Applying NoSQL Databases for Integrating Web Educational Stores - An Ontology-Based Approach

Reem Qadan Al Fayez(✉) and Mike Joy

Department of Computer Science, University of Warwick,
Coventry CV4 7AL, UK
{r.qadan-al-fayez,m.s.joy}@warwick.ac.uk

Abstract. Educational content available on the web is playing an important role in the teaching and learning process. Learners search for different types of learning objects such as videos, pictures, and blog articles and use them to understand concepts they are studying in books and articles. The current search platforms provided can be frustrating to use. Either they are not specified for educational purposes or they are provided as a service by a library or a repository for searching a limited dataset of educational content. This paper presents a novel system for automatic harvesting and connecting of medical educational objects based on biomedical ontologies. The challenge in this work is to transform disjoint heterogeneous web databases entries into one coherent linked dataset. First, harvesting APIs were developed for collecting content from various web sources such as YouTube, blogging platforms, and PubMed library. Then, the system maps its entries into one data model and annotates its content using biomedical ontologies to enable its linkage. The resulted dataset is organized in a proposed NoSQL RDF Triple Store which consists of 2720 entries of articles, videos, and blogs. We tested the system using different ontologies for enriching its content such as MeSH and SNOMED CT and compared the results obtained. Using SNOMED CT doubled the number of linkages built between the dataset entries. Experiments of querying the dataset is conducted and the results are promising compared with simple text-based search.

Keywords: Linked Data · Web databases · Ontologies

1 Introduction

The use of web content is increasing because of its accessibility at any time and from any place. Online libraries have started to support open access and professionals are increasingly using Web 2.0 technologies to spread their knowledge. In medical education, although its content should be of a high quality and provided by authorized sources, the web had played an important role in providing such content. Medical communities have a high awareness of the range of educational content available and show substantial interest in using such resources [1].

S. Maneth (Ed.): BICOD 2015, LNCS 9147, pp. 29–40, 2015.
DOI: 10.1007/978-3-319-20424-6_4

Searching for relevant educational content on the web can be challenging for its users. The vast amount of information available and the diversity of its types makes the search process time consuming. The content of any website is usually stored in a relational database with different fields used for describing its records. Therefore, integrating web databases into one data store is a challenging issue. New practices for publishing web content using Linked Data are being adopted by an increasing number of content providers leading to the web of data [2].

In this paper, we present a novel system that adopts Linked Data practices for automatic linking of heterogeneous web stores into one dataset based on biomedical ontologies enrichment. The developed system links some of the high quality User Generated Content (UGC), published on *YouTube* and *blogs* by medical educators and organizations, with content from online medical libraries. Using biomedical ontologies, we enriched the content of these databases by annotating free-text descriptions provided in their metadata records. Ontology-based annotation allows the system to discover keyword terms in web database content and builds dynamic linkages between them. The final linked dataset is represented in RDF/XML format and URIs are used for describing the dataset content.

Researchers in the field of e-learning refer to online educational resources that can be used in the learning process as Learning Objects (LOs). Learning Objects as defined in [3] can be of different types -images, videos, or text-, and differ in granularity from books to small web pages. Since the application domain of this work is medical education, we refer to the educational resources retrieved from the web and used in this system as *Educational Medical Objects (EMOs)*. The result of our work is a linked dataset of EMOs named the *LEMO dataset* and a system for managing them called the *LEMO system*.

The paper is structured as follows. Section 2 presents background and related work about the subject. Section 3 describes the processes applied for harvesting distributed web stores and building the LEMO RDF Triple Store. Section 4 provides more details about the ontologies used in the LEMO system, and explains the use of these ontologies in the annotation and enrichment process. Furthermore, a detailed description of the NoSQL RDF Triple Store components are presented in this section. Section 5 details a comparative analyses for using the LEMO system with the MeSH and SNOMED CT ontologies, and discusses experiments conducted for querying this dataset using ontological-based queries. Finally, Sect. 6 presents the conclusions and future work.

2 Background and Related Work

Using Linked Data and ontologies for data enrichment have been researched heavily in the recent years. The enrichment methods can happen at the server-side or client-side of a system. Both variations have been tested in [4] and the advantages and disadvantages were compared. Enriching queries is another method applied at the server-side of the system, and the work presented in [5] investigated enriching queries made to a collection of medical images provided by one library. The queries have been expanded after enriching the text with MeSH ontology terms.

