

Università degli Studi di Torino

Corso di Laurea Magistrale in Informatica



Semantic Web e Linked Data

Litetary Work Topics

Relatore: Prof. Damiano Rossana

Candidato: Enrico Chiesa

Mat.918206

Candidato: Davide Vitturini

Mat.997078

30 Gennaio 2025

Contents

1	Motivazioni	3
2	Requisiti	6
3	Descrizione del dominio	10
4	Competency Questions	13
5	Documentazione	15
5.1	Documentazione sul Dominio Letterario	15
5.1.1	Fonti Informali e Standard Esistenti	15
5.1.2	Dati di Esempio	18
5.1.3	Validatori	21
6	Visualizzazione	24

6.1	Esempio di Libro	25
7	Query	26
8	Sparql Anything	35
8.1	Fonti di dati	35
8.2	Triplificazione dei dati con Sparql Anything	37
8.2.1	Triplificazione da JSON: estrazione strutturata . . .	37
8.2.2	Triplificazione da CSV: gestione delle intestazioni . .	41
8.3	Popolamento dell'ontologia	43
9	Data Availability Statement	44

Chapter 1

Motivazioni

L'idea di creare un'ontologia sugli argomenti letterari nasce dalla necessità di disporre di uno strumento unificato per descrivere, analizzare e condividere le tematiche ricorrenti all'interno delle opere. Questi argomenti, come l'amore, il potere, il conflitto o il viaggio, rappresentano un elemento chiave sia per interpretare correttamente i testi, sia per valorizzarne la fruizione in contesti diversi.

*Dal punto di vista **culturale** e **formativo**, classificare e collegare le opere in base ai loro temi fornisce un quadro di riferimento per studenti, docenti e ricercatori che desiderano studiare la circolazione di determinate idee o motivi letterari. In questo modo, è più semplice riconoscere come uno stesso argo-*

mento possa essere declinato in epoche, generi o tradizioni letterarie differenti, mettendo in luce evoluzioni storiche, adattamenti e innovazioni stilistiche.

A livello **professionale**, un modello ontologico potrebbe rappresentare standard di sistemi di classificazione già esistenti nelle grandi catene librerie (come Mondadori) o negli ambienti bibliotecari. Sebbene sia già prassi comune organizzare i testi in base a generi e filoni tematici, un approccio strutturato e coerente a livello semantico offre la possibilità di mettere in relazione gli argomenti in modo dinamico e flessibile, facilitando sia la creazione di cataloghi tematici sia l'organizzazione di eventi culturali (come mostre, rassegne o letture guidate). Inoltre, anche i lettori non specialisti e gli studiosi occasionali traggono vantaggio da un sistema che permette di esplorare le opere seguendo i temi di interesse, a prescindere dalla loro appartenenza a un preciso genere o periodo storico, potendo così scoprire connessioni e percorsi di lettura spesso invisibili nelle tradizionali classificazioni.

Inoltre, la struttura proposta rimane aperta a futuri ampliamenti e integrazioni, come suggerito dall'approccio NeOn: lo stesso schema ontologico, se opportunamente esteso, potrebbe risultare utile non solo per la letteratura in senso stretto, ma anche per altre forme artistiche (pittura, cinema, teatro) che condividono tematiche, motivi o simboli. Questa prospettiva permette di indi-

viduare e collegare gli stessi argomenti in contesti espressivi diversi, offrendo un catalogo multi-modale di entità ricollegabili ad un unico concetto.

Chapter 2

Requisiti

*La progettazione dell'ontologia sugli argomenti letterari ha come **finalità generale** quella di fornire una struttura unificata e condivisa per descrivere, ricercare e collegare i temi ricorrenti all'interno delle opere. L'obiettivo è rendere la classificazione e l'accesso a tali argomenti più agevole e accurato, sia per i professionisti del settore (bibliotecari, editori, curatori) sia per un pubblico eterogeneo (studenti, docenti, semplici appassionati).*

*Dal punto di vista dei **task specifici**, l'ontologia è pensata per:*

- **Consultazione:** *gli utenti possono individuare opere, autori o periodi storici correlati a un determinato argomento (ad esempio, “viaggio” o “memoria”). Ciò semplifica sia la ricerca bibliografica sia la*

preparazione di liste di lettura mirate (per corsi universitari, rassegne tematiche, mostre o iniziative culturali).

- **Reference e catalogazione:** *l'ontologia può fungere da base per la costruzione di sistemi di tagging semantico nei cataloghi di librerie e biblioteche. Attraverso un vocabolario controllato e relazioni ben definite, si riducono le ambiguità e si rafforza la coerenza terminologica.*
- **Verifica e integrazione:** *grazie a questa struttura, è possibile validare la coerenza dei dati catalografici (ad esempio, controllando se un argomento è trattato in un'opera indicata) e integrare fonti diverse in modo progressivo, senza che ciò comporti una revisione radicale dello schema.*
- **Contesto inter-disciplinare:** *l'ontologia, per sua natura, si presta a essere estesa verso altre forme d'arte o domini correlati (cinema, pittura, fumetto), favorendo la creazione di percorsi tematici tra media diversi.*

*Il **contesto** in cui si colloca potrebbe includere biblioteche pubbliche e accademiche, librerie fisiche o online, archivi digitali e piattaforme di editoria elettronica. In tutti questi ambienti, l'esigenza di organizzare i contenuti secondo criteri semantici è crescente: l'ontologia risponde a tale esigenza fornendo uno strumento flessibile che può essere adattato alle specifiche esigenze del singolo*

ente o progetto.

