Investigation of Ontologies in Software-Engineering-Meta-Research

Dmitrii Seletkov

Institute for Program Structures and Data Organization (IPD)
Advisor: Dipl.-Inform. Angelika Kaplan

Abstract.

1 Introduction

- The definition of Ontology in Philosophy and Computer Science

- The definition of Meta Research

- Above-named terms Ontology and Meta-research in Software Engineering(SE)

- Help of ontology-based systems in scientific search

Generally, the content (what it will be about) of the following proseminar work is summed up.

2 Foundations

In this section the required knowledge foundations are introduced and set up.

2.1 Ontology in Computer Science

2.1.1 Ontology languages

The short historical survey of ontology languages approaches, based on lectures notes of course "Ontology and knowledge representation from Prof. Boris Konev Head of Department of Computer Science, the University of Liverpool, UK and member of the Knowledge Representation Research Group [9]. In this sequence

- Resource Description Framework Schema: the first standard of W3C for ontologies [4], its semantic power and drawbacks

- Description Logic: introduction in Description Logic, EL language, its architecture and semantic

- Web Ontology Language: the newest standard of W3L based on Description Logic [2]

2.1.2 Expediency and reasons for using of Ontologies

Here is discussed, when and where the onologies should be used and when it is superfluous

2.1.3 Examples of Ontologies

Here come the examples of ontologies in medicine, then in Meta-research and SE.

2.2 Meta-research

2.2.1 Motivation

Why research on research. Problem is given: Several million new research papers are published annually, and the number of publishing authors exceeded → challenge existing notions of data sharing, data ownership, research planning, collaboration, and replication.

2.2.2 Ares of Meta-research

Based on Categorization of J. Ioannidis [7] For each term in the categorization initiatives over the world as examples are given. Also, it is discussed in what fields they are and in what are missing or cannot be so classified.

1. Methods "peforming research" e.g. study design

2. Reporting reporting standard, study registrations

3. Reproducibility "verifying research"

4. Evaluation Prepublication, peer review

5. Incentives rewards and penalties for individual research

3 Ontology in Software-Engineering-Meta-Research

Based on reference-based search or so-called «Snowballing» and the suggested papers as seeds the other papers were found and will be analyzed further in this section. Firstly, divided using the classification of studies in **Primary** and **Secondary**, which were shown in the J. Biolcchini's work [1] at high level and, secondly, using the template «Problem, Objectives, Suggested method and Future works», which was used by me for the reading of papers and can be used for comparison of approaches that were presented in different papers.

3.1 Ontology to support systematic reviews in Software Engineering

As introduction to this subsection could be the inspired by medicine Evidence-based SE [8] (for secondary studies) - what works, when and where, in terms of software engineering practice, tools and standards and whose main instrument are Systematic reviews

- Problem: Produce knowledge that can be based on sound scientific methodology, has become one of the major challenges to strengthen the foundations of SE.

- Objectives
  - Discussing the significance of experimental studies, particularly **systematic reviews**(specific methodology of research, to gather and evaluate the available evidence to a focused topic), and their use in supporting software processes
  - Present a template designed to support systematic reviews in SE
  - Introduce development of ontologies to describe knowledge regarding such experimental studies

- Suggested Method the Ontology

- Result

- Future Work

3.2 Ontologies for Controlled Experiments on Software Engineering

Based on R. E. Garcia's works [12] and [6], where many ontologies for supporting of controlled experiments were presented, they are analyzed using the above-introduced schema:

- Problem: Share of knowledge among research groups requires Replication of Controlled Experiments in SE. The generated Knowledge during these experiments is registered in Lab Packages(procedures, the results and conclusions). However, researchers face difficulties reviewing the lab package and suffer from the lack of standardization, what leads to problems in share knowledge among research groups.

- Objectives

– Present a workflow to generate lab packages based on EXPEROntology

– Present a Possibility to evolve the Ontology iteratively

– Through iterative evolution achieve a standard that is able to accommodate different lab packages → facilitate to review and understand their content

• Suggested Method Suggested Ontologies

• Result

• Future Work

Then, it is discussed another approach, based on Siy's and Wu's work [13] Ontology for support empirical studies, in particular the design of software engineering experiment is presented and compared (also considered like an extension to the above introduced ones)

4 Ontology-based systems in scientific search

During the searching and reading of the suggested and found by reference-based search papers the two facts appeared in sight: Ontologies are used in contemporary search engines and the the scientific world is using ontologies successfully for a big amount of tasks. Therefore, based on database search (because this was not the original accent of the suggested literature list) the papers about Ontology-based systems in scientific search were found and will be analyzed further in this section.

