



HAL
open science

A Survey on Educational Ontologies and Their Development Cycle

Abdelghani Karkar, Jihad Mohamad Al Ja'am, Sebti Foufou

► **To cite this version:**

Abdelghani Karkar, Jihad Mohamad Al Ja'am, Sebti Foufou. A Survey on Educational Ontologies and Their Development Cycle. 12th IFIP International Conference on Product Lifecycle Management (PLM), Oct 2015, Doha, Qatar. pp.649-658, 10.1007/978-3-319-33111-9_59 . hal-01377493

HAL Id: hal-01377493

<https://inria.hal.science/hal-01377493>

Submitted on 7 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A Survey on Educational Ontologies and their Development Cycle

AbdelGhani Karkar, Jihad Mohamad Al Ja'am, Sebti Foufou

Department of Computer Science and Engineering, Qatar University, Doha, Qatar

a.karkar@qu.edu.qa

jaam@qu.edu.qa

sfoufou@qu.edu.qa

Abstract— Nowadays, the grid of Internet has demonstrated to be plentiful and tremendous data source of information, where diverse domains can be reached and mined. Semantic web is part of the Internet grid where knowledge is provided and has a predefined sense. People can use the big quantity of accessible information for entertainment, exploring knowledge, and learning. In this paper, we provide a survey of educational ontologies, their development life cycle, and the tools used for their implementation. The classification outcomes are beneficial not only for practicality purposes but also for building educational ontologies and their reusability, since it provides a framework for selecting the suitable methodology to be used in specific context, depending also on the requirements of the application itself.

Keywords—*Semantic Web, Ontology; Methodology; Development life cycle; Engineering Education.*

1. Introduction

Ontology has gained the attention of people in both industrial and academic fields. The term ontology has been defined in various ways [1]. Initially, it originates from philosophy, where it refers to the basic characteristics of semantic existence in the word. Different domain ontologies are applied to provide a formal concept for knowledge structure, such as, management, nutriments, animals and medicines. Domain knowledge can provide meaning based and requirement statements in a particular domain [2]. Domain concepts in a domain and their semantic relations can lead to discover new facts.

Educational content is created by an instructor in order to provide his students with useful learning materials. It is usually divided into different complementary sections where each section covers a particular area that can be related to the previous one. Any educational content can have general information like domain, title, description, chapter, etc. Educational ontologies are developed to improve the learning concepts in a particular area. They can be used in different domains to facilitate the access to information and generate new knowledge through reasoning. Different educational ontologies for different domains have been created and published [3, 4, 5]. However, the reusability aspect of these ontologies in a complete educational system is still missing. This study is a survey of the existing educational ontologies. It provides an initial step to reuse and maintain different domain models for learning and training purposes.

This paper is organized as follows: section 2 presents the literature background of educational and learning ontologies. Section 3, surveys the methodologies used to construct ontologies. Section 4 reviews the used tools for developing the ontologies and some technical details of about them. Section 5 elucidates the significance of reusing and maintaining educational ontologies for educational and training purposes. Finally, section 6 concludes the paper.

2. Background

An ontology can be constructed to define specific type of information that will be used by communities. Researchers have been constructing conceptual ground for developing efficient solutions to daily problems based on shared and reusable knowledge components. The construction of ontology denotes clear concepts, relations and instances which define the characteristics of conceptualization. In another words, an ontology is the backbone to define concepts specifications [5].

Sawsaa and Lu [5] have developed an ontology of information science (OIS) that can be used for information science education. The formulation of concepts depends on recognizing the information science (IS) concepts and coinciding them into a hierarchical structure based on their classifications. The study presents formal semantic elucidation for IS meta-data. The formal semantic elucidation involves the strength of reading information and processing them by Artificial Intelligence (AI) systems [6]. Wang et al. [7] proposed an image processing system to verify automatically road signs depending on pre-defined road sign regulations (RSRs) and simulate the generation of new road signs process when new roads are constructed. The system is based on semantic-enabled road sign management (SeRSM), where information are integrated semantically using the Large Knowledge Collider (LarKC) platform [8]. LarKC is a huge distributed data used for enormous knowledge classification, reasoning, representation querying, and it has been excessively applied in the fields of urban computing, biomedical, and sciences. The main goal of LarKC projects is to develop a reasoning platform using huge amounts of diverse information. The platform contains a pluggable interface to achieve heuristics and techniques from different areas such as machine learning, databases and semantic web.