Per quanto riguarda il **tipo di utenza**, la struttura ontologica è stata pensata per:

- **Professionisti dell'informazione** (bibliotecari, documentaristi, architetti dell'informazione), che hanno necessità di creare cataloghi tematici o arricchire quelli esistenti.
- **Studiosi e ricercatori**, interessati a confrontare o correlare gli argomenti letterari in periodi e culture diverse tramite analisi comparative.
- **Appassionati di lettura** e utenti generici, che vogliono esplorare in modo intuitivo i temi più ricorrenti in letteratura o scoprire nuove opere partendo da un argomento di interesse.
- **Strumenti di raccomandazione automatica**, la classificazione in generi letterari potrebbe non tenere conto di argomenti trasversali di interesse dell'utente.

In questo modo, l'ontologia si propone di affiancare le pratiche di catalogazione già in uso. L'obiettivo è quello di fornire uno strumento per la classificazione e la consultazione dei temi letterari, aiutando a rendere più

intuitiva la scoperta e la condivisione dei contenuti, sia tra i professionisti sia tra gli appassionati.

Chapter 3

Descrizione del dominio

Il dominio letterario è estremamente vario e articolato, poiché comprende un'ampia gamma di forme testuali (narrativa, poesia, teatro, saggistica) e di generi (dal romanzo storico alla fantascienza, dal giallo al memoir). All'interno di questa varietà, gli argomenti – come l'amore, il conflitto, l'identità o il viaggio – si presentano come elementi trasversali che possono emergere in maniera più o meno esplicita a seconda del contesto storico-culturale e dello stile dell'autore.

Per identificare i temi più comuni e le modalità con cui vengono catalogati, sono state considerate diverse fonti, tra cui:

- *Open Library Subjects, che fornisce liste di opere organizzate per argo-*

mento, mostrando come i lettori e i bibliotecari categorizzano il materiale,

- Goodreads Lists, un'area del noto social network di lettori che aggrega liste tematiche basate sulle preferenze e le segnalazioni degli utenti,*
- Mondadori Store e laFeltrinelli, piattaforme online di due grandi catene librerie che propongono categorie e filoni di lettura in chiave commerciale.*

La nostra scelta non è certamente l'unica interpretazione possibile: l'interesse principale è quello di modellare l'argomento letterario come "chiave di lettura" in grado di:

- collegare tra loro opere molto diverse che condividono un tema (ad esempio, la figura dell'eroe tragico o la memoria storica),*
- fornire suggerimenti di lettura a chi, appassionato di un argomento, desidera esplorarne le varie declinazioni nella letteratura,*
- integrare future estensioni verso altre forme d'arte (come il cinema o la pittura) che affrontano gli stessi temi.*

Alcuni testi letterari mettono l'argomento al centro della loro struttura

narrativa (basti pensare ai romanzi a tesi o alla letteratura di denuncia), mentre altri lo toccano solo marginalmente, oppure si limitano a suggerirlo in modo simbolico. La volontà di rappresentare queste diverse sfumature all'interno di un'unica ontologia ha orientato la definizione di classi e relazioni specifiche per "catturare" sia la presenza principale di un argomento, sia la sua comparsa secondaria o occasionale.

Considerazioni su ontologie esistenti. *Esistono già modelli bibliografici e ontologie di dominio (ad esempio FABIO, FRBR, BIBO, DOLCE) che coprono vari aspetti relativi alle opere letterarie e alla loro descrizione. Il nostro progetto, essendo consapevole di questi strumenti, si concentra sulle tematiche e sulla loro rappresentazione semantica in un'ottica essenziale e didattica.*

In sintesi, come nel caso di altri domini complessi, anche in ambito letterario esiste il rischio di tralasciare o semplificare eccessivamente alcuni aspetti. Consapevoli di queste criticità, la scelta progettuale adottata mantiene un'impostazione volutamente snella, così da poter essere estesa o riadattata in base a specifiche necessità di catalogazione, ricerca o divulgazione. In questo modo, si garantisce la flessibilità necessaria per integrare gradualmente ulteriori risorse e perfezionare il modello qualora ve ne fosse l'esigenza.

Chapter 4

Competency Questions

Le seguenti domande delineano gli obiettivi essenziali dell'ontologia e forniscono un quadro dei requisiti che essa dovrebbe soddisfare per risultare utile e completa:

- *Quali sono le saghe che trattano un determinato argomento?*
- *Di quali argomenti parla una data saga?*
- *Quali autori affrontano uno specifico argomento?*
- *Quali sono i temi principali di un certo autore?*
- *In quale data è stato pubblicato il primo libro che tratta di un determinato argomento?*

- *Esistono opere che affrontano contemporaneamente due (o più) argomenti specifici?*
- *In quali periodi storici è prevalente un determinato tema letterario?*
- *Quali argomenti risultano più frequenti nella letteratura di un particolare paese o area culturale?*
- *Quali autori hanno contribuito maggiormente alla diffusione di un argomento?*

Chapter 5

Documentazione

5.1 Documentazione sul Dominio Letterario

Il dominio letterario copre una vasta gamma di argomenti, generi e forme testuali, rendendolo uno dei campi più ricchi e complessi da modellare. Per sviluppare un'ontologia che rappresenti al meglio gli argomenti letterari, è fondamentale basarsi su fonti di riferimento strutturate e su esempi concreti che illustrino le relazioni tra opere, autori e temi.