Data enrichment has also been used with UGC content on the web, because user generated tags or folksonomies describing *YouTube* videos may be poorly chosen. The tag expansion and raking applied in [6] has been shown to enhance the description of the videos on *YouTube*. Enriching the content of a single dataset has been heavily researched, especially in the medical field. This is due to having mature and well maintained biomedical ontologies [7].

Linked Data principles have been adopted in education. Projects have been developed for supporting the use of web educational data [8]. Efforts for linking different educational ICT tools registries are presented in [9]. Another project for publishing datasets of educational medical materials in Linked Data has been developed in [10], which focused on providing a data model for exposing various types of educational materials to the web in order to enhance their interoperability and reuse across the web. It is clear that Linked Data will have a potential in the education field. A project presented in [11] developed a curriculum from Open Educational Resources tailored for teaching and training practitioners to use Linked Data. These days, educational organizations and universities are considering storing and publishing data using a Linked Data format [12].

3 LEMO Triple RDF Store

Before integrating web educational stores, we need to harvest and model distributed Educational Medical Objects (EMOs) into one data model. Our goal is to integrate different types of EMOs into one linked data set that is searchable and queryable.

In order to accomplish this goal, we developed two harvesting endpoints. In the first one, we incorporated the OAI-PMH protocol [13]. The other endpoint is basically an RSS feed reader storing web feeds from websites that provide them. Many online libraries expose their metadata using an OAI-PMH protocol such as the *PubMed* library. Using these harvesting endpoints, developed in the LEMO system, we harvested articles from the *PubMed* library and videos and blogs from *YouTube* and *blogging platforms*. The resulted dataset consisted of 2720 medical educational objects divided into 1000 articles from PubMed library and 1720 blogs and videos harvested from five different blogging websites and 6 YouTube channels. The chosen blogs and YouTube channels are maintained by either medical academics or journals and dedicated to educational purposes.

The harvested metadata records are retrieved in XML formats. The OAI service in PubMed supports DCMI metadata, therefore we can set the format parameter in the OAI requests produced by LEMO to be DCMI for harvesting content. On the other hand, blogs and video RSS feeds are structured XML documents which are machine interpretable and provide access to parts of the website entries such as title, authors, publication date, and content [14]. A fragment of the XML files harvested is illustrated in Fig. 1 along with the processes needed to build LEMO RDF Triple Store.

The heterogeneous metadata structures for all EMOs retrieved are mapped into the LEMO data model using XSLT techniques. The LEMO data model has

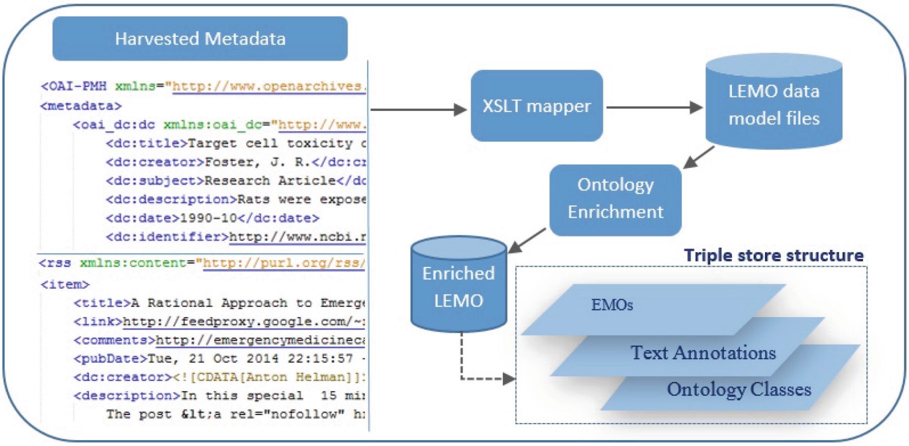


Fig. 1. LEMO RDF Triple Store structure and development process

been proposed in [15] at an earlier stage of developing the LEMO system after conducting a comparative study of existing data model in medical education. It is based on the DCMI metadata schema and implemented in RDF/XML formats. New LEMO properties were introduced for describing the enriched resources in LEMO store which will be discussed in detail in Sect. 4. The mapped files are then sent to the ontology enrichment process which annotates the free-text of EMOs, and discovers possible subjects to categorize them. This will result in having an enriched LEMO Triple Store which consists of EMOs, terms annotated in EMOs, and ontology classes used for annotation, as shown in Fig. 1.

