Journal of Web Engineering, Vol. 15, No.3&4 (2016) 339-360

MULTIMEDIA DATA RETRIEVING BASED ON SOA ARCHITECTURE

SID AHMED DJALLAL MIDOUNI [1,2]
[1]*Département d'informatique, Université Abou Bekr Belkaid - Tlemcen*
B.P.230- Tlemcen 13000, Algérie
d_midouni@mail.univ-tlemcen.dz
sid-ahmed-djallal.midouni@insa-lyon.fr

YOUSSEF AMGHAR [2]
[2]*Université de Lyon. CNRS, INSA-Lyon. LIRIS. UMR5205, F-69621. France*
youssef.amghar@insa-lyon.fr

AZEDDINE CHIKH [1,3]
[3]*Information Systems Department*
College of Computer Science, King Saud University, Riyadh, Saudi Arabia
az_chikh@ksu.edu.sa

Received July 8, 2015
Revised January 25, 2016

In our past research we have already defined a full service approach to compose MaaS services for multimedia data retrieving. This approach is based on a four phases process: description; filtering; clustering; and restitution. In this work, we are especially interested in the description and filtering phases of this process. Our contribution is two-fold. First, we propose to extend for the MaaS description the W3C recommendation on semantics for web services (SAWSDL). To do so, we use two types of ontologies: a Domain Ontology encompassing concepts that define semantics of the related business domain and a Multimedia Ontology encompassing concepts that define a set of annotation properties of the multimedia content. Second, we show how this extension allows addressing the problem of matching between MaaS services and user needs.

Key words: semantic web services, information retrieval, service description, SAWSDL, service matching.
Communicated by: B. White & O. Pastor

1 Introduction

Multimedia systems have become vital on the web. Thus, we find it quite natural, now, to integrate video, sound or music in various applications: educational, scientific, technical, medical, administrative, commercial, cultural and artistic.

Accessing to multimedia data in distributed systems poses new challenges due to many system parameters: volume, diversity of interfaces, representation format, location, etc. In addition, the growing needs of users and applications to incorporate semantics in the information retrieval pose new issues. Current solutions, which are based on integration or interoperability, are often unsatisfactory. They only partially meet these needs and the requirements of globally controlling the multimedia content flows.

In our past research, we have already proposed a full service approach [26]. It overcomes some missing issues in accessing and searching multimedia data in the context of distributed and heterogeneous data systems. A new pattern of services is defined, that is called multimedia web services (MaaS: Multimedia as a Service). A MaaS service is a specific data web service that accesses to multimedia data. The proposed approach is based on SOA architecture. It gives a relevant answer to the user, based on a four phases process: description; filtering; clustering; and restitution. This research, which falls within this dynamic, is a continuation of our previous contribution. We explain in detail how MaaS services are described and matched to the user needs.

1.1 Motivating Example

As a running scenario, let us consider the following scenario from the medical domain. Assume the student Bob would like to study the different diagnostic methods of lung cancer disease. He wants to have as a result all videos provided by "cancercenter" hospital in "mpeg" format. Assume he has at his disposal the MaaS services in Table 1. S_1 and S_2 return diagnostic methods of lung cancer, knowing that fiberoptic bronchososcope is an imaging examination that can diagnose a patient case. Both services have a similar semantics (or functionalities) but return different types of files and have a different providers. S_3 returns videos about treatments of lung cancer, it hasn't a same semantic (or functionality) as S_1 and S_2 and it is provided by a "Medicine University". Bob can use the MaaS services of table 1 to meet his needs. In the first step, he compared each service functionality with his request functionality. He eliminated S_3 and retained S_1 and S_2 services because their semantics match the semantic of his needs. In the second step, he retained only S_1 service because its multimedia properties meet the multimedia properties of his needs.

Services	**Service Functionality**	**Provider**	**Format**
S_1	Returns videos showing the progress of a fiberoptic bronchososcope of a patient suffering from lung cancer	Cancercenter Hospital	mpeg
S_2	Returns images visualizing the different examinations for lung cancer	Imagines Laboratory	jpeg
S_3	Returns courses videos on treatments of lung cancer	Medicine University	mp4

Table 1 MaaS services used in the motivating example.