5.1.1 Fonti Informali e Standard Esistenti

- *Open Library Subjects:*

- *Open Library organizza le opere letterarie per argomenti in una sezione chiamata Subjects, fornendo una lista strutturata basata sulle preferenze di bibliotecari e lettori. Questi argomenti spaziano da temi generali (es. Amore) a sotto-categorie più specifiche (es. Relazioni interpersonali in contesti storici).*

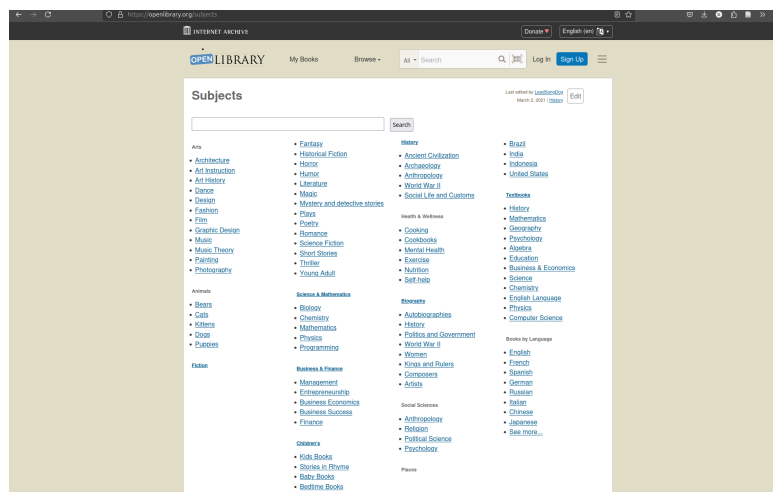


Figure 5.1: vista Subjects su Open Library con elenchi di generi e temi disponibili.

- **Goodreads Listopia:**

- *Goodreads, un social network per lettori, include una sezione denominata Listopia. Gli utenti creano liste tematiche che aggregano libri basati su argomenti o preferenze personali (es. Libri sulla*

resilienza o Romanzi epistolari).

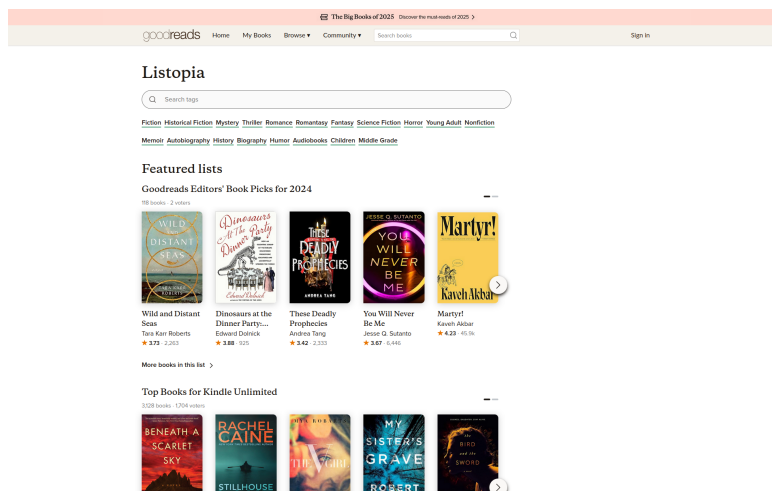


Figure 5.2: Vista di Listopia con una selezione di liste popolari create dagli utenti.

- **Piattaforme di librerie online:**
 - *Mondadori Store e laFeltrinelli* presentano categorie come *Filoni di lettura* e *Percorsi tematici*. Queste sezioni sono progettate per aiutare gli utenti a scoprire nuovi libri partendo da interessi specifici (es. *Classici della letteratura distopica*).
- **Ontologie Esistenti:**
 - **FABIO (FRBR-aligned Bibliographic Ontology):** Modello per la rappresentazione delle entità bibliografiche.

- **FRBR (Functional Requirements for Bibliographic Records):**
Modello che definisce le relazioni tra opere, espressioni, manifestazioni e item.
- **BIBO (Bibliographic Ontology):** *Ontologia focalizzata su riferimenti e citazioni bibliografiche.*
- **DOLCE:** *Ontologia fondazionale che può supportare la definizione di relazioni semantiche complesse.*

5.1.2 Dati di Esempio

Per illustrare il dominio, consideriamo un esempio tratto dalla letteratura:

- **Camera dei Segreti:** *02-07-1998 di J.K Rowling*
 - **Autore:** *J.K Rowling*
 - **Temi principali:** *Bacchette Magiche, Amicizia.*
 - **Genere:** *Fantasy*
 - **Edizioni:** *Varie edizioni in diverse lingue, pubblicate tra il 1998 e oggi.*

Questo esempio può essere rappresentato nell'ontologia con una classe per

goodreads
Home
My Books
Browse
Community



Harry Potter #2

Harry Potter and the Chamber of Secrets

J.K. Rowling, Mary GrandPré (Illustrator)

★★★★☆ 4.43 4,210,240 ratings · 87,918 reviews

Ever since Harry Potter had come home for the summer, the Dursleys had been so mean and hideous that all Harry wanted was to get back to the Hogwarts School for Witchcraft and Wizardry. But just as he's packing his bags, Harry receives a warning from a strange implish creature who says that if Harry returns to Hogwarts, disaster will strike.

Show more

Genres [Fantasy](#) [Fiction](#) [Young Adult](#) [Magic](#) [Childrens](#) [Audiobook](#) [Middle Grade](#) [...more](#)

352 pages, Hardcover

First published July 2, 1998

Book details & editions

Want to read
Shop this series

☆☆☆☆
Rate this book

Figure 5.3: Libro specifico su

l'opera letteraria, collegata alle classi per autore, temi, genere ed edizione tramite object properties.

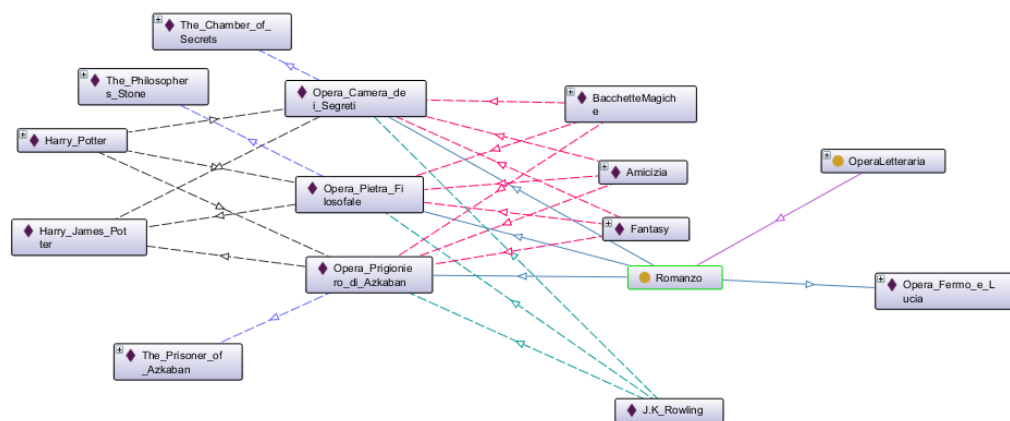



Figure 5.4: Vista di OntoGrf espanso sull'esempio specifico

5.1.3 Validatori

Sono stati usati inoltre due validatori, un validatore RDF per assicurare la bontà dell'esportazione ed evitare file malformati, e il validatore OOPS.