4.1 Semantic Web

Definition, change from syntax Web (WWW) to the semantic Web, motivation and importance of that [3]

4.2 Semantic Search Engine

Semantic search is required by a search engine to properly interpret the meaning of a user's query and the inherent relations among the terms that a document contains with respect to a specific domain. Traditional approaches do not care about the semantic relations among the terms that a document contains with respect to a specific domain. How to make it possible for the search engine to work even in case that there are no available domain ontologies for user requests [5]

4.3 Examples

In medicine: Semantic search engine for cancer [11] with Objective:
to build an Ontology, using Medical Knowledge Base, in order to analyze the knowledge about

cancer, its categories, cause, symptoms ets., thereby translating the human natural language in a machine and human readable language

In general (with evaluation in medicine): information extraction from scientific abstracts [10] with Objective:
to describe a novel ontology-based interactive information extraction (OBIIE) framework and how this system enables life scientists to make ad hoc queries similar to using a standard search engine specific OBIIE system (evaluating for extracting co-factors from EMBASE and MEDLINE (medical databases))

5 Conclusion

References

[1] Jorge Calmon de Almeida Biolchini et al. "Scientific research ontology to support systematic review in software engineering". In: *Adv. Eng. Informatics* 21.2 (2007), pp. 133–151.

[2] Sean Bechhofer et al. *OWL Web Ontology Language Reference.* 2004. URL: http://www.w3.org/TR/2004/REC-owl-ref-20040210/.

[3] Tim Berners-Lee, James Hendler, and Ora Lassila. "The Semantic Web". In: 284 (2001), pp. 34–43. ISSN: 0036-8733. DOI: 10.1038/scientificamerican0501-34.

[4] Dan Brickley and R.V. Guha. *RDF Vocabulary Description Language 1.0: RDF Schema.* W3C Recommendation. World Wide Web Consortium, Feb. 2004. URL: http://www.w3.org/TR/2004/REC-rdf-schema-20040210/.

[5] Wei-Dong Fang et al. *Toward a semantic search engine based on ontologies.* 2005. DOI: 10.1109/icmlc.2005.1527258.

[6] Rogério Eduardo Garcia et al. "An Ontology for Controlled Experiments on Software Engineering". In: *Proceedings of the Twentieth International Conference on Software Engineering & Knowledge Engineering (SEKE'2008), San Francisco, CA, USA, July 1-3, 2008.* Knowledge Systems Institute Graduate School, 2008, pp. 685–690.

[7] John P. A. Ioannidis et al. "Meta-research: Evaluation and Improvement of Research Methods and Practices". In: 13 (), e1002264. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.1002264.

[8] Barbara A. Kitchenham, Tore Dyba, and Magne Jorgensen. "Evidence-Based Software Engineering". In: *26th International Conference on Software Engineering (ICSE '04).* Edinburgh, Scotland, May 2004, pp. 273–281.

[9] Boris Konev. *Lecture notes of course Ontology and knowledge representation.* 2010. URL: https://www.lektorium.tv/speaker/2680.

[10] David Milward et al. "Ontology-Based Interactive Information Extraction From Scientific Abstracts". In: 6 (2005), pp. 67–71. ISSN: 1531-6912. DOI: `10.1002/cfg.456`.

[11] Syam RajBS and Sarumathi S. "Ontology based Semantic Search Engine for Cancer". In: (2014), pp. 39–43. ISSN: 0975-8887. DOI: `10.5120/16594-6308`.

[12] Lilian Passos Scatalon, Rogério Eduardo Garcia, and Ronaldo Celso Messias Correia. "Packaging Controlled Experiments Using an Evolutionary Approach Based on Ontology(S)". In: *Proceedings of the 23rd International Conference on Software Engineering & Knowledge Engineering (SEKE'2011), Eden Roc Renaissance, Miami Beach, USA, July 7-9, 2011.* Knowledge Systems Institute Graduate School, 2011, pp. 408–413.

[13] Harvey Siy and Yan Wu. *An Ontology to Support Empirical Studies in Software Engineering.* en. Mar. 2012. URL: `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.216.8899;%20http://csalpha.ist.unomaha.edu/~hsiy/research/icc09.pdf`.