CIT. Journal of Computing and Information Technology, Vol. 24, No. 4, December 2016, 383–400
doi: 10.20532/cit.2016.1003228

Integrating Ontological Data Sources Using Viewpoints-Based Approach

Bouchra Boulkroun, Fouzia Benchikha and Chahinez Bachtarzi

LIRE Laboratory, Constantine 2 University Abdelhamid Mehri, Constantine, Algeria

Within the development of Internet and intranets, information integration from various data sources becomes increasingly important and more challenging issue. Recently, the trend in data integration has favored the semantic integration using ontologies. However, the existing ontology-based approaches do not support the aspect of data multi-representations, which is important in the development of multi-user applications. The motivation of this paper is to address a novel semantic integration approach based on ontologies and viewpoints paradigms. This contribution combines the advantages of existing ontology-based integration approaches while avoiding their drawbacks. The proposed integration approach is evaluated using query processing. Profiles are introduced to offer answers to users according to their viewpoints and choices.

ACM CCS (2012) Classification: Information systems → Information integration → Mediators and data integration

Keywords: integration, viewpoint, multi-viewpoint, mediation, ontologies, profiles

1. Introduction

Nowadays, due to the success of the World Wide Web and the rapid growth of information exchange over it, there has been an increasing interest in accessing, relating and combining data from multiple heterogeneous data sources. Indeed, developing information integration systems becomes a crucial need. An information integration system provides a uniform interface to efficiently access and use these heterogeneous data sources.

Several works focusing on heterogeneous data integration have been proposed in the literature. Most of them are based on the common mediation architecture [1], [2]. The main challenge consists on resolving data heterogeneity problems on both structural and semantic levels. The structural heterogeneity is often resolved by defining a common data model. The semantic heterogeneity represents the major problem and ontologies have been recently proposed to deal with it.

Several ontology-based integration approaches have been developed [3], [4]. They can be classified into two categories: the "a priori" and the "a posteriori" approaches [5]. In the former, the concepts of local ontologies are a priori articulated with those of a pre-existent global ontology and the integration process is completely automated. The latter attempts to integrate data sources that possess their own local ontology conceived independently. The integration process, in this case, requires human intervention to define the correspondences between the concepts of different ontologies.

On the other hand, another particularity of data integration is that it is a collaborative task and may involve many experts with different degrees of knowledge in the same application domain [6]. However, each expert can have his own vision of the domain according to his own standpoint, expectations and point of view. Unfortunately, existing integration systems do not support the aspect that data, in a particular viewpoint, can be more relevant to answer user's query.

In this paper, we propose an original ontology-based integration approach that benefits from the advantages of both "a priori" and "a posteriori" approaches. It aims to integrate a large number of ontology-based data sources considering at the same time the existing viewpoints and perspectives in the domain. The integration framework of our approach holds three principal characteristics: (*i*) there is a domain ontology that represents the domain vocabulary as well as the experts' viewpoints, (*ii*) the data sources keep their autonomy since they are conceived independently from the shared ontology, and (*iii*) the integration process is performed in an automatic way.

We propose a query processing method where SPARQL queries expressed on the multi-viewpoint ontology are reformulated into sub-queries on the viewpoint ontologies. This reformulation makes use of the notion of viewpoints and the mappings between the different levels of the integration system defined in an algebraic formalism.

The remainder of the paper is organized as follows. In the next section, we present some research works on ontology-based integration and viewpoints. An overview of the proposed approach is presented in Section 3. In Section 4, we detail the viewpoint-based integration methodology with its advantages, followed by the mediation architecture in Section 5. In Section 6, we give some validation examples and experiments. Finally, as conclusion, we discuss some prospects and future works.

2. Background

In this section, we present some related works dealing with semantic data integration, then we discuss the utilization of the viewpoint paradigm in different research works and in the context of ontologies in particular.

2.1. Ontology-Based Integration

Information integration has been studied in different domains and contexts to reach several goals. Each domain has its particular requirements that lead to new approach for data integration. In [7], the authors outline some of these new ways such as: event-based integration, stream-based integration. Some researchers have been focused on quality assessment where several works have adopted different quality definitions such as information quality and data quality, but these techniques are generally applied at the query answering stage. In this area, we can quote [8], where the authors propose methods for measuring the quality of information retrieved from the data sources. In [9], the authors propose three quality criteria explicitly for the data integration context: schema completeness, data type consistency, schema minimality. WASSIT [10] is an example of ontology-based integration frameworks that exploits data quality as a criterion for data sources selection. In [6], the authors describe a quality framework for data integration that is able to represent different quality requirements arising from different stakeholders.

In this paper, we are particularly interested in the integration of heterogeneous data sources to ensure their interoperability and to provide a unified access to them. However, the first generation of integration systems was represented by the manual integration [2]. In all these propositions, efforts have been essentially focused on automating the syntactic interoperability of data; the semantic conflicts are resolved in a manual way. According to the point of view of Hacid *et al.* [11], the integration raises several problems such as the expressiveness of the operation and the lack of an intelligent interface that assists users in the formulation of their queries.

For this purpose, several researchers have developed new integration methods based on the notion of ontology called "Ontology-based integration approaches". In the proposed approaches, different types of ontologies are used to deal with data heterogeneity. In the MOMIS project [12], a linguistic ontology is employed to semi-automatically integrate data of structured and semi-structured sources. In other research works such as: PICSEL [13], OBSERVER [14], KRAFT [15], designers articulate the semantics of their data sources with a conceptual ontology.

Ontology-based integration approaches are classified in two main categories: a posteriori semantic integration and a priori semantic integration.

- *The a posteriori semantic integration* is adopted when data sources possess their own local ontology conceived independently of the developed integration system. In this approach, the integration is performed in a manual or semi-automatic way by establishing correspondences between the concepts of the ontologies and those of the domain ontology [14], [15]. The main advantage of such an approach is the large autonomy of the participants' data sources. However, human intervention is required to reference explicitly the shared ontology and to define the semantic relations between the concepts of local ontologies and the shared ontology. Another drawback of this approach lies in the evolution of the integration system, since adding a new data source may require changes in the global ontology.

- In the *a priori semantic integration*, each data source builds its local ontology by a priori taking concepts from the pre-existent domain ontology. Thus, the local ontologies can be seen as subsets of the pre-existent global ontology. The integration in this approach is naturally automatic since the semantic relations between the local ontologies and the global ontology are a priori established [13], [16], [17]. Such an approach allows easily integrating a new data source into the system if the semantics of this source are defined in the global ontology. On the other hand, this aspect limits the autonomy of data sources. In [17], authors propose an integration approach where each data source is allowed to make its own extensions of the shared ontology.

2.2. Viewpoints and Ontologies

In knowledge acquisition and modeling areas, the viewpoint paradigm has largely contributed to master the complexity of the design of complex systems. Viewpoints help to model knowledge by giving different perceptions of an object with respect to the observer's position. By opposition to the mono-viewpoint approach, the multi-viewpoint approach allows modeling the same reality according to different viewpoints. Researches on viewpoints were carried out in the seventies with the works of [18]. They considered that an object could be seen by different observers according to diverse viewpoints. Later, the viewpoint concept has been used with diverse senses in various domains such as in databases [19], [20], knowledge representation [21], [22], semantic web [23], [24], [25], etc.

In this paper, we are particularly interested in the use of viewpoints in ontology's modeling and exploitation. Indeed, authors in [26] were the first to show the duality between ontologies and viewpoints: while ontologies are shared models, viewpoints are local models. They highlight the importance of contextualizing ontologies and propose an extension of the OWL language using contexts. In [25], the author also raises the issue of reusing ontologies in a context where only parts of the originally encoded aspects are relevant. He proposes different viewpoints of the same ontology that can be used for multi-viewpoint reasoning.

