An Approach to Creating Domain Ontologies for Higher Education in Economics

Josip Mesarić
J.J. Strossmayer University, Faculty of Economics
mesaric@efos.hr
Branimir Dukić
J.J. Strossmayer University, Faculty of Economics
bdukic@efos.hr

Abstract. *Owing to development of ICT, knowledge representation and possibilities of its usage, sharing, storing and reusing, increasingly rely on ontologies as formal representation of (knowledge) conceptualization in different domain. Economics being the widest area of human personal and organizational activity as well as countless (structured and unstructured) data has no unified upper ontology and taxonomy. Economic curricula on faculties of economic in the Republic of Croatia reflect the problem of taxonomic shortcomings in economics as academic discipline. Numerous courses in economic curricula show the trends to specialization of economic knowledge but that also point out on decreasing capabilities of complex problem solving.*

The main idea in this paper is to research and define appropriate approach to ontology development useful for:
- representing the institutionl economic curricula which can serve to potential students in making the choice of their desirable courses,
- representing academic discipline,
- documenting the data,
-creating meta data about learning and management systems .
The approch will will be based on multiphase process of ontology creation and multilevel approach to common vocabulary creation and it will be tested on faculties of economics at universities in Croatia.

Keywords. Economic ontology, higher education.

1. Introduction

Economics being the widest area of human personal and organizational activity (personal economics, economics of family life, business systems, national economies, world economy), scientific area with numerous theories, models, practical solutions and interdependencies among them, as well as countless (structured and unstructured) data, and as such it represents the widest area of modern technologies application for creating, sharing and achieving interoperability of data, information and knowledge.

In the higher education system in Croatia, economics and closely related disciplines can be studied on all Croatian universities as well as on non university higher schools of economics and management. Economic curriculla at faculties of economics and management in Croatian universities, cover over 600 different courses included in undergraduate and graduate levels. 15 of that courses represent so called core courses, 6 represent basic (fundamentals) and rest are result of specialization in different eonomic areas. There is obvious trend to specialization in narrower areas which gives students opportunities to model their (optimal ?) curriculum on the one side, but also results in knowledge fragmentation that could decrease students abilities for knowledge integration needed in complex problem solving.

List of courses shows that there are no unified taxonomic categories (course is method, course is a set of method, course is broader area, course is a process...) and managerial paradigms (management of everything). Besides, different conceptual paradigm and different thesaurus could be found for the same course title (which could be the matter of choice, concern and interpretation the things in domain but also could indicate lack of criteria conciliation in narrower domain curriculum.) Knowledge sources for the same courses are heterogeneous and contents of different goal oriented courses are overlapped, or divergent for the same goal oriented courses.

University curricula are coarse knowledge representation by themselves and they represent the basis for finer and detailed knowledge semantics and taxonomy, relations among the terms, methods and concepts. Knowledge representation, which serves for new knowledge creation useful, shareable and interoperable for humans and computers, evolved from simple stand alone applications of mathematical logic (variants of first order logic), frame based systems, heuristic systems, expert systems, neural networks and hybrid applications to web based systems, where the knowledge and meta knowledge are structured in ontologies and intelligent agents.

Success of some method used in knowledge representation depends on how the knowledge is represented according to user's requirements, needs and possibilities for reusing on the one side, and effectiveness of technology used for searching, manipulating, sharing, representing and reusing knowledge on the other side.

Due to ICT development, (www, standardized languages for data and meta data processing, knowledge and meta knowledge representation, reusability and interoperability among the different computer systems) ontological knowledge representation becomes superior and more and more eligible in knowledge representation. „The general consensus is that ontologies are able to improve communication, sharing and reuse" [20].

Economic ontologies range from structural economics and financial ontologies to ontology of economic sectors (agriculture, [1] energetic, different industrial sector), financial institutions, (banks, stock exchanges), business information systems, to small limited ontologies for learning and tutoring.

