Process of Building Reference Ontology for Higher Education

L. Zemmouchi-Ghomari and A. R. Ghomari

***Abstract*—Most ontologies are application ontologies, that are not reusable and are difficult to link together as they are too specific. Reference ontology is able to contribute significantly in reducing the issue of ontology applications specificity.**
Particularly considering higher education domain, we think that a reference ontology dedicated to this knowledge area, can be regarded as a valuable tool for several stakeholders interested in analyzing the system of higher education as a whole, especially in a context of academic systems diversity all over the world.
Motivated by this potential application and even more, we decided to build a reference ontology called HERO ontology, which stands for "Higher Education Reference Ontology".
In this paper we explain HERO ontology building process from requirements specification until ontology evaluation using NeOn methodology.

***Index Terms*—Reference ontology, higher education ontology, ontology engineering, NeOn methodology**

I. INTRODUCTION

THE majority of available ontologies are too specific and do not stand the test of large applications [1]. Consequently, constructing ontologies from scratch to support domain applications requires a great deal of effort and time [2].

Reference ontology is able to contribute significantly in resolving or at least reducing the issue of ontology applications specificity and might provide significant advantages over domain and application ontology previously used [3][4].

Besides, reference ontology for higher education domain can be considered as a relevant instrument for sharpening an institution's mission and profile [5][6][7][8][9][10]. By focusing on the relevant constituents of the ontology, the institutions indicated that they would be able to strengthen their strategic orientation and develop and communicate their profile [5].

In addition, institutions indicated that they would be highly interested in identifying and learning from other institutions comparable to them on a number of relevant dimensions and indicators

Manuscript received March 1, 2013.

Leila Zemmouchi-Ghomari is with UMBB, M'hammed Bouguerra Boumerdès University, Boumerdès, Algeria, (e-mail:l_ghomari@umbb.dz).

Abdessamed Réda Ghomari is with LMCS Laboratory, E.S.I, National Superior School of Computer Science, Algiers, Algeria, (e-mail: a_ghomari@esi.dz).

In this perspective, we decided to build reference ontology for higher education area that we called HERO ontology which stands for "Higher Education Reference Ontology".

We adopted NeOn methodology [11] to construct HERO ontology since it is based on famous ontology engineering methodologies such as: METHONTOLOGY [12], On-To-Knowledge [13] and DILIGENT [14] combined with good practices and feedback from previous experiences of NeOn consortium members.

In this paper, we explain HERO ontology building process from requirements specification until ontology evaluation throughout section 3 and section 4. Section 2 is intended to present the related work. We conclude our work in section 5 by presenting the main results and what follows from our work.

II. RELATED WORK

There exist some university domain ontologies that can be qualified as good representations with regard to aspects, such as: correctness of syntax language, satisfactory coverage degree of university domain. Nonetheless, these ontologies do not meet the criteria of reference ontologies defined in [4], in short: to be heavyweight (to include axioms that limit ontology primitives' interpretations), to contain only central concepts (to be core ontology) and to not be intended for specific applications.

A. University Ontology

The author is Jeff Heflin of Lehigh University. The current version dates from 2000 and is no longer maintained. This ontology defines elements for describing universities and the activities that occur at them. It includes concepts such as departments, faculty, students, courses, research, and publications. This ontology is a lightweight ontology (no inference rules are defined).

B. Univ-Bench

LUBM, Lehigh University Benchmark, the author is Zhengxiang from Lehigh University. The current version dates from 2004. This ontology has been developed to facilitate the evaluation of Semantic Web repositories in a standard and systematic way. The benchmark is intended to evaluate the performance of those repositories with respect to extensional queries over a large data set that commits to a single realistic ontology. This ontology has been designed for a specific application, i.e.: to provide synthetic data for test queries and performance metrics.

ISBN: 978-988-19252-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

C. Academic Institution Internal Structure Ontology (AIISO)

The authors are Rob Styles and Nadeem Shabir from Talis. The current version dates from 2008. The Academic Institution Internal Structure Ontology (AIISO) provides classes and properties to describe the internal organisational structure of an academic institution. AIISO is designed to work in partnership with Participation Schema (http://purl.org/vocab/participation/schema), FOAF (http://xmlns.com/foaf/0.1/) and AIISO-roles (http://purl.org/vocab/aiiso-roles/schema) to describe the roles that people play within an institution. This ontology focuses on structural perspective of a university (this is reflected by the small number of classes: 15).