However, ontology corresponds to a specific domain modeling and requires that the members of the concerned community make a commitment to use it. According to [21], the maintenance of such a consensus becomes a very difficult task. Therefore, an ontology should be built in a multi-viewpoint environment to take into account the diversity of the data sources and the diverse categories of users by keeping certain level of consensus. The constructed ontology in such a way is called "multi-viewpoint ontology". In [24], the authors define a multi-viewpoint ontology (MVPO) as an ontology in which a concept can be associated to several definitions, each corresponding to a particular viewpoint. The place of the concept in the hierarchy depends on its definition. Viewpoints have been used for different purposes to solve various problems. In the context of our work, we use the viewpoint paradigm to integrate a large number of distributed heterogeneous data sources.

3. Approach Overview

Currently, we agree to recognize the interest of the viewpoints in the conception and the development of multi-user applications that require the cooperation of several experts, each with his interests and knowledge. In [27], the authors show the benefits of viewpoints to allow data integration in an intuitive way.

In this section, we first introduce the basic principles of the proposed viewpoint-based integration approach and then we discuss the purpose of using viewpoints for data integration.

3.1. Basic Principles

The proposed approach is based on the viewpoint paradigm. It allows the integration of numerous distributed heterogeneous ontological data sources in an innovative manner. The integration is held in three levels (see Figure 1): *local level*, *Viewpoint level and Multi-Viewpoint level*, detailed in the following.

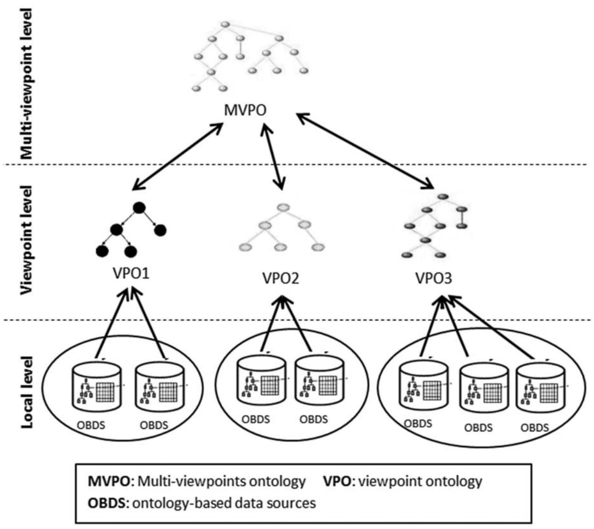

Figure 1. Ontology-based integration using viewpoints.

a) ***Local level:*** at this level, a set of N data sources that participate in the integration system are considered, where:
   - Each data source is an ontology-based data source. This means that it defines its terms in its own ontology, called local ontology (LO).
   - Each data source represents data of the same domain according to a particular viewpoint.

b) ***Viewpoint level:*** it is an intermediate integration level which holds the set of viewpoints ontologies (VPO_i, $i = 1, p$). The data sources are clustered into categories. Each category is related to a particular viewpoint, and:
   - Data sources of the same category are associated to the same VPO,
   - Each VPO represents one viewpoint (VP_j) and holds the ontological concepts description according to the considered viewpoint.

c) ***Multi-viewpoint level:*** it is the global integration level that holds the multi-viewpoint ontology (MVPO), where:
   - The MVPO represents the shared ontology and provides a conceptual description of the domain by a set of consensual terms.
   - The MVPO provides a unified vision of the different VPO. It is used as a reference for the classification of the data sources into viewpoint categories.

3.2. Why Using Viewpoints?

In this section, we essentially motivate the choice of the viewpoint paradigm in our approach and present its main advantages. However, existing ontology-based integration approaches have the ability to respond to various problems related to the semantic heterogeneity but they do not take into account the divergence of the experts' and end users' viewpoints. Therefore, the development of an integration approach that exploits both ontologies and viewpoints appears as a revolution in this context. Moreover, the use of viewpoints allows avoiding the problem of conflicts related to the different interpretations that may have the same object in the real world. Other advantages are offered:

- Viewpoints can be used as a means for the interpretation of initial knowledge of the domain ontology.
- Viewpoints allow filtering the access to data sources according to the users' preferences and needs as a way to improve system performances.
- Viewpoints can separate the initial knowledge at different granularity levels. This separation promotes the division of the domain into several sub-domains, which are easier to manage.

4. The Proposed Integration Methodology

In this section, we present the proposed viewpoint-based integration approach, then we compare it with existing techniques and systems.

4.1. The Viewpoint-Based Integration Approach

The integration methodology proposed in this paper allows designers to exploit existing ontology-based data sources for the realization of a unified and integrated system. It breaks the integration process into four steps (see Figure 2): MVPO definition, VPO extraction, correspondences detection and data sources classification. In the next section, we give more details about each step, and provide formal definitions of the concepts. These definitions aim to reduce the specification ambiguity.

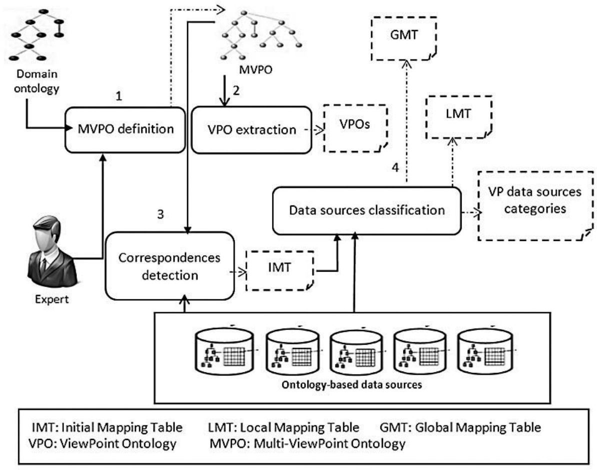

Figure 2. Ontology-based integration using viewpoints methodology.

4.1.1. MVPO Definition

Rather than developing a MVPO from scratch, it is more convenient and effortless to reuse existing domain ontology. The goal of this phase is to exploit domain ontology and transform it into a MVPO. Indeed, to achieve a multi-viewpoint representation for different experts' perspectives, the MVPO must contain all the knowledge and terminologies of their viewpoints. In the following, a formal definition of a MVPO is given. The MVP-OWL language [21] is adopted for the MVPO representation. In the annex, we succinctly present this model.

Definition 4.1. (Multi-View Point Ontology): a MVPO allows expressing in the same ontology multiple data descriptions according to different experts' viewpoints, and the consensus between them. A MVPO is a sextuple: $\langle C, VP, C_{VP}, P, R, I\rangle$ where:

- C: is the set of classes/concepts, defined by the primitives owl:Thing and owl:nothing
- VP: is the set of viewpoints represented in the ontology. They are defined by the primitive vp:viewpoint
- C_{VP}: is the set of concepts defined according to the experts' viewpoints. $C_{VP} = \bigcup_{i=1,\dots,n} C_{VP_i}$, where C_{VP_i} is the set of concepts defined according to a particular viewpoint VP_i.
- P: is the set of properties, $P = P_D \cup P_I$ where:
  P_D: the set of properties having a data for value.
  P_I: the set of properties having an individual for value.
- R: the set of relations between the concepts of different viewpoints.
- I: the set of individuals.

Figure 3 depicts a fragment of the MVPO for a "Manufacturing company" expressed using the MVP-OWL model. It contains three viewpoints "Accounting", "Commercial" and "Manufacturing" viewpoints.

4.1.2. VPO Extraction

In this phase, after the identification of the existing viewpoints, the VPOs are generated. They are conceived by a priori extracting concepts, properties and instances which are pertinent in the considered viewpoint from the MVPO, leading to its partitioning into a set of sub-ontologies. The formal definition of a VPO is given in the following.

