

# Discovering Mappings Between Ontologies

**Vikram Sorathia**

*Dhirubhai Ambani Institute of Information and Communication Technology, India*

**Anutosh Maitra**

*Dhirubhai Ambani Institute of Information and Communication Technology, India*

## INTRODUCTION

Knowledge Representation is important part of AI. The purpose is to reveal best possible representation of the Universe of Discourse (UoD) by capturing entities, concepts and relations among them. With increased understanding of various scientific and technological disciplines, it is possible to derive rules that governs the behaviour and outcome of the entities in the UoD. In certain cases, it is not possible to establish any explicit rule, yet through experience or observation, some experts can define rules from their tacit knowledge in specific domain.

Knowledge representation techniques are focused on techniques that allows externalization of implicit and explicit knowledge of expert(s) with a goal of reuse in absence of physical presence of such expertise. To ease this task, two parallel dimensions have developed over period of time. One dimension is focused on investigating more efficient methods that best suit the knowledge representation requirement resulting in theories and tools that allows capturing the domain knowledge (Brachman & Levesque, 2004). Another development has taken place in harmonization of tools and techniques that allows standard based representation of knowledge (Davies, Studer, & Warren, 2006).

Various languages are proposed for representation of the knowledge. Reasoning and classification algorithms are also realized. As an outcome of standardization process, standards like DAML-OIL (Horrocks & Patel-Schneider, 2001), RDF (Manola & Miller, 2004) and OWL (Antoniou & Harmelen, 2004) are introduced. Capturing the benefit of both developments, the tooling is also came in to existence that allows creation of knowledgebase.

As a result of these developments, the amount of publicly shared knowledge is continuously increasing. At the time of this writing, a search engine like Swoogle (Ding et al., 2004)-developed to index publicly available

Ontologies, is handling over 2,173,724 semantic web documents containing 431,467,096 triples.

While the developments are yielding positive results by such a huge amount of knowledge available for reuse, it have become difficult to select and reuse required knowledge from this vast pool. The concepts and their relations that are important to the given problem could have already been defined in multiple Ontologies with different perspectives with specific level of details. It is very likely that to get complete representation of the knowledge, multiple Ontologies must be utilized. This requirement has introduced a new discipline within the domain of knowledge representation that is focused on investigation of techniques and tools that allows integration of multiple shared Ontologies.

## BACKGROUND

The problem of Ontology integration is not completely new. Schema Matching is a similar problem being addressed in the context of enterprise integration. But, in Ontology matching, the scale and complexity is much higher and requires special considerations. (Shvaiko & Euzenat, 2006) highlights the key similarities and differences between both the techniques. In schema matching, the semantics of the given term is guessed whereas the ontology matching methods relies on deriving the semantics from explicit representation of concepts and relations in given Ontology. Numerous methods and approaches have been proposed that attempt to solve the problem targeting specific aspects of the represented knowledge (Ehring, 2007).

Apart from standards that guide the languages used for the development of Ontology, some standard Ontologies have also been defined. The role of these Ontologies is to provide framework of vary basic elements and their relations, based on which complex domain knowledge can be developed. SUO (Niles & Pease,

2001), SUMA(Niles & Pease, 2003), OpenCyc(Sicilia et al., 2004) are examples of the same. SWEET(Raskin, 2003) provides standard Ontologies in environmental science domain. Hence, the levels in Ontology also address important dimension in knowledge engineering through integrating available Ontologies.

### ONTOLOGY MAPPING TECHNIQUES

Research in integration of multiple Ontologies have resulted in various techniques and tools that have successfully demonstrated capabilities in producing the required results(Noy, 2004b). The ontology integration is addressed as Ontology mapping, matching, merging, transforming and other such activities. The integration is achieved by focusing on finding similarities among the concepts of separate Ontologies. The similarity or nearness can be established by employing various techniques, and numerous such approaches have been published demonstrating the suitability of single or hybrid approaches. The taxonomic overview of existing methodology is provided in many survey papers that provides a reasonable entry in to the domain of Ontology integration. (Kalfoglou et al., 2005) provides comprehensive survey of Ontology mapping approach and classify them on Semantic Intensity Spectrum. (Noy, 2004a) (Kalfoglou & Schorlemmer, 2005) and (Predoiu et al., 2006) provides comprehensive survey discussing state-of-the-art of present research efforts.