III. HERO DEVELOPMENT PROCESS

A. Selected scenarios

Ontology reuse is recommended by default in current methodologies and guidelines as a key factor to develop cost-effective and high-quality ontologies. The underlying principle is that reusing existing and already consensuated terminology allows saving time and money in the ontology development process, and promotes the application of good practices [2].

Due to the complexity of the domain of interest on one hand and the need of a broad coverage of the reference ontology, we decided to combine development from scratch with reuse-oriented engineering strategy, performed according to the following phases:

- During the specification phase: reuse of non ontological resource, such as academic classifications (e.g. Carnegie Classification), domain-related sites and documents: our search strategy consists in focusing on reliable and well recognized information sources, like: official governmental websites, university Associations, academic reports.
  Then, every time we encounter an interesting statement that is supposed to be encoded in the ontology we consider it as a potential answer to a competency question.
- During the conceptualization phase: reuse of ontologies via Watson tool (NeOn Toolkit plug in) which integrates the search capabilities of the Watson Semantic Web gateway within the environment of the ontology editor (the NeOn Toolkit) and finds, in online ontologies, statements that are relevant to extend the description of a particular ontology entity. The statements selected by the user are integrated in the ontology.

B. HERO building phases

In this section, we present Hero building process from specification to implementation.

Specification phase

Neon methodology provides very precise guidelines in the specification phase throughout the Ontology requirement specification document (ORSD) [11] and the competency questions technique [15] (Table I).

TABLE I. HIGHER EDUCATION REFERENCE ONTOLOGY REQUIREMENTS SPECIFICATION DOCUMENT (EXCERPT OF HERO ORSD)

Purpose
The purpose of building the Reference Ontology is to provide a consensual knowledge model of university domain that can be considered as a basis to derive more specific university domain ontology. This reference ontology is named " HERO" which stands for "Higher Education Reference Ontology"
Scope
The ontology has to describe several aspects of university domain such as organisational structure, staff (academic and administrative), roles (teaching, research), and incomes. The level of granularity which is determined by the level of concept specificity must be relevant or at least convenient to describe any university. Considering the level of formality, the reference ontology, has to be a formal and heavyweight ontology.
Implementation Language
The reference ontology has to be implemented in OWL 2
Intended End-Users
Users of this ontology might be: 1. Ontology developers in collaboration with domain experts interested in building domain or application ontologies by deriving more specific ontologies from the reference ontology; 2. Ontology aligners and evaluators which may consider this reference ontology as a gold standard in the ontology alignment and/or the evaluation process; 3. Academic Committees, higher education governmental structures, independent accreditation organisms which are interested in analyzing higher education issues, such as: university funding, universities ranking, higher education policy in general.
Intended Uses
The main use of a reference ontology describing higher education domain is to provide a consensual knowledge about the domain of interest in order to be shared and reused among different users, different communities and different universities.
Ontology Requirements
a. Non-Functional Requirements
These requirements refer to syntactical aspects of the resulted ontology such as the language in which the ontology is described, namely: English language; Another aspect is the selected terminology that will be used to describe the ontology. The designated terminology has to be consensual or at least, used in most higher education institutions.
b. Functional Requirements
Refer to the particular knowledge that has to be represented by the ontology, or what knowledge the ontology must contain. This specification is achieved by using the technique of competency questions. The adopted strategy for identifying these questions is the Middle-Out approach: we start by selecting the most important questions regarding to the ontology goals further we will study the possibility to decompose (and obtain more concrete or simpler questions) or on the contrary, to gather some of them (and obtain more abstract or complex questions). The competency questions have to be classified by category and optionally by priority. This categorization facilitates the highlight of the different ontology modules belonging to the same domain and hence increases their reusability. Additionally, sorting these questions into categories present some advantages such as to highlight key concepts (the most information-rich concepts) deduced by two indicators, the first one is the intersection between categories which overlap because of their dependencies, for example: a person can be a student and a teacher at the same time, a dean is an administrator and a research project

manager, on the other hand, the frequency of the terms extracted from the questions and their answers (the frequency of a given term is proportionally accorded to its importance).

We opted for six categories inspired by higher education domain classification of ACE (American Council on Education) reports, which is a leading organism in analyzing higher education domain, these categories are:

1. Faculty, appointments and research area
2. Student and their life
3. Administration
4. Degrees and curriculum programs
5. Finance
6. Governance

As a first attempt to detect HERO ontology requirements, we have identified eighty one (81) Competency questions (CQs) in the specification phase of HERO ontology development process, some examples are provided by table II.