1.2 Challenges

The user (i.e. the student in the example) in this task is confronted with the two following challenges:

i. **Describing the semantics of MaaS services:** Bob needs to explore the MaaS services repository and understand the semantics of each individual service in order to identify the services that may contribute to the resolution of his request. Many MaaS services may have the same semantic in terms of their functionalities, but completely different multimedia properties. For example, the services S_1 and S_2 have the same semantic about their functionalities (having the same input and output) but they are different about their multimedia properties (type, provider and format). The challenge here is how to describe the MaaS services to enhance expressiveness and precision of the solution. The semantic description task must take into account all specificities and features of MaaS services. This description will be about two aspects: the business domain aspects and multimedia aspects of MaaS services.

ii. **Selecting relevant MaaS services:** Let us assume now that Bob is able to describe and understand the semantics of available MaaS services. The next step would be to identify relevant MaaS services to meet his needs, i.e. how to find MaaS services satisfying his query in the best way. Such an evaluation mechanism is called *matching*. The challenge here is to show how MaaS services, with an enriched description, are matched to meet user needs. Bob should realize that only the service S_1 satisfies his needs in terms of business domain aspects and multimedia aspects at the same time.

1.3 Contributions

The aforementioned challenges motivate the need for an approach whose purpose is to find a solution that best meets the user's needs. We summarize bellow our major contributions in this paper:

1) Semantic description of MaaS services

Defining and describing the semantics of MaaS services is a key requirement for our approach. In literature, many approaches of semantic web services description have been defined such as OWL-S (OWL for Services) [23], WSMO (Web Services Modeling Ontology) [8] and SAWSDL (Semantic Annotations for WSDL and XML Schema) [20]. These approaches use ontologies to provide semantic description of web services. We are interested in extending these approaches for the description of MaaS services. Our extension, which is based especially on SAWSDL standard, uses two types of ontologies: (1) A Domain Ontology encompassing concepts that define semantics of the related business domain (e.g. health, education, tourism ...). (2) A Multimedia Ontology encompassing concepts, that define a set of annotation properties of the multimedia content (format, location, creation information, etc.), and relations between these concepts. Using this new manner of MaaS description, we enhanced the precision of the proposed solution. The services S_1 and S_2 will be different, they have the same domain description but dissimilar multimedia description.

2) Filtering and matching MaaS services with user query

In order to address the problem of matching between MaaS services and user needs, we propose a new matching mechanism according to the MaaS description method. Different approaches for matching semantic web services have been developed during the last years. The aim of these approaches is to identify a degree of similarity between a query and web services. Our matching mechanism is twofold: a domain matching and a multimedia matching. The second matching is performed if and only if the first one has succeeded. For the domain matching, we propose an algorithm based on the calculation of similarity degree between (1) semantic concepts annotating MaaS services; and (2) those annotating the query. For the multimedia matching, we propose an algorithm able to compare multimedia description of MaaS services and query. The multimedia description is defined as a SPARQL query over multimedia ontology.

The remaining of this paper is organized as follows: Section 2 presents the related works on semantic web services description and matching. Section 3 describes briefly the most important multimedia ontologies. Then, we describe in section 4 our new approach for the MaaS services retrieval problem. Section 5 presents the experiments conducted to assess and validate the proposed approach. The last section is devoted to the conclusion and future works.

2 State of the art

In this section, we present the most related works to our contribution and review the approaches for describing and matching web services. We first present a review of existing research on web service description.

2.1 Web service description

WSDL is an XML format for describing the web services and only describes the syntactic interface of web services. To overcome the lack of semantics in WSDL, many languages and approaches have been developed with the aim to describe semantic web services by using ontologies and other semantic models. We distinguish two main classes of these approaches. Approaches of the first class are based on using of high-level ontology such as OWL-S [23] and WSMO [8]. Approaches of the second class are based on adding annotations, such as USDL [30] and SAWSDL [20]. The figure 1 presents our classification of web service description approaches.

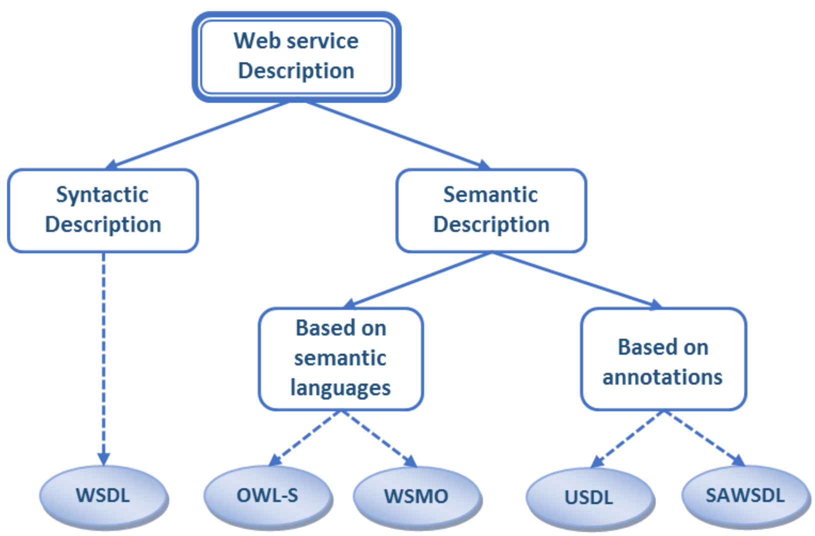


Figure 1 Web service description approaches.

