

Chapter 99

State of the Art on Ontology Alignment

Maria Vargas-Vera
Universidad Adolfo Ibanez, Chile

Miklos Nagy
The Open University, UK

ABSTRACT

Ontology mapping as a semantic data integration approach has evolved from traditional data integration solutions. The core problems and open issues related to early data integration approaches are also applicable to ontology mapping on the Semantic Web community. Therefore, in this review the authors present the related literature, starting from the traditional data integration approaches, in order to highlight the evolution of data integration from the early approaches. Once the roots of semantic data integration have been presented, the authors proceed to introduce the state-of-the-art of the ontology mappings systems including the early approaches and the systems that can be compared through the Ontology Alignment Initiative (OAEI).

INTRODUCTION

Ontology mapping is considered to be one of the major roadblocks towards successful data and information integration on the Semantic Web. This is also supported by the fact that a significant amount of research effort has been invested in this area during the last 14 years. The invested research effort has brought numerous ontology mapping systems into various operational stages. The most successful approaches have participated in the OAEI (Ontology Alignment Evaluation Initiative), which is an international benchmarking

effort in this domain. The participation level of the mapping systems have varied during the past 8 years but the flagship approaches, including our ontology mapping approach DSSim (Nagy & Vargas-Vera, 2011; Nagy et. Al., 2008, 2006, 2005), have dominated the group of consistent participants in terms of mapping quality and the number of mapping tasks being performed. The main advantage of participating in the OAEI is that systems can be compared and evaluated under the same conditions (Euzenat, 2002, 2004, 2006; Euzenat et al., 2005, 2007, 2009; Euzenat & Shvaiko, 2007).

DOI: 10.4018/978-1-4666-8751-6.ch099

IMPORTANCE OF THE PROBLEM

The problem of ontology mapping is important because its solution have an immediate application in search engines, question answering systems, etc. In fact the Information Retrieval community could in principle profoundly benefit from this proposed work. Based on the systems currently available at competitive level, we have visualized that the key challenges are as follows:

- a. **Robustness across domains:** Most systems use multiple techniques such as heuristics, machine learning or Natural Language Processing (NLP) in order to transform the information in the ontologies into their internal representation. For example, *ASMOV* (Jean-Mary et al., 2009) uses domain specific background knowledge whereas *RiMOM* (Zhang et al., 2009) applies pre-defined rules to assess similarities. *Anchor-Flood* (Seddiqui & Aono, 2009) and *TaxoMap* (Hamdi et al., 2009) have been designed to exploit large textual descriptions of the ontology concepts, which is an assumption that cannot be satisfied across domains. These techniques have the problem that could impact domain independence negatively because they require a-priori knowledge from a designer. It is important to emphasize that ontology designers will always have the freedom to model their domain according to their need. In the same way, database designers can come up with different models for the same problem. To overcome this problem existing systems utilize various types of domain knowledge (heuristic rule or training data set).
- b. **Uncertain reasoning:** Some ontology alignment systems provide limited reasoning support, in order to derive new information from the input ontologies. Unfortunately, not

enough emphasis is placed on the reasoning part in spite of the fact that it has the potential to provide an added value to the systems. Furthermore, the uncertain reasoning possibility is completely missing from many of the existing systems.

- c. **Managing conflicts:** Conflict detection is only provided as a post-processing functionality. However, conflicts that normally appear during the mapping process are not treated properly by the current solutions. Managing conflicting information does have the potential to improve the end results slightly or moderately depending on the domains.
- d. **Mapping optimization:** Only two systems (*RiMOM*, *TaxoMap*) consider optimization of the mapping process, while the rest of the systems do not even consider it as a problem at this stage. This can be explained by the fact that most of the systems have not faced the problem of processing large-scale real world ontologies. While it is true that optimization issues can be addressed later on, it is important that the mapping solutions are conceived with scalability options.
- e. **Mapping visualization:** Each system presents the mapping result to the users, although little emphasis has been placed on how these results are presented. Most systems show the results as a list of mapping pairs, and only some employ two-dimensional graph-based visualization. Additionally, there is no way to examine how the system produced these results as only the end results are kept.

This paper discusses the state of the art in ontology alignment presenting the problem of alignment from two perspectives Databases and Ontologies research. Firstly, the paper presents an overview of data integration. Secondly, it shows relevant work on semantic data integration (from

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/state-of-the-art-on-ontology-alignment/138378

Related Content

Causal and Total Order in Opportunistic Networks

Mihail Costea, Radu-Ioan Ciobanu, Radu-Corneliu Marin, Ciprian Dobre, Constandinos X. Mavromoustakis and George Mastorakis (2016). *Emerging Innovations in Wireless Networks and Broadband Technologies* (pp. 221-262).
www.irma-international.org/chapter/causal-and-total-order-in-opportunistic-networks/148597

Cooperative Space Time Coding for Semi Distributed Detection in Wireless Sensor Networks

Mohammad A. Al-Jarrah, Nedal K. Al-Ababneh, Mohammad M. Al-Ibrahim and Rami A. Al-Jarrah (2012). *International Journal of Wireless Networks and Broadband Technologies* (pp. 1-15).
www.irma-international.org/article/cooperative-space-time-coding-for-semi-distributed-detection-in-wireless-sensor-networks/85002

Reviewing Mobile Marketing Research to Date: Towards Ubiquitous Marketing

Dimitris Drossos and George M. Giaglis (2012). *Wireless Technologies: Concepts, Methodologies, Tools and Applications* (pp. 181-207).
www.irma-international.org/chapter/reviewing-mobile-marketing-research-date/58788

IoT, IR 4.0, and AI Technology Usability and Future Trend Demands: Multi-Criteria Decision-Making for Technology Evaluation

Mujahid Tabassum, Sundresan Perumal, Saju Mohanan, P. Suresh, Sunitha Cheriyan and Waqar Hassan (2021). *Design Methodologies and Tools for 5G Network Development and Application* (pp. 109-144).
www.irma-international.org/chapter/iot-ir-40-and-ai-technology-usability-and-future-trend-demands/271578

A Survey on Wireless Sensor Networks

Homero Toral-Cruz, Faouzi Hidoussi, Djallel Eddine Boubiche, Romeli Barbosa, Miroslav Voznak and Kamaljit I. Lakhtaria (2015). *Next Generation Wireless Network Security and Privacy* (pp. 171-210).
www.irma-international.org/chapter/a-survey-on-wireless-sensor-networks/139430