Challco et al. [9] have developed an ontology for gamifying collaborative learning (CL) scenarios (OntoGaCLeS). The term gamification exceeds the meaning of playing games only; it conducts the design of game elements to motivate individual requirement depending on personality characteristics. "Human desires" is take into consideration the needs of individual's motivation through the integration of diverse game techniques to provide sufficient environment [10]. Rezgui et al. [11] have analyzed various competency modeling approaches and proposed an ontology to formally define the characteristics of competency-related and learning resources. The ontology intends to provide visual presentation for competency information management aspects. Thus, it serves lifelong competency development within learning networks. Chung and Kim [3] have defined an ontological structure for syllabus with their semantic relationships. They have integrated the classification based on ACM/IEEE computing curriculum, learning activity, formalization of learning goals, and they performed learning evaluation using Bloom's taxonomy to ameliorate syllabus usability [12]. They have proposed an efficient method for improving students' learning effect using the constructed subject ontology, knowledge sharing, and visual presentation.

Quinn et al. [13] have proposed a personalized approach to improve patient education. They presented three models, health conditions, user characteristics, and educational content. The three models can be used to create educational content that supports patient understanding and concentrates on specific patient's health concerns. The educational content is presented in a web-based application that provides the patients with details about their diagnosis and how they can improve their health. Yoon et al. [14] suggested a model to provide protective technique to access multimedia content in the mobile cloud. The model manages different content forms, and builds the multimedia ontology to improve reliability during the retrieval process. The ontology can solve conflict problems caused by mobile device meta-information and it manages the content in different forms. Meta information includes name, store location, production day, etc. Conflict problems, such

as meta-information collision and semantic collision, can be resolved through mapping between content of instances. Kim et al. [15] constructed a spine ontology that represents disease information and spine anatomical structure that is compatible with KISTI Simulation model. The ontology concentrates on the persistent diseases that concern Koreans. It contains methods of treatment, classification, and cause related with spine. The spinal ontology can be used for education purpose by medical students, biomedical engineers, and physicians.

3. Review of Used Methodologies

Sawsaa and Lu [5] developed an ontology of information science (OIS) according to the IEEE principle for the process of software development life cycle. The formulation of concepts depends on recognizing the Information Science (IS) notions, and coinciding them into a hierarchical structure view based on their classifications. Wang et al. [7] investigated a data integration solution to provide a ground for intelligence road sign management system instituted over LarKC platform. The solution carries the processing of huge amount of lined geo-data, data modeling, road sign data, reasoning and scalable querying. The data is associated using mediation ontology. SeRSM Geographic data is declared originally by three elements (i.e., node, way and relation). Way elements are accentuated as a range of nodes, and a road on the map is formed from ways. The node is modeled in three types: 1) generic nodes that can distinguish either a point on way or a junction between multiple roads. 2) Road sign (RS) nodes which identify the location of a road sign. And 3) Zhenjiang POIs (ZJPOI) that identify Point of Interests (POI) from (Open Street Map¹) Open Street Map (OSM) and Baidu map. SeRSM system exploits large scale of semantic data. Semantic data was transformed to a uniform representation using XSLT and other technologies to facilitate the handling process and reasoning between them.