Definition 4.2. (Viewpoint ontology): Each VPO is considered as a sub-set of the MVPO and is defined as a quadruplet VPO: $\langle C_{VP_i}, P_{VP_i}, R_{VP_i}, I_{VP_i}\rangle$ where:

- $C_{VP_i} \subseteq C_{VP}$: is the set of concepts defined according to the viewpoint VP_i.
- $P_{VP_i} \subseteq P$: is the set of properties defined according to the viewpoint VP_i.
- $R_{VP_i} \subseteq R$: is the set of relations between the concepts of the same viewpoint VP_i.

```
1 <rdf:RDF xmlns="http://www.owl-
ontologies.com/Manufacturing.owl#"
    xml:base="http://www.owl-ontologies.com/Manufacturing.owl"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:swrl="http://www.w3.org/2003/11/swrl#"
    xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
    xmlns:xsp="http://www.owl-
ontologies.com/2005/08/07/xsp.owl#"
    xmlns:p1="http://www.owl-ontologies.com/mvp-owl#"
    .....
  <owl:imports rdf:resource="http://www.owl-ontologies.com/mvp-
owl"/>
   </owl:Ontology>

2  <vp:Viewpoint rdf:ID="Commercial-VP"/>
   <vp:Viewpoint rdf:ID="Accounting-VP"/>
   <vp:Viewpoint rdf:ID="Manufacturing-VP"/>

3 <p1:ClassWithViewpoint rdf:ID="AccountingCustomer">
     <rdfs:subClassOf rdf:resource="&owl;Thing"/>
     <rdfs:subClassOf rdf:resource="#Customer"/>
     <p1:onClass rdf:resource="# Customer "/>
     <p1:onViewpoint rdf:resource="#Accounting-VP"/>
   </p1:ClassWithViewpoint>

4 <p1:ClassWithViewpoint rdf:ID="CommercialCustomer">
     <rdfs:subClassOf rdf:resource="&owl;Thing"/>
     <rdfs:subClassOf rdf:resource="# Customer "/>
     <p1:onClass rdf:resource="#Customer"/>
     <p1:onViewpoint rdf:resource="#Commercial-VP"/>
   </p1:ClassWithViewpoint>

5 <p1:ClassWithViewpoint rdf:ID="AccountingProduct">
     <rdfs:subClassOf rdf:resource="&owl;Thing"/>
     <rdfs:subClassOf rdf:resource="#Product"/>
     <p1:onClass rdf:resource="#Product"/>
     <p1:onViewpoint rdf:resource="#Accounting-VP"/>
   </p1:ClassWithViewpoint>

6 <p1:ClassWithViewpoint rdf:ID="CommercialProduct">
     <rdfs:subClassOf rdf:resource="&owl;Thing"/>
     <rdfs:subClassOf rdf:resource="#Product"/>
     <p1:onClass rdf:resource="#Product"/>
     <p1:onViewpoint rdf:resource="#Commercial-VP"/>
   </p1:ClassWithViewpoint>

7 <owl:DatatypeProperty rdf:ID="Designation">
     <rdfs:domain rdf:resource="#AccountingProduct"/>
     <rdfs:range rdf:resource="&xsd;double"/>
  </owl:DatatypeProperty>

8 <owl:DatatypeProperty rdf:ID="Average-margin">
     <rdfs:domain rdf:resource="#AccountingProduct"/>
     <rdfs:range rdf:resource="&xsd;double"/>
  </owl:DatatypeProperty>

9 <owl:DataProperty rdf:ID="Sales">
     <rdfs:domain rdf:resource="#CommercialProduct"/>
     <rdfs:range rdf:resource="&xsd;double"/>
  </owl:DatatypeProperty>

10 <owl:ObjectProperty rdf:ID="Sold-to">
     <rdfs:domain rdf:resource="#AccountingProduct"/>
     <owl:inverseOf rdf:resource="#Buy">
     <rdfs:range rdf:resource="#AccountingCustomer"/>
  </owl:ObjectProperty>

 11 <owl:DatatypeProperty rdf:ID="Address">
     <rdfs:domain rdf:resource="#Customer"/>
     <rdfs:range rdf:resource="&xsd;string"/>
  </owl:DatatypeProperty>
.......
</rdf:RDF>
```

Figure 3. Extract of the MVPO described in MVP-OWL language.

- $I_{VP_i} \subseteq I$: is the set of individuals defined according to the viewpoint VP_i

The set of VPOs is formally represented using an operational language (OWL). Figure 4 represents an extract of two VPOs: "Accounting" and "Manufacturing". The global mapping table (GMT) is generated to define the correspondences between the MVPO and the VPOs.

4.1.3. Correspondences Detection

The main objective of this step is to identify, through similarity measures, the relations between entities of the MVPO and those of the local ontologies. These relations will be used later for data sources classification. To this end, we exploit the ASCO1 algorithm proposed by [21]. This algorithm allows deducting structural similarities between the entities of different ontologies. We propose the ASCO-VP as an extension of the ASCO1 algorithm. In ASCO-VP, we add some measures to consider both semantic and structural similarities. The Initial Mapping Table (IMT), containing mapping information between the local ontologies and the MVPO, is generated as result.

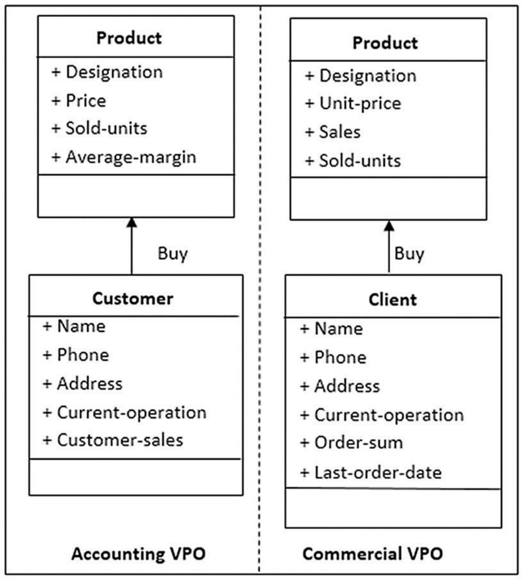

Figure 4. An extract of the Accounting VPO and Commercial VPO.

The ASCO-VP algorithm is given in Figure 5: all the extensions are represented in bold characters.

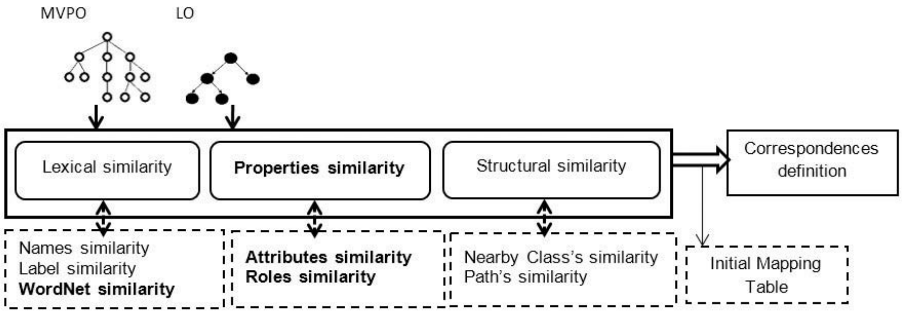

Figure 5. Correspondences detection algorithm $ASCO_{VP}$.

WordNet similarity. WordNet is a semantic network, where every node is a set of synonyms, called "synset" representing a real-world concept. The synsets are connected by links that describe the relation between different concepts. The idea is that two concepts are semantically close if their synsets are at least connected by a path. The WordNet similarity is given by S_{WN}:

$$S_{WN}\left(\mathrm{C}_1,\mathrm{C}_2\right)=\begin{cases}1 & \text{if } \mathrm{C}_1 \text{ and } \mathrm{C}_2 \text{ are synonyms or } \exists \text{ at least a terminological relation in WordNet between the synsets of } \mathrm{C}_1 \text{ and } \mathrm{C}_2\\ 0 & \text{in the opposite case}\end{cases}$$

Properties similarity. The properties similarity value is calculated from two proportions of similarity based on the internal structure of the concepts. These proportions concern the attributes distance and the roles, which are defined using the linguistic similarity LSim of ASCO1.

