

PEC3 asignatura Web de Datos

Instrucciones

Habr  tres niveles de dificultad:

- Enunciado 1: ejercicio b sico
- Enunciado 2: ejercicio nivel medio (opcional)
- Enunciado 3: ejercicio nivel avanzado (opcional)

Teniendo que realizar al menos el ejercicio b sico (Enunciado 1) para superar cada PEC con un 5.

Enunciado 1: Recuperando datos de la web de datos (5 puntos)

A partir de la informaci n disponible en [la LOD Cloud](#) e informaci n embebida en las p ginas HTML se proceder  a enriquecer las entidades que tenemos en nuestro dataset local (instancia1, instancia3, instancia 4).

Enunciado 1.1

- 1) Crear un repositorio en RDF4J llamado "SocialNetwork". A adir el fichero "Entidades.n3" (adjunto con el enunciado) al repositorio "SocialNetwork".

Para la instancia1, instancia3 e instancia 4 se proceder  a explorar unos datasets y posteriormente a enriquecer las instancias locales con informaci n de esos datasets remotos. La informaci n se recuperar program ticamente sirvi ndose de fichero "PEC3_1.py" (adjunto con el enunciado).

Enunciado 1.2

- 2) Para ello primero vamos a explorar los siguientes datasets:
 - BBPedia_ES
 - ~~SPARQL:~~ [SPARQL: http://dbpedia.org/sparql](http://dbpedia.org/sparql)
 - ~~Ejemplo de recurso:~~ http://www.dbpedia.org/page/Alicia_Keys
 - MDB:
 - SPARQL: <http://data.linkedmdb.org/sparql>
 - Ejemplo de recurso: <http://data.linkedmdb.org/page/film/300>
 - BNE SPARQL: <http://datos.bne.es/sparql>
 - Ejemplo de recurso: <http://datos.bne.es/resource/XX968899>

Para cada uno de los "Ejemplos de recursos" anteriores se proceder  a navegar bien directamente en la p gina que se muestra (DBPEDIA o MDB) o bien explorando el formato RDF-turtle que aparece en la parte inferior de la p gina (BNE) e indicar lo siguiente:

- Qu  tipo de entidades est n describiendo cada uno de los recursos. Nota 1.
- Propiedad usada para indicar el t tulo o "name" de cada recurso. Nota 2
- T tulo o name que tiene los recursos
- Para cada uno de los recursos anteriores obtener:
Propiedad usada para describir el a o de nacimiento o el a o de creaci n del artista, pel cula o autora

Nota 1: El tipo de cada recurso venia determinado por la propiedad del vocabulario RDF: <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>. Si tiene varias descripciones con esta propiedad indicar aquellas de vocabularios m s conocidos.

Nota 2: Un recurso puede tener varias propiedades describiendo el título (title), el nombre (name), su etiqueta (label).

El resultado de este punto es solo informativo para realizar el punto siguiente, es decir no se espera ningún documento ni nada explicando la información encontrada.

Enunciado 1.3

3) Una vez que hemos identificado la información anterior vamos a proceder a enriquecer nuestro repositorio local. Para ello habrá que incluir en el script las consultas SPARQL necesarias, (aparece indicado con comentarios donde se debe incluir las consultas SPARQL o modificar los endpoints). Para este nivel básico bastara con imprimir la URI de los recursos de los repositorios remotos que coinciden con el label de las instancias locales, esto es:

- a. Para el caso de dbpedia, obtener aquel o aquellos recursos cuyo nombre coincida con el label de la instancia 1.
- b. Para el caso de MDB, obtener aquel o aquellos recursos cuyo nombre coincida con el label de la instancia 3.
- c. Para el caso de BNE, obtener aquel o aquellos recursos cuyo nombre coincida con el label de la instancia 4.

Pista: habrá que hacer uso de la propiedad titulo o "name" que hayáis identificado en el punto anterior

Outcomes: PEC3_1.py completado

Enunciado 2: Obteniendo información de repositorios remotos (+2,5)

Enunciado 2.4

4) Partiendo del punto anterior, se procederá a obtener alguna información extra (alguna propiedad) de los recursos obtenidos remotos. Por ejemplo año de nacimiento o publicación.

Outcomes: PEC3_2.py completado

Enunciado 3: Incorporando tripletas de repositorios remotos en nuestro repositorio (+2,5)

Enunciado 3.5

5) Insertar programáticamente las tripletas obtenidas con el enquirer en el repositorio SocialNetwork . Esto implica hacer una extensión/modificación del enquirer.py para insertarlas directamente en el repositorio mediante la librería SPARQLWrapper o mediante una consulta SPARQL UPDATE-LOAD

Outcomes: PEC3_3.py conteniendo el código del volcado