4 Ontology-Based Annotations

Biomedical ontologies are formal representations of knowledge with definition of concepts and their relations [16]. Such ontologies have been used for indexing data produced by researchers in the field to ease its integration [17]. They are also used for indexing articles in medicine libraries such as the use of MeSH ontology for indexing PubMed library articles. In the LEMO system, we use ontologies to annotate free-text in the harvested EMO metadata such as titles and descriptions. Annotating the free-text enables us to discover relations between non related objects on the web. In the LEMO system, we adopt the Linked Data format for building the LEMO Triple Store which is considered the best practice for connecting semantically related data that was not previously linked [2].

The application domain of the LEMO system is medicine education. The BioPortal[1] open repository for biomedical ontologies is used to explore possible ontologies to integrate them with the LEMO system. BioPortal provides access to ontologies developed in different formats via web services which enable its

[1] http://bioportal.bioontology.org/.

users to get information about the ontologies or its content [18]. The LEMO system uses additional web services such as the annotator service provided by Bioportal for annotating and linking objects in the dataset. The ontologies used in the LEMO system so far are the Systematized Nomenclature of Medicine - Clinical Terms SNOMED CT and the Medical Subject Headings MeSH.

4.1 SNOMED CT and MeSH Ontologies

The SNOMED CT ontology has been developed by specialized organizations in both the USA and the UK and offers a standardized healthcare terminology which is comprehensive, scientifically validated, with relationships built into its core concepts [19]. This ontology was released in 2002 and since then new versions of it have been released semi-annually. SNOMED CT has been designated as the preferred clinical terminology to use in 19 countries [20]. Its application in medical information systems is expected to increase. The popularity and broad use of this ontology in the field of medicine was the main reason for applying SNOMED CT in the LEMO system.

As for the second ontology applied in the LEMO system, MeSH has been used for indexing PubMed Library content. Therefore, we applied it in the LEMO system to annotate and link EMOs based on its classes and relations. MeSH has been developed by the National Library of Medicine (NLM) in the USA and is considered the controlled vocabulary set used for indexing its articles. It consists of a set of terms naming descriptors organized in a hierarchical structure from general to more specific descriptors [21].

4.2 EMOs Annotation Enrichment

The LEMO RDF Triple Store consists of collections of linked resources describing its content and organized as illustrated in Fig. 2. Each resource is identified by a unique URI and a set of predicates to describe its properties. The major component in the LEMO RDF Triple Store is *EMOs*. After applying ontology enrichment, additional *term annotations* resources are added to the collection of EMOs' *title* and *description* resources to enrich them. The smallest component of the LEMO store is *ontology classes*. Each annotation made in the free-text points to a class in the ontology used for annotation. The classes of an ontology are arranged in a graph structure where relations exist to identify the class hierarchy. The collection of classes used in annotating the LEMO RDF Triple Store forms a subset of the original ontology graph.

The resources of the LEMO RDF Triple Store are described using the LEMO data model. The model is based on DCMI properties enhanced with new properties proposed for representing the annotations in LEMO. Such new properties are defined using the prefix "*lemo*". As shown in the XML fragment of Fig. 2, EMOs are described using only DCMI properties. The values of their **title** and **description** predicates are new resources created, rather than textual values, used for linking the EMOs to annotations using the *lemo:lemoTitleAnnotation* and *lemo:lemoDescAnnotation* properties. The *term annotations* are described

Fig. 2. Snippet of LEMO RDF Triple Store components

in detail using LEMO properties that store details about the original text annotated, its indices and content along with details about the class it was annotated to using a specific ontology, its ID, label, definition, and synonyms if they exist. The terms' annotated classes are nodes in the original ontology used for enriching the dataset. Hence, a sub graph of the original ontology can be built using the collection of *ontology classes* used for annotating its terms. The class relations are stored using the *lemo:adjacentTo* property. These class resources will enable further processing of the LEMO Triple Store to discover subjects or categories for EMOs and build dynamic linkages between its resources.