- *Attributes distance* (AD): The attributes distance AD measures the similarity between two attributes Pi and Pj of two concepts C_1 and C_2, respectively.

$$\mathrm{AD}\left(\mathrm{C}_1.\mathrm{P}_i,\mathrm{C}_2.\mathrm{P}_j\right)=\begin{cases}1 & \text{if } \mathrm{P}_i \text{ and } \mathrm{P}_j \text{ are identical or synonymic}\\ 0 & \text{in the opposite case}\end{cases}$$

- *Roles distance* (RD): The roles distance RD measures the similarity between two roles Ri and Rj of two concepts C_1 and C_2, respectively.

$$\mathrm{RD}\left(\mathrm{C}_1.\mathrm{R}_i,\mathrm{C}_2.\mathrm{R}_j\right)=\begin{cases}1 & \text{if } \mathrm{R}_i \text{ and } \mathrm{R}_j \text{ are identical or synonymic}\\ 0 & \text{in the opposite case}\end{cases}$$

- *Combination of properties similarity*: The final results of attributes and roles distances are organized in two matrices ADM (Attribute Distance Matrix) and RDM (Roles Distance Matrix), respectively. These matrices are used to calculate the similarity of two concepts according to their attributes and roles.

Definition 4.3. (ADM and RDM): Let ADM and RDM be two matrices that contain all the similarity distance proportions between the attributes/roles of two concepts C_1 and C_2. $\lambda(i)$ is the function that returns a similarity value (1 or 0) according to the sum of AD/RD results between each attribute/role of C_1 and m attributes/roles of C_2. The similarity average value for n attributes or roles between C_1 and C_2 is respectively returned by ||ADM|| or ||RDM||

$$\left\|\mathrm{ADM}\right\|=\frac{\sum_{i=1}^{n}\lambda(i)}{n},$$

where

$$\lambda(i)=\begin{cases}1 & \text{if } \sum_{i,j=1}^{m}\mathrm{AD}\left(\mathrm{C}_1.\mathrm{P}_i,\mathrm{C}_2.\mathrm{P}_j\right)>0\\ 0 & \text{in the opposite case}\end{cases}$$

$$\left\|\mathrm{RDM}\right\|=\frac{\sum_{i=1}^{n}\lambda(i)}{n},$$

where

$$\lambda(i) = \begin{cases} 1 & \text{if } \sum_{i,j=1}^{m} \mathrm{RD}\left(\mathrm{C}_1.\mathrm{R}_i, \mathrm{C}_2.\mathrm{R}_j\right) > 0 \\ 0 & \text{in the opposite case} \end{cases}$$

Definition 4.4. (Properties similarity): Let A and B be two classes of two ontologies O_1 and O_2, respectively. The properties similarity between A and B is a similarity function of signature $S_{Property}$: $O_1 \times O_2 \rightarrow [0, 1]$, which is calculated as the average of the RDM and the ADM:

$$\mathrm{S}_{\mathrm{Property}} = \mathrm{Average}\left(\|\mathrm{ADM}\| + \|\mathrm{RDM}\|\right)$$

Correspondences generation. The result of the $\mathrm{ASCO}_{\mathrm{VP}}$ algorithm is a set of correspondences between the entities of the shared ontology and every local ontology. These correspondences are represented by triplets (A, B, S_{Total}), where A and B are two entities of the same type in both ontologies, and S_{Total} is the similarity value between them. Thus, the correspondences are generated according to the last result of similarity S_{Total}, where S_{Total} represents the sum of $S_{Linguistic}$, $S_{Structural}$ and $S_{Property}$ values.

Definition 4.5. (Total similarity): Let A and B be two entities of the same type of two ontologies O_1 and O_2, respectively. The total similarity between these two entities is a similarity function S_{Total}: $O_1 \times O_2 \rightarrow [0, 1]$ such as:

$$\begin{aligned} \mathrm{S}_{\mathrm{Total}}(\mathrm{A},\mathrm{B}) = {} & \mathrm{S}_{\mathrm{Linguistic}}(\mathrm{A},\mathrm{B}) \times \mathrm{W}_{\mathrm{Linguistic}} \\ & + \mathrm{S}_{\mathrm{Property}}(\mathrm{A},\mathrm{B}) \times \mathrm{W}_{\mathrm{Property}} \\ & + \mathrm{S}_{\mathrm{Structural}}(\mathrm{A},\mathrm{B}) \times \mathrm{W}_{\mathrm{Structural}} \end{aligned}$$

where $W_{Linguistic} + W_{Property} + W_{Structural} = 1$.

Let $S_{Similarity} \succeq 0$ be a predefined threshold of similarity. Two entities A and B are considered similar, if and only if: $S_{Total}(A, B) \succeq S_{Similarity}$.

At the end of this process, a mapping table is generated (Initial Mapping Table IMT). It contains mapping information between local ontologies and the shared ontology including the defined viewpoints. This latter will be used in the definition of correspondences between VPO and local ontologies.

4.1.4 Data sources classification

This step aims to classify data sources into viewpoints categories; each category represents a particular viewpoint. To this end, the similarity results obtained in the previous step are used. We also propose a similarity measure called S_{VP} to calculate the similarity between the concepts of a local ontology and the concepts representing a particular viewpoint in the MVPO. The result of this stage is represented by basic similarity matrix (SM), which will be used later to build the IMT. The SM is a set of several sub-matrices (S-SM), each one corresponding to a particular viewpoint (see Figure 6).

MVPO

		Accounting VP			Commercial VP	
		Product	Professional-product	Refundable-product	Commercial-Product	....
	Merchandise	0.90	0.12	0.12	0.45	....
LO_i	Professional-Merchandise	0.12	0.92	0.12	0.12	....
	...					

S-CM

Figure 6. Basic similarity matrix.

Definition 4.6. (Viewpoints similarity): Let the SM be the matrix representing the similarity between the concepts of the two ontologies MVPO and LO_i, VP is the set of viewpoints represented in the global ontology, and S-SM is the sub matrix that contains the similarity values between the concepts of a local ontology and those of a particular viewpoint. The viewpoint's similarity S_{VP} between a local ontology and a particular viewpoint is a similarity function: S_{VP}: $LO_i \times VP_i \rightarrow [1, 0]$ such as:

$$\mathrm{S}_{\mathrm{VP}}\left(\mathrm{LO}_i, \mathrm{VP}_i\right) = \sum_{i,j=1}^{n} \frac{\mathrm{S}_{\mathrm{Total}}\left(\mathrm{C}_i, \mathrm{C}_j\right)}{\mathrm{Max}\left(\mathrm{C}_i, \mathrm{C}_j\right)}$$

where $S_{Total} \geq 0.5$.

Where Max (C_i, C_j): is the maximal number of compared concepts between the local ontology LO_i and a given viewpoint VP_j.

The algorithm that allows data sources classification is given in the following Algorithm 1.

Algorithm 1. Data sources classification.

Input VP :set of existing viewpoints
O :set of local ontologies
Begin
n := the number of local ontologies
m := the number of viewpoints
For $i = 1, ..., m$ do
Create Cat_j
$S_{Comparison} \succeq 0$ (: = 0.6)
For $i = 1, ..., n$ do
If $S_{VP} \succeq S_{Comparison}$ then
Add O_i to Cat_j
End for
Return Cat //Cat: is the set of viewpoint categories
End

At this stage, both the GMT and the IMT are used to build the local mapping table (LMT) which represents the possible links between the concepts of the LO_i in a same category and those in the correspondent VPO.

4.2. Discussion

Through the literature review of semantic data integration approaches, the advantages and the inconvenience of a posteriori semantic integration and a priori semantic integration are in duality. The a posteriori semantic integration guarantees the autonomy of the local ontologies, but it is performed in a manual or semi-automatic way and thus requires more efforts. In contrast, the a priori semantic integration provides a low autonomy degree for the data sources, but the integration process is completely automated.

Our approach brings an innovative solution that presents the advantages of the two approaches while avoiding their drawbacks (see Table 1). It provides two kinds of ontologies: (1) the VPO, which allows taking into account a limited and controlled consensus sufficient for the integration of a set of data sources related to a particular viewpoint, and (2) the MVPO, which offers the consensual knowledge shared by all the data sources; it ensures the semantic compatibility between the different viewpoint descriptions offered by the VPOs. With its specific characteristics, our integration approach offers the following advantages:

- it helps the exploitation of a unique and shareable ontology by accessing it according to several viewpoints and using it in a decentralized way,
- it ensures a completely automated integration process based on similarity measures used for the classification of the data sources,

Table 1. Ontology-based integration approaches comparison.

