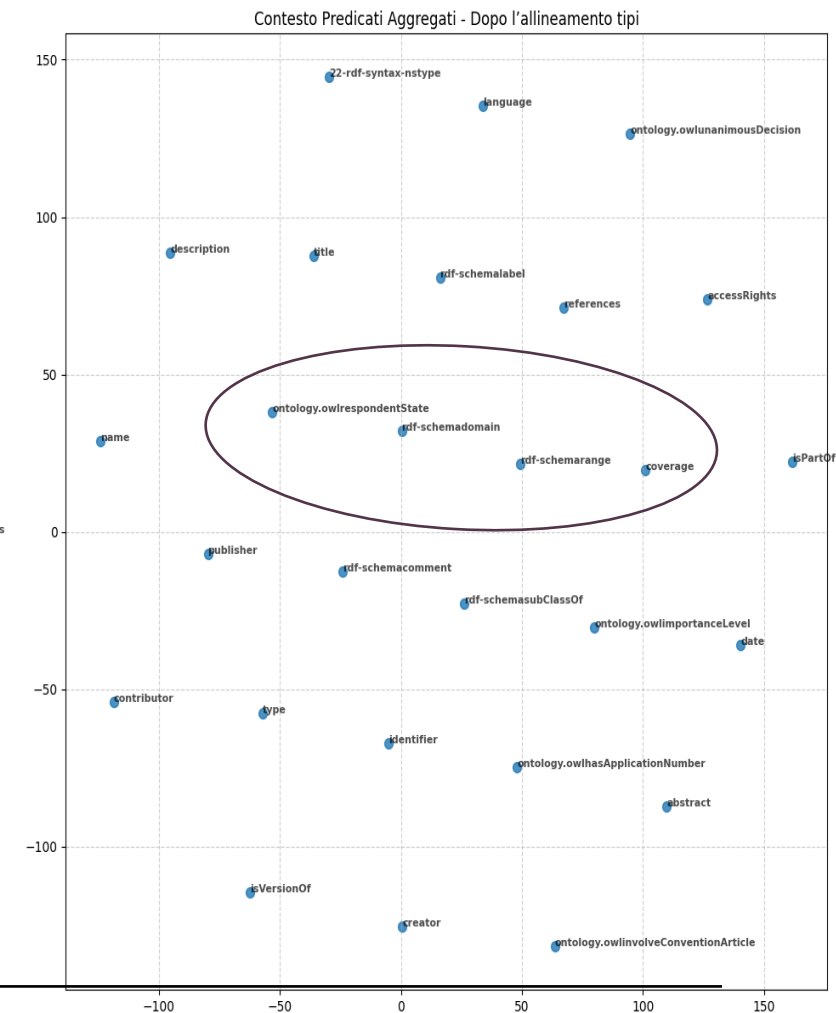
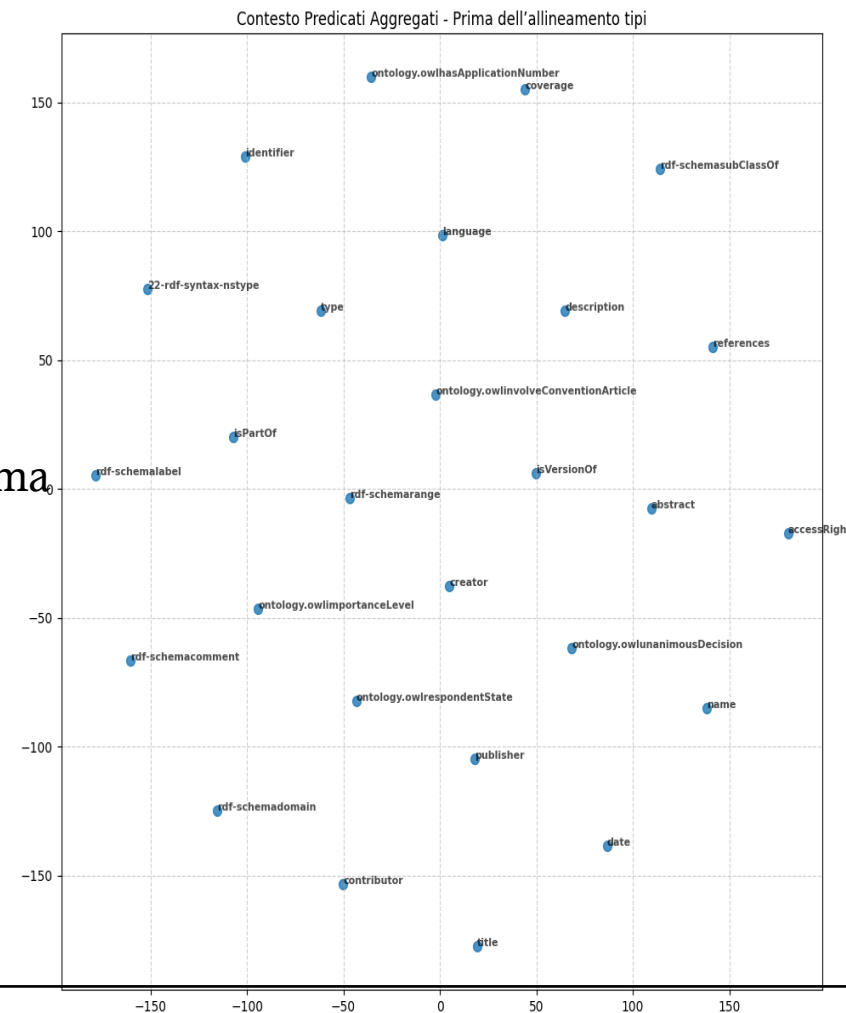
È stato quindi deciso di utilizzare e mantenere i sinonimi ritrovati con *GPT*:

```
(judgment, LegalJudgment)
(court decision, LegalJudgment)
(legal case, LegalCase)
(StrasbourgCaseLaw, LegalCase)
(DomesticLaw, LegalArticle)
(InternationalLaw, LegalArticle)
(written work, LegalArticle)
```

```
=== Risultato Analisi Riduzione Tipi ===
Numero di tipi unici prima dell'allineamento: 100
Numero di tipi unici dopo l'allineamento: 93
Riduzione assoluta del numero di tipi: 7
Riduzione percentuale: 7.00%
Tipi unificati (sostituiti): ['DomesticLaw', 'InternationalLaw', 'StrasbourgCaseLaw', 'court decision', 'judgment', 'legal case', 'written work']
=====
```


VALUTAZIONE QUALITATIVA

Grafico sulla prossimità dei predicati prima e dopo l'allineamento dei tipi:



CONCLUSIONI E SVILUPPI FUTURI

Conclusioni:

- è stata investigata l'integrazione di KGs tramite AutoAlign applicando al dominio legale concentrandosi sulla creazione del grafo di prossimità dei predicati;
- eseguita un'analisi qualitativa che ha mostrato dei primi risultati positivi ma contenuti considerando il dominio complesso.

Sviluppi futuri:

- ampliare gli step successivi di AutoAlign;
- utilizzo di una valutazione quantitativa sull'integrazione dei due knowledge graphs su ampia scala;
- ulteriori LLM.

**GRAZIE
PER
L'ATTENZIONE**