

Софийски университет “Св. Климент Охридски”
Факултет по математика и информатика

ПРОЕКТ

**Представяне и моделиране на знания,
зимен семестър 2023/2024**



Име: Надежда Францева

Фн: 8MI3400357

Специалност: Изкуствен Интелект

Онтология за книги

1. Идея на проекта

Идеята на проекта е да се изгради онтология в предметната област на книгите. Крайните цели са:

- да улесни класифицирането на нови издадени книги
- да улесни намирането на книга по зададени критерии от потребител.

Част от въпросите, които ще можем да задаваме върху онтологията с книги като потребители са:

1. **За книгите:**

- Кои книги са написани от даден автор?
- Кои са всички книги в даден жанр?
- Кои са новите книги за последните 6 месеца?

2. **За авторите:**

- Кой е написал дадена книга?
- Кои автори са писали в даден жанр?

3. **За издателствата:**

- Кои книги са издадени от конкретно издателство?
- Кои издателства са издавали книги в последните 3 години?
- Кои са най-популярните издателства в даден жанр?

4. **За библиотеките и книжарниците:**

- Кои книги са налични в конкретна библиотека?
- Къде мога да намеря определена книга в близост до мен?

5. **За рейтингите:**

- Кои са текущите книги с най-висок рейтинг?
- Кои автори имат най-висок рейтинг?

6. **За преводите:**

- Кои книги са преведени на даден език?

2. Преглед и анализ на съществуващи и генерирани от интелигентен агент решения

2.1. Съществуващи решения

Сред най-известните онтологии, свързани с книги, Dublin Core и Schema.org са вероятно най-широко използваните и признати стандарти.

- Dublin Core Metadata Element Set (DCMES):

<https://www.dublincore.org/specifications/dublin-core/usageguide/elements/>

Dublin Core е широко приет и използван стандарт за метаданни в цялата информационна област. Той предоставя базови елементи за описание на ресурси, включително книги. DCMES включва стандартизирани елементи като заглавие, автор, издателство, дата на публикация, и т.н.

- Schema.org

<https://schema.org/Book>

Schema.org е проект, поддържан от големите търсачки като Google, Bing и Yahoo. Той предоставя структурирани данни за уеб страници и включва подробни маркировки за различни типове обекти, включително книги. Този стандарт често се използва за обогатяване на метаданните и подпомагане на търсачките.

2.2. Генерирани решения

Ето проста примерна онтология за книги, ограничена до 5 класа, 5 свойства и 5 индивида, генерирана от ChatGPT:

Класове (Types):

- Книга (Book)
- Автор (Author)
- Издателство (Publisher)
- Жанр (Genre)
- Език (Language)

Свойства (Properties):

- Има Заглавие (hasTitle)
- Има Автор (hasAuthor)
- Има Издателство (hasPublisher)
- Има Жанр (hasGenre)
- Има Език (hasLanguage)

Индивиди (Individuals):

- Книга1 (Book1)
 - hasTitle: "Пътеводител за Онтологии"
 - hasAuthor: "Иван Иванов"
 - hasPublisher: "Университетско Издателство"
 - hasGenre: "Научна литература"

- hasLanguage: "Български"
- Книга2 (Book2)
 - hasTitle: "Роман за Летен Лагер"
 - hasAuthor: "Мария Петрова"
 - hasPublisher: "Детско Издателство"
 - hasGenre: "Детска литература"
 - hasLanguage: "Български"
- Автор1 (Author1)
 - hasName: "Иван Иванов"
 - hasBirthYear: 1980
- Автор2 (Author2)
 - hasName: "Мария Петрова"
 - hasBirthYear: 1995
- Издателство1 (Publisher1)
 - hasName: "Университетско Издателство"

Предимства:

Онтологията може да бъде лесно адаптирана, ако се налага добавяне на допълнителни атрибути или класове. Предоставя базова структура за представяне на информацията за книгите.

Недостатъци:

Ако има нужда от по-голяма детайлност или допълнителни аспекти на книгите, може да бъде необходимо добавяне на нови свойства или класове. Не са включени всички възможни атрибути или връзки, като например допълнителни детайли за авторите, книгите и т.н.