In the annotation process, free-text of EMOs is sent for the BioPortal annotator service and an ontology is specified in the request parameter. Then, the response is read and terms' annotated resources are created and linked to EMOs stored in the LEMO RDF Triple Store. After the annotation process, each EMO is represented by a set of keywords which are the terms annotated in its title and description. Each set of keywords representing an EMO forms a smaller sub graph of linked ontology classes based on their adjacency lists stored in the LEMO Triple Store. For discovering subjects for an EMO, we apply a simple term filtering technique to identify a smaller set of keywords which represent the EMO subject property and stored as the value of the *dc:subject* predicate for that EMO. In term filtering, we assign weights for the keywords based on their position and co-occurrence in an EMO term annotation set. Then, the accumulated weight for each keyword is calculated based on its hierarchical level in ontology. If the term annotated class is a parent of many terms annotated for the same EMO, then it will be more important than a term that is leaf in the ontology. The final weight value for each term annotated is stored in the *rdf:value* property of the term annotation resource. After normalizing the weights of the keywords, the top ranked keywords are selected as *subjects* of an EMO.

We tested the LEMO system against two biomedical ontologies: MeSH and SNOMED. Comparison of annotation results, term filtering, and linkage discovery of these two experiments are detailed in the following section. The results of

Table 1. Number of terms annotated for the set of EMOs using different ontologies

		MeSH annotations			SNOMED CT annotations		
Type of EMOs	Number of EMOs	Title	Description	Total	Title	Description	Total
Article	1000	3887	12192	16079	6166	29859	36025
Video	1259	3027	4304	7331	3677	5710	9387
Blog	461	754	4720	5474	1572	9756	11328
Total	2720	7668	21216	28884	11415	45325	56740

the comparisons helps to decide which ontology to use in LEMO system based on the larger number of annotations created.

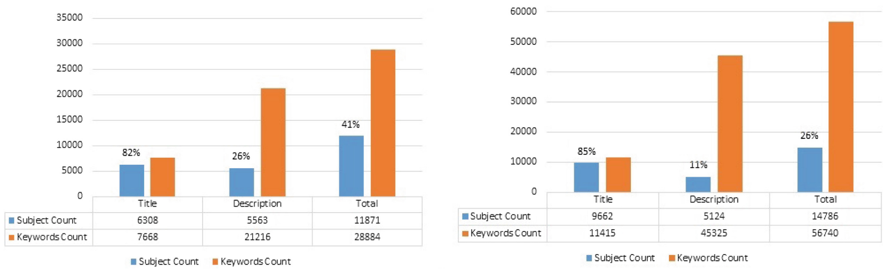
5 Results and Discussion: MeSH vs. SNOMED Ontology

The components of the dataset, harvested from the web for testing this system, are detailed in Table 1. The table details the numbers of resources harvested grouped by its type. It also details the number of keywords discovered after annotating its textual content whether annotated in the *title* or the *description* based on MeSH and SNOMED CT ontologies.

We notice that the number of terms annotated using the SNOMED CT ontology is greater than the number of keywords annotated using the MeSH ontology. The difference is not significant for video and blog EMOs compared to article EMOs. This is due to the short text provided in the metadata of blogs and videos compared to the longer text provided for articles in the online libraries. The collection of terms annotated for the dataset is used for building linkages between the EMOs and discovering subjects for categorizing the EMOs.

5.1 Discovering Subjects Using Ontologies

After processing the keyword set for each EMO, *subjects* were selected from the keywords annotated for categorizing each EMO. The resulting sets of subjects selected for each EMO are variable in size. We calculated the total numbers of subjects selected in the LEMO dataset and compared them against the keywords set sizes. The results are detailed in Fig. 3a and b for the MeSH and SNOMED CT subject selections respectively. The percentage of keywords chosen as subject terms from the SNOMED CT annotated terms is less than the percentage from the MeSH annotated collection. In both experiments, the subjects discovered are mainly chosen from the keywords annotated in the titles of EMOs. This indicates that the term filtering techniques have succeed in this matter. We can notice that in SNOMED CT, only 11 % of the keywords annotated in the description were chosen as a subject, compared to a higher percentage of 26 % keywords in the description using MeSH.



