INTERNATIONAL JOURNAL OF ENGINEERING EDUCATION (IJEE)

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

Efficacy of portals on the principles of ontology and through a portal HETOP and replace it with Bio Portal (BIOPORTAL)

Hamid Ghazizadeh, PhD (Faculty of Information Sciences **payame Noor University** of Iran**)**

Mahdiye Khazaneha (PhD student of Information Sciences of Iran)

Moluk Sadat Hussein Beheshti, PhD (Faculty of Department of IRANDOC of Iran)

Fateme Motamedi, PhD (Assistant Professor in Department of Medical Library Sciences Kerman)

Abstract

Purpose: the present research aims at more accessing to the ontology terminology in health and hygiene through more effective portals.

Methodology: this applied research uses the investigation and comparisons method for determining and recognizing kinds of terminology in the portals of special field, for example anatomy, uncommon diseases, medical systems. The society is the terminology in two famous portals of ontology, HETOP and BIOPORTAL. The dada has been collected by counting the existent terminology on the basis of lexicon list in each portal.

Findings: on the basis of research findings, since T/O is used worldwide for different kinds of applications, so HETOP portal makes it possible to access more than 70 ontological terminologies in health and hygiene for the users. Also, HETOP portal has two protocols of SOAP and REST that causes the users more and more effective utilization of this web server.

Conclusion: HETOP portal has higher capability, in comparison with BIOPORTAL, because of its more effective protocols of SOAP and REST and Oracle through which, the users will be able to access the sources, medicines and different meaning of the related expression in different languages.

Keywords: ontology, access, HETOP, BIOPORTAL, terminology.

INTERNATIONAL JOURNAL OF ENGINEERING EDUCATION (IJEE)
ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org
Article Number: ISSN: 0949-149X /5493392901

Introduction

Biomedicine data are increasing constantly by new technologies and internet media. Smart methods necessitate that by indexing makes it possible the words with controlled language and proper structure for saving and recovering. One of the key aspects of creating the ability of semantics cooperation for biosciences data is the use of terminology or ontology (T/O) as a common port with data structure. (T/O) classification is the knowledge description of special domain with their relations and concepts. Ontology is more complicated than terminology since we can define functions and rules for recognizing the structure and knowledge (Walls, & Athena**,** 2012).

Since T/O usually is used in information systems and especially in hygiene, its usage has been changed to a unique port in hygienic systems. Generally, its usage is a real challenge for managing the information and hygiene.

In fact, most of ontologies have been used for different purposes in recent decays, such as document indexing, annotating, knowledge organizing, facts deductions (Hunter, 2003).

Most of daily ontologies such as ICD-104, MeSH3 and CCAM are useful in hospitals or laboratories for recovering the data. T/O has not been organized because of lack of standard and formatting. Also, syntactic semantics and cooperation ability between T/O is a major challenge for linking the systems and knowledge (Hall, & McAuliffe, 2016). Various tools are necessary for searching and saving and using several T/O simultaneously such as UMLS (Unified Medical Language System), EBI, ontology of American green card (Hanna, Joseph, & Hogan, 2013), BIOPORTAL, NCBO, and CISMeF HETOP that have been created by NCBO(National Center for Biomedical Ontology, Stanford University) and CISMeF team (Catalogue et Index des Sites Medicaid de Langue Françoise)[1]. On the basis of T/O the most important of them are BIOPORTAL, HETOP, and their services for searching the documents on the web.

Both portal groups, BIOPORTAL and HETOP provide different services, for example indexing, cataloguing, virtualizing, and annotating. Regarding to different applications of these two portals, sharing the sources and discovering new sources on the basis of T/O are general purposes to provide services for the medical society in using the newest sources.

In this article, at first we introduce HETOP portal, and then compare it with BIOPORTAL.

Methodology

[1] Catalog and Index of French-Speaking Medical Sites

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

By using the site of HETOP and BIOPORTAL and also the previous studies the data has been collected. Finally, the architectures, contents, and specialty field of these portals have been investigated and compared.

Introducing the software HETOP:

This ontology portal provides the access to more than 70 terminologies/ ontology (T/O) in health and hygiene. T/O, at present, is used worldwide for different applications, such that these terminologies are related to special subjects such as anatomy, uncommon diseases, and medical systems and so on. Web service of HeTOP consists of protocols SOAP and REST, and its call web service is too. This web service is accessible through its protocols but it is necessary to sign up. Protocol REST has compatible network technologies; therefore, it is easy to use it even in network browsers.