As the chronologically first approach for semantic web service, OWL-S defines an upper ontology for semantically annotating Web services. OWL-S is an OWL ontology that includes three primary subontologies: service profile; process model; and grounding. The service profile is used to describe *what the service does*; the process model is used to describe *how the service is used*; and the grounding is used to describe *how to interact with the service*. The service profile and process model are thought of as abstract characterizations of a service, whereas the grounding makes it possible to interact with a service by providing the necessary concrete details related to message format, transport protocol, and so on.

WSMO is a conceptual model for relevant aspects related to semantic web services. It provides an ontology based framework, which supports the deployment and interoperability of semantic web services.

The WSMO has four main components: Ontologies; Goals; Web Services; and Mediators. *Ontologies* are described in WSMO at a meta-level. A meta-ontology provides the terminology used by other WSMO elements. *Goals* are defined in WSMO as the objectives that a client may have when consulting a Web service. *Web Services* provide a semantic description of services on the web, including their functional and non-functional properties, as well as other aspects relevant to their interoperation. *Mediators* in WSMO are special elements used to link heterogeneous components involved in the modeling of a Web service. They define the necessary mappings, transformations and reductions between linked elements. [33]

SAWSDL is a 2007 published technical recommendation of W3C in the context of semantic web framework. It is based primarily on the earlier work on WSDL-S [1]. Its main advantage remains in its extensibility and compatibility with the WSDL standard. SAWSDL provides a standard means by which WSDL documents can be related to semantic descriptions, such as those provided by OWL-S and WSMO.

SAWSDL defines how to add semantic annotations to various parts of a WSDL document, it defines extension attributes that can be applied to elements in both WSDL and XML Schema in order to annotate WSDL interfaces, operations and their input and output messages. There are three extensibility attributes to enable semantic annotation of WSDL components: *modelReference*, *liftingSchemaMapping* and *loweringSchemaMapping*. The *modelReference* specifies the association between a WSDL component and a concept in some semantic model. This *modelReference* attribute can be used especially to annotate XML Schema type definitions, element declarations, and attribute declarations as well as WSDL interfaces, operations, and faults. The *liftingSchemaMapping* and *loweringSchemaMapping* attributes are added to XML schema element declarations and type definitions for specifying mappings between semantic data and XML. The addition of these attributes requires no other changes to existing WSDL and XML Schema documents, or the manner in which they had been used previously [25].

In the remainder of this section, we present some works that are based on these languages.

In [14] and [19], the authors propose object-oriented languages, the DIANE Service Description (DSD) and the DIANE Elements (DE), in order to put into practice additional requirements that are not fulfilled by semantic service description such us WSMO and OWL-S. DE is a general ontology language with specific features to enhance semantic web service description. DSD is a service description language that uses the specialized constructs provided by DE to describe and match services offers and requests.

Authors in [24] explain what OWL-S constructs are appropriate for use with the various SAWSDL annotations. They provide a rationale and guidelines for their use. The idea is to continue employing the OWL-S constructs and to adopt a SAWSDL-based perspective. Finally, the authors give a set of recommendations for using OWL-S constructs as referents of *modelReference* attribute of every element within WSDL. In others words, how the *modelReference* attribute of every WSDL element can refer an OWL-S constructs.

In [6], authors propose an extension of SAWSDL. This extension is composed of a technical ontology aimed at describing service concepts including non-functional properties, and a domain ontology which describes the semantics of the service. Similarly, authors in [2] propose another extension of SAWSDL for intentional service. The semantic annotations added to the descriptor are based on intentional service ontology. This ontology is built upon the intentional service model and the goal model; it contains all necessary concepts for defining the goal and the intentional service.

Another effort in enhancing Web services description with semantics, related to WSMO, is WSMO-Lite [32, 9]. WSMO-Lite is a lightweight approach to semantic Web service description, evolved from the WSMO framework. WSMO-Lite defines an ontology for service semantics, used directly in SAWSDL to annotate WSDL-based services. Two types of annotations are used in WSMO-Lite, the *reference annotations* and the *transformation annotations*. A reference annotation points from any WSDL element to a WSMO-Lite semantic concept (equivalent to *modelReference* extension attribute in SAWSDL). A transformation annotation specifies a data transformation called *lifting* from a component of XML schema to an element of ontology, and a reverse transformation (from ontology to XML) called *lowering* (equivalent to *liftingSchemaMapping* and *loweringSchemaMapping* extension attributes in SAWSDL).