**Validation Service**

[Skip Navigation](#) [Home](#)
[Documentation](#)
[Feedback](#)

Jump To:

- [Source](#)
- [Triples](#)
- [Messages](#)
- [Graph](#)
- [Feedback](#)
- [Back to Validator Input](#)

Validation Results
Your RDF document validated successfully.

Triples of the Data Model

Number	Subject	Predicate	Object
1	https://litetaryworktopics.com/litetaryworktopics2025#	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#Ontology
2	https://litetaryworktopics.com/litetaryworktopics2025#	http://purl.org/dc/elements/1.1/creator	"Davide Vitturini"@en
3	https://litetaryworktopics.com/litetaryworktopics2025#	http://purl.org/dc/elements/1.1/creator	"Enrico Chiesa"@en
4	https://litetaryworktopics.com/litetaryworktopics2025#	http://purl.org/dc/elements/1.1/date	"2025-01-29"
5	https://litetaryworktopics.com/litetaryworktopics2025#	http://purl.org/dc/elements/1.1/description	"Questa ontologia ha l'obiettivo di fornire una rappresentazione generale sulle principali tematiche trattate dalla letteratura, con il particolare obiettivo di rendere possibile il suggerimento di nuove da parte di esperti o amatori."@it
6	https://litetaryworktopics.com/litetaryworktopics2025#	http://purl.org/vocab/vann/preferredNamespacePrefix	"lwt"
7	https://litetaryworktopics.com/litetaryworktopics2025#	http://purl.org/vocab/vann/preferredNamespaceUri	"https://litetaryworktopics.com/litetaryworktopics2025#@en
8	https://litetaryworktopics.com/litetaryworktopics2025#	http://web.resource.org/cc/license	"MIT"
9	https://litetaryworktopics.com/litetaryworktopics2025#	http://www.w3.org/2000/01/rdf-schema#label	"litetaryworktopics2025"@en
10	http://purl.org/dc/elements/1.1/Date	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#AnnotationProperty
11	http://purl.org/dc/elements/1.1/Title	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#AnnotationProperty
12	http://purl.org/dc/elements/1.1/creator	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#AnnotationProperty
13	http://purl.org/dc/elements/1.1/date	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#AnnotationProperty
14	http://purl.org/dc/elements/1.1/description	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#AnnotationProperty
15	http://purl.org/vocab/vann/preferredNamespacePrefix	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#AnnotationProperty
16	http://purl.org/vocab/vann/preferredNamespaceUri	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#AnnotationProperty
17	http://schema.org/relatedLink	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/2002/07/owl#AnnotationProperty

Figure 5.5: Vista del validatore RDF.

Per quanto riguarda i risultati forniti secondo dal validatore, si può asserire che nel nostro caso di modellazione le segnalazioni P10 e P11 non siano critiche in quanto la mancanza di disgiunzioni favorisce l'adattabilità alla rappresentazione di elementi che non sono inquadrabili in una categoria fissa.

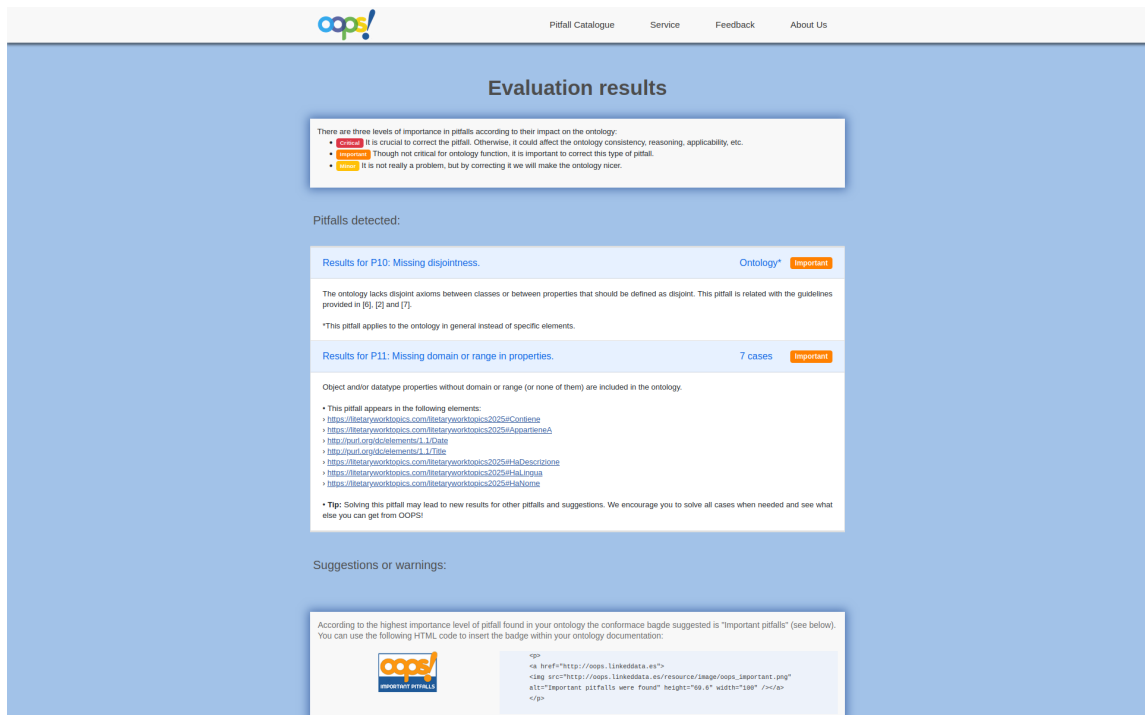


Figure 5.6: Vista del validatore OOPS.