Three steps of creating an account:

1. Creating an account and password
2. The account and password must be coded on the basis of Base64
3. A certificate by the title of "Basic" is necessary.

Concepts of REST and SOAP

We can design the REST concept on the same basis of web architecture and concept design. Its fast and easy development and its long life are proper architectures for unexpected development and changes in long-time. In REST-based systems we consider all things as a series of sources that are specified by URIs. Indeed, these URIs are not fixed necessarily, and don't point to a real source, rather they are abstract and REST uses them. According to REST, these URIs are selected by the server and can be changes (client should receive this URI form the server itself). There is no fixed case in this field. Xml is used for providing the sources. We use http and methods of PUT, POST, HEAD, GET and DELETE for performing the operations. URI is used for ID of a web service, and the response has the format of xml.

SOAP is one of the most common standards which are used in web services. According to the observations, it has been introduced by Developer Mentor, User Land Company and Microsoft in 1998 for the first time, and its first version has been published in 1999. The main purpose of SOAP is creating a method for sending data among systems that are distributed on the network. When a program starts to make a connection with the web service, SOAP messages are tools for linking and transferring data between them. A SOAP message is sent to the web service, and a function implements it, i.e. this message wants the web service to do a task. The web service uses the message content and starts its operations. In the end, it sends the results with another SOAP message to the main program.

Search method in HETOP

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

There are two search method in this portal network service, search (query, langs, options) and get Bylds (cismefldds, langs, options), that confirm additional efficiencies for the below purposes:

1. Shorter time because of removing some steps
2. Improving XML output
3. Deleting the unrelated classifications
4. There is an alternative in the option section that is similar to urn parameters.

HETOP network service recovers a specialized format-based XML file by Boolean logic. XML is a hypertext language as HTML, and tags and signs are used for grouping and transferring the information in it. Necessary labels for programming it are presented in the below table.

Table 1: necessary labels for programming in HETOP

Labels	Meanings
cis:dboResp	Concept list
cis:dbo	Concept
cis:id	Concept identifier
cis:ti	Metadata identifier (concept, data property or relation property)
cis:lb	Concept or metadata label
cis:an	Concept or metadata annotation
cis:dps	Data property list
cis:dp	Data property
cis:hies	Hierarchy list
cis:hie	Hierarchy
cis:bt	Broader concept
cis:nt	Narrower concept
cis:path	Hierarchy path
cis:ops	Relation list
cis:op	Relation between concepts
cis:tar	Target concept id

Data in HETOP is on the basis of metadata model. Metadata means information about information, such that each file containing information has an explanation wit itself. Metadata may explain the structure of file, or has some information about the file content. The ID structure of CISMeF used by HETOP is itself a metadata model.

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

What is CISMeF?

CISMeF is the abbreviation of Catalog and Index of French-speaking medical sites. In December 2007 the number of indexed sources has been more than 41300 by the average of 80 new sources in each week. CISMeF has been started by RUH in February 1995. In 2007, CISMeF emerged in information world as a source information for health. It's a health portal and the most important source and information of organizational health in France because of its controlled quality and index. The purpose of CISMeF is to help the hygiene specialists and consumers to search electronic information in the internet. CISMeF has used two standard tools for organizing the information: midline, thesaurus of bibliographic database, and several metadata elements such as Dublin Core. By the title, author, subject, keywords, explanations, publisher, date, source type, format, ID, and language, the sources in CISMeF have been described. The sources indexes in CISMeF have been organized by the use of four different concepts of meta-term, subheading, keyword, and resource typeCISMeF that consist of subject list including medical specialties and alphabetic list. HON codes, evaluation criteria for hygienic information quality on the internet, are important for CISMeF. CISMeF is an important tool for French-speaking hygienic society.

The purpose of HETOP portal is not only to use the ontological rules and operations and the existing data models in this portal, but also to access the terminologies of informal, natural and native languages.

It means that real ontology decomposed the terminologies to be usable. BIOPORTAL in ontology is used for conceptual maps and is one of the oldest portal in ontology portals for medical informatics and other cases.

Comparison of HETOP and BIOPORTAL

In this study, BIOPORTAL and HETOP have been compared on the basis of many criteria. This comparison helps to understanding the philosophy and properties of both portals. Furthermore, it allows find the advantages and weak points of BIOPORTAL and HETOP, and use the portal with more advantages in the ontology tools.