TABLE II. EXAMPLES OF COMPETENCY QUESTIONS AND THEIR EXPECTED ANSWERS

CQs' Examples	Corresponding Answers
CQ3. Must a university teacher be a researcher?	Nearly all faculty members are expected to engage in research
CQ4. What is expected from university teachers?	Teachers are expected to (1) teach; (2) engage in research and scholarly or artistic activity; and (3) serve their campus, scholarly association, and local communities.
CQ53.What high education degrees exist?	Undergraduate degrees are: associate and bachelor degrees, graduate degrees are master and doctorate degrees.

From the requirements in form of competency questions and their respective answers, we extracted the terminology (names, adjectives and verbs) that will be formally represented in the ontology by means of concepts, attributes and relations [16] as illustrated in table III.

TABLE III. GLOSSARY OF TERMS AND THEIR FREQUENCY (EXCERPT)

Faculty, appointments and research area	**Student and their life**	**Degrees and curriculum programs**
Faculty member: 12	Service: 11	Associate degree: 10
Faculty: 36	Student: 74	Bachelor degree: 21
Research: 39	**Administration**	Course: 33
Teacher tenure: 11	Admission: 11	Credit: 29
Teacher:12	Department: 14	Degree: 73
University: 17	Higher education institution: 13	Grade: 19
Finance	Organisation: 11	Grading: 11
Financial aid: 5	President: 10	Program: 34
Tuition: 5	Private: 10	

Conceptualization phase

The conceptualization phase includes the concepts supposed to exist in the world and their relationships. This step integrates the following intermediate representation techniques: Data Dictionary (DD) (table IV), Concepts hierarchy (figure 1), Attributes Classification Trees (figure 2) and Object properties table (table V).

TABLE IV. DATA DICTIONARY (EXCERPT)

Attributes of concept: Course
1. CourseCategory
2. CourseClassHours
3. CourseCode
4. CourseCreditsNumber
5. CourseGradingSystem
5.1. CriterionReferencedGrading
5.2. GradingOnCurve
5.3. NonGradedEvaluation
5.4. PassFailSystem
6. CourseLevel
7. CourseMaterial
8. CoursePrerequisites
9. CourseRoom
10. CourseSessionCode
10.1.SessionTiming
10.2.SessionType
11. CourseSyllabus
11.1.CourseDescription
11.2.CourseObjectives
12. CourseTitle
13. Lecture
13.1.LectureRoom
13.2.LectureSchedule

Once the ontology builder has almost done the DD, the next step is to develop Concepts Classification trees. Given all the concepts of the DD, a concepts classification tree usually organizes the domain concepts in a class/subclass taxonomy in which concepts are linked by subclass-of relations.

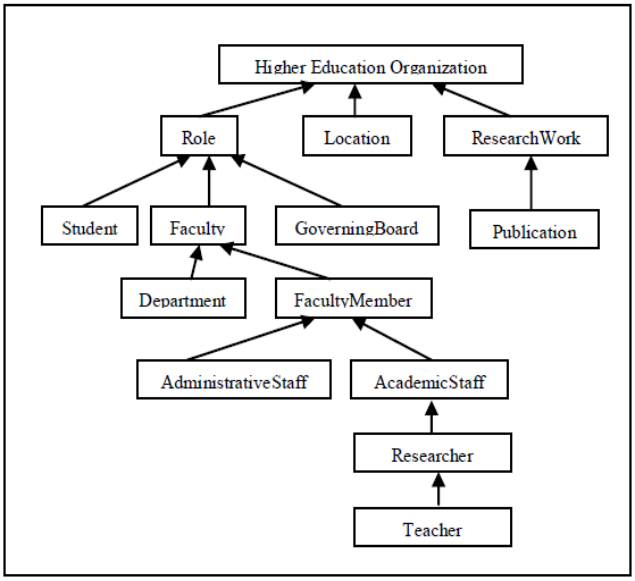

Fig 1 Key Concepts tree of HERO ontology

ISBN: 978-988-19252-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

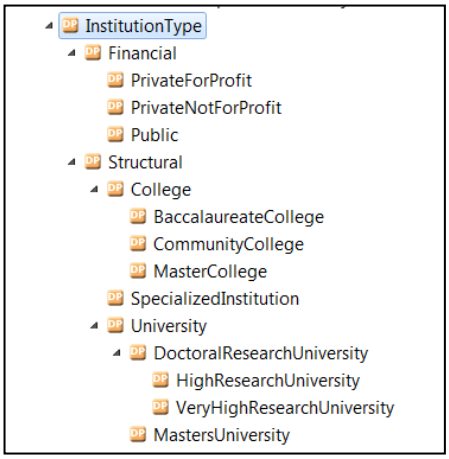

Fig 2 Attributes Classification Tree (institution type example)

