Accepted Manuscript

Building an effective and efficient background knowledge resource to enhance ontology matching

Amina Annane, Zohra Bellahsene, Faiçal Azouaou, Clement Jonquet

PII: S1570-8268(18)30017-9
DOI: https://doi.org/10.1016/j.websem.2018.04.001
Reference: WEBSEM 458

To appear in: *Web Semantics: Science, Services and Agents on the World Wide Web*

Received date : 16 June 2017
Revised date : 23 March 2018
Accepted date : 24 April 2018

Please cite this article as: A. Annane, Z. Bellahsene, F. Azouaou, C. Jonquet, Building an effective and efficient background knowledge resource to enhance ontology matching, *Web Semantics: Science, Services and Agents on the World Wide Web* (2018), https://doi.org/10.1016/j.websem.2018.04.001

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

*Manuscript
Click here to view linked References

ACCEPTED MANUSCRIPT

Building an effective and efficient background knowledge resource to enhance ontology matching

Amina Annane[1,2], Zohra Bellahsene[2], Faiçal Azouaou[1], Clement Jonquet[2,3]

[1] *Ecole nationale Supérieure d'Informatique, BP 68M, 16309, Oued-Smar, Alger, Algérie*

[2] *LIRMM, Université de Montpellier, CNRS, Montpellier, France*

[3] *Center for BioMedical Informatics Research (BMIR), Stanford University, USA*

Abstract

Ontology matching is critical for data integration and interoperability. Original ontology matching approaches relied solely on the content of the ontologies to align. However, these approaches are less effective when equivalent concepts have dissimilar labels and are structured with different modeling views. To overcome this semantic heterogeneity, the community has turned to the use of external background knowledge resources. Several methods have been proposed to select ontologies, other than the ones to align, as background knowledge to enhance a given ontology-matching task. However, these methods return a set of complete ontologies, while, in most cases, only fragments of the returned ontologies are effective for discovering new mappings. In this article, we propose an approach to select and build a background knowledge resource with just the right concepts chosen from a set of ontologies, which improves efficiency without loss of effectiveness. The use of background knowledge in ontology matching is a double-edged sword: while it may increase recall (i.e., retrieve more correct mappings), it may lower precision (i.e., produce more incorrect mappings). Therefore, we propose two methods to select the most relevant mappings from the candidate ones: (1) a selection based on a set of rules and (2) a selection based on supervised machine learning. Our experiments, conducted on two Ontology Alignment Evaluation Initiative (OAEI) datasets, confirm the effectiveness and efficiency of our approach. Moreover, the F-measure values obtained with our approach are very competitive to those of the state-of-the-art matchers exploiting background knowledge resources.

Keywords: Ontology matching, Ontology alignment, Background knowledge, Indirect matching, External resource, Anchoring, Derivation, Background knowledge selection, Supervised machine learning.

1. Introduction

Ontologies provide conceptual models to represent and share knowledge. Some of the data management challenges for which ontologies are often used include interoperability [55] and data integration [35]. In recent years, because of the explosion of the number of ontologies, especially in domains that produce and manage large amounts of data (such as biomedicine [43]), these challenges have become increasingly complex. To achieve interoperability and integration, one solution is to identify mappings (correspondences) between different ontologies of the same domain. This process is known as *ontology matching* or *ontology alignment*.

Ontologies are heterogeneous because they have been designed independently, by different developers, and following diverse modeling principles and patterns. This heterogeneity makes the matching process complex [20]. The first ontology matching methods were based only on the lexical and structural content of the ontologies to align; this is known as *direct matching* or *content-based matching*. To that end, many syntactic and structural similarity measures have been developed [8, 42, 20]. However, direct matching is less effective to find correspondences between concepts that are equivalent, but described with dissimilar labels and structured with different modeling views [1, 50].

To overcome this semantic heterogeneity, the community has turned to the exploitation of external knowledge resource(s), commonly called background knowledge resources. In contrast to direct matching, this approach is known as *indirect matching*, *BK-based matching* or *context-based matching* [37], as it exploits external resources to identify mappings between the ontologies to align.

The BK-based matching approach raises two main issues: (i) how to select (or build) background knowledge resource(s) for a given ontology matching task? and (ii) how to concretely use the selected background knowledge resource(s) to enhance the quality of the matching result?

In the literature, several works have addressed these

Email addresses: a_annane@esi.dz (Amina Annane[1,2]), bella@lirmm.fr (Zohra Bellahsene[2]), f_azouaou@esi.dz (Faiçal Azouaou[1]), jonquet@lirmm.fr (Clement Jonquet[2,3])

Download English Version:

https://daneshyari.com/en/article/6950427

Download Persian Version:

https://daneshyari.com/article/6950427

Daneshyari.com