

Ficha rpcw2024-recurso

Avaliação: Representação e Processamento de Conhecimento na Web

11 de Junho de 2024, 14h, Ed.1 - sala 1.17

Mestrado em Engenharia Informática (1º ano)

Sinopsis

O objectivo principal deste teste é avaliar os conhecimentos obtidos durante as aulas na especificação e desenvolvimento de ontologias e de aplicações que tiram partido destas.

Antes de começares, lê atentamente até ao fim para ficares com uma percepção do todo que se pretende. Vais ver que tomarás decisões mais acertadas depois de uma leitura completa.

Os resultados finais deverão ser enviados ao docente da seguinte forma:

- Enviar por email para: jcr@di.uminho.pt;
- Colocar no subject/assunto: **RPCW2024::Recurso::Axxxxx**, em que **Axxxxx** corresponde ao número do aluno (nalguns casos **PGxxxxx**);
- Enviar ao docente um link do github para um repositório novo criado especificamente para o exame com o seguinte conteúdo (esta preparação poderá valer 1 valor do exame):
 - O repositório no GitHub deverá chamar-se **RPCW2024-Recurso**;
 - Dentro do repositório deverá haver um ficheiro, **PR.md**, contendo uma descrição do que fez e como fez, contendo instruções se necessário, para o docente poder replicar os cenários na correção.
 - Dentro do repositório deverão existir duas pastas: **ex1**, onde colocarão a aplicação desenvolvida para responder ao primeiro exercício e, **ex2**, onde colocarão a aplicação desenvolvida para responder ao segundo exercício.

Exercício 1: Vamos extrair conhecimento de uma história

Neste exercício, deverás especificar uma ontologia OWL, podes usar o **Protégé**, para a história abaixo e os requisitos que lhe seguem.

História

O Eduardo tem 21 anos e é um estudante da Universidade do Minho oriundo do Porto. Ele fala fluentemente português, espanhol e inglês.

Neste semestre decidiu aprender alemão. Juntamente com os seus amigos, o Carlos e a Ana, inscreveu-se no curso de Alemão lecionado pelo professor Helmut Ratz, que será lecionado todas as sextas feiras no centro de línguas da universidade.

O professor Ratz é alemão e docente de carreira da Escola de Letras, Artes e Ciências Humanas da Universidade do Minho.

O Eduardo tem muita facilidade na aprendizagem de línguas estrangeiras e faz rápidos progressos. No entanto, o prof. Helmut repara que ele precisa de praticar mais a oralidade e apresenta-o a Hanna, uma estudante alemã que está na universidade a frequentar o curso de biotecnologia.

Hanna é alemã e está ansiosa por falar português. O Eduardo e a Hanna conhecem-se e tornam-se parceiros nesta aventura de aprender uma nova língua. A experiência revela-se frutuosa e ambos aprendem a falar fluentemente alemão e português.

A Ontologia (3 val.)

Especifica agora a ontologia que dê suporte ao conhecimento nesta história:

1. Define as Classes;
2. Define os atributos de cada Classe (**data properties**);
3. Define as relações que se podem estabelecer entre indivíduos da mesma classe ou de classes diferentes (**object properties**);
4. Cria os indivíduos com os seus atributos e as suas relações;
5. Deixa a ontologia num ficheiro TTL: **historia.ttl**.

As Queries (2 val. = 0.5 * 4)

Adicionalmente especifica queries em SPARQL para responder às seguintes questões (coloca as queries, devidamente identificadas, num ficheiro de texto designado por **queries.txt**):

1. Quantas línguas fala o Eduardo?
2. Quem se inscreveu no curso de alemão?
3. Quantos indivíduos existem na tua ontologia?
4. Quem é Hanna? (recupera o máximo de informação deste personagem)

Exercício 2: Era uma vez... uns livros interessantes (15 val.)

Considera o dataset em anexo a este teste (**dataset.json**).

Os seus registos de informação têm a seguinte forma:

```
{
  "bookId": "2767052-the-hunger-games",
  "title": "The Hunger Games",
  "series": "The Hunger Games #1",
  "author": "Suzanne Collins",
  "rating": "4.33",
  "description": "WINNING MEANS FAME AND FORTUNE.LOSING MEANS CERTAIN DEATH.THE HUNGER GAMES HAVE BEGUN. . . .In the ruins of a place once known as North America lies the nation of Panem, a shining Capitol surrounded by twelve outlying districts. The Capitol is harsh and cruel and keeps the districts in line by forcing them all to send one boy and once girl between the ages of twelve and eighteen to participate in the annual Hunger Games, a fight to the death on live TV.Sixteen-year-old
```