(a) Subjects based on MeSH **(b)** Subjects based on SNOMED

Fig. 3. Subject selection

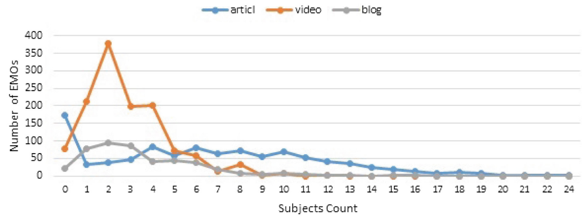
The LEMO dataset consists of different types of EMOs. Video and blog EMOs usually have shorter descriptions in their metadata fields if any. This affects the number of keywords annotated for EMOs from such types and that reflects on the number of subjects selected. Using MeSH and SNOMED CT annotations, the results of discovering subjects for video and blog EMOs are not enhanced in both experiments. Figure 4a and b illustrate the relation between the counts of subjects discovered in MeSH and SNOMED CT annotated EMOs and their types. In both experiments, video and blog EMOs have low subject counts. This is due to the low numbers of terms annotated in this type of EMOs. Comparing the subject selection process based on MeSH and SNOMED CT, we notice that in the SNOMED CT based dataset, very few EMOs did not have any subject count, while in the MeSH based dataset, more than 150 EMOs from articles, videos or blog types have subject counts equal to zero.

5.2 Links Analysis

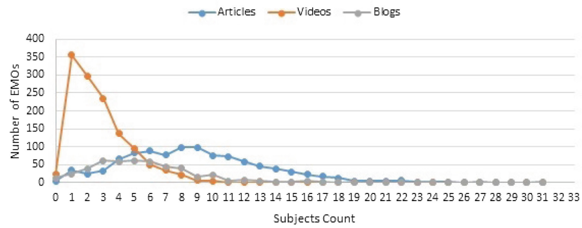
After the subject selection process, we analysed and compared the dynamically built links in the LEMO dataset. We consider that there exists a link between two EMOs if they have a similar annotated class in their subjects or keywords set. Also, we count a link between two EMOs as directed links. Therefore, if there is a link from node a to node b the link count will be two not one. We compared the number of links built in the dataset in the two experiments conducted. Table 2 illustrates the number of links built in LEMO dataset in both experiments; MeSH and SNOMED CT annotations. As detailed in the table, the number of links based on SNOMED CT ontology is greater than ones based on MeSH ontology. The results are almost doubled in the links count. This is due to the large number of annotation discovered using SNOMED CT ontology.

5.3 Ontology-Based Access

We provided a user interface for querying the LEMO Triple Store enriched with SNOMED CT ontology classes since it resulted in higher number of annotations



(a) EMOs annotated in MeSH and their subjects



(b) EMOs annotated in SNOMED and their subjects

Fig. 4. Relation between Subjects count and EMOs types

Table 2. Links count in LEMO dataset based of MeSH And SNOMED CT ontologies

	Number of Links based on		
Based on Ontology	Title	Desc	Subject
MeSH	352029	867636	248704
SNOMED CT	464876	1443667	418782

in LEMO content. The system binds the user with choosing ontological classes rather than writing a free-text in the search box. Figure 5 illustrates the auto-complete feature presented in LEMO for ontological based access. The auto-complete text box retrieves SNOMED CT ontology classes used in LEMO store. Algorithm 1 explains the ontological-based technique for searching and ranking the results of searching for a selected class.

The algorithm developed in LEMO is based on the NoSQL structure for LEMO store explained previously in Fig. 2. As explained in the ontology-based query algorithm, the search process starts with one ontology class Q. Then, a query vector is built based on the class adjacency properties stored in LEMO store. Now, we start searching for EMOs annotated with any of the ontology classes related to the query class Q. So far the search results are not ranked according to its relevance to the query initiated. Hence, the related classes retrieved are weighted according to their co-occurrences with Q class in the search result set. Then, the weights are normalized according to the length of the search result size and the class Q in the vector will have a weight of 1. For each EMO in the search result, the weights of its annotations found in QVecor

Algorithm 1: Ontology-based Query

```
inputs :
        Ontology class to be queried Q, LEMO dataset LEMO
output:
        Ranked Search Result set of EMOs
RelClasses ← getRelatedClasses(Q)            ▷Stores related classes to Q
foreach c ∈ RelClasses do
 |  qResults ← getEMOsAnnotatedWith(c)
 |  add qResults to ResSet              ▷ResSet is the final search results
end
QVector ← weightQVector(RelClasses)          ▷Weight related classes to Q
foreach d ∈ ResSet do
 |  dVector ← weightDVector(d)  ▷Weight d annotations based on QVector
end
foreach d ∈ ResSet do
 |  calculatedEucildeanDist(dVector,qVector)
end
Sort(ResSet)                                      ▷Sort results ascendingly
```