Inoltre la mancanza di Domain o Range per gli altri elementi è motivata dal fatto che siano relazioni generiche, applicabili quindi ad ogni elemento rappresentato.

Come altro strumento di visualizzazione e documentazione è stato utilizzato pyLODE, una libreria di Python che permette di documentare in maniera automatica ontologie.

Questa documentazione è fornita in formato HTML nella Repository pubblica.

Chapter 6

Visualizzazione

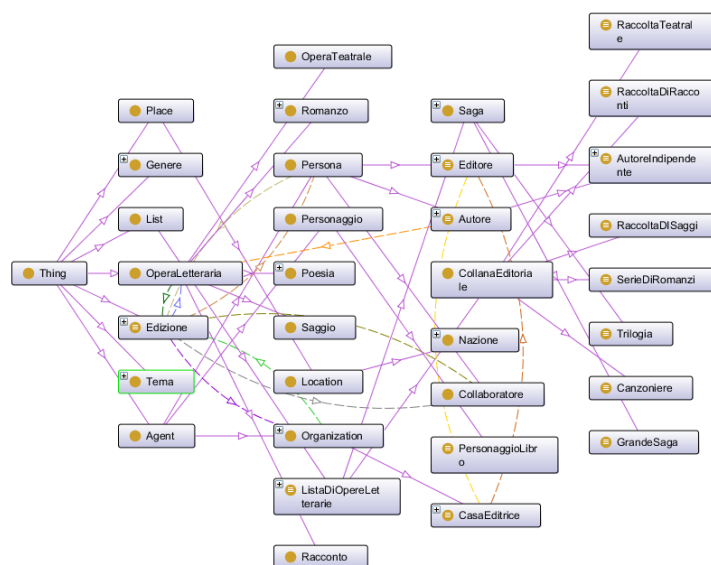


Figure 6.1: Struttura della tassonomia dell'ontologia.

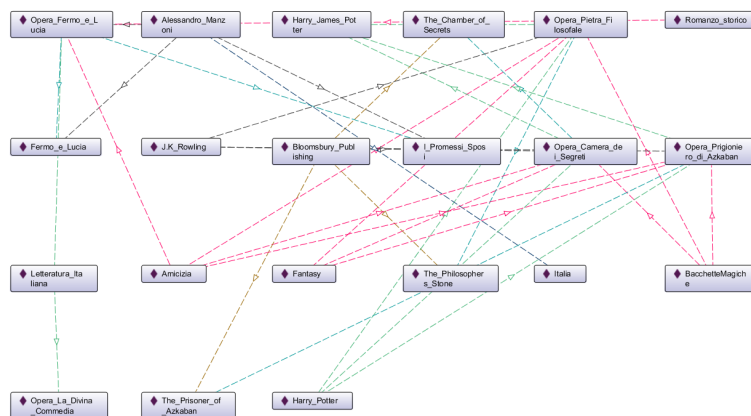


Figure 6.2: A-Box con OntoGraf.

6.1 Esempio di Libro

Soggetto	Predicato	Oggetto
Harry James Potter	AppartieneA	Opera Camera dei Segreti
Opera Camera dei Segreti	HaEdizione	The Chamber of Secrets
The Chamber of Secrets	PubblicataDa	Bloomsbury Publishing
Opera Camera dei Segreti	ScrittoDa	J.K. Rowling
Opera Camera dei Segreti	VieneClassificatoCome	Amicizia

Chapter 7

Query

Per ogni competency questions siamo andati a realizzare una query SPARQL per poter interrogare con l'ontologia.

I seguenti prefissi sono una parte di tutte le query, per migliorare la leggibilità sono riportati solo una volta

```
PREFIX : <https://litetaryworktopics.com/litetaryworktopics2025#>
```

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
```

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
```

Domande a cui risponde l'ontologia:

- *Quali sono le saghe che trattano un determinato argomento?*

```
SELECT DISTINCT ?nomeSaga ?tema

WHERE {

    ?tema a :Tema .                                # Trova i temi

    ?tema :haNome ?nomeTema .                      # Collega il tema al suo nome

    FILTER(?nomeTema = "Amicizia") .                # Filtra per il nome del tema "Ami

    ?tema :VieneUtilizzatoIn ?opera .              # Trova le opere che utilizzano il

    ?opera :AppartieneA ?saga .                    # Trova la saga a cui appartiene l

    ?saga :haNome ?nomeSaga

}

ORDER BY ?nomeSaga
```

- *Di quali argomenti parla una data saga?*

```
SELECT DISTINCT ?nomeSaga ?nomeTema

WHERE {

    ?saga :haNome ?nomeSaga . #trova il nome della saga

    FILTER(?nomeSaga = "Harry Potter") .          # Filtra per il nome del tema
```

```

?saga :Contiene ?opera .

?opera :VieneClassificatoCome ?tema .

?tema a :Tema .                                # non considera i generi

?tema :haNome ?nomeTema .                        # Collega il tema al suo nome
}

ORDER BY ?nomeSaga