Katniss Everdeen regards it as a death sentence when she steps forward to take her sister's place in the Games. But Katniss has been close to dead before\and survival, for her, is second nature. Without really meaning to, she becomes a contender. But if she is to win, she will have to start making choices that weight survival against humanity and life against love.",

```

    "language": "English",
    "isbn": "9780439023481",
    "genres": "['Young Adult', 'Fiction', 'Dystopia', 'Fantasy',
'Science Fiction', 'Romance', 'Adventure', 'Teen', 'Post Apocalyptic',
'Action']",
    "characters": "['Katniss Everdeen', 'Peeta Mellark', 'Cato (Hunger
Games)', 'Primrose Everdeen', 'Gale Hawthorne', 'Effie Trinket', 'Haymitch
Abernathy', 'Cinna', 'President Coriolanus Snow', 'Rue', 'Flavius',
'Lavinia (Hunger Games)', 'Marvel', 'Glimmer', 'Clove', 'Foxface',
'Thresh', 'Greasy Sae', 'Madge Undersee', 'Caesar Flickerman', 'Claudius
Templesmith', 'Octavia (Hunger Games)', 'Portia (hunger Games)']",
    "bookFormat": "Hardcover",
    "edition": "First Edition",
    "pages": "374",
    "publisher": "Scholastic Press",
    "publishDate": "09/14/08",
    "firstPublishDate": "",
    "awards": "['Locus Award Nominee for Best Young Adult Book
(2009)', 'Georgia Peach Book Award (2009)', 'Buxtehuder Bulle (2009)',
'Golden Duck Award for Young Adult (Hal Clement Award) (2009)', \"Grand
Prix de l'Imaginaire Nominee for Roman jeunesse \u00e9tranger (2010)\",
'Books I Loved Best Yearly (BILBY) Awards for Older Readers (2012)',
\"West Australian Young Readers' Book Award (WAYRBA) for Older Readers
(2010)\", \"Red House Children's Book Award for Older Readers & Overall
(2010)\", 'South Carolina Book Award for Junior and Young Adult Book
(2011)', 'Charlotte Award (2010)', 'Colorado Blue Spruce Young Adult Book
Award (2010)', 'Teen Buckeye Book Award (2009)', \"Pennsylvania Young
Readers' Choice Award for Young Adults (2010)\", 'Rhode Island Teen Book
Award (2010)', \"Dorothy Canfield Fisher Children's Book Award (2010)\",
'Evergreen Teen Book Award (2011)', 'Soaring Eagle Book Award (2009)',
'Milwaukee County Teen Book Award Nominee (2010)', 'Sakura Medal for
Middle School Book (2010)', 'Michigan Library Association Thumbs Up! Award
(2009)', 'Florida Teens Read (2009)', 'Deutscher Jugendliteraturpreis for
Preis der Jugendjury (2010)', 'Iowa High School Book Award (2011)', 'New
Mexico Land of Enchantment Award for Young Adult (2011)', 'Eliot Rosewater
Indiana High School Book Award (2010)', 'The Inky Awards for Silver Inky
(2009)', 'California Young Readers Medal for Young Adult (2011)', 'Lincoln
Award (2011)', 'Kinderboekwinkelprijz (2010)', 'Missouri Truman Readers
Award (2011)', 'CYBILS Award for Young Adult Fantasy & Science Fiction
(2008)', 'Literaturpreis der Jury der jungen Leser for Jugendbuch (2010)',
'The Inky Awards Shortlist for Silver Inky (2009)', 'Prix Et-lisez-moi
(2011)', 'Missouri Gateway Readers Award (2011)', 'Oklahoma Sequoyah Award
for High School and Intermediate (2011)', 'Premio El Templo de las Mil
Puertas for Mejor novela extranjera perteneciente a saga (2009)',
\"Rebecca Caudill Young Readers' Book Award (2011)\", 'LovelyBooks
Leserpreis for Fantasy (2009)', 'LovelyBooks Leserpreis for Bestes
Cover/Umschlag (2009)', 'Premi Protagonista Jove for Categoria 13-14 anys
(2010)']",

```

```

    "numRatings": "6376780",
    "ratingsByStars": "['3444695', '1921313', '745221', '171994',
'93557']",
    "likedPercent": "96",
    "setting": "['District 12, Panem', 'Capitol, Panem', 'Panem
(United States)']",
    "coverImg": "https://i.gr-
assets.com/images/S/compressed.photo.goodreads.com/books/1586722975l/27670
52.jpg",
    "bbeScore": "2993816",
    "bbeVotes": "30516",
    "price": "5.09"
  }

```

Analisa o dataset com atenção e realiza as seguintes operações:

1. [3 val.] Cria uma ontologia OWL para suportar a informação do dataset (especificação estrutural: classes, relações e atributos) e grava-a num ficheiro denominado `books_base.ttl`;
2. [1 val.] Limpa e prepara o dataset para a operação seguinte: pensa qual deverá ser o `id` dos indivíduos, há campos que no dataset estão como string e são realmente listas: `author`, `genres`, `awards`, `characters`, etc;
3. [3 val.] Cria um script (numa linguagem à tua escolha) e com ela povoa a ontologia com a informação do dataset (acrescenta os indivíduos). Grava a nova ontologia (estrutura e indivíduos) num novo ficheiro denominado `books.ttl`;
4. [6* 1 val.] Depois de carregares a tua ontologia no GraphDB especifica as seguintes queries SPARQL:
 - Quantas livros estão presentes na ontologia?
 - Quais os títulos dos livros em que um dos autores tem apelido "`Austen`"? (lista de títulos ordenada alfabeticamente)
 - Qual a distribuição de livros por género? (lista de géneros sem repetições e ordenada alfabeticamente em que a cada género está associado o número de livros pertencente a esse género)
 - Produz uma lista de pares, nome do personagem e a lista de livros onde esse personagem aparece, ordenada alfabeticamente por nome de personagem;
 - Qual a distribuição de livros por autor? (lista de autores sem repetições e ordenada alfabeticamente em que a cada autor está associado o número de livros escritos por esse autor);
 - Quais os títulos dos livros pertencentes ao género "`Adventure`"? (lista ordenada alfabeticamente de títulos)
5. [1 val.] Cria uma query CONSTRUCT que associe os co-autores (autores do mesmo livro) pela relação `hasCoAuthor`. No fim, acrescenta estes triplos à ontologia;
6. [1 val.] Cria um query SPARQL que produza um índice dos autores e respetivos co-autores, ou seja, dá como resultado uma lista de pares (`autor`, `lista de co-autores`).

Bom trabalho e boa sorte jcr