TABLE V. HERO OBJECT PROPERTIES (EXCERPT)

ObjectProperty	Domain	Range
AppointedTo	Teacher	Department
BelongsTo	Researcher	Research Group
Composed Of	Research Group	Researcher
Cooperates With	Researcher	Researcher
Enrolled By	Student	Higher Education Organisation
Organises	Laboratory	Seminar
Provides Financial Aid To	Higher Education Organisation	Student
Studies At	Student	Department
Supervised By	Student	Teacher
Supervises	Teacher	Student
Writes	Researcher	Publication

Formalization & Implementation phase

Formal ontology must include axioms or axiomatic theories using formal language to constrain the possible interpretations of the ontology components. Since OWL is based on Description Logics, we used it to express property restrictions. These restrictions are used to limit the individuals belonging to a single class and contain anonymous classes that satisfy those limits.

An example of HERO property restrictions is provided beneath:

Restriction: Doctoral degree is necessarily preceded by a research master degree

TABLE VI. EXAMPLE OF A RESTRICTION

```
<owl:Class rdf:about="http://www.UniversityReferenceOntology.org/
HERO#Doctorate">
<rdfs:subClassOf>
<owl:Restriction>
<owl:onProperty rdf:resource="http://www.UniversityReferenceOntology.or
g/HERO#HasDegree"/>
<owl:allValuesFrom rdf:resource="http://www.UniversityReferenceOntolog
y.org/HERO#ResearchMaster"/>
</owl:Restriction>
</rdfs:subClassOf>
</owl:Class>
```

When all classes, properties, relations and restrictions are created, the next task to perform in the ontology development process is to document these previously mentioned components. This kind of information is named annotations or metadata. Besides, ontology metadata helps users to understand, maintain and update the ontology knowledge and its consistency.

Finally, HERO ontology has been implemented in OWL 2 DL profile produced by Neon Toolkit editor. The resulting ontology is available at: http://sourceforge.net/projects/heronto/?source=directory.

More information about the ontology is provided in HERO ontology website [17].

IV. HERO ONTOLOGY EVALUATION

In this section, we measure ontology quality with regard to three main groups of dimensions: structural, functional and usability-related dimensions [18].

A. Structural Evaluation

Structural evaluation considers the logical structure of the ontology, usually depicted as a graph of elements which focuses on syntax and formal semantics of ontology graph.

Several language-dependent ontology verification tools and ontology platforms, such as Protégé, NeOn toolkit with Pellet[19], FaCT++ [20], Hermitt [21] and Racer [22], can be used in order to evaluate these ontologies. Such tools focus on detecting inconsistencies and redundancies in concept taxonomies.

After submitting HERO ontology to the previously mentioned reasoners, neither inconsistency nor redundancy has been discovered.

B. Functional Evaluation

Functional evaluations focus on the usage of the ontology, how well it matches the intended conceptualization or a set of contextual assumptions about a world. This evaluation can include: expert agreement, user satisfaction, task assessment, and topic assessment.

In our case, we focused on expert agreement and task based assessments that are explained below.

Evaluation by domain experts

This type of evaluation is done by domain experts who try to assess how well the ontology meets a set of predefined criteria, standards and requirements [23].

With the aim of achieving this investigation we chose to use an online questionnaire [24][25] proposed to higher education domain experts who include: researchers, teachers, administrators and students (current and alumni).

The survey has been divided into several parts, namely: verification of the five levels of the ontology; verification of restrictions; verification of relations between concepts and verification of descriptive attributes of concepts.

Domain experts answer to the questionnaire and make several comments on the knowledge encoded in the ontology, such as: a student could be enrolled in an undergraduate program and in a graduate program at the same time while we have declared undergraduate student class and graduate student class as disjoint classes (no common instances).

ISBN: 978-988-19252-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

As a result of this evaluation, the ontology has been updated according to opinions' experts who obtained the majority (more than 50%) because the purpose of reference ontology is to materialize specialists' consensus.