Characteristics \ Approaches	A priori approach	A posteriori approach	Viewpoint-based integration approach
Local ontologies	Created	Pre-existing	Pre-existing
Integration automaticity	Automatic	Manual/ Semiautomatic	Automatic
Sources autonomy	Less autonomy	High autonomy	High autonomy
System dynamicity	Easy integration of a new data sources	Difficult integration of a new data sources	Easy integration of a new data sources

- it offers a high level of autonomy for the participant sources as the addition of a new data source doesn't require either antecedent conditions or a priori articulations,
- the integration process is effortless since the integration of a new data source does not affect the global ontology; instead, this task is supported by an existing VPO,
- it ensures the coherence of the system by integrating different involved viewpoints in the field. The consistency of the system can be kept only by the integration of different viewpoints in a single representation (in our approach the MVPO construction).

Our approach is compared with existing semantic integration systems. Table 2 gives some comparison criteria. Note that, in OBSERVER [14], data sources are also grouped. A single local ontology is used for each group. Furthermore, the concept of global ontology is not supported and a unified access is not provided.

After presenting the proposed viewpoint-based integration approach, we outline in the following section the global architecture of the system.

5. Viewpoint-Based Mediation Architecture

The architecture supporting the viewpoint-based integration approach is represented as an arborescence of mediators. In this architecture, three levels are distinguished (see Figure 7): local level including ontology-based data sources, viewpoint-based integration level including a set of viewpoint mediators, and the global level containing the global mediator and the user's interface. In addition, the viewpoint-based architecture includes an integration module.

- **Integration module:** The integration module is responsible for the classification of the different data sources for their integration. During the integration process, the MVPO is first constructed using the MVP-OWL language and the ontology editor Protégé. The participant local ontologies are then imported into the system using the charging module. This latter is implemented using the Jena library that contains Java classes allowing the development of Semantic Web applications. The different parts of the $ASCO_{VP}$ algorithm are implemented using Java language and its evaluation is achieved by a set of pertinence measures, such as the precision, recall and F-measure. The VPOs extraction is performed by a specific module that is based on a set of algorithms exploiting the basics of the MVP-OWL language.
- **Global mediator:** The goal of this mediator is the management and the processing of the end users queries. When a user query is expressed via the user interface, it is sent to the global mediator which uses the GMT to rewrite and optimize the query execution plan (QEP) using the rewriting and optimization modules successively. Afterward, the sub-queries are transferred to the concerned viewpoint mediators.
- **The viewpoint mediator:** This mediator uses the evaluation and execution modu-

Table 2. Ontology-based integration systems comparison.

Systems	Data type	Ontology architecture	Semantic integration	System architecture
OBSERVER	HTML, Data bases, Files	Multiple	A posteriori	Adapter, ontology server
KRAFT	Data bases	Hybrid	A posteriori	Agent-based
SIMS	Data bases	Unique	A priori	Ontology-based mediation
PICSEL	Data bases, XML	Hybrid	A priori	Ontology-based mediation
Proposed system	OBDB	Hybrid	Hybrid	Ontology-based mediation

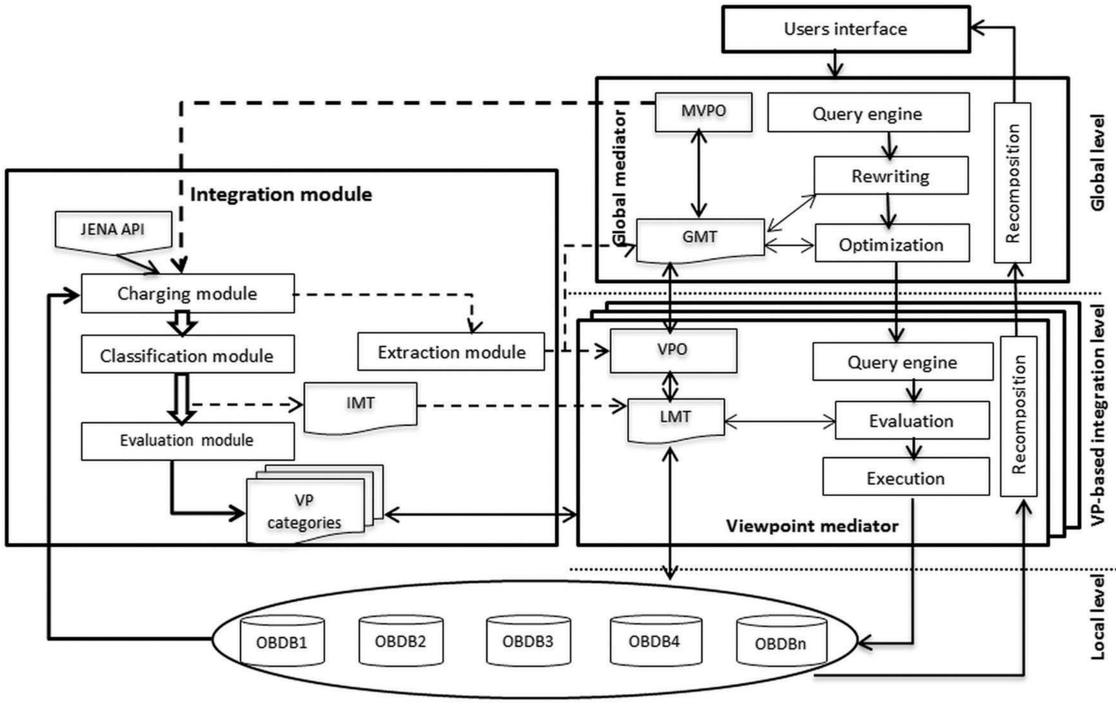

Figure 7. Viewpoint-based mediation architecture.

les to rewrite and execute optimized sub -queries over the pertinent data sources. Finally, the recomposition module is used at the viewpoint-based level as well as the global level to reformulate the answers according to the QEP.

Between the different levels of the architecture, mappings (IMT, LMT and GMT) that represent possible links between the different ontologies are created and managed. In order to define a generic process for SPARQL query treatment, the mappings are translated into the SPARQL algebra [28]. An example of both global mappings and local mappings represented in the form of SPARQL assertions is given in Table 3.

6. Evaluation and Validation

In this section, we consider both qualitative and quantitative evaluation of our approach. For the former, we present a simple case study applying the query processing method. For the latter, we give some experiments, which will show that the proposed approach may reduce the execution time of the users' queries.

Table 3. Correspondences assertions.

SGMT assertions	LMT assertions
/*the namespace prefixes "mvpo:", "vpoa:" and "vpoc:" refer to the vocabularies of the MVPO, Accounting VPO and Commercial VPO, respectively*/ **Class' assertions** 1. (?p rdf :type mvpo:AccountingProduct) ← (?p rdf :type vpoa: Product) 2. (?p rdf :type mvpo:CommercialProduct) ← (?p rdf :type vpoc: Product) **Property' assertions:** 1. (?pmvpo:Average-margin ? Average-margin) ← (?p vpoa : Average-margin ? Average-margin) 2. (?p mvpo:Sales ?Sales) ← (?pvpoc: Sales ? Sales)	/*the namespace prefixes "vpoc:", "s1:" and "s2:" refer to the vocabularies of the Manufacturing VPO and data sources S1 and S2 respectively*/ **Class' assertions:** 1. (?p rdf :type vpoa: Product) ← (?p rdf :type S1:Product) 2. (?prdf :type vpoa: Product) ← (?prdf :type S2:Produit) **Property' assertions:** 1. (?pvpoa:Sales ?Sales) ← (?p S2:Chiffre-affaire ?Chiffre-affaire)

6.1. Qualitative Validation

The query processing method proceeds in four steps: query rewriting, optimization, evaluation and execution (see Figure 8).