The facts listed above were the motiv for researching and creating taxonomies – common vocabulary (thesaurus) for creation different goal oriented ontologies of economy in higher education. Studing the ontologies in higher education Milam, [11] point out that ontologies for higher education are created for:

- representing the institution,
- representing academic discipline
- documenting the data
- dreating meta data about learning and management systems
- describing the nature of higher education enterprise
- creating online resources for training materials and teaching.

Ontologies which we plan to build cover the purposes mentioned above with main initial focus on elements of vocabulary and their taxonomy.

2. What is ontology?

Comprehension of reality, causal relationships between entities of cognition, and the way the reality is described have changed over time being influence by philosophical ideas, but also by technology used to describe, change, connect and/or permanently save representations of knowledge. Thus definitions and concepts in different contexts, as well as interpretations of identical concepts in different interest professional societies (philosophers and the IT community) are taken in different ways with inevitable criticism of the others' views, but also of the efforts to find the fundamental starting point and common ground on which to build new insights. An interesting clarification of the philosophical term ontology is provided by Guarino and Giareta [7]. They summarized several common definitions of ontology and they tried to elaborate further the main consideration that ontology is a specification of a conceptualization. According to [7] ontology may be viewed as:

- a philosophical discipline
- an informal conceptual system
- a formal semantic account
- a specification of a "conceptualization"
- as a representation of a conceptual system via a logical theory characterized by specific formal properties characterized only by its specific purposes
- as the vocabulary used by a logical theory
- as a (meta-level) specification of a logical theory.

Here we shall accept Poli's view [16] of ontology which seeks to close the gap between the philosophical and AI idea of ontology: "Ontology is the theory of objects. And it is so of every type of object, concrete and abstract, existent and non-existent, real and ideal, independent and dependent. Whatever objects we are or might be dealing with, ontology is their theory. *Object* is used in this sense as synonymous with the traditional term *being*." Poli introduce three key hypotheses for right comprehension of ontology:

- an ontology is not a catalogue of the world, taxonomy, a terminology or a list of objects, things or whatever else.
- an ontology is not reducible to pure cognitive analysis
- there is nothing to prevent the existence of several ontologies, in the plural.

Broadly accepted definition of ontology in AI community is the one proposed by Gruber [4] who pointed out that an "ontology is explicit specification of conceptualization" with respect of notion of Guarino [6] who seeks out to have in mind what is and how something can be described as explicit by the language of formal logic and what the conceptualization is in sense of what was intended (the scope and with the inner purpose) in making relations among the things and concerning the constraints for that relations. Pragmatically, we accept that ontology is „formal explicit description of concepts in domain discourse (classes or concepts), properties of each concept describing various features and attributes of the concept (roles or properties) and restrictions." [13]

3. An approach to ontology development

There is no unified approach to ontology development (creation). Ushold and Gruninger [20] found that "many uses of ontologies have emerged with important differences in sophistication and objectives. They distinguish among three broad categories of ontology usage:
- as knowledge dictionary (explicitly records the meaning of the domain concepts, the relationships between concepts, and the constraints that apply to concepts.),
- as a support for conceptual design (ontologies offer important guidance for construction of application models in a specific domain.) and
- in operational use (the concepts, relationships between concepts and constraints are explicitly recorded, and these then become part of the applications themselves)

These uses are neither exhaustive nor exclusive. In the literature, as well as in practical development of ontology different approaches were proposed. Geerts, [3] point out that approaches in ontology creation could be seen as pragmatical, theoretical or empirical. Pragmatically approach defines ontological constructs by solving problems, theoretical derives conceptualizations from existing theories and empirical is oriented to testing the ontological theories embodied in human cognition.