Evaluation via Competency Questions technique

In order to achieve the translation of natural language competency questions into SPARQL queries (The entire set of competency questions and their corresponding SPARQL queries are available at: http://herontology.esi.dz/content/downloads), we proposed a new approach [26] even though inspired by the guidelines proposed in [27].

This approach can be summarized in five steps:

1) Identifying competency questions categories according to expected answers' types
2) Determining the expected (perfect or ideal) answer;
3) Extracting Entity or Entities from questions and their corresponding expected answers identified in 2;
4) Identifying answer entity type (class, data property, object property, annotation, axiom, instance) and its location in the ontology;
5) Constructing the appropriate SPARQL query that gives the closest answer to the ideal answer: based on question type identified in 1 and questions'/answers' characteristics extracted from 3 and 4, namely: entity, entity type and its location in the ontology.

In table VII, we present some competency questions with their corresponding SPARQL queries.

TABLE VII. SOME COMPETENCY QUESTIONS WITH THEIR CORRESPONDING SPARQL QUERIES

Competency Question	SPARQL QUERY
CQ3. Must a university teacher be a researcher?	ASK { HERO:Teacher rdfs:subClassOf HERO:Researcher .}
CQ4. What is expected from university teachers?	SELECT ?prop ?range WHERE {?prop rdfs:domain HERO:Teacher ; rdfs:range ?range ; a owl:ObjectProperty .}
CQ53.What high education degrees exist?	SELECT * WHERE { ?subclass rdfs:subClassOf HERO:Degree }

As a result to HERO evaluation via competency questions technique, we can confirm that knowledge encoded in HERO ontology is sufficient to respond to SPARQL queries translated from this set of natural language competency questions.

C. Usability issues

This evaluation dimension depends on the level of annotation of the evaluated ontology. How easy is for users to recognize ontology properties? How easy is for users to find out which ontology is more suitable for a given task? [18].

Motivated by annotations' benefits described previously, big effort has been spent in annotating HERO ontology (classes and properties). Based on input resources described in selected scenarios (III.A), we have documented our ontology by 97 annotations, as follows: 48 definitions, 37 comments and 12 labels, in order to allow better understanding of ontology components.

V. CONCLUSION

This work was undertaken to construct Higher Education Reference Ontology by following the guidelines indicated by NeOn methodology.

In fact, reference ontology for higher education domain can serve as an Instrument for university profiling and strategy development in addition to providing a non discriminatory ranking tool.

Based on these potential applications, we undertook the construction of a reference ontology for higher education ontology (HERO ontology), by combining three scenarios among the nine scenarios proposed by NeOn methodology, namely: development from scratch with reuse of ontological and non ontological resources in order to achieve a broad coverage of relevant concepts describing the knowledge related to the domain of interest.

These concepts have been related to each other via hierarchical links and associations, described by properties and bounded in their interpretations by some axioms.

HERO ontology has been evaluated according to several perspectives: structural, functional and usability issues.

Besides, assessment should not stop at this point, since reference ontology objective is to reach the broadest possible agreement of domain experts.

REFERENCES

[1] D. Oberle, A. Ankolekar, P. Hitzler, P. Cimiano, M. Sintek, M. Kiesel, B. Mougouie, S. Baumann, S. Vembu, M. Romanelli, P. Buitelaar, R. Engel, D. Sonntag, N. Reithinger, B. Loos, H. Zorn, V. Micelli, R. Porzel, C. Schmidt, M. Weiten, F. Burkhardt, and J. Zhou, "DOLCE, ergo SUMO: On Foundational and Domain Models in the SmartWeb, Integrated Ontology (SWIntO)", *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 5, N° 3, 2007, pp. 156–174.

[2] L. Zemmouchi-Ghomari and A. R. Ghomari, "Terminologies versus Ontologies from the perspective of ontologists", *International Journal of Web Science*, Inderscience Publisher, Vol. 1, N° 4, 2013, pp. 315-331.

[3] B. Smith and C. Welty, "Ontology: Towards a New Synthesis", *in Proc. of Formal Ontology in Information Systems*, New York, 2001, pp. 3–9.

[4] L. Zemmouchi-Ghomari and A. R. Ghomari, "Reference ontology", *in Proc. of International IEEE Conference on Signal-Image Technologies and Internet-Based System*, Marrakech, Morocco, 2009, pp. 485-491.

[5] F. A. Vught, *Mapping the Higher Education Landscape: Towards a European Classification of Higher Education*, Springer, 2009.