Challco et al. [9] examined gamifying CL scenarios to make the learning experience more meaningful and enjoyable. Learners will be highly motivated while performing their proposed tasks. They will enjoy using different game techniques like point system, social connection, leaderboards, etc. Challco et al. introduced different terms to support the games mechanics' personalization in CL scenarios: 1) *I-mot goal* is the increase of individual motivation to achieve the goal at specific stage. 2) $Y \leq I\text{-mot}$ goal is the increase of motivational strategy to attain the goal. 3) You-player role and I-player role are the active player role. 4) *I-gameplay* is the behavior of strategy employed at runtime of the person in focus (I). The ontology contains the player's roles concepts, where two prerequisites are denoted as desired and necessary conditions. The desired conditions define learner satisfaction to attain full interest of a player that are defined in the playing style, while the necessary conditions are important for a learner to play the game which also refer to psychological needs and motivational stages. They have defined in their ontology only one restriction for socializer which can work only with other socializer. Rezgui et al. [11] defined several concepts in their competency ontology to formalize the terms *competency profile* and *competency*. Competency profiles in real applications illuminate most evident aspects of competency modeling. They provide support to various tasks, like creating job/project competency profile for hiring candidates, creating personal competency profiles to set the highlight on certain capabilities, creating prerequisite competency profiles to show basic requirements for specific programs, and so. They defined their competency model according to the definition of Paquette [16]. Paquette [16] defined the competency as a statement having different relationships to all achieved skills and knowledge at specific performance level. For instance, "apply all processes to construct a use case diagram"

¹ Open Street Map: <http://www.openstreetmap.org/>

is a competency statement in which the term apply indicates the usage of previous knowledge for construction.

Chung and Kim [3] developed their syllabus ontology to show learning concepts and relations between them. The syllabus is written by the instructor for his/her students to present how learning materials will be provided. The syllabus does not provide semantic concepts and essential relations among different information (e.g., title, summary, grading criteria, etc.). It can be used only to organize information in particular order. Syllabus ontology demonstrates an efficient method to improve students' learning effect where it conceptualizes teaching contents. It is designed to cover many sorts of e-learning environment knowledge from curriculum. The curriculum can be structured as a set of courses' description and syllabus. The curriculum ontology defines the concepts of curriculum-related knowledge (e.g., Course, Program, Level, etc.) with their semantic relation between each course and its syllabus ontology. The syllabus ontology defines the structure of syllabuses where it defines the core concept, called *Syllabus*. The *Syllabus* has different properties (e.g., titleOfCourse, gradingPolicy, description, etc.), which describe the characteristics of the course. The subject ontology is composed from learner-based ontologies and teacher-based ontologies. Learner-based ontologies provide knowledge of learning materials created by students. Teacher-based ontology provides knowledge of learning materials that will be studied in class. When a teacher displays learning subjects, students conduct displayed subjects to extract required concepts and knowledge.

During the ontology design phase proposed by Quinn et al. [13], different sources were surveyed to identify which characteristics should be checked for the person and health conditions items. These sources comprised Diabetes UK, the UK National Health Service, The American Diabetes Association, and academic publications. The proposed personalized framework is composed of two layers; 1) Modelling & Management ontology layer: it has three main entities; the patient, medical conditions, and education content. Information is modeled in an ontology to share understanding of domain interest. Ontology data ensures that semantic information is not ambiguous and it is appropriate for performing reasoning technologies. Information in this layer includes patients' characteristics collected from their medical records and activity specifications inspected by a physician; and 2) Personalization & Contextualization layer: it includes rules and reasoning engine for personalization. The reasoning component can use the ontology and rules to generate inferences and their semantic relations. For instance, a rule can be defined to identify the relation between a set of symptoms denoted by a patient including specific health involvement.

Kim et al. [15] are professional researchers from different domains, which include computer professional, medical informatics, and clinical experts. They collaborated together to design and construct the spinal ontology. They passed through five phases: 1) review existing ontologies in the domain where required information are gathered after the classifying of diseases and analyzing the ontology of Rat anatomy; 2) selection of the diverse spine diseases including the research subject, where the selection is done according to three criteria: (a) most spinal diseases occurring among Koreans, (b) the disease must happen in particular area rather than the whole spine, and (c) the disease should have a computerized representation; 3) development and review of ontology information, 4) construction of the ontology relying on the feedback of clinical experts, and 5) review the ontology by specialists. The represented results allow users to access 3D images where they can rotate, move, and change the zoom level of the image.