Methodological frames for ontology development vary from complex meta theories (TOGA, [2]) with generaly accepted strategies, axioms and paradigms to explicit specification of phases which make the life cycle of ontology development. Lenat and Gluha [9] proposed threephased metodology in which the first phase include knowledge definition and coding the part of knowledge from resources by hand. The second and the third phase are comprised of new knowledge discovery by using natural language and computer tools.
Ontology development can be seen as a process in which a set of problem-, goal- and limitation-oriented activities takes place. These activities are called life cycle phases [15]. The set of activities is understood as the stages in the ontology life cycle (specification, conceptualization, formalization and implementation. Uschold and King [20] proposed four phased process: identification the purpose of ontology, ontology construction, ontology evaluation and documentation.
Gruninger and Fox [5] proposed methodology inspired by knowledge based systems created by using first order logic: they started with creation of main scenarios in using the ontology. After that the questions about the domain are raised which serves to define the scope of domain and define the terminology. Elements of terminology are formalized by the notion of logic of first order. Pinto and Martens [15] proposed life cycle methodology divided in five phases. The set of activities is understood as the stages in the ontology life cycle (specification, conceptualization, formalization, implementation and maintenance). There are also activities that should be performed during the whole life cycle: knowledge acquisition, evaluation and documentation. In the pragmatic sense, ontology development comprises the following [13]:

- determine the domain and scope of the ontology including competency question,
- consider reusing of existing ontology,
- enumerate important terms in the ontology.

After this specification, conceptual analysis and choice of tools, concrete tasks are as follows: defining the classes and their hierarchy, class characteristics, values, determine cardinality, create instances and check the ontology consistency.

Tool at our disposal for building ontologies (Protégé, [17]) can have a shortened or a more detailed development process which normally corresponds with the activities listed above in life cycle phases and the phases of pragmatic realization of ontologies on some of the tools.

Our approach to ontology creation is based on ideas of SUMO [18] and SENSUS ontology development and multilevel approach to element definition of taxonomy (common vocabulary). At the beginning we shall identify the terms (entities) relevant to our domain. After that terms are mapped to upper taxonomy (SUMO) and can be followed to root terms. If term does not exist in upper ontology it is added to upper domain and previous step is repeated. At the end for the nodes through which most paths are passing, the new sub trees are defined.

SUMO ontology is appropriate upper ontology because some middle level ontologies (economics, finance and accounting) are already created based on SUMO upper ontology. According to SUMO, entities are defined as objects, processes, abstracts and quantities and this conception enable creation most of terms used in economics.

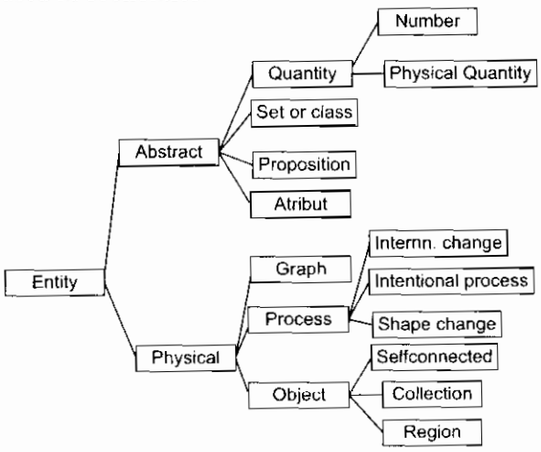

Figure 1. Part of SUMO classes

3.1. Creation of common vocabulary

Based on what (theoretical premises, actual paradigms, axioms, models, methods, conceptual criteria and constraints) thesaurus for some knowledge domain should be developed? Every field of science has its breadth and depth of cognition, usability and relation with other areas. Thus the knowledge creation in dictionary (thesaurus) could be viewed as four level activity process, [12] as it is shown on figure 2.