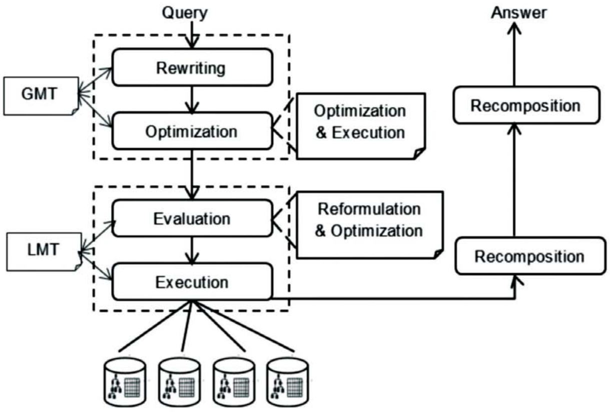

Figure 8. Query treatment process.

Let's consider a query Q expressed in SPARQL over the MVPO given in Figure 4. The query Q searches the designation, the average margin and the sales of products sold to customers who live in Constantine (see Figure 9).

```
PREFIX mvpo:<http://owl-ontologies.com/
      manufacturing_company.owl#>
PREFIX rdf:<…>
SELECT ?Designation ?Average-margin ?Sales
WHERE{
?p mvpo:Designation ?Designation.
?p rdf:type mvpo:Product.
?p mvpo:Average-margin ?Average-margin.
?p mvpo:Sales ?Sales.
?p mvpo:sold-to ?Customer.
?Customer mvpo:Address ?Address
FILTER regex (?Address, "Constantine")
}
```

Figure 9. Query Q expressed in SPARQL syntax.

6.1.1. Query Rewriting

The query Q is firstly rewritten into a query tree describing the involved entities, properties and eventually conditions. Then, using the GMT containing the correspondences assertions presented in Table 4, the query tree is expanded to a combination of sub-queries, where the MVPO concepts are replaced by the VPOs correspondent concepts (see Figure 10). The property *Average-margin* searched by the user characterizes the class *Product* in the "Accounting VPO", and the property sales characterizes the class *Product* in the "Commercial VPO".

6.1.2. Optimization

In this step, the optimal query execution plan is generated using the cost model. In the considered example, the execution plan of Q contains two sub-queries SQ_1 and SQ_2, which can be executed in parallel on the "Accounting VPO" and the "Commercial VPO". SQ_1 extracts designation and average margin of pro-

Table 4. User query and profile.

	User 1	**User 2**
Query	**SELECT** ?Designation ?Average-margin ?Sales **WHERE** {?p mvpo: Designation ?Designation. ?p rdf: type mvpo:Product. ?p mvpo:Average-margin ?Average-margin. ?p mvpo:Sales ?Sales. ?p mvpo:supplied-by ?Supplier. ?Supplier mvpo:Address ?Address. **FILTER regex** (?Address, "Constantine") }	**SELECT** ?Designation ?Price **WHERE** {?p mvpo:Designation ?Designation. ?p rdf: type mvpo:Product. ?p mvpo:Price ?Price. ?p mvpo:supplied-by ?Supplier. ?Supplier mvpo:Address ?Address. **FILTER regex** (?Address, "Constantine") }
Profile	(Product; 0.50, Supplier; 0.50, AccountingSupplier; 0.80, CommercialSupplier; 0.80,AcountingPorduct; 0.80, CommercialProduct; 0.80)	(Product; 0.80, Supplier; 0.80)

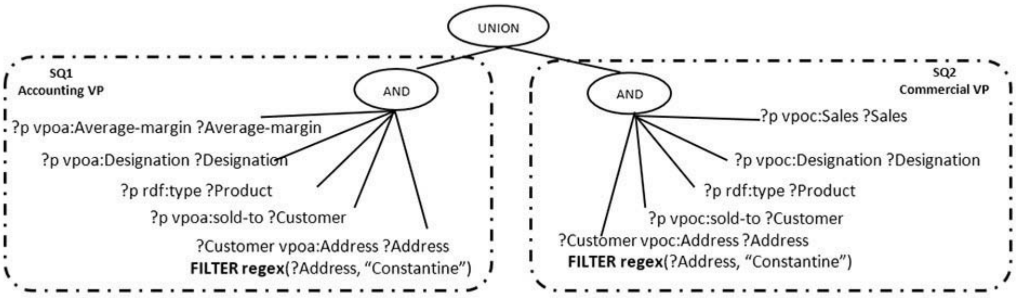

Figure 10. The query' expanded tree representation.

```
SQ1:
PREFIX vpoa :<http://owl-ontologies.com/Accounting_vpo.owl#>
PREFIX rdf:<…>
SELECT ?Designation ?Average-margin
WHERE{
?p vpoa:Designation ?Designation.
?p rdf:type vpoa:Product.
?p vpoa:Average-margin ?Average-margin.
?p vpoa:sold-to ?Customer.
?Customer vpoa:Address ?Address.
FILTER regex (?Address, "Constantine")
}
-----------------------------------------------------------------
SQ2:
PREFIX vpoc :<http://owl-ontologies.com/Commercial_vpo.owl#>
PREFIX rdf:<…>
SELECT ?Designation ?Sales
WHERE{
?p vpoc:Designation ?Designation.
?p rdf:type vpoc:Product.
?p vpoc:Sales ?Sales.
?p vpoc:sold-to ?Client.
? Client vpoc:Address ?Address
FILTER regex (?Address, "Constantine")
}
```

Figure 11. SPARQL sub-queries representing the expanded tree of query Q.

ducts sold to customers from Constantine, and SQ2 extracts designation and sales of products sold to customers from Constantine. The generated sub-queries are represented in Figure 11, and the execution plan is the union of both sub-queries ($SQ_1 \cup SQ_2$).

6.1.3. Evaluation

In this step, the relevant queries over the data sources are obtained. However, using the LMT, sub-queries sent to the target viewpoint mediators are rewritten in terms of the local ontologies of the same viewpoint. A set of optimal QEPs representing the local QEPs is thus generated. For example, Figure 12 presents the decomposition of SQ_2 using LMT in Table 3 defined between the "Commercial VPO" and the local ontologies.

6.1.4. Execution and Recomposition

After the execution of the sub-queries over the data sources, results are sent back to the combination module at the viewpoint level. This one builds the answers received form data sources according to the defined QEP. These intermediary results are then sent to the global mediator to be combined into a final answer for the end user (see Figure 13).

By the decomposition of the query processing and the selection of relevant data sources ac-

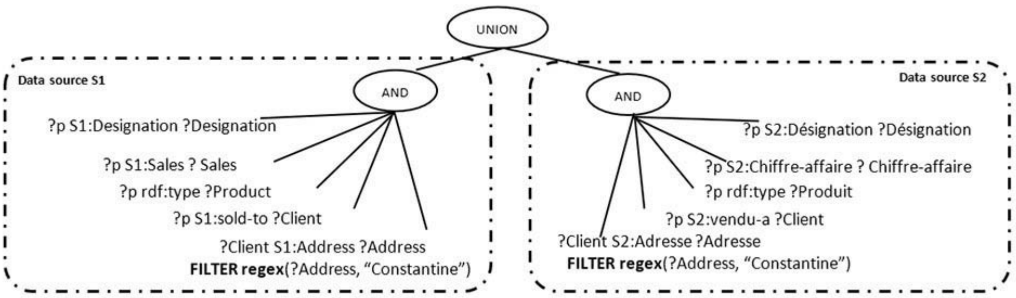

Figure 12. Tree representation of SQ_2.

