Migración del DeCS a RDF

El DeCS se desarrolla a partir del MeSH, aportando la traducción de los términos en español y portugués, además de nuevas categorías que se adaptan mejor a la realidad de latinoamérica y el Caribe.

La migración puede realizarse de dos formas:

1. utilizar los archivos RDF de MeSH (que se encuentran disponibles como datos abiertos) y agregarle las traducciones al espanol y portugués
2. encuestar al DeCS y representar sus datos con el modelo de datos del MeSH RDF

Para la migración del DeCS se propone:

1. utilizar los archivos RDF de MeSH que se pueden descargar en el FTP de NLM para guardar los datos en la BD MySQL.
   * Descargado el fichero mesh.nt, con todos los datos del mesh en formato n-triples, que tiene un tamaño de 1,8 GB
2. Convertir los datos de n-triples a json

* Se identificaron Aplicaciones para convertir de N-Triple a JSON-LD

<http://rdf-translator.appspot.com/> (esta bloqueado para cuba)

<https://sourceforge.net/projects/rdfconvert/>

<http://www.easyrdf.org/converter>

* Se descargo rdfconverter y funciona bien, pero no puede convertir el fichero de MeSH.nt pq es demasiado grande (1,8 GB). Hay que dividirlo en ficheros de alrededor de 100000 lineas (menos de 14 MB) para que funcione.
* Para dividir el fichero mesh.nt se utiliza el editor sublime text, que puede leer ficheros grandes.
* los ficheros se dividen de forma que contenga todas las triples de la ultima entidad, para que todos los datos de una entidad estén en un mismo fichero
* los ficheros se guardan con extensión json

1. Guardar datos de fichero json en la BD como json-ld
   * Se parsea el json, para dividir las entidades descritas en un fichero json, pq la función json\_decode da error cuando el json es muy grande.
   * A cada entidad se le asigna la uri del mesh en la propiedad sameAs
   * Se identifica el tipo de entidad y se guarda cada entidad en la tabla correspondiente
   * A cada entidad se agrega la URI del DeCS, a partir del identificador autonumerico del registro.
   * Se asignan URIs del DeCS en las propiedades de una entidad que hacen referencia a otras entidades. Por ejemplo un descriptor hace referencia a calificadores, términos, conceptos, números de arbol
2. Encuestar al DeCS para obtener todos los campos traducidos y agregar las traducciones a los datos del MeSH
   * Para Descriptores y Calificadores se pueden obtener:
     + Label
     + scopeNote
     + annotation
     + considerAlso

Para otras entidades como Concept y Term no hay métodos para recuperar sus datos en el DeCS

* + Existen otros textos traducidos que no son devueltos
  + Solo label puede obtenerse en los 3 idiomas de una vez, los demás hay que pedirlos para cada idioma.

1. Migrar las categorías del DeCS que no existen en el MeSH al modelo de datos de MESH RDF

Aunque en GitHub están las transformaciones XSLT que permiten pasar del MESH XML a MESH RDF. No pueden utilizarse pq el XML del DeCS no tiene el mismo formato que el del MESH XML, el DeCS no se migra a partir de MeSH XML, sino de versiones anteriores

Implementacion de la migración:

1. decsDB.php : creación de las tablas a utilizar
   * Descriptor
   * Qualifier
   * SupplementaryConceptRecord
   * DescriptorQualifierPair
   * TreeNumber
   * Concept
   * Term
2. get\_content.php : parsea fichero con