

Graduation Project
Sahar Mojtahedi
June 2012
VU University Amsterdam

A Study on Ontology Design for University Programmes

Contents

INTRODUCTION	2
A Brief History of Ontology	2
Ontology Use in Information Science	2
Ontologies for University	3
PROBLEM STATEMENT	4
Purpose	4
Research questions	4
APPROACH	5
Scope	5
Concept Highlighting	6
TIMELINE	7
REFERENCE	8

INTRODUCTION

A Brief History of Ontology

The term 'Ontology' has historical roots in ancient Greece and refers to the Study of Existence. There has been a wide philosophical discussion about the term. According to Barry Smith's article [1], Ontology seeks not explanation but rather a description of reality in terms of a classification of entities that is exhaustive in the sense that it can serve as an answer to questions. In the same paper, it is discussed about how in history the term was used and evolved. In some time after Aristotle, it was used as a way to define taxonomy to describe and categorise things. Later, ontologists engaged in study of folk ontologies development and the degree to which elements of folk ontologies imitate universal features of the human cognitive system. Then we reach the stage of technology and how ontologies got used in Information Science and the Web.

In 2001, Tim B. Lee et. al. [2], the Semantic Web was introduced as an extension to the Web which would enable machines to "understand" the concepts provided in web pages. The Semantic Web aims to evolve the World Wide Web from being a medium for sharing information to a platform for processable data. To do this, the idea of Knowledge Representation was implemented on the Web documents using XML and RDF languages where XML brought structure and RDF provided the means to represent meanings through triples. But then the issues arose when information needed to be shared. This is when ontologies enter into the story to help define structures and provide a platform to align them and let computers compare and manipulate data further.

Ontology Use in Information Science

In the field of Information Science the term "ontology" is used to refer to a representation of knowledge as a set of concepts within a domain, and the relations between those concepts [3]. Different from hierarchical views of concepts such as taxonomies, ontologies can have linked or networked structures. In a hypothetically multi-way, concepts can be related to other concepts in a series of relationships [4]. Nowadays, ontologies are widely used in various fields. They are used to describe knowledge in different domains in a structured way to be shared online. Ontologies in Information systems are used to manage knowledge. Also, using ontologies helps

Web search to be conducted with higher accuracy due to the semantics behind their design. While Web search engines perform their tasks on the base of syntax and term-matching [5], the introduction of the Semantic Web and use of ontologies to represent data on the Web has brought many efforts on creating ontologies in various domains. Using ontologies, search engine can search keywords that are conceptually linked instead of just similarity of the words used [6].

According to Odell [4], main uses of ontologies can be listed as the followings:

- 1) For communication purposes
- 2) To structure and organize knowledge in reusable artifacts
- 3) To facilitate capturing semantics
- 4) To enable computational inferences
- 5) To drive software engineering

In this research ontology is used for the third and fourth uses on the list.

Ontologies for University

One of the domains for which ontologies have been developed, and still are in an on-going stage, is for University. Not surprisingly, there are many ontologies created in this field. But, due to biased design of ontologies in Information Science, none of them fully satisfies requirements of other projects. Although the ultimate goal might be to have one ontology that can contain all concepts in a domain, but in every project the concentration is more or less different from the other ones. Therefore, there are various ontologies introduced in every field including university domain. Examples of ontologies intended to describe and categorise concepts in universities are Institutional Ontology [7] and Academic Institution Internal Structure Ontology (AIISO) [8]. They include the basic structure of concepts in universities and can be used by other researchers to apply more concepts to them and make them specific for their own purpose.

Ontologies represent concepts and concepts can be defined according to multiple viewpoints. Therefore, before using an existing ontology one needs to study and understand the viewpoints according to which data and concepts are meant to be in it. Then, even multiple ontologies can be aligned and used for one specific purpose.

In this paper, the knowledge about ontologies as discussed in this section is used to study a current problem.

PROBLEM STATEMENT

Most (if not all) of prospective Master's students spend a lot of time searching for a programme to match with their interests, skills and future plans. There are online search engines that are specified for the educational programmes such as 'School Finder' [9] but they are not working based on semantics. This is while concepts to be covered in programmes are as important (if not more) as factors including world rank of university, location and length of programmes. Therefore, not knowing a specific programme and the concepts included in it would make the process of searching harder for a student.

There have been efforts to design and represent structure of universities through ontologies. But not much work is done focusing on programme contents. In this paper, the problem of representing programmes based on their objectives and concepts is studied.

Purpose

Studying Information Science at VU University Amsterdam has given me the insight to think about such problem in a new way. Some factors supported my thoughts for choosing a related topic for my graduation thesis. The factors are mainly learning about Semantic webs and ontology designs, studying about the power of implemented semantic search engines such as 'Yummly' [10] - which is an online recipe search engine and is in use by many users now -, and also observing the use of ontologies and semantics in various areas. Doing so would also help me to improve the skills I have learned in the Ontology Engineering course and also let me get use of the concepts I have learned during the whole programme. To conduct an experiment that covers what is mentioned I need to define research questions which are going to be the focus of my thesis.

Research questions

To sum up the problem and the goal of this paper, the content of university programmes have significant value for prospective students and programme coordinators spend some valuable time to design a programme. Due to lack of semantic representation of programmes, available

search engines retrieve data from university programmes syntactically. Respectively, this paper focuses on three research questions given below;

1. What material is used by prospective Master's students to review a programme? What material do universities prepare to represent a programme?
2. What are the semantics of educational programmes in universities?
3. How can the semantics of individual programmes be modelled into an ontology?