4. Review the Used Tools and Technical Details

The OIS ontology proposed by Sawsaa and Lu [5] has been developed using Protégé where fourteen facets have been defined (e.g., Actors, Methods, Practice, etc.). The evaluation of the ontology showed adequate results where it has been assessed by domain's experts to inspect different criteria, which include avoiding concepts duplication, inconsistent relationships, clarity and excessiveness. The dataset proposed by Wang et al. [7] contains 3 million triples representing streets of Zhenjiang which are collected by OSM, 0.8 million triples represent road signs which are collected from Baidu map, and 0.1 million triples represent road signs which are collected by the team members. 32 Chinese road sign regulations related to naming and positioning are converted to SPARQL queries. The presentation of data is a web based interface that displays extracted information directly on the map. If the settled threshold has too much information in one road sign, the system will show warning information for the users.

Challco et al. [9] have developed a semantic web utility that serves in the design of Collaborative Learning (CL) based on game design, instructional design, and learning theories principles. They have developed their ontology using Hozo Ontology editor. They have defined eight gamified CL scenarios using their developed ontology. Each scenario is associated with one player's role to satisfy the psychological needs, motivation stages and playing style. The competency ontology proposed by Rezgui et al. [11] is constructed using Protégé. OntoViz plugin has been used to visualize the ontology as a graph. It reuses some concepts defined semantic web vocabularies for taxonomy representation (e.g., SKOS Core ontology [17] and annotation of content (e.g., Dublin Core) [18]). The concept *competency* has different properties (e.g., *isComposedOf*, *requires*, *hasProficiencyLevel*, etc.) and it is related with different concepts: 1) *Skill* denotes the selected skill from skills of learning-domain taxonomy; 2) *Knowledge* specifies the correspondent view of a subject matter; 3) the context which is basically identified from the domain ontology; and 4) *ProficiencyLevel* which identifies individual proficiency record.

The syllabus ontology proposed by Chung and Kim [3] provides diverse semantic relationships through its main class named *Syllabus*. Learning elements will be displayed as sequentially connected nodes for students inside a learning graph. Each node contains the lectures, learning activities, goals and assessments. The Syllabus also can be presented in the learning graph to provide more information about the course (e.g., description, lectures, learning materials, etc.) and to show the order between multiple syllabuses. The teacher can define 1) a learning goal in the format of $\langle \text{goal}_p, C_i, A_j, S_k \rangle$, like $\langle \text{"Understanding functions in JAVA"}, C_3, A_2, S_3 \rangle$, where p -th denotes the learning goal, i^{th} for understanding complexity level, j^{th} for the attitude level, and k^{th} for the skills level respectively; 2) a learning activity in the format of $\langle \text{LA}_p, C_i, A_j, S_k \rangle$, where LA_p denotes the learning activity (e.g., (R)eading, (P)resentation, pr(A)ctice, etc.); and 3) a learning assessment in the format of $\langle \text{QE}_p, \text{sentence}_g \rangle$ where QE_p is the learning activities (e.g., quiz, assignment, exam, etc.), and sentence_g denotes the description for the learning assessment.

The design of the ontology proposed by Quinn et al. [13] is developed with Protégé. Domain concepts are defined, declared and arranged in a hierarchical structure of superclass-subclass relationships. The data in the user model are separated into three classes, *HealthProfile*, *PersonalProfile*, and *EducationProfile*. *HealthProfile* contains all aspects of activity objectives and patient's health. *PersonalProfile* contains personal information of the patient, such as, gender, age, language, and ethnicity. *EducationProfile* stores information used to identify the readability level of the patient. The educational content includes graphical and text components, such as, diagrams and illustrations. The educational content consists of three classes, *Educational-Content*, *Text*, and *Image*. Class *Text* shows textual information that is related to the obesity and diabetes. Its properties include *hasSubject*, *hasLanguage*, and *hasReadabilityLevel*. These properties can be associated with the *PersonalProfile* and *EducationProfile* to guarantee the language preferences and readability needs of the patient. Class *Image* properties include *hasGender*, *hasEthnicity*, and *hasAge*. They provide relations to infer associations with comparable characteristics recorded about the patient.