Ontologies consists of concepts and elements. The integration process that establishes the similarity among concepts consists of three dimensions (Shvaiko & Euzenat, 2006). The input dimension is related to underlying data model and can operate at schema level or instance level. Second is the process dimension that classifies approach as exact or approximate determination. Third dimension deals with output in the form of Cardinality, type of relation and the confidence. Integration can be done by identification of Alignment.

#### Concept Level Approaches

Concept level approaches are restricted only to the name of the concept and employ various methods to match whole or part of the concept names that belong to different Ontologies. Though these syntax oriented approaches proves to be less efficient when applied in isolation, they are generally employed in pre-integra-

tion preparation phase (or normalization phase) of more complex semantic oriented approaches. Many of the Schema Matching techniques are directly applicable for concept level approaches.

#### String Level Concept Matching

It is based on the simple assumption that concept having similarity is represented with same name in different Ontologies. Upon identification of such string level similarity the source Ontologies can either mapped or merged. PROMPT(Noy & Musen, 2000) Ontology Merging tool employs string level concept matching approach.

#### Sub-String Level Concept Matching

Approaches that brakes the input concepts in to smaller segments on the basis of prefix, suffix and other structures. Another approach establishes the similarity by identifying the *Edit Distance*. For example if Nikon and NKN are under consideration, the Edit Distance is a number of insertion, deletion and substitution of characters that will be required in Nikon and NKN to transform one into the other. N-gram technique is employed for deriving a set of substrings by selecting n number of characters from input string. For example trigram of NIKON results in NIK, IKO and KON. The derived set can further be subjected to simple string matcher for finding similarities.

#### Lexical Matching

Lexical approaches are employed to identify and extract tokens from the input string. This is particularly useful when concept name are created using mix of alphanumeric characters that can be processed to separate operators, numbers, punctuations and other types of token to reveal processable substrings. LOM(Li, 2004)- a Lexicon based Ontology Mapping tool employs strategy to determine similarity by matching the whole term, word constituent, synset, and type matching (Choi et al., 2006). OLA(Euzenat et al., 2005) and Cupid(Madhavan et al., 2001) also employs lexical techniques for finding similarity among concepts.

5 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/discovering-mappings-between-ontologies/10292](http://www.igi-global.com/chapter/discovering-mappings-between-ontologies/10292)

## Related Content

---

### How Intelligent Are Ambient Intelligence Systems?

María J. Santofimia, Francisco Moya, Félix J. Villanueva, David Villa and Juan C. López (2010). *International Journal of Ambient Computing and Intelligence* (pp. 66-72).

[www.irma-international.org/article/intelligent-ambient-intelligence-systems/40351](http://www.irma-international.org/article/intelligent-ambient-intelligence-systems/40351)

### Incremental Load in a Data Warehousing Environment

Nayem Rahman (2012). *Insights into Advancements in Intelligent Information Technologies: Discoveries* (pp. 161-177).

[www.irma-international.org/chapter/incremental-load-data-warehousing-environment/64375](http://www.irma-international.org/chapter/incremental-load-data-warehousing-environment/64375)

### What Does Artificial Life Tell Us About Death?

Carlos Gershenson (2013). *Investigations into Living Systems, Artificial Life, and Real-World Solutions* (pp. 17-22).

[www.irma-international.org/chapter/does-artificial-life-tell-death/75916](http://www.irma-international.org/chapter/does-artificial-life-tell-death/75916)

### The Influence of Culture on Sentiments Expressed in Online Reviews of Eco-Friendly Hotels: The Case Study of Amsterdam

Estefania Ballester Chirica, Carla Ruiz-Mafé and Natalia Rubio (2023). *Philosophy of Artificial Intelligence and Its Place in Society* (pp. 115-141).

[www.irma-international.org/chapter/the-influence-of-culture-on-sentiments-expressed-in-online-reviews-of-eco-friendly-hotels/332603](http://www.irma-international.org/chapter/the-influence-of-culture-on-sentiments-expressed-in-online-reviews-of-eco-friendly-hotels/332603)

### The Impact of Emerging Technologies and Social Media on Different Business(es): Marketing and Management

Smita Agrawal (2017). *Maximizing Business Performance and Efficiency Through Intelligent Systems* (pp. 37-49).

[www.irma-international.org/chapter/the-impact-of-emerging-technologies-and-social-media-on-different-businesses/178295](http://www.irma-international.org/chapter/the-impact-of-emerging-technologies-and-social-media-on-different-businesses/178295)