```

- *Quali autori affrontano uno specifico argomento?*

```

[language=SPARQL]

SELECT DISTINCT ?nomeAutore ?nomeTema

WHERE {

    ?tema a :Tema .

    # Trova i temi

    ?tema :haNome ?nomeTema .

    # Collega il tema al suo nome

    FILTER(?nomeTema = "Amicizia") .

    # Filtra per il nome del tema "Amicizia"

    ?tema :VieneUtilizzatoIn ?opera .

    # Trova le opere che utilizzano il tema

```

```

?opera :ScrittoDa ?autore .

# Trova gli autori delle opere

?autore :haNome ?nomeAutore .

# Collega l'autore al suo nome

}

ORDER BY ?nomeAutore

```

- *Quali sono i temi che tratta un certo autore?*

```

SELECT DISTINCT ?nomeAutore ?nomeTema

WHERE {

    ?autore :haNome ?nomeAutore .

    # Trova l'autore dal nome

    FILTER(?nomeAutore = "J.K. Rowling") .

    # Filtra per il nome dell'autore specifico

    ?autore :AutoreDi ?opera .

    # Trova le opere scritte dall'autore

    ?opera :VieneClassificatoCome ?tema .

    # Collega le opere ai temi

    ?tema a :Tema .

```

```

# Assicura che siano temi

?tema :haNome ?nomeTema .

# Collega i temi ai loro nomi
}

ORDER BY ?nomeTema

```

- *In quale data è stato pubblicato il primo libro che tratta di un determinato argomento?*

```

SELECT ?nomeOpera ?dataPubblicazione

WHERE {

    ?tema a :Tema .

    # Trova i temi

    ?tema :haNome ?nomeTema .

    # Collega il tema al suo nome

    FILTER(?nomeTema = "Amicizia") .

    # Filtra per il tema specifico

    ?tema :VieneUtilizzatoIn ?opera .

    # Trova le opere che trattano il tema

    ?opera :HaEdizione ?edizione .

```

```

        ?edizione :haDataPubblicazione ?dataPubblicazione .

        # Trova la data di pubblicazione

        ?edizione :haNome ?nomeOpera
    }

ORDER BY ASC(?dataPubblicazione)

```

- *Esistono opere che affrontano contemporaneamente due (o più) argomenti specifici?*

```

SELECT ?nomeOpera

WHERE {

    ?tema1 a :Tema .

    ?tema1 :haNome ?nomeTema1 .

    FILTER(?nomeTema1 = "Amicizia") .

    # Primo argomento specifico

    ?tema2 a :Tema .

    ?tema2 :haNome ?nomeTema2 .

    FILTER(?nomeTema2 = "Bacchette Magiche") .

    # Secondo argomento specifico

    ?opera :VieneClassificatoCome ?tema1 .

```



```

?opera :VieneClassificatoCome ?tema2 .

# La stessa opera è associata a entrambi i temi

?opera :haNome ?nomeOpera .

}

```

- *In quali periodi storici è prevalente un determinato tema letterario?*

```

SELECT DISTINCT ?DataPubblicazione ?nomeTema

WHERE {

    ?tema a :Tema .

    # Trova i temi

    ?tema :haNome ?nomeTema .

    # Collega il tema al suo nome

    FILTER(?nomeTema = "Amicizia") .

    # Filtra per il tema specifico

    ?tema :VieneUtilizzatoIn ?opera .

    # Trova le opere associate al tema

    ?opera :HaEdizione ?edizione .

    # Trova le edizioni delle opere

    ?edizione :haDataPubblicazione ?DataPubblicazione .

```

```

        # Collega l'edizione al periodo storico
    }

```

```

ORDER BY ?DataPubblicazione

```

- *Quali argomenti risultano più frequenti nella letteratura di un particolare paese o area culturale?*

```

SELECT DISTINCT ?nomeNazione ?nomeTema (COUNT(?opera) AS ?numeroOpere)
WHERE {
    ?paese a :Nazione .

    ?paese :haNome ?nomeNazione .

    FILTER(?nomeNazione = "Italia") .

    # Filtra la nazione specifica

    ?autore :HaProvenienza ?paese .

    ?autore :AutoreDi ?opera .

    ?opera :VieneClassificatoCome ?tema .

    ?tema a :Tema .

    ?tema :haNome ?nomeTema .

}

GROUP BY ?nomeNazione ?nomeTema

```

```
ORDER BY DESC(?numeroOpere)
```

- *Quali autori hanno contribuito maggiormente alla diffusione di un argomento?*

```
SELECT DISTINCT ?nomeAutore ?nomeTema (COUNT(?opera) AS ?numeroOpere)
```

```
WHERE {
```

```
    ?tema a :Tema .
```

```
    ?tema :haNome ?nomeTema .
```

```
    FILTER(?nomeTema = "Amicizia") .
```

```
    # Filtra per il tema specifico
```

```
    ?tema :VieneUtilizzatoIn ?opera .
```

```
    ?opera :ScrittoDa ?autore .
```

```
    ?autore :haNome ?nomeAutore .
```

```
}
```

```
GROUP BY ?nomeAutore ?nomeTema
```

```
ORDER BY DESC(?numeroOpere)
```

Chapter 8

Sparql Anything

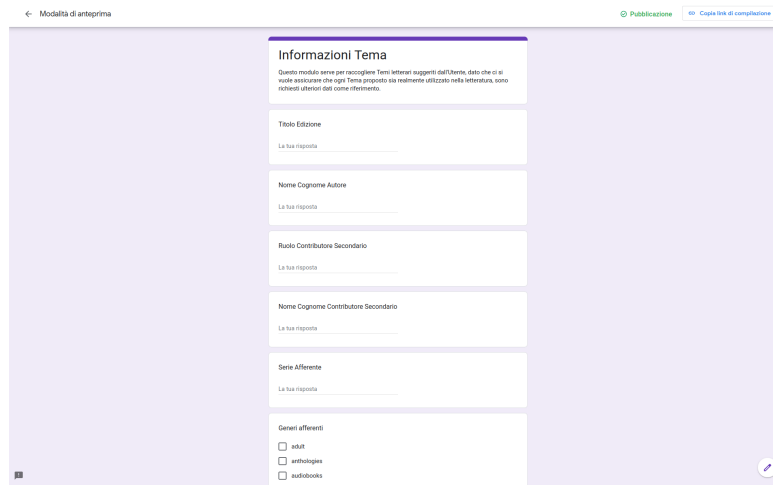
Per la realizzazione del Knowledge Graph abbiamo utilizzato Sparql Anything, uno strumento che permette di triplificare dati provenienti da fonti eterogenee in modo dinamico, garantendo la possibilità di interfacciarsi con formati diversi e trasformarli in RDF interrogabile.