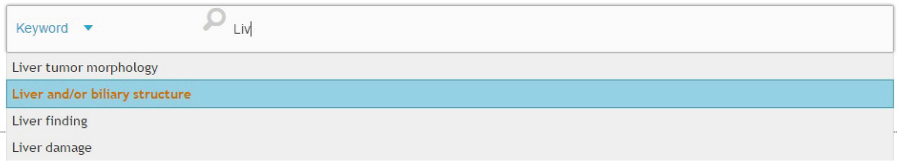


Fig. 5. The ontology-based LEMO search user interface

are retrieved from **rdf:value** in LEMO store which is used to store the weight of an annotated class as explained in the previous sections. In order to have more accurate weights for EMOs' classes, the weights are normalized based on the length of their annotation list. Finally, the search results are ranked based on its euclidean distance from the query vector as shown in the algorithm. In order to test this algorithm, we conducted random queries of 5 classes found in LEMO Triple Store. The results are shown in Table 3 and compared with exact text matching search.

From the sample of classes queried in this experiments, we can notice that the ontological-based search always retrieved higher number of results than text-based access. The overlap coefficient always indicated a percentage higher than 90 % for all the queries tested. In other words, the ontological access covers almost all the search results of text-based access. We calculated the Jaccard Similarity coefficient to emphasize the case in the last query "*RENAL DISEASE*". In this query, the text-based results are only 4 EMOs since we used exact text matching. Hence, the jaccard similarity is very low between the two search results. The query vector resulted from searching for "*RENAL DISEASE*"

Table 3. Ontological-based vs. Text-based search results in LEMO Store

Query Class	Size of Ontology-based Result set (O)	Size of Text-based Result set (T)	Overlap Coefficient $O \cap T$	Jaccard Similartiy Coef
HEPATITIS	27	21	100 %	0.78
INFLUENZA	30	25	92 %	0.71
MUSCLE	66	65	95 %	0.89
BRAIN	61	49	100 %	0.80
RENAL DISEASE	36	4	100 %	0.11

included 24 other classes related to it based on SNOMED CT ontology as stored in LEMO dataset. The list of related classes include: (Renal vascular disorder, Nephritis, nephrosis and nephrotic syndrome, Renal impairment, Infectious disorder of kidney and, $\cdots$, others). The algorithm of ontological-based query gives higher weight for the class queried Q which results in having EMOs containing the class Q ranked at the top of the search result set.

6 Conclusions and Future Work

In this paper, we present a system for linking Educational Medical Objects (EMOs) harvested from distributed web databases. The aim of the system was to bridge the gap between UGC content, provided by YouTube and blogging platforms, and online medical libraries such as PubMed. We have tested the system against a sample dataset and compared the results of using MeSH and SNOMED CT ontologies in the enrichment process. The final dataset consisted of 2720 linked EMOs which are annotated, linked, and accessed by the system developed. Using LEMO dataset enriched with SNOMED CT, we tested accessing the dataset using ontological-based approach vs. simple text-based matching. The results indicated the efficiency of ontological-based access in LEMO dataset and the overlapping coefficient between the search results of the two approaches presented values above 90 % in all queries tested. In the future, a more developed user interface will be built with more advanced features for browsing and querying the LEMO dataset presented.

References

1. Sandars, J., Schroter, S.: Web 2.0 technologies for undergraduate and postgraduate medical education: an online survey. Postgrad. Med. J. **83**(986), 759–762 (2007)
2. Bizer, C., Heath, T., Berners-Lee, T.: Linked data-the story so far. Int. J. Seman. Web Inform. Syst. **5**(3), 1–22 (2009)
3. McGreal, R.: Learning objects: a practical definition. Instr. Technol. **1**, 21 (2004)
4. Ritze, D., Eckert, K.: Data enrichment in discovery systems using linked data. In: Spiliopoulou, M., Schmidt-Thieme, L., Janning, R. (eds.) Data Analysis, Machine Learning and Knowledge Discovery, pp. 455–462. Springer, Heidelberg (2014)