[6] M. Bucos, B. Dragulescu, and M. Veltan, "Designing a Semantic Web Ontology for Elearning in Higher Education", *in Proc. of IEEE 9th International Symposium on Electronics and Telecommunications*, Timisoara, Romania, 2010, pp. 415 – 418.

[7] A. Laoufi, S. Mouhim, E. H. Megder, C. Cherkaoui, "An Ontology Based Architecture to Support the Knowledge Management in Higher Education", *in Proc. of International Conference on Multimedia Computing and Systems*, Ouarzazate, Morocco, 2011, pp. 1–6.

[8] J. Mesaric and B. Dukic, "An Approach to Creating Domain Ontologies for Higher Education in Economics", *in Proc. of 29th International Conference on Information Technology Interfaces*, Cavtat, Croatia, 2007, pp. 75 – 80.

[9] A. Burgun, "Desiderata for domain reference ontologies in biomedicine", *Journal of Biomedical Information*, Elsevier Publisher, Vol. 39, N° 3, 2006, pp. 307-313.

[10] J. Milam, *Ontologies in Higher Education*, Information Science Publishing, 2006, Ch. 3, pp. 34-62.

[11] M. C. Suarez-Figueroa, "NeOn methodology for building ontology networks: Specification, scheduling and reuse", Ph.D. dissertation, Univ. Politécnica, Madrid, 2010.

[12] A. Gomez-Perez, M. Fernandez, and A. De Vicente, "Towards a method to conceptualize domain ontologies", *Workshop on Ontological Engineering, ECAI '96*, Budapest, Hungary, 1996, pp. 41–52.

[13] Y. Sure and R. Studer, *On-To-Knowledge methodology*, J. Wiley and Sons, 2002.

[14] H. C. Pinto, C. Tempich, and S. Staab, "Diligent Towards a fine-grained methodology for distributed, loosely-controlled and evolving engineering of ontologies", *in Proc. of the 16th European Conference on Artificial Intelligence*, Valencia, Spain, 2004, pp. 393–397.

[15] M. Gruninger and M. S. Fox, "Methodology for the design and evaluation of ontologies", *in Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, Canada, 1995, pp. 6.1–6.10.

[16] M. Suarez-Figueroa, A. Gomez-Perez, and M. Fernandez-Lopez, *the NeOn Methodology for Ontology Engineering*, Springer, Berlin Heidelberg, 2008, Ch. 1, pp. 9-34.

[17] http://herontology.esi.dz/, HERO ontology website (accessed on march 1st 2013)

[18] A. Gangemi, C. Catenacci, M. Ciaramita, and J. Lehmann, "Ontology evaluation and validation, an integrated formal model for the quality diagnostic task", Technical report, Laboratory for Applied Ontology, Trento, Italy, 2005.

[19] http://clarkparsia.com/pellet/, logical reasoner Pellet (accessed on march 1st 2013)

[20] http://owl.man.ac.uk/factplusplus/, logical reasoner Fact++ (accessed on march 1st 2013)

[21] http://hermit-reasoner.com/, logical reasoner Hermit (accessed on march 1st 2013)

[22] http://www.sts.tu-harburg.de/~r.f.moeller/racer/, logical reasoner Racer (accessed on march 1st 2013)

[23] A. Lozano-Tello and A. Gomez-Perez, "Ontometric: A method to choose the appropriate ontology", *Journal of Database Management*, IGI-Global Publisher, Vol. 15, N° 2, 2004, pp. 1–18.

[24] http://herontology.esi.dz/content/questionnaire-hero-ontology, HERO ontology online questionnaire (accessed on march 1st 2013)

[25] L. Zemmouchi-Ghomari, A. R. Ghomari, "A New Approach for Human Assessment of Ontologies", *in Proc. of the third international conference of information systems and technologies*, Tangier, Morocco, 2013.

[26] L. Zemmouchi-Ghomari, A. R. Ghomari, "Translating Natural Language Competency Questions into SPARQL Queries: a Case Study*", in Proc. of The First International Conference on Building and Exploring Web Based Environments*, Seville, Spain, 2013, pp. 81-86.

[27] A. Ben Abacha and P. Zweigenbaum, "Medical Question Answering: Translating Medical Questions into SPARQL Queries", *in Proc. of the 2nd ACM SIGHIT International Health Informatics Symposium*, Miami, USA, 2012, pp. 41-50.

ISBN: 978-988-19252-9-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)