Table1. Comparison of HETOP and BIOPORTAL(Cane, & Michael, 2014)

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

Property	BioPortal	HeTOP
T/O metrics	n(T/O) = 368 n(concepts) = 5 960 457 n(terms) = 6 600 000 n(relations) = ? n(mappings) = 5 000 000	n(T/O) = 56 n(concepts) = 1 951 834 n(terms) = 6 636 654 n(relations) = 8 023 181 n(mappings) = 1 340 855
T/O source	From the UMLS From the OBO Foundry Directly from registered users	From the UMLS From different official sources (see HeTOP terminologies list for details) From other research teams
T/O formats	OWL, OBO, RDF, RRF	No automatic imports
Handling T/O initial model	No	Yes
T/O documentation	Documentation available on specific pages	Depending on the T/O, on a single page
T/O update	Automatic for OBO Foundry and UMLS. Available form for custom T/O	Depending on the T/O and its update frequency, some are automatic (4)
T/O versioning	Yes, with a persistent `ontology_version_id`	No, only the lastest version available
T/O languages	Some "views" of English T/O	n = 23, coverage depending on the T/O
T/O added content	User annotations New automatic mappings	Opened to the community, reviewed by curators New content from automatic/manual tools
T/O organization	T/O are categorized in groups and categories	N/A
T/O quality leverage	No	Yes, on the vast majority of T/O
T/O interoperability	Automatic and manual mappings	Automatic and manual mappings

Table 2: comparison of properties of two portals BIOPORTAL and BIOPORTAL

Comparison of contents in table 2 shows that there is a difference between BIOPORTAL and HETOP in information volume that is very important (respectively, 3 t0 6 ratio). In single-language BIOPORTAL hasn't the proper comparison index for concepts and synonyms in relation to Multilanguage HETOP portal. The more records and statistics are related to the total number and the relations number. In fact, BIOPORTAL and HETOP almost have the same words number (about 6600000). But unfortunately calculating the relations number in BIOPORTAL is not easy and possible. The comparison formats in these two portals are completely different.

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

Table2: comparison formats between two portals of BIOPORTAL and HETOP(Boaco, 2015)

Property	BioPortal	HeTOP
T/O choices	UMLS ontologies and community ontologies	Reference T/O and projects T/O
T/O access	All ontologies are public except those set to private by users	22 T/O are freely available, 16 more when registered and the others are restrited to research projects users (e.g. industrial partners T/O)
T/O edition	No edition available but one can add notes to concepts	No edition available (except for curators) but one can suggest translations or synonyms for specific T/O. CISMeF team curators can valid them.
Downloadable content	Yes for public ontologies and public mappings	No
User options	Users can own ontologies and manage versions. Projects creation is possible.	A query history and a selection history are available.
	The user interface is mostly the same for every ontologies.	A special effort is brought to adapt the user interface to the T/O content. Especially for metadata labels (attributes, relations, etc.) and for a multilingual use.
User documentation	Yes, in a Help tab	No
Registered users	n = 3 017	n = 1 518
Traffic	17 500 hits/day	15 000 hits/day (500 users/day)
Research projects	215	HeTOP is a support a research projects and integrates some productions such as VCM icons, mappings, interface terminologies, etc.

BIOPORTAL concentrates on the odontology. It is time consuming when using programs and standard formats. But HETOP is the host of heterogeneous formats such as Microsoft excels files, XML files and or database (Arp, 2015).

Table 3: heterogeneous files of Microsoft excel, XML files in two portals (Cobb & Graham, **2011)**

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

Property	BioPortal	HeTOP
search engine	Exact match terms (concepts preferred terms and synonyms) and advanced options	Exact and partial match terms (concepts preferred terms, synonyms, ids, definitions,...), and advanced options
results display	In a dedicated page, organized by T/O	Integrated into a unique view, limited to 500 results and organized by T/O and concept types
concept view	In a description tab	In a description tab
hierarchy view	Always visible (if provided), no poly-hierarchy	In a dedicated tab, with poly-hierarchy
relations view	Flat, uncommented relation types	Organized and understandable and multilingual relations types
mappings view	In a dedicated tab	Embedded in the relations tab
resources access	Not accessible from the concept page	Dedicated tab to access the InfoRoute tool
cross-lingual navigation	No	Yes by clicking on the flags

In spite of similarity of basic tools for the two portals, there is some trivial differences (table 3), and there is a direct result for the ultimate users. BIOPORTAL search engine provides the search only in English (that has special relations and synonyms). But HETOP is able to add general lexicons to the search words in two languages simultaneously. This affects directly the users search. For example, if we search "myopathy" in English in NCIT10, BIOPORTAL provides 5 words, while HETOP recalls 25 ones. HETOP engine has many words, for instance, it recalls "*myopathy" * and words such as "cardio-myopathy". Available tools for BIOPORTAL and HETOP are not the same.