8.1 Fonti di dati

Abbiamo utilizzato diverse fonti di dati per garantire una copertura ampia del dominio letterario e automatizzare il popolamento della nostra ontologia:

1. **Modulo Google:** Un questionario che raccoglieva dati dagli utenti sulle

loro preferenze letterarie, creando una fonte strutturata di dati etnografici.



The image shows a Google Form titled "Informazioni Tema". At the top, there are navigation links: "← Modalità di anteprima" and "Pubblicazione" (with a green checkmark icon). Below the title, there is a paragraph explaining the form's purpose: "Questo modulo serve per raccogliere Temi letterari suggeriti dall'utente, dato che ci si vuole assicurare che ogni Tema proposto sia realmente utilizzato nella letteratura, sono richiesti ulteriori dati come riferimento." The form contains several input fields: "Titolo Edizione" (with a placeholder "La tua risposta"), "Nome Cognome Autore" (with a placeholder "La tua risposta"), "Ruolo Contributore Secondario" (with a placeholder "La tua risposta"), "Nome Cognome Contributore Secondario" (with a placeholder "La tua risposta"), and "Serie Affinente" (with a placeholder "La tua risposta"). At the bottom, there is a section for "Generi affinenti" with checkboxes for "adult", "anthologies", and "audiobooks", followed by a "..." option. A small icon of a person is visible in the bottom right corner of the form area.

Figure 8.1: Vista del Modulo Google realizzato.

2. **Scraping di Goodreads:** Seguendo una traccia avanzata suggerita dal prof. Daga, abbiamo estratto informazioni da Goodreads, includendo opere, autori, saghe e classificazioni tematiche.
3. **Dati da CSV e JSON:** Abbiamo lavorato con dati in formato CSV e JSON, rappresentanti liste di libri, temi letterari e generi, trasformandoli in triple RDF per il Knowledge Graph.

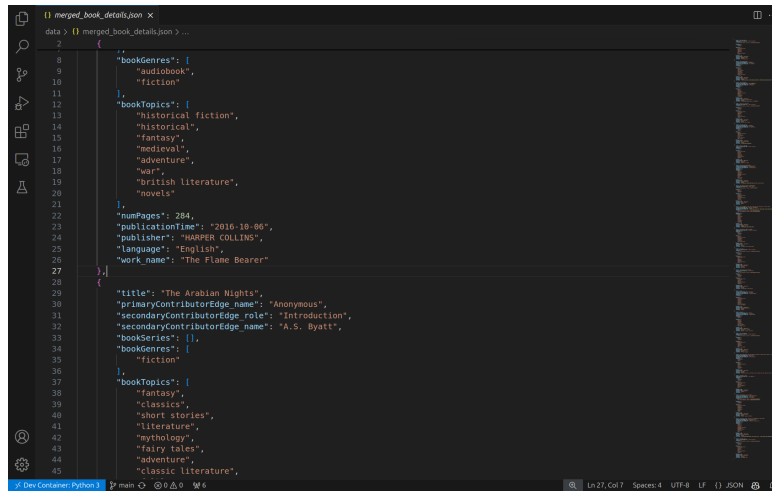


Figure 8.2: Esempio di json estratto da Goodreads.

8.2 Triplificazione dei dati con Sparql Anything

Abbiamo sviluppato diverse query di triplificazione per convertire dati da strutture complesse in RDF, ottenendo individui e classi in maniera dinamica. Un aspetto chiave del nostro lavoro è stata la generazione automatica di classi a partire da liste, un'operazione avanzata in Sparql Anything.

8.2.1 Triplificazione da JSON: estrazione strutturata

Uno dei file JSON che abbiamo triplificato conteneva informazioni su opere letterarie, con dati strutturati su autori, generi e tematiche. La seguente query ha estratto e trasformato i dati in RDF, generando automaticamente classi e

istanze:

PREFIX fx: <http://sparql.xyz/facade-x/ns/>

PREFIX schema: <http://schema.org/>

PREFIX lwt: <https://litetaryworktopics.com/litetaryworktopics2025#>

CONSTRUCT {

 ?bookMyNs a lwt:OperaLetteraria .

 ?bookMyNs lwt:HaNome ?title .

 ?bookMyNs lwt:ScrittoDa ?primaryContributorMyNs .

 ?bookMyNs lwt:AppartieneA ?seriesMyNs .

 ?bookMyNs lwt:VieneClassificatoCome ?genreMyNs .

 ?bookMyNs lwt:VieneClassificatoCome ?topicMyNs .

 ?primaryContributorMyNs a lwt:Autore .

 ?primaryContributorMyNs lwt:HaNome ?primaryContributorName .

 ?seriesMyNs a lwt:Saga .