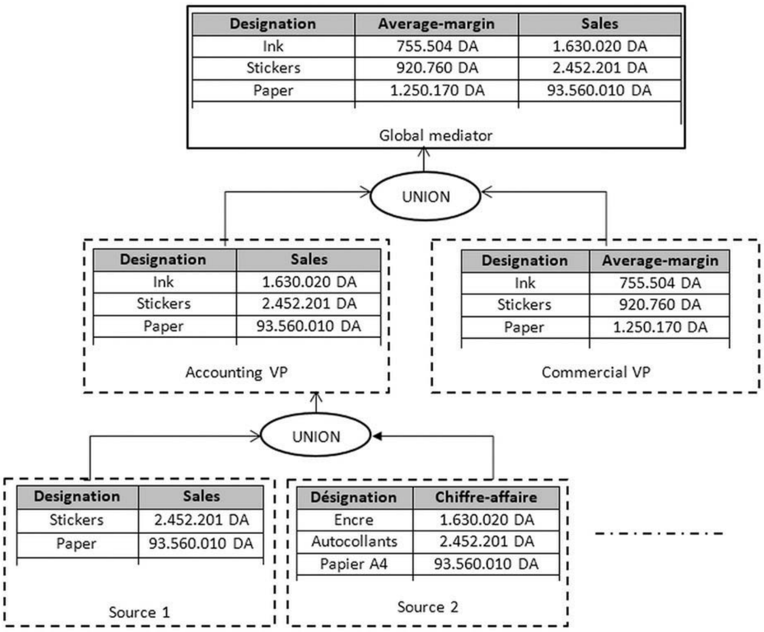

Figure 13. Example of answer composition.

cording to particular viewpoints, the response time is improved. However, mediators at both global level and viewpoint level deal only with a limited number of data sources.

6.2. Quantitative validation

The major purpose of mediation systems is to offer a better access to a multitude of data sources. However, the traditional integration systems are not able to give answers according to the users' viewpoints and choices. For this, we introduce the notion of profiles to measure the efficacy of the system. Unlike existing approaches, the proposed mediation system takes into account the users' preferences to select the most relevant data sources (see Figure 14).

The most common representation of a profile is a set of keywords. With each keyword is associated a weight giving a digital representation of the user's interests. For facility purpose, we consider two profiles at the global level as follows:

- *MVPO profile:* represented in a vertical way and divided in categories; each category describes a particular viewpoint.

$$\text{MVPO: } VP_1\ (C_1, P_1; C_2, P_2; \ldots; C_n, P_n)$$
$$VP_2\ (C_1, P_1; C_2, P_2; \ldots; C_n, P_n)$$

where P = frequency C_i(MVPO) * log (N/N_i) and (N/N_i) = (VP number in (MVPO)/ number of VP that contains C_i)

- *User profile:* is generally built by handling the history of a user's searches in the system. In the mediation system, the user's history is represented by the queries sent over the system. The user's history is used to define his viewpoint.

To illustrate the difference between the query processing with and without viewpoint, we consider in Table 4 the description of two users' preferences: the first one specifies his proper viewpoint in his query, and the second one does not specify any viewpoint.

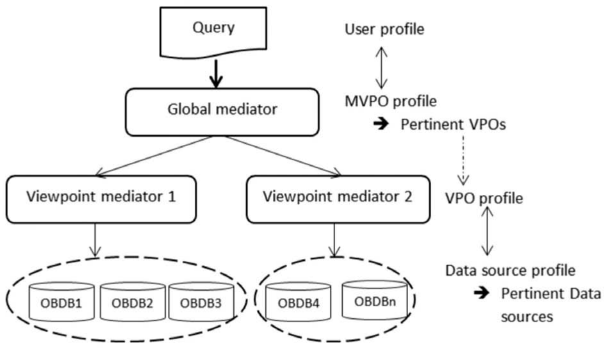

Figure 14. Query processing using viewpoints and profiles.

Let us consider a part of the MVPO where all the concepts have the same weight:

MVPO:

Commercial_VP (CommercialProduct ; 0.50, CommercialSupplier ; 0.50, CommercialClient ; 0.50)

Accounting_VP (AccountingProduct ; 0.50, AccountingSupplier ; 0.50, AccountingCustomer ; 0.50)

Manufacturing_VP (Manufacturing-Product; 0.50, ManufacturingSupplier; 0.50)

To read easily these statistics, we homogenize them in the same table (see Table 5). Absent concepts are added in some profiles, while affecting a weight of zero.

To measure the similarity between the various profiles, we use the inner product. Thus, we define a score of similarity to rank the VPOs from the most relevant to the less relevant according to the user profile. We assume that the VPO with the highest score is more relevant to represent the users' preference and choice.

Definition 6.1. (Inner product): The inner product is algebraically defined between two vectors A $[a_1, a_2, ..., a_n]$ and B $[b_1, b_2, \dots, b_n]$ with the following formula:

$$\mathrm{A.B} = \sum_{i=1}^{n} a_i.b_i = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$

Let's calculate, in Table 6, the similarity scores between the users and all the viewpoints vectors of the previous example.

We observe that the most relevant viewpoints for User1 are the "Accounting VP" and the "Commercial VP" and his query will be sent directly to the specific mediators. Lacking information about User2' viewpoints, his query will be sent to all viewpoint mediators. By specifying the viewpoints notion, the number of data sources is consequently limited and the quality of answers is better. Furthermore, the query processing time is reduced.

7. Conclusion

In this paper, we have presented a novel ontology-based integration approach using the paradigm of viewpoints. The main objective is to benefit from viewpoints to overcome the drawbacks of the "a priori" and the "a posteriori" semantic integration approaches. The proposed

Table 5. Homogenizing profile values.

	User 1	User 2	Accounting_VP	Commercial_VP	Manufacturing_VP
Product	0.50	0.80	0	0	0
Supplier	0.50	0.80	0	0	0
Accounting-Supplier	0.80	0	0.50	0	0
Commercial-Supplier	0.80	0	0	0.50	0
Manufacturing-Supplier	0	0	0	0	0.50
Accounting-Product	0.8	0	0.50	0	0
Commercial-Product	0.8	0	0	0.50	0
Manufacturing-Product	0	0	0	0	0.50
Commercial-Client	0	0	0	0.50	0
Accounting-Customer	0	0	0.50	0	0

Table 6. Similarity scores between users and viewpoints profiles.

Similarity score	Accounting_VP	Commercial_VP	Manufacturing_VP
User 1	0.8	0.8	0
User 2	0	0	0

Table 7. MVP-OWL: new primitives' description.

Class name	rdfs:subClassOf	Description
vp:Viewpoint	rdfs:class	Define classes of type: viewpoint
vp:ClassWithViewpoint	owl:Class	To define classes with viewpoints
Property name	**Type**	**Description**
vp:belongToViewpoint	owl:objectProperty	Define the individuals' viewpoint
vp:characterisedBy	owl:objectProperty	Define a viewpoint' properties and criterion
vp:onClass	owl:objectProperty	To define classes with viewpoints
vp:onViewpoint	owl:objectProperty	

approach is based on two kinds of mediators, each governed by a kind of ontologies. First, the viewpoint mediators use viewpoint ontologies (VPO) to represent a specific viewpoint of data sources and fulfill the mediation between them. The second kind represented by the global mediator uses a multi-viewpoint ontology (MVPO) as a shared ontology to represent the different viewpoints that exist in the considered domain. Additionally, the integration process is completely automatic.

We currently intend to improve the query optimization process by using the notion of profiles at both levels of the integration system. Future works will especially be concerned with the development of a new data source presenting a new viewpoint that is not considered in the global ontology and with proving that the pile of mediators in the mediation architecture have no impact on the system performances.

Annex

In our research, we adopt the model proposed in [21] for the representation of the MVPO. The MVP-OWL model contains an extension of the OWL language. In this model, new primitives are added to the OWL language to integrate the viewpoint notion in ontologies (see Table 7).

References

[1] G. Wiederhold, "Mediators in the Architecture of Future Information Systems", *IEEE Computer*, vol. 25, no. 3, pp. 38–49, 1992.

[2] H. Garcia-Molina *et al.*, "The TSIMMIS Approach to Mediation: Data Models and Languages", *J. of Intell. Inform. Syst.*, vol. 8, no. 2, pp. 117–132, 1997. http://dx.doi.org/10.1023/A:1008683107812

[3] H. Wache *et al.*, "Ontology-Based Information Integration: A Survey of Existing Approaches", in *Int. Joint Conf. on Artificial Intelligence*, Seattle, Washington, USA, 2001.

[4] A. Buccella *et al.*, "Ontology-Based Data Integration Methods: A Framework for Comparison", *Revista Colombiana de Computacion*, vol. 6, no. 1, 2005.