Table 4: comparison of recalls between two portals

Property	BioPortal	HeTOP
Multilingual display	No	Yes
Contextual links to other portals/browsers	No	Yes, to several other portals (including BioPortal, MeSH, LOINC, OMIM browsers)
Navigation display	Hierarchy and concept links	Hierarchies, concept links and cross-lingual navigation
Search engine results (execution times/number)	5.57 sec. / 19.7	3.17 sec. / 359.4
Search engine results with wildcard	Not applicable on search tab	4.46 sec / 501.3
Concept page access time	4.1 sec.	less than 1 second

Response time and response operations in the portals have been compared.

Table 5: comparison of the portals architectures

INTERNATIONAL JOURNAL OF ENGINEERING EDUCATION (IJEE)

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

Property	BioPortal	HeTOP
Data model	RDF (ontologies are loaded in a triple store)	Meta-model for T/O which encapsulate specific T/O models
Database implementation	4Store triple store	Oracle 11g r2, with partionning & domain indexes options
Web services	Accessible in REST (http://www.bioontology.org/wiki/index.php/Resource_Index_REST_Web_Service_User_Guide)	Accessible in SOAP and REST (http://cispro.chu-rouen.fr/CISMeFhetopservice/)
Other API	CTS2 (OMG Standard)	Yes but not accessible
Technical documentation	NCBO wiki: full support documentation about technologies and developments	Yes but not accessible
Web site technologies	Ruby on Rails, Javascript, Spring/Hibernate, Protégé, LexGrid, Rainbow	Vaadin, JEE, Infinispan, CXF
License	open source	proprietary
Reusability	NCBO Virtual appliance is available to install and run its own version of BioPortal locally	N/A

In table 5, methods, technologies, sources and their productions are compared. It is notable that in portal HETOP, oracles are used, representing the terminologies is hierarchical, and the response time for the terminologies and concepts is shorter than BIOPORTAL.

Conclusion

Produced HETOP for many T/O, metadata and labels (attributes, relations and so on) have been well translated in several languages. Through HETOP we can access to the sources, medicines, identifiers, different meanings of the related terminology to the various languages. But BIOPORTAL is provided only in French. Since in HETOP organization a team of specialists of NCBO, CISMeF and LIRMM have cooperated, it is more powerful in terminology field.

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

By studying the contents of two portals in this research, we got various results. Noting the data of the research tables that has been collected from the contents of two portals of HETOP and BIPORTAL, the results showed that HETOP is more capable than BIOPORTAL in the case of terminologies, format type, heterogeneous structure of Microsoft excel files, XML, fast response.

Regarding abovementioned matters, it is suggested that the users when utilize the medical ontological terminologies, instead of common portal of BIOPORTAL that is in lower level than HETOP in format, lexicons, Microsoft files consistency, prefer the portal of HETOP to get more ontological terminologies.

References

Arp, R., (**2015**). Building ontologies with basic formal ontology. Cambridge: The MIT Press.

Boaco, F. A., (**2015**). Correlational effect size benchmarks. Journal of Applied Psychology, 100, 431–449.

Cane, J. & Michael, S. (**2014**). Validation of the theoretical domains framework for use in behavior change and implementation research. Implementation Science, 7, 37.

Cobb, N. K., & Graham, A. (2011). Online social networks and smoking cessation: A scientific research agenda. Journal of Medical Internet Research, 13, e119.

Hall, D. & McAuliffe, A. (2016). Sharing heterogeneous data: The national database for autism research. Nauru informatics, 10, 331–339.

Hanna, J., Joseph, E., & Hogan, W. R. (2013). Building a drug ontology based on Rexnord and other sources. Journal of Biomedical Semantics, 4, 44. Retrieved from: http:// www.jbiomedsem.com/content/4/1/44

Hunter, J. (2003). Enhancing the semantic interoperability of multimedia through core ontology. IEEE Transactions on Circuits and Systems for Video Technology, 13, 49–58.

ISSN: 0949-149X July-August2017 Vole 9, No 4 www.sweetmaxwell.org

Article Number: ISSN: 0949-149X /5493392901

Walls, R. & Athena, E. (**2012**). Ontologies as integrative tools for plant science. American Journal of Botany, 99, 1263–1275.

Wong, W., Liu, W. (2012). Ontology learning from text: A look back and into the future. ACM Computing Surveys (CSUR), 44, 20.

Health Organization. (2013). How to use the ICF: A practical manual for using the international classification of functioning. Disability and health (ICF). Exposure draft for comment.