Discussion on the initial assumptions and methods to cover the goal of this paper is given in the 'Approach' section.

APPROACH

In this research, a glossary will be represented that is intended to be non-controversial. Primary emphasis is on identifying the main structure of a comprehensive ontology for university programmes. It is intended to be an accurate record of how terms are currently used to refer to main objective of programmes and how concepts are distributed in programmes and in relation to other programmes. The aim is to reflect commonly used terms to represent the objectives and content. The glossary can then be used as a complementary to an existing university. Such ontology can be applied to search engines as a source to help prospective students find their best suited university programme and let them explore concepts flowing inside and between programmes.

At the early stage, we need to define the scope of the research and to decide about how definitions are to be chosen for the purpose of this research. Then we focus on the method to highlight the concepts.

Scope

The chosen scope for this research is focused on Master's programmes. The major motive for this selection is that prospective students of Master's programme pay assiduous attention to the details, objectives and focus of the programmes. Various articles such as the article of Ann van der Merwe [11] have discussed about graduation studies and the importance of choosing the right Master's programme. In Merwe's article, she talked about the considerations before deciding if you want to participate in a graduation school. Such articles confirm that values of university programmes are more important to Master's students who mostly look into Master's

degree as a way to specialise their skills in a specific area. Therefore, it is of immense importance for them to understand the intension of the programme into which they would invest their time and money.

Narrowing down the scope to the Master's level, we chose to concentrate the research at the VU University Amsterdam. This decision was made due to the ease of access to people in charge of programmes and to current students of programmes. But in order to have a work on an ontology to represent the programmes, another university is also chosen. For the second university to study, University of Amsterdam (UvA) is chosen due to the possibility to access the programme departments and also the fact that these universities work closely in different research areas and also make some courses available for the other university students. Therefore, it is of importance to make sure both universities have the same understanding and definition of concepts introduced in programmes.

To meet the boundaries of resources considered for this research, the specification continues to precisely two programmes. Artificial Intelligence (AI) and Computer Science are two programmes that are chosen to be studied in this paper. Both programmes are at Master's level and are offered by both universities.

The focus of this research is to study programmes' content structure - how subjects that are going to be covered are distributed in a programme and the relations between included concepts.

Concept Highlighting

In order to understand the semantics of programmes I have planned to study the programme pages on the website of VU and UvA for the targeted programmes. Also, the course descriptions given for each programme would be studied to pick the main focus of the programmes and highlight the main concepts covered in each. In this way, I would also need the help of programme coordinates to confirm if my interpretations are in the right direction and to check if the concepts are picked properly according to programmes objective which articulate the knowledge and skills they want students to acquire by the end of the programme. Later also student would be involved in evaluating the defined concepts and the relations in the design.

To structure concepts and map them into an ontology, vocabularies such as ACM Computing Classification System (CCS) [12] will be used as reference. Classification of the programme according to standard vocabularies will enable us to later share the result with even more universities. Additionally, for the design purposes 'Protégé' – a free, open source ontology editor and knowledge-base framework [13] – will be used. Later, based on the time schedule and requirements of the thesis I need to evaluate the design. For the evaluation of the design, a questionnaire will be prepared and shared with students of the field and also university staff. Some ethical problems can be assumed to appear in this regard.

TIMELINE

The time table for this thesis is defined as the following.

Task	Status
Literature Study/Problem Statement	Done
Thesis Proposal Submission	In progress
Study ACM CCS and other possible vocabularies	In progress
Concept breakdown based on available material	In progress
Ontology design	In progress
Define ontological relations	In progress
Ontology evaluation (Questionnaire)	
Define sample uses of the ontology.	

The steps mentioned above are going to be done in parallel and in iterations. An initial estimation of two months is considered for the project.

REFERENCE

- [1] Smith, B. (2009). Ontology and Information Systems. *Science*, (1964), 1-93. Retrieved from <http://eiah.org:8084/arch/handle/123456789/1212>
- [2] Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*. Retrieved from http://www.debitdedata.net/sites/debitdedata.net/files/fichier_pdf/The-semantic-web_scientific-american_0.pdf
- [3] Ontology (Information Science), Wikipedia, retrieved on 2/06/2012 from [http://en.wikipedia.org/wiki/Ontology_\(information_science\)](http://en.wikipedia.org/wiki/Ontology_(information_science))
- [4] Odell, J. (2011). Ontology White Paper. *CSC CATALYST*, 2011-07-15.
- [5] Bonino, D., Corno, F., & Farinetti, L. (2004). Ontology driven semantic search. *WSEAS Transaction on*. Retrieved from <http://www.wseas.us/e-library/conferences/venice2004/papers/472-334.pdf>
- [6] Esa, A. M., & Taib, S. M. (2010). Prototype of semantic search engine using ontology. *2010 IEEE Conference on Open Systems (ICOS 2010)*, (Icos), 109-114. IEEE. doi:10.1109/ICOS.2010.5720073
- [7] Institutional Ontology, Protégé Ontology Library, retrieved on 15/05/2012 from http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library#OWL_ontologies
- [8] Styles, R., Shabir, N. (2008), Academic institution internal structure ontology (AIISO), Retrieved from <http://www.vocab.org/aiiso/schema>
- [9] <http://www.schoolfinder.com/>
- [10] <http://www.yummly.com/>
- [11] van der Merwe, A. (2012). Deciding if graduate school is right for you. Retrieved from <http://www.gradschools.com> on 12/May/2012. Accessible at: <http://www.gradschools.com/article-detail/should-i-go-to-graduate-school-1940>
- [12] <http://www.acm.org/about/class>
- [13] <http://protege.stanford.edu/>