The multimedia content ontology proposed by Yoon et al. [14] is developed with Protégé. It is based on the OWL language. Extracted information can be classified by *instance*, *category*, *position*, and *value partition* sectors. The *instance* stores the semantic relation between contents. The *category* applies Multimedia Description Schema (MDS) of MPEG-7. The *position* access *instance* information and stores its location. It facilitates the access to particular content by a user. Eventually, *value partition* stores the assortment of the values obtained by *instance*. Collisions might occur between the content stored on mobile devices and meta-information. They can be resolved by the mapping of instances between meta-information. For instance, meta-information for an audio file on device A cloud includes artist, name, media, size, long date; whilst meta-information of the same audio file on device B cloud includes author, title, media size, short date; both meta-information refer to the same information where *synonym*, *format convert*, and *mapping* can be used to solve this collision.

The spinal ontology proposed by Kim et al. [15] was developed using Protégé and Altova SemanticWorks. The ontology includes the connections to sample disease images and images from KISTI. It contains detailed information about a disease that can occur in particular region. Anatomical information were classified into two categories, anatomical location (e.g., spine:isPartOf) and anatomic properties (e.g., spine:isPartOf, rdfs:label, rdfs:subClassOf, etc.). Disease related information were classified into five categories: anatomical location (e.g., spine:hasSite, etc.), property of the disease (e.g., rdfs:description, spine:diagnosis, spine:hasCause, etc.), method of treatment (e.g., spine:hasNonSurgicalTreatment, spine:hasSurgicalTreatment, etc.), symptom/sign (e.g., spine:hasSymptom, spine:hasSign, etc.), and image (e.g., spine:hasImageBeforeTx, spine:hasImageAfterTx).

5. Ontology Usage in Education and Professional Training

Large composite applications can be costly, difficult and controversial especially when diverse forms describe the same concept. However, extracting information from previous existing ontologies remains viable since they have common high-level structure. One of the intelligent tasks that can be performed with ontologies is the semantic relation extraction (SRE), which includes finding new semantic relations (Inference²) to optimize the taxonomic of reasoning. These tasks cannot be done in database models since they rely on static predefined structure. SER can identify entity acquaintance, relationship with other entities, and its types. It can be used also for knowledge extraction (i.e., extracting entity names from text content, etc.) that is required during the phase of text processing. Educational ontologies can be used for: 1) education to teach students (e.g., processing a child story to show a lion is carnivorous and eats rabbit; if we know that a tiger is carnivorous, we can infer that the tiger can also eat rabbit), and 2) professional trainings to provide profound information in particular area, for instance, to provide a representation about electro-mechanical assemblies require the extraction of required information from an ontology that illuminates the usage of each part and the relation between each other (e.g., a Ring Gear can hold a Sun Gear, an Output Shaft requires both Planet Gears and Planet Gear Pins to be connected [4], etc.).

6. Conclusion

Based on our survey, it can be seen that there is a growing interest in the field of educational ontologies. Our paper provided a survey of some available educational ontologies from different domains of application

² Inference: <http://www.w3.org/standards/semanticweb/inference>

and discussed their development methodologies as well as used tools. This work can be considered as a tentative effort to manage and use them in a unified aspect in order to provide a complete semantic knowledge representation for the diverse domains. Therefore, it is recommended to design a system that can manage the reusability of existing ontologies to extract relevant information and provide learners with suitable knowledge.