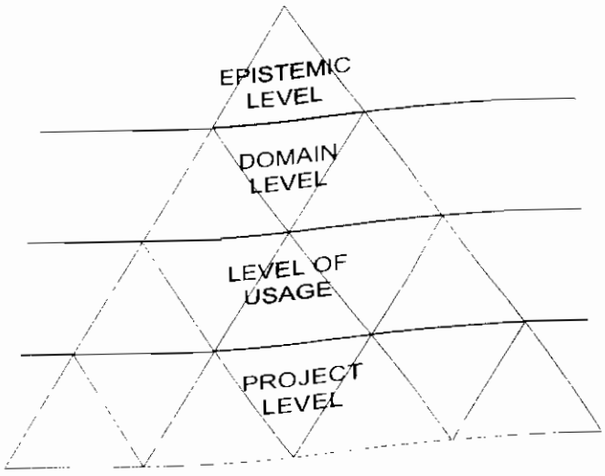

Figure 2. Four level approach to dictionary creation

On the epistemic level the set of most generalized terms and relations among them are defined. The question what economics is and which terms deseribe it should be answered. Domain level describes the information structures specified for the domain, documented by theories, models and axioms. On this level domain curriculum could be developed. Level of usage enhances knowledge from domain level in order to describe specific segment in domain. On this level content of specific courses could be defined.

The level of project is enhancement of previous levels to specified implementation. On this level learning objects repository for tutoring systems could be defined.

Methodology for vocabulary creation proposed by Kim [9] has two phases:

- phase of empirical investigations (Identification of taxonomy on highest level, identification of key terms in domain, taxonomy creation for terms (entities) in vocabulary, testing of knowledge description)
- phase of computer implementation (creation of thesaurus, and final assembling of vocabulary and taxonomy)

(For each particular phase authors suggest methods and evaluation procedure.

During the first phase we collect five curricula from faculty of economics in Croatia and put them in excel table. Using Excel analytical tool

we derive all particular courses and common and repetitive modules and courses which will serve for coarse taxonomy of economics on universities in Croatia. Detailed analysis of course contents is based on course description. In this very first phase experimental comparative analysis of terms (including course title and terms from course description) with terms and categories of economics used in JEL classification of economic literature. We found out that classification categories used JEL could not be directly mapped to classification categories used in course description. Analysis showed that course could vary from broad economics fields through very narrow economic field, method, and even a model. Some key economic terms founded in JEL classification are neither the particular course name nor the part of course content. This raise some questions about creation of knowledge content represented in course and curriculum: what are the paradigms, axioms and goals followed in course creation?

3.2. Economic thesaurus and taxonomies

As we mentioned in introduction, our knowledge about ontology usage will be grouped around next conceptual terms (entities):
- representation of institutional curricula,
- representation of academic discipline,
- documenting the data and metadata,
- creating meta data about learning and management systems,

Creating online resources for training materials and teaching.

There is no unified taxonomy of economics. Simple taxonomy of economics proposed for curriculum development on faculties of economics in Republic of Croatia is to coarse and incomplete and does not support mapping into, for instance, widely accepted classification of economic literature (JEL). At the same JEL classification and contents of particular courses do not aloud bidirectional mapping too. English economic thesaurus need to be translated for use in economic ontology developed for Croatian language. Some narrow specific vocabularies in economics (accounting, marketing) are of limited usefulness.

4. Conclusion and recommendation

Ontology development is a complex task which needs eligible approach. The nature of ontology and goals of ontology creation define the intensity of phases followed in chosen approach. Usually this is combination of top-down and bottom-up steps based on adoptive paradigms, axioms, constraints and goals. Any particular domain ontology may be considered as a part of some higher (upper) ontology and knowledge represented in that ontology. The purpose and goals of economic ontologies which we plan to create should represent the structure of economic knowledge in economic curricula at universities in Croatia. They will facilitate desired personal curriculum creation, enable precise insight in academic disciplines, document the data about economic knowledge, and create web based learning systems. The approach to ontology creation for such goals is divided in few levels (epistemic, domain, level of usage and project level) and phases. Initial phase in which our project is, requires intensive analytical investigations for vocabulary elements definition. Classification paradigm of upper ontology is adoptive from SUMO upper ontology. Taxonomic categorization is searched in available economic taxonomies and thesaurus, which make them limited for our purpose.

