# Chapter 98
# Semantic Mediation in MedPeer: An Ontology-Based Heterogeneous Data Sources Integration System

**Naïma Souâd Ougouti**
*Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf, Algeria*

**Haféda Belbachir**
*Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf, Algeria*

**Youssef Amghar**
*INSA of Lyon, France*

## ABSTRACT

*Peer-to-Peer (P2P) infrastructure is an emerging paradigm that offers new opportunities for the development of large-scale distributed systems. This architecture combined with the new techniques introduced by semantic web as ontologies encouraged the emergence of new multi-source data integration possibilities for sharing information. A challenging problem in such systems is to find correspondences between concepts of their different ontologies. This is a necessary step before locating peers that are relevant with respect to a given query. In this paper, the authors propose a new ontology alignment method which deals with both linguistic and semantic characteristics of concepts and uses graph structure to explore multiple depth levels of neighborhood in calculation of semantic similarity which is the most important part of their global similarity measure. This function is implemented into their new P2P heterogeneous and distributed data integration system MedPeer.*

## INTRODUCTION

The expansion of data sources present on the web has generated a new type of applications that aim at the integration of data from these different sources. Nowadays, the user would like to spend a minimum of time and resources to have the necessary information. However, this information is often located in independent heterogeneous and distributed data sources. In order to effectively share that kind of in-

DOI: 10.4018/978-1-5225-5191-1.ch098

formation, several solutions have been proposed. Among them, we can cite peer-to-peer systems (P2P) (Daswani et al., 2003) combined with semantic web techniques (Berners-Lee, 2005).

(Staab & Stuckenschmidt, 2006) gave the following definition to these two technologies:

*The Semantic Web and Peer-to-Peer are two technologies that address a common need at different levels:*

- The Semantic Web addresses the requirement that one may model, manipulate and query knowledge and information at the conceptual level rather than at the level of some technical implementation. Moreover, it pursues this objective in a way that allows people from all over the world to relate their own view to this conceptual layer. Thus, the Semantic Web brings new degrees of freedom for changing and exchanging the conceptual layer of applications.
- Peer-to-Peer technologies aim at abandoning centralized control in favor of decentralized organization principles. In this objective they bring new degrees of freedom for changing information architectures and exchanging information between different nodes in a network.
- Together, Semantic Web and Peer-to-Peer allow for combined flexibility at the level of information structuring and distribution."

One of the most important tools of the Semantic Web is Ontology. It provides a vocabulary that describes a domain of interest and a specification of the meaning of terms used in the vocabulary (Euzenat & Shvaiko, 2016).

In this context, we have introduced a new P2P ontology-based heterogeneous data sources integration system named MedPeer. This system has a super-peer architecture based on a peer grouping by media type (texts, images, relational databases, semi-structured data, etc.). Each super-peer must have its own specific domain ontology. When peers have different schemas to manage, semantic matching is needed to overcome semantic heterogeneity between them.

(Giunchiglia et al., 2007) consider semantic matching as a fundamental technique which applies in many areas such as resource discovery, data integration, data migration, query translation, peer to peer networks, agent communication, schema and ontology merging. Semantic matching is a type of ontology matching technique that relies on semantic information encoded in lightweight ontologies to identify nodes that are semantically related. It operates on graph-like structures and has been proposed as a valid solution to the semantic heterogeneity problem, namely managing the diversity in knowledge.

In this paper, we present a global similarity measure between domain ontology concepts and those of local ontologies. We assume that peers contain relational databases. In a first step, local ontologies are generated to describe the schema of these databases. This is achieved by a new method that we have proposed in a previous work named Relationnal.OWL2E (Ougouti et al., 2015), allowing from a relational schema to generate automatically the corresponding OWL2 based ontology. In the second step, the semantic mediation process may start by comparing the concepts in order to find similarities which will be stored and used in the semantic routing process and the rewriting queries one.

The core contribution of this paper with regard to similar works cited in the literature is that our similarity measure is complete and relies on all information encoded in ontologies. It also uses graph structure to explore multiple depth levels of the neighborhood in calculation of semantic similarity which is the most important part of our global similarity function. When other methods stop the neighborhood to direct links, our approach considers direct and indirect ones.