3. Елементи на онтологията

3.1. Класове

Address \sqsubseteq owl:Thing
Book \sqsubseteq owl:Thing
Building \sqsubseteq owl:Thing

Library \sqsubseteq Building
Publisher \sqsubseteq Building
Cover \sqsubseteq owl:Thing
Genre \sqsubseteq owl:Thing
Biograpohy \sqsubseteq Genre
Children\'s \sqsubseteq Genre
Classic \sqsubseteq Genre
Comedy \sqsubseteq Genre
Comic \sqsubseteq Genre
Comic \sqsubseteq Genre
Comic \sqsubseteq Genre
Drama \sqsubseteq Genre
Fantasy \sqsubseteq Genre
Historical \sqsubseteq Genre
Horror \sqsubseteq Genre
Mystery \sqsubseteq Genre
Poetry \sqsubseteq Genre
Romance \sqsubseteq Genre
Science \sqsubseteq Genre
Person \sqsubseteq owl:Thing
Author \sqsubseteq Person

3.2. Свойства

3.2.1. Свойства на данните

Property	Domain	Range	Characteristics
hasStreet	Address	xsd:string	NonFunctional
hasEmail	Author	xsd:string	NonFunctional
hasISBN	Book	xsd:string	Functional
hasLanguage	Book	xsd:language	NonFunctional
hasName	Person	xsd:string	Functional
hasPagesCount	Book	xsd:integer	Functional
hasPrice	Book	xsd:decimal	NonFunctional

hasPublicationDate	Book	xsd:dateTime	Functional
hssRate	Book	xsd:decimal	Functional
hasTitle	Book	xsd:string	Functional

3.2.2. Свойства на обектите

Property	Domain	Range	Characteristics
hasCountry	Address	Country	Asymmetric and Irreflexive
hasCity	Address	City	Asymmetric and Irreflexive
hasAddress	Building	Address	Functional, Asymmetric and Irreflexive
hasAuthor	Book	Author	Asymmetric, Irreflexive and Inverse Of hasBook
hasBook	Author	Book	Asymmetric, Irreflexive and Inverse Of hasAuthor
hasCover	Book	Cover	Functional, Asymmetric and Irreflexive
hasBooks	Library	Book	Asymmetric and Irreflexive
hasPrequel	Book	Book	Transitive, Asymmetric and Irreflexive

hasPublisher	Book	Publisher	Functional, Asymmetric and Irreflexive
hasSameAuthorLike	Book	Book	Symmetric and Reflexive