Problem that arises with usage of ontology come from the fact that its usage is highly depended of natural language of users and thus the conceptualization and vocabulary should be available in the natural language. (Automatic language translators may help in solving this problem, but there are developed mainly for big world languages)

Economic curricula on faculties of economics and management in Republic of Croatia are, like elsewhere, results of anticipated knowledge of graduated students (needed in environment where they usually find their jobs) and organizational advantages or constraints of particular faculty. We think that ontological representation of economic knowledge will help in creation better body of knowledge and that it will represent good basis for documenting economic knowledge and creating learning objects for web based learning systems.

5. References

[1] Agrovoc. http://www.fao.org/aims/ag_intro.htm

[2] Gadomski, A.M., Systemic Approach for the SOPHOCLES Global Specification, Technical Report, (Extended & Revisioned Version Nov. 2002), Rome 2002

[3] Geerts, G., The Ontological Foundation of REA Enterprise Information Systems, The University of Delaware William E. McCarthy, Michigan State University, 2005

[4] Gruber, T., A translation approaches to portable ontology specifications. In: Knowledge Acquisition 5, (1993) pp. 199-220)

[5] Gruninger, M., Fox, M.S., Methodology for the design and evaluation of ontologies, Workshop on Basic Ontological Issues in Knowledge Sharing; Montreal; 1995.

[6] Guarino, N., Formal ontology and information systems, In: N. Guarino (ed.), Formal Ontology in Information Systems. Proceedings of the First International Conference, Trento, Italy, 6-8 June 1998. IOS Press, 1998 p. 4

[7] Guarino N., Giareta, P., Ontologies and knowledge bases: Towards a terminological clarification, In Mars, N., Editor, IOS Press, p. 25-32.

[8] Journal of economie literature Clasification System. http://www.aeaweb.org/journal/jel_ class_system.htm

[9] Kim, Ahmed, S., S., Wallace, K. A methodology for creating ontologies for engineering design, Proceedings of DTM 2005; ASME 2005 Design, Theory and Methodology Conference; Long Beach, California, USA; 2005

[10] Lenat, D.B., Guha, R.V., Building Large Knowledge-Base Systems: Representation and Interfence in the Cyc Project, Addison-Wesley; Boston; 1990.

[11] Milam, J., Ph.D. Ontologies in Higher Education, HigherEd.org, Inc.

[12] Mehkilef, M., Bourey, J.P., Bigand, M., An UML modelling of an architecture for knowledge documentation"; 14th International Conference on Engineering Design, ICED 03, Stockholm, Sweden, 2003.

[13] Noy, F.N., McGuinnes, D.L., Ontology Development 101: A Guide to Creating Your First Ontology, http://protege.stanford.edu/publication/ontology101-noy-mcguiness.html, October 6th, 2005. http://dip.semanticweb.org/documents/D10-7-Stock-Market-Ontology.pdf

[14] Partridge C., A new foundation for accounting: Steps towards the development of referenee ontology for accounting, Technical Report 23/02, LADSEB-CNR, Padova, Italy, December 2002. http://www.boroprogram.org/bp_pipex/ladsebreports/ladscb_t_r_23-02.pdf

[15] Pinto, H.S., Martins, P. Ontologies, How can they be Built. Knowledge and Information Systems (2004) 6: 441-464): (13)

[16] Poli, R., Ontology for knowledge organization, in Rebecca Green (ed.), Knowledge organization and change, Index, Frankfurt, 1996, pp. 313-319.

[17] Protégé http://protege.stanford.edu/

[18] SUMO (Suggested Upper Merged Ontology - www.ontologyportal.org)

[19] Štorga, M., Model rječnika za računalnu razmjenu informacija u distribuiranom razvoju proizvoda, Disertacija, Fakultet strojarstva i brodogradnje, Zagreb, 2005.

[20] Uschold, M., King, M., Towards a Methodology for Building Ontologies, IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing; Montreal, 1995.