 ?seriesMyNs lwt:HaNome ?seriesName .

```

?genreMyNs a lwt:Genere .

?genreMyNs lwt:HaNome ?genreName .


?topicMyNs a lwt:Tema .

?topicMyNs lwt:HaNome ?topicName .


} WHERE {

    SERVICE <x-sparql-anything:location=books.json> {

        ?book xyz:title ?title .

        BIND (IRI(CONCAT("https://litetaryworktopics.com/...

        ?book xyz:primaryContributorEdge_name ?primaryContributorName .

        BIND (IRI(CONCAT("https://litetaryworktopics.com/...

    OPTIONAL {

        ?book xyz:bookSeries ?seriesContainer .

        ?seriesContainer fx:anySlot ?seriesName .

        BIND (IRI(CONCAT("https://litetaryworktopics.com/

```



```

        litetaryworktopics2025#", REPLACE(?seriesName, " ", "_", "i")))

        AS ?seriesMyNs)

    }

    ?book xyz:bookGenres ?genreContainer .

    ?genreContainer fx:anySlot ?genreName .

    BIND (IRI(CONCAT("https://litetaryworktopics.com/

    litetaryworktopics2025#", REPLACE(?genreName, " ", "_", "i"))))

    AS ?genreMyNs)

    ?book xyz:bookTopics ?topicContainer .

    ?topicContainer fx:anySlot ?topicName .

    BIND (IRI(CONCAT("https://litetaryworktopics.com/

    litetaryworktopics2025#", REPLACE(?topicName, " ", "_", "i"))))

    AS ?topicMyNs)

}

}

```

Questa query ha consentito di trasformare ogni libro in un'istanza di `lwt:OperaLetteraria`, associandola agli autori, generi e temi in modo com-

pletamente dinamico, riconoscendo le liste e creando automaticamente le istanze corrispondenti.

8.2.2 Triplificazione da CSV: gestione delle intestazioni

Abbiamo anche triplificato dati CSV contenenti informazioni dettagliate sulle edizioni dei libri. Un aspetto critico era la gestione delle intestazioni della tabella, poiché Sparql Anything le considera di default come numeri di colonna. Per ovviare a questo problema, abbiamo attivato l'opzione `csv.headers=true`:

```
PREFIX fx: <http://sparql.xyz/facade-x/ns/>
```

```
PREFIX lwt: <https://litetaryworktopics.com/litetaryworktopics2025#>
```

```
CONSTRUCT {
```

```
    ?bookMyNs a lwt:Edizione .
```

```
    ?bookMyNs dc:Title ?cleanEditionTitle .
```

```
    ?bookMyNs lwt:ScrittoDa ?authorMyNs .
```

```
    ?bookMyNs lwt:PubblicatoDallOrganizzazione ?publisherMyNs .
```

```
    ?bookMyNs lwt:HaDataPubblicazione ?publicationDate .
```

```
    ?authorMyNs a lwt:Autore .
```

```

?authorMyNs lwt:HaNome ?cleanAuthorName .

?publisherMyNs a lwt:CasaEditrice .

?publisherMyNs lwt:HaPubblicato ?bookMyNs .

} WHERE {

    SERVICE <x-sparql-anything:location=book_details.csv,csv.headers=true> {

        ?row <http://sparql.xyz/facade-x/data/Titolo%20Edizione> ?editionTitle .

        BIND (REPLACE(?editionTitle, "\"", "") AS ?cleanEditionTitle)

        BIND (IRI(CONCAT("https://litetaryworktopics.com/

litetaryworktopics2025#", REPLACE(?cleanEditionTitle, " ", "_", "i"))))

        AS ?bookMyNs)

        ?row <http://sparql.xyz/facade-x/data/Nome%20Cognome%20Autore> ?authorName

        BIND (REPLACE(?authorName, "\"", "") AS ?cleanAuthorName)

        BIND (IRI(CONCAT("https://litetaryworktopics.com/

litetaryworktopics2025#", REPLACE(?cleanAuthorName, " ", "_", "i"))))

        AS ?authorMyNs)

        ?row <http://sparql.xyz/facade-x/data/Data%20di%20pubblicazione> ?publicat

```

```

    ?row <http://sparql.xyz/facade-x/data/Casa%20Editrice> ?publisherName .

    BIND (REPLACE(?publisherName, "\"", "") AS ?cleanPublisherName)

    BIND (IRI(CONCAT("https://litetaryworktopics.com/
    litetaryworktopics2025#", REPLACE(?cleanPublisherName, " ", "_", "i")))
    AS ?publisherMyNs)

  }

}

```

8.3 Popolamento dell'ontologia

Dopo la triplificazione, i dati RDF sono stati caricati automaticamente nella nostra ontologia utilizzando un notebook Python. Grazie a questo processo, abbiamo popolato automaticamente la nostra ontologia con centinaia di individui, rendendo interrogabile il Knowledge Graph con query SPARQL.

Chapter 9

Data Availability Statement

L'intero progetto è disponibile nella Repository GitHub ed è usufruibile sotto licenza <https://mit-license.org/>.

Bibliography

- [1] Aldo Gangemi et al. *DOLCE: A Descriptive Ontology for Linguistic and Cognitive Engineering*. Tech. rep. Ontologia di alto livello per concetti complessi. WonderWeb Deliverable D18, 2002. URL: <http://www.loa.istc.cnr.it>.
- [2] International Federation of Library Associations and Institutions (IFLA). *Functional Requirements for Bibliographic Records (FRBR)*. Struttura concettuale per la rappresentazione delle entità bibliografiche. IFLA, 1998.
- [3] Sébastien Mottelet. *The Bibliographic Ontology (BIBO)*. Modello per rappresentare dati bibliografici. Accesso: gennaio 2025. URL: <http://bibliontology.com>.

- [4] *Ontologia di Goodreads*. Accesso: gennaio 2025. URL: <https://www.goodreads.com/genres>.
- [5] Silvio Peroni and David Shotton. “FaBiO: The FRBR-aligned Bibliographic Ontology”. In: *Semantic Web Journal* (2012). Ontologia bibliografica allineata al modello FRBR. URL: <http://purl.org/spar/fabio>.
- [6] Mari Carmen Suárez-Figueroa et al. *NeOn Methodology for Building Ontology Networks*. Guida metodologica per costruzione e integrazione di reti ontologiche. Springer, 2012.