References

1. Sowa, J.F.: *Ontology, metadata, and semiotics*. In: *Conceptual structures: Logical, linguistic, and computational issues*, pp. 55-81. Springer (2000)
2. Kaiya, H., Saeki, M.: *Using domain ontology as domain knowledge for requirements elicitation*. In: *Requirements Engineering, 14th IEEE International Conference*. IEEE (2006)
3. Chung, H.-S., Kim, J.-M.: *Semantic Model of Syllabus and Learning Ontology for Intelligent Learning System*. In: *Computational Collective Intelligence. Technologies and Applications*, pp. 175-183. Springer (2014)
4. Rachuri, S., Han, Y.-H., Foufou, S., Feng, S.C., Roy, U., Wang, F., Sriram, R.D., Lyons, K.W.: *A model for capturing product assembly information*. Vol. 6 (1), pp. 11-21. *Journal of Computing and Information Science in Engineering* (2006)
5. Sawsaa, A., Lu, J.: *Building an Advance Domain Ontology Model of Information Science (OIS)*. Vol. 4 (2), pp. 258-266. *International Journal of Digital Information and Wireless Communications (IJDIWC)* (2014)
6. Marquardt, W., Morbach, J., Wiesner, A., Yang, A., *OntoCAPE: A re-usable ontology for chemical process engineering*. Springer Science & Business Media (2009)
7. Wang, D., Huang, Z., Liu, Q., Zhang, X., Xu, D., Wang, Z., Li, N., Zhang, J., Zhang, D.: *Using Semantic Technology for Consistency Checking of Road Signs*. In: *Web Information Systems Engineering–WISE 2013 Workshops*. Springer (2014)
8. Fensel, D., van Harmelen, F., Andersson, B., Brennan, P., Cunningham, H., Della Valle, E., Fischer, F., Huang, Z., Kiryakov, A., Lee, T.-I.: *Towards LarKC: a platform for web-scale reasoning*. In: *Semantic Computing, 2008 IEEE International Conference on*. IEEE (2008)
9. Chalco, G.C., Moreira, D., Mizoguchi, R., Isotani, S.: *Towards an Ontology for Gamifying Collaborative Learning Scenarios*. In: *Intelligent Tutoring Systems*. Springer (2014)
10. Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J.-J.: *Gamifying learning experiences: Practical implications and outcomes*. Vol. 63, pp. 380-392. *Computers & Education* (2013)
11. Rezgui, K., Mhiri, H., Ghédira, K.: *An Ontology-Based Approach to Competency Modeling and Management in Learning Networks*. In: *Agent and Multi-Agent Systems: Technologies and Applications*, pp. 257-266. Springer (2014)
12. Forehand, M.: *Bloom's taxonomy*. pp. 41-47. *Emerging perspectives on learning, teaching, and technology* (2010)
13. Quinn, S., Bond, R., Nugent, C.D.: *An Ontology Based Approach to the Provision of Personalized Patient Education*. In: *Ambient Assisted Living and Daily Activities*, pp. 67-74. Springer (2014)
14. Yoon, C.-P., Moon, S.-J., Hwang, C.-G.: *MCSOSA: multimedia content share using ontology and secure access agent in mobile cloud*. Vol. 71 (2), pp. 667-684. *Multimedia Tools and Applications* (2014)
15. Kim, G.-H., Kang, M.-J., Jung, C.Y., Jung, J.-Y., Jung, S.E., Kim, J.-S., Jeong, J.-S., Kim, D.-H., Yoo, K.-H., Seo, D.: *Development of Ontology for the Diseases of Spine*. In: *Advances in Computer Science and its Applications*, pp. 1171-1178. Springer (2014)
16. Paquette, G.: *An ontology and a software framework for competency modeling and management*. Vol. 10 (3), pp. 1-21. *Educational Technology & Society* (2007)
17. Bechhofer, S., Miles, A.: *SKOS simple knowledge organization system reference*. W3C recommendation, W3C (2009)
18. DC, Dublin Core Metadata Element Set, Version 1.1: Reference Description <http://dublincore.org/documents/2003/02/04/dces/> (visited on 14/5/2015)

APPENDIX

Table 1: Review the existing ontologies, domain, technologies

Number	Domain	Description	Used Technologies	Reference
1	Information Science (OIS)	Provides Hierarchical view of Information Science concepts	Protégé, FaCT++	Sawsaa et al. 2014
2	Road Sign / Transportation	Provides semantic information about Road signs	LarKC, XSLT, SPARQL	Wang et al. 2014
3	Ontology for Gamify Collaborative Learning Scenarios (OntoGaCLeS)	Provides collaborative learning scenarios based on theory's games	Hozo Ontology editor	Challco et al. 2014
4	Competency ontology	It models competency information aspects to support learning networks	Protégé, OntoViz Plugin	Rezgui et al. 2014
5	Syllabus Ontology	Covers different e-learning curriculum and learning materials	-	Chung et al. 2014
6	Ontology for Patients	Provides personalized patient education, medical conditions, and patient profile	Protégé, Web Pages	Quinn et al. 2014
7	Multimedia Content Ontology	Provides secure multimedia content in mobile cloud	Protégé	Yoon et al. 2014
8	Ontology for the Diseases of Spine	Spinal ontology can be used for education of medical students	Altova SemanticWorks, Protégé	Kim et al. 2014