[5] D. N. Xuan, "Autonomous Data Sources Integration by a Priori Articulation of Ontologies: Application to the Catalogs of Industrial Components", PhD dissertation, Doctoral School of Sciences for Engineer, Poitiers University, France, 2006. (in French)

[6] N. Martin *et al.*, "A methodology and Architecture Embedding Quality Assessment in Data Integration", *J. of Data and Inform. Quality (JDIQ)*, vol. 4, no. 4, pp. 17, 2014. http://dx.doi.org/10.1145/2567663

[7] M. Mohania and M. Bhide, "New Trends in Information Integration", in 2*nd Int. Conf. on Ubiquitous Information Management and Communication*, Suwon, Korea, pp. 74–81, 2008. http://dx.doi.org/10.1145/1352793.1352810

[8] U. Leser *et al.*, "Data Integration in the Life Sciences", in *3rd Int. Workshop on Data Integration for the Life Sciences*, Hinxton, UK., Springer-Verlag Berlin Heidelberg, 2006. http://dx.doi.org/10.1007/11799511

[9] M. D. C. M. Batista and A. C. Salgado, "Information Quality Measurement in Data Integration Schemas", in *5th Int. Workshop on Quality in Databases*, Vienna, Austria, pp. 61–72, 2007.

[10] F. Wadjinny *et al.*, "Integrating Disparate Digital Libraries Using the WASSIT Mediation Framework", in Digital Libraries – Methods and Applications, Dr. Kuo Hung Huang (Ed.), InTech. http://dx.doi.org/10.5772/15823

[11] M. Hacid *et al.*, "Data Sources Integration", *J. of I3: Information – Interaction – Intelligence*, vol. 4, no. 2, pp. 4, 2004. (in French)

[12] D. Beneventano and S. Bergamaschi, "The MOMIS Methodology for Integrating Heterogeneous Data Sources", in *26th Int. Conf. on Very Large Data Bases*, Cairo, Egypt, pp. 611–614, 2004. http://dx.doi.org/10.1007/978-1-4020-8157-6_3

[13] M. C. Rousset *et al.*, "Construction of Mediators to Integrate Multiple and Heterogeneous Information Sources: The PICSEL project", *J. of I3: Information-Interaction-Intelligence*, vol. 2, no. 1, 2002. (in French)

[14] E. Mena *et al.*, "OBSERVER: An Approach for Query Processing in Global Information Systems Based on Interoperation Across Pre-existing Ontologies", *Int. J. of Distributed and Parallel Databases*, vol. 8, no. 2, pp. 223–271, 2000. http://dx.doi.org/10.1023/A:1008741824956

[15] A. Preece *et al.*, "KRAFT: An Agent Architecture for Knowledge Fusion", *Int. J. of Cooperative Inform. Syst.*, vol. 10, no. 01n02, pp. 171–195, 2001. http://dx.doi.org/10.1142/S0218843001000308

[16] Y. Arens *et al.*, "Retrieving and Integrating Data From Multiple Information Sources", *Int. J. of Cooperative Inform. Syst.*, vol. 2, no. 2, pp. 127–158, 1993. http://dx.doi.org/10.1142/S0218215793000071

[17] L. Bellatreche *et al.*, "An a priori Approach for Automatic Integration of Heterogeneous and Autonomous Databases", in *15th Int. Conf. on Database and Expert Systems Applications*, Zaragoza, Spain, pp. 475–485, 2004. http://dx.doi.org/10.1007/978-3-540-30075-5_46

[18] M. Minsky, "A Framework for Representing Knowledge", in the Psychology of Computer Vision. Massachusetts Institute of Technology Cambridge, MA, USA, Tech. Rep. 1974.

[19] C. Bachtarzi and F. Benchikha, "View-OD: a View Model for Ontology-Based Databases", *Int. J. of Intell. Inform. and Database Syst.*, vol. 7, no. 4, pp. 295–323, 2013. http://dx.doi.org/10.1504/IJIIDS.2013.056309

[20] F. Benchikha and M. Boufaïda, "The Viewpoint Mechanism for Object-Oriented Databases Modelling, Distribution and Evolution", *J. of Computing and Inform. Technology*, vol. 15, no. 2, pp. 95–110, 2007. http://dx.doi.org/10.2498/cit.1000692

[21] L. Bach Thânh, "Construction of a Multi-Viewpoint Semantic Web", PhD dissertation, The School of Mines ParisTech at Sophia Antipolis, Paris, France. 2006. (in French)

[22] M. Ribiere and R. Dieng-Kuntz, "A Viewpoint Model for Cooperative Building of an Ontology", in *10th Int. Conf. on Conceptual Structures, Borovets*, Bulgaria, pp. 220–234, 2002. http://dx.doi.org/10.1007/3-540-45483-7_17

[23] D. Benslimane *et al.*, "Contextual Ontologies", in *4th International Conf. Advances in Information Systems*, Izmir, Turkey, pp. 168–176, 2006. http://dx.doi.org/10.1007/11890393_18

[24] G. Falquet and C-L. Mottaz, "A Model for the Collaborative Design of Multi Point of View Terminological Knowledge Bases", in *Knowledge Management and Organizational Memories*, Springer US, pp. 193–202, 2002. http://dx.doi.org/10.1007/978-1-4615-0947-9_17

[25] H. Stuckenschmidt, "Toward Multi-Viewpoint Reasoning with OWL Ontologies", in *3rd European Semantic Web Conf.*, Budva, Montenegro, pp. 259–272, 2006. http://dx.doi.org/10.1007/11762256_21

[26] P. Bouquet *et al.*, "Contextualizing Ontologies", *J. of Web Semantics: Sci., Services and Agents on the World Wide Web*, vol. 1, no. 4, pp. 325–343, 2004. http://dx.doi.org/10.1016/j.websem.2004.07.001

[27] B. Bouchra *et al.*, "Data Sources Integration Using Viewpoint-Based Approach", in *1st Int. Conf. on Intelligent Information Processing, Security and Advanced Communication*, Batna, Algeria, pp. 25, 2015. http://dx.doi.org/10.1145/2816839.2816877

[28] M. Arenas *et al.*, "On the Semantics of SPARQL", *Semantic Web Information Management: A Model-Based Perspective*, Springer Berlin Heidelberg, pp. 281–307, 2010. http://dx.doi.org/10.1007/978-3-642-04329-1_13

Received: November 2015
Revised: December 2016
Accepted: December 2016

Contact addresses:
Bouchra Boulkroun
LIRE Laboratory
Software Technologies and Information Systems Department
Constantine 2 University Abdelhamid Mehri
Route d'Ain-El Bey, 25000
Constantine, Algeria
e-mail: bouchra.boulkroun@univ-constantine2.dz

Fouzia Benchikha
LIRE Laboratory
Software Technologies and Information Systems Department
Constantine 2 University Abdelhamid Mehri
Route d'Ain-El Bey, 25000
Constantine, Algeria
e-mail: f_benchikha@yahoo.fr

Chahinez Bachtarzi
LIRE Laboratory
Software Technologies and Information Systems Department
Constantine 2 University Abdelhamid Mehri
Route d'Ain-El Bey, 25000
Constantine, Algeria
e-mail: bachtarzi_chahinez@live.fr

BOUCHRA BOULKROUN is pursuing her PhD in Computer Science at Constantine 2 University Abdelhamid Mehri, Algeria. Her research interests include distributed information systems, ontology design, and multi -viewpoint representations. She focuses her doctoral research on applying the viewpoint paradigm to integrate distributed ontology-based data sources using an original ontology-based integration approach..

FOUZIA BENCHIKHA is a Professor in Department of Software Technologies and Information Systems at Constantine 2 University Abdelhamid Mehri, Algeria. Her current research activities are conducted at the LIRE Laboratory from the same university. Her research interests are data modeling and integration, information systems interoperability, multi-viewpoint representations, and semantic web technology.

CHAHINEZ BACHTARZI is a Doctor in Department of Software Technologies and Information Systems the Constantine 2 University Abdelhamid Mehri, Algeria. She received her PhD from Constantine 2 University Abdelhamid Mehri, Algeria in 2014. Her current research activities are conducted at the LIRE Laboratory from the same university. Her research interests are ontology design, database technology, and multi-viewpoint ontology.