# Ejercicio Final Modulo Web de Datos

A partir de la información disponible en la LOD Cloud e información embebida en las páginas HTML se procederá a enriquecer las entidades que tenemos en nuestro dataset local (instancia1, instancia2, instancia3, instancia 4).

1. Crear un repositorio en Sesame “SocialNetwork”
2. Añadir el fichero “Entidades.n3” (mostrado más abajo) al repositorio “SocialNetwork”.
3. Partiendo del ejercicio 5\_1, y según la lista de datasets sugeridos (ver más abajo), se determinará que propiedades nos interesa incorporar a nuestras entidades.
4. Para la instancia1, instancia3 e instancia 4 se procederá a enriquecer la información con información de los datasets previamente recomendados de LOD Cloud. La información se recuperar programáticamente sirviéndose de fichero “enquierer.py” que hay en la carpeta de Ejercicio Final.
   1. En él se muestra los pasos inconclusos para la recuperación de información de dbpedia para enriquecer la instancia1.
   2. Para proceder a enriquecer otras instancias se procederá de manera similar añadiendo nuevas funciones en las que cambiaran el endpoint endpoints y las queries.
5. Para superar la prueba bastara con imprimir las tripletas obtenidas. Outcome el fichero enquirer.py incluyendo en el main las llamadas a todas las funciones auxiliares que se hayan creado. De modo que al ejecutar el enquirer.py se obtenga toda la información por pantalla.
6. Para la instancia 2 se procederá a enriquecer la información con información embebida en páginas HTML. Partiendo de la información que se habrá recopilado en el archivo “manuChao.n3”, proceder a volcarlo en el repositorio de SocialNetwork y hacer una query que liste los “MusicRecording”.
7. (Opcional). Insertar las tripletas obtenidas con el enquierer en el repositorio SocialNetwork:
   1. Volcar las tripletas a un fichero y volcarlas al repositorio SocialNetwork via el workbench
   2. O insertarlas directamente en el repositorio mediante la librería SPARQLWrapper.

# Fecha de Entrega

25 de septiembre 2015

# Outcomes

* Enquirer.py completado
* Fichero txt incluyendo la query SPARQL para la instancia 4
* Opcional: Enquirer\_con\_volcado.py conteniendo el código del volcado. Y fichero socialNetword\_enriched.n3 conteniendo una exportación (a través del workbench) del repositorio.

# Fichero de entidades (Entities.n3)

# Default graph

@prefix sn: <http://ciff.curso2015/ontologies/owl/socialNetwork#> .

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

sn:instancia1 a sn:Person;

rdfs:label "Alicia Keys"@en .

sn:instancia2 a sn:Person;

rdfs:label "Manu Chao"@en-gb .

sn:instancia3 a sn:Film;

rdfs:label "Batman" ;

rdfs:label "Batman"@es ;

rdfs:label "Madrid"@en .

sn:instancia4 a sn:Guide;

rdfs:label "Un vino cosmopolita" ;

rdfs:label "Un vino cosmopolita"@es ;

rdfs:label "Un vino cosmopolita"@en .

**Dataset Recomendados**

* BBPedia\_ES (<http://es.dbpedia.org/sparql>)
  + http://www.dbpedia.org/page/Alicia\_Keys
* MusicBrainz (<http://dbtune.org/musicbrainz/sparql>)
  + <http://dbtune.org/musicbrainz/resource/artist/704acdbb-1415-4782-b0b6-0596b8c55e46>
  + <http://dbtune.org/musicbrainz/page/artist/a992aada-7108-455d-9747-0b7b6a089e8d>
* Web n+1 el viajero: (<http://webenemasuno.linkeddata.es/sparql>)
  + <http://webenemasuno.linkeddata.es/page/elviajero/resource/Guide/20060513ELPVIALBV_5.TES>
* MDB: (<http://data.linkedmdb.org/sparql>)
  + http://data.linkedmdb.org/page/film/300