5. Díaz-Galiano, M.C., García-Cumbreras, M.Á., Martín-Valdivia, M.T., Montejo-Ráez, A., Ureña-López, L.A.: Integrating MeSH ontology to improve medical information retrieval. In: Peters, C., Jijkoun, V., Mandl, T., Müller, H., Oard, D.W., Peñas, A., Petras, V., Santos, D. (eds.) CLEF 2007. LNCS, vol. 5152, pp. 601–606. Springer, Heidelberg (2008)
6. Choudhury, S., Breslin, J.G., Passant, A.: Enrichment and ranking of the YouTube tag space and integration with the linked data cloud. In: Bernstein, A., Karger, D.R., Heath, T., Feigenbaum, L., Maynard, D., Motta, E., Thirunarayan, K. (eds.) ISWC 2009. LNCS, vol. 5823, pp. 747–762. Springer, Heidelberg (2009)
7. Hoehndorf, R., Dumontier, M., Gkoutos, G.V.: Evaluation of research in biomedical ontologies. Briefings Bioinform. (2012)
8. Domingue, J., Mikroyannidis, A., Dietze, S.: Online learning and linked data: lessons learned and best practices. In: Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion, pp. 191–192 (2014)
9. Ruiz-Calleja, A., Vega-Gorgojo, G., Asensio-Perez, J.I., Bote-Lorenzo, M.L., Gomez-Sanchez, E., Alario-Hoyos, C.: A data approach for the discovery of educational ICT tools in the web of data. Comput. Educ. **59**(3), 952–962 (2012)
10. Mitsopoulou, E., Taibi, D., Giordano, D., Dietze, S., Yu, H.Q., Bamidis, P., Bratsas, C., Woodham, L.: Connecting medical educational resources to the linked data cloud: the meducator RDF schema, store and API. In: Proceedings of Linked Learning 22 (2011)
11. Mikroyannidis, A., Domingue, J., Maleshkova, M., Norton, B., Simperl, E.: Developing a curriculum of open educational resources for linked data. In: Proceedings of 10th Annual OpenCourseWare Consortium Global Conference (OCWC) (2014)
12. d'Aquin, M.: Putting linked data to use in a large higher-education organisation. In: Proceedings of the Interacting with Linked Data (ILD) workshop at Extended Semantic Web Conference (ESWC) (2012)
13. Lagoze, C., Van de Sompel, H.: The making of the open archives initiative protocol for metadata harvesting. Library Hi Tech **21**(2), 118–128 (2003)
14. Gkotsis, G., Stepanyan, K., Cristea, A.I., Joy, M.: Self-supervised automated wrapper generation for weblog data extraction. In: Gottlob, G., Grasso, G., Olteanu, D., Schallhart, C. (eds.) BNCOD 2013. LNCS, vol. 7968, pp. 292–302. Springer, Heidelberg (2013)
15. Al Fayez, R.Q., Joy, M.: A framework for linking educational medical objects: connecting web2.0 and traditional education. In: Benatallah, B., Bestavros, A., Manolopoulos, Y., Vakali, A., Zhang, Y. (eds.) WISE 2014, Part II. LNCS, vol. 8787, pp. 158–167. Springer, Heidelberg (2014)
16. Rubin, D.L., Shah, N.H., Noy, N.F.: Biomedical ontologies: a functional perspective. Briefings Bioinform. **9**(1), 75–90 (2008)
17. Jonquet, C., LePendu, P., Falconer, S., Coulet, A., Noy, N.F., Musen, M.A., Shah, N.H.: NCBO resource index: ontology-based search and mining of biomedical resources. Web Seman. Sci. Serv. Agents World Wide Web **9**(3), 316–324 (2011)
18. Noy, N.F., Shah, N.H., Whetzel, P.L., Dai, B., Dorf, M., Griffith, N., Jonquet, C., Rubin, D.L., Storey, M.A., Chute, C.G., et al.: BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res. **37**, W170–W173 (2009)
19. Elevitch, F.R.: SNOMED CT: electronic health record enhances anesthesia patient safety. AANA J. **73**(5), 361 (2005)
20. Lee, D., de Keizer, N., Lau, F., Cornet, R.: Literature review of SNOMED CT use. JAMIA J. Am. Med. Inform. Assoc. **21**, e11–e19 (2013)
21. Lipscomb, C.E.: Medical subject headings (MeSH). Bull. Med. Libr. Assoc. **88**(3), 265 (2000)