3.2.3. Индивиди

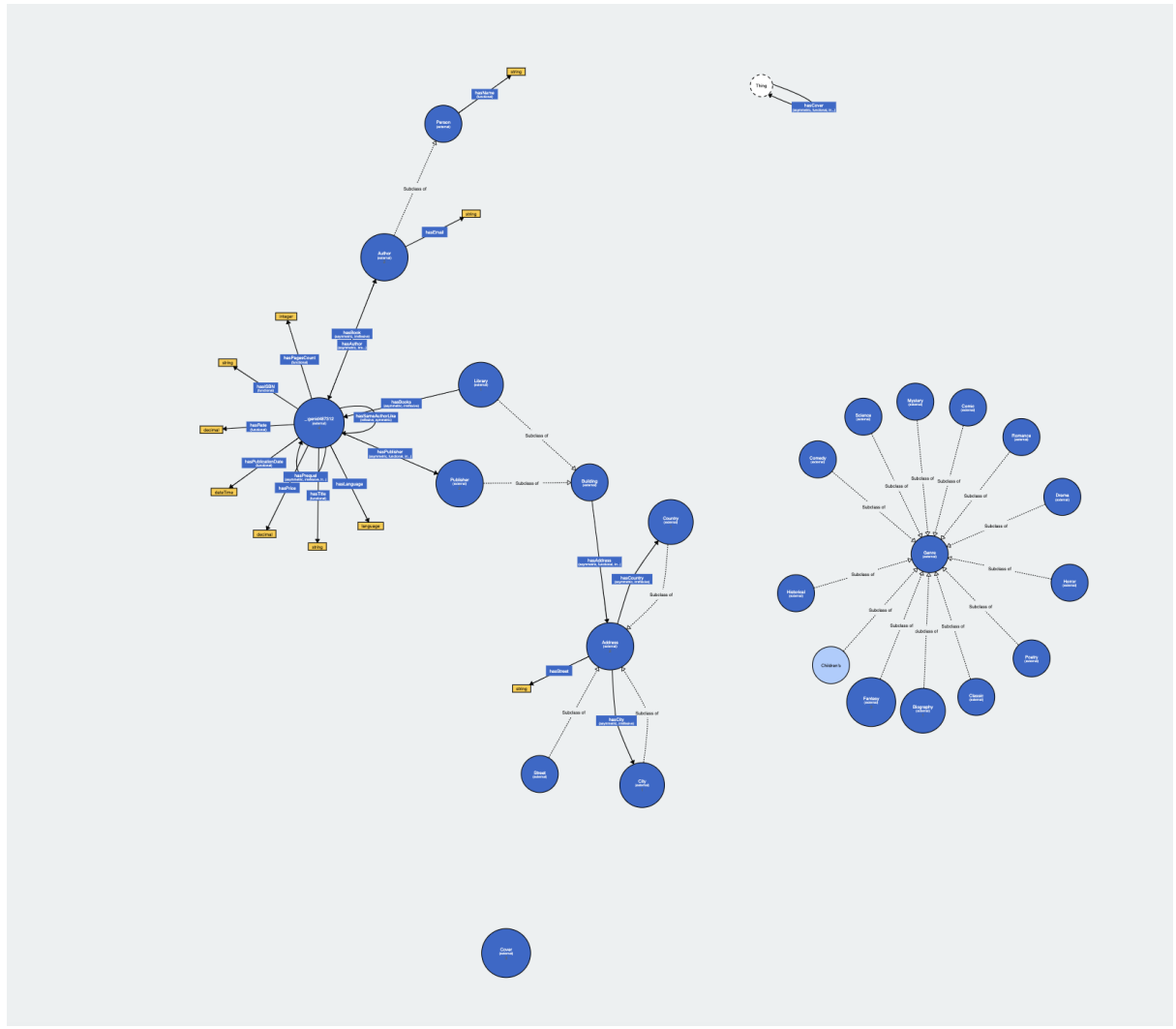
:AddressBloomsbury rdf:type owl:NamedIndividual , :Address	Адресът на издателство Bloomsbury, без подробности за улица, град, държава
:AddressBulgarianNationalLibrary rdf:type owl:NamedIndividual , :Address ; :hasCity :Sofia ; :hasCountry :Bulgaria ; :hasStreet "Vasil Levski Blvd."	Адресът на Българската Национална Библиотека – улица, град, държава
:Bloomsbury rdf:type owl:NamedIndividual , :Publisher ; :hasAddress :AddressBloomsbury	Издателство Bloomsbury с адрес – обекта AddressBloomsbury
:Bulgaria rdf:type owl:NamedIndividual , :Country	Страна България
:EBook rdf:type owl:NamedIndividual	Електронна книга, прост индивид с цел да участва в Enumerated Class Cover
:Hardcover rdf:type owl:NamedIndividual	Твърда корица на книга, прост индивид с цел да участва в Enumerated Class Cover
:HarryPotter1 rdf:type owl:NamedIndividual , :Book , :Fantasy ; :hasAuthor :JKRowling ; :hasCover :Hardcover ; :hasPublisher :Bloomsbury ; :hasSameAuthorLike :HarryPotter2 , :HarryPotter3 ;	Книгата „Хари Потър и философският камък“ с описание – автор, корица, издателство, ISBN, език, брой страници, цена, дата на публикуване, оценка, пълно заглавие

<pre> :hasISBN "978-0590353427" ; :hasLanguage "English"@en ; :hasPagesCount 223 ; :hasPrice 9.66 ; :hasPublicationDate "1997-06-26"^^xsd:dateTime ; :hasRate 4.7 ; :hasTitle "Harry Potter and the Philosopher's Stone" . </pre>	
<pre> :HarryPotter2 rdf:type owl:NamedIndividual , :Book , :Fantasy ; :hasAuthor :JKRowling ; :hasCover :Hardcover ; :hasPrequel :HarryPotter1 ; :hasPublisher :Bloomsbury ; :hasSameAuthorLike :HarryPotter1 , :HarryPotter3 ; :hasISBN "0-7475-3849-2" ; :hasLanguage "English"@en ; :hasPagesCount 251 ; :hasPrice 37.99 ; :hasPublicationDate "1998-07-01"^^xsd:dateTime ; :hasRate 4.8 ; :hasTitle "Harry Potter and the Chamber of Secrets" . </pre>	<p>Книгата „Хари Потър и Стаята на тайните“ с описание – автор, корица, издателство, ISBN, език, брой страници, цена, дата на публикуване, оценка, пълно заглавие</p>
<pre> :HarryPotter3 rdf:type owl:NamedIndividual , :Book , :Fantasy ; :hasAuthor :JKRowling ; :hasCover :Hardcover ; :hasPrequel :HarryPotter2 ; :hasPublisher :Bloomsbury ; :hasSameAuthorLike :HarryPotter1 , :HarryPotter2 ; :hasISBN "0-7475-4215-5" ; :hasLanguage "English"@en ; :hasPagesCount 317 ; :hasPrice 9.66 ; :hasPublicationDate "1999-07-08"^^xsd:dateTime ; :hasRate 4.7 ; :hasTitle "Harry Potter and the Prisoner of Azkaban" </pre>	<p>Книгата „Хари Потър и затворникът от Азкабан“ с описание – автор, корица, издателство, ISBN, език, брой страници, цена, дата на публикуване, оценка, пълно заглавие</p>

:JKRowling rdf:type owl:NamedIndividual , :Author ; :hasBook :HarryPotter1 , :HarryPotter2 , :HarryPotter3 ; :hasEmail "info@jkrowling.com" ; :hasName "Joanne Rowling"	Авторката Джоан Роулинг с описание – кои са книгите, които е написала и ги има в нашата онтология, валиден емайл, пълно име
:NationalLibraryStStCyrilAndMethodius rdf:type owl:NamedIndividual , :Library ; :hasAddress :AddressBulgarianNationalLibrary ; :hasBooks :HarryPotter1 , :HarryPotter2 , :HarryPotter3	Национална библиотека „Св. Св. Кирил и Методий“ с описание – адрес, книги, които присъстват в нея и ги има в нашата онтология
:Paperback rdf:type owl:NamedIndividual	Мека корица на книга, прост индивид с цел да участва в Enumerated Class Cover
:Sofia rdf:type owl:NamedIndividual , :City	Град София
:SteveJobs rdf:type owl:NamedIndividual , :Biography , :Book ; :hasAuthor :WalterIsaacson ; :hasPublisher <http://www.semanticweb.org/fnadezhda/ontologies/2024/0/book_ontology#Simon&Schuster> ; :hasISBN "1-4516-4853-7" ; :hasLanguage "English"@en ; :hasPagesCount "630"^^xsd:int ; :hasPrice 23.29 ; :hasPublicationDate "2011-10-24"^^xsd:dateTime ; :hasRate 4.7 ; :hasTitle "Steve Jobs by Walter Isaacson"	Книгата за биографията на Стив Джобс с описание – автор, издателство, ISBN, език, брой страници, цена, дата на публикуване, оценка, пълно заглавие
:WalterIsaacson rdf:type owl:NamedIndividual , :Author ; :hasBook :SteveJobs ; :hasEmail "wisaacson@gmail.com" ; :hasName "Walter Isaacson"	Авторът Уолтър Айзъксън с описание – кои са книгите, които е написала и ги има в нашата онтология, валиден емайл, пълно име
<http://www.semanticweb.org/fnadezhda/ontologies/2024/0/book_ontology#Simon&Schuster> rdf:type owl:NamedIndividual ,	Издателство Simon & Schuster

:Publisher	
------------	--

Визуализация:



4. Използвани технологии

- Проектът е реализиран чрез системата Protégé.
- Онтологията е проверена за неконсистентности чрез reasoner ELK версия 0.5.0:
<https://eclipse.dev/elk/downloads/releasenotes/release-0.5.0.html>
- За информация са ползвани сайтовете wikipedia и сайтове за продажба на книги като Amazon
- За визуализиране схемата на онтологията е използван инструментът WebVOWL:
<http://vowl.visualdataweb.org/webvowl-old/webvowl-old.html#file=onto.Owl>

5. Бъдещо развитие

- Добавяне на още класове – например клас читатели
- Добавяне на още индивиди - книги, автори, библиотеки и издателства
- Разширяване с характеристики на книгите – резюме, описание, брой на копията, ключови думи и т.н., с характеристики на авторите – възраст, телефон и т.н., на библиотеките – посещаемост, карта и т.н., на и издателствата – хора, които работят за издателството, услуги и т.н.
- Механизъм за автоматизирано попълване и синхронизиране чрез използване на wikipedia или transfermarkt. Някои от данните са на разположение в машинно-четим формат от различни източници. Чрез crawling могат да бъдат извлечени и онтологията да бъде поддържана в крак с времето.

6. Външни източници

- A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition 1.3:
https://learn.fmi.uni-sofia.bg/pluginfile.php/468239/mod_resource/content/1/ProtegeOWLTutorialP4_v1_3.pdf
- Protege Documentation:
<http://protegeproject.github.io/protege/>