WSMO-Lite defines four types of service semantics: information model; functional; non-functional; and behavioral. The information model defines the meaning of the information exchanged with the service. The functional semantics is a static description of the service capability, i.e. what the service can offer to its clients when it is invoked (what the service does). The non-functional semantics defines any incidental details specific to the implementation or running environment of a service, such as its price, location or quality of service. The behavioral semantics specifies the protocol that a client needs to follow when consuming a service (how to interact with the service).

Another work related to WSMO is presented in [5]. The authors of this work present a solution based on WSMO ontology to semantically describe the web services interface supporting multimedia indexing. The solution considers the possibility to combine several services to get richer descriptors. The idea was to use a generic XML format that covers the formats of existing multimedia metadata to describe the functionality of such a web service in terms of metadata provided after indexing.

2.2 Web service matching

We present in this section, some important research about service matching. Service matching is the act of finding relevant service for a user request. Generally, web service matching is similar to the matching problems of other areas, such as database matching, text matching and software pattern matching [34]. But these matching problems are still different from web service matching.

Accordingly, the research achievements on these areas are not suitable for our research context. The matching approaches depend on the parts of the service description to match. Some approaches focus on service process; some on service profile (functional, non-functional, etc.); and some others on both of them.

In the literature, we identify three categories of approaches of web service matching. They depend of the way to perform the matching that can be logic-based or not. The first category, which is called logic-based matching, is based on deductive approach. The second category, which is called non logic-based matching, is based either on text similarity measurement, on structured graph matching, or on path-length-based similarity of concepts. The last category, which is called hybrid matching, is based on a combination of logic and non-logic mechanisms. In the rest of this section, we present the most recent works of logic-based, non logic-based and hybrid matching. The figure 2 presents our classification of web service matching approaches.

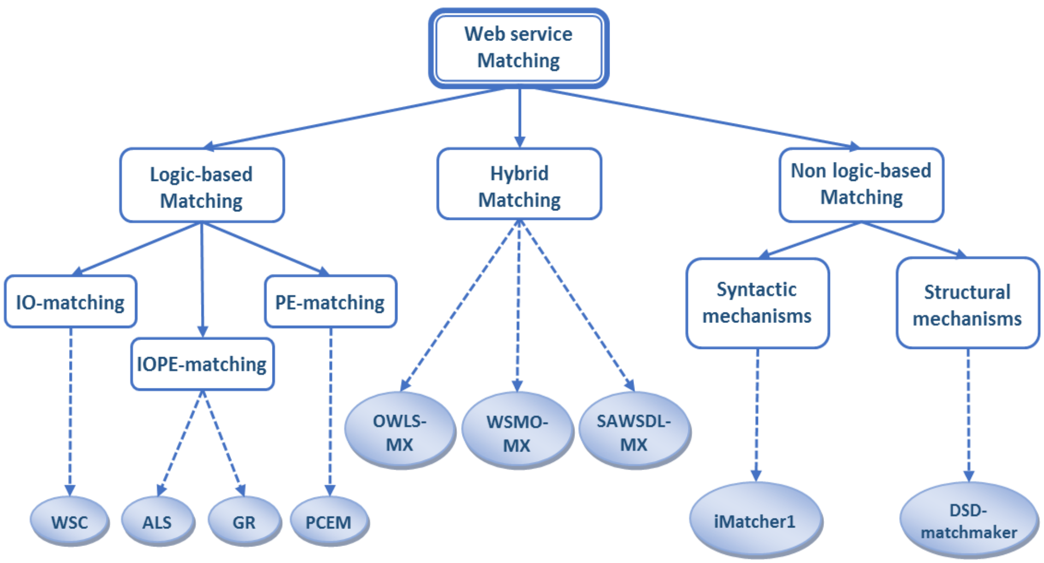


Figure 2 Web service matching approaches.

The logic-based matching approaches use ontological concepts and logical rules. Matching degrees are defined differently depending on semantics of matched description elements. There are mainly three matching approaches [7]:

- **IO-matching:** also called "service profile IO-matching". This type of match is determined from semantic data service parameters: inputs (I) and outputs (O). WSC (Web Services Capabilities) [27] is an example of IO-matching.

- **PE-matching:** determined from matching on pre-conditions (P) and effects (E) of services and queries. PCEM (Pre-conditions and Effects Matchmaker) [17] is an example of PE-matching.

- **IOPE-matching:** determined from matching semantic data of inputs (I), outputs (O), pre-conditions (P) and effects (E) of services and queries. ALS (Automatic Location of Services) [12] and GR (Graded Relevance) [18] are is an examples of IOPE-matching.

The non logic-based matching approaches use syntactic and structural mechanisms like syntactic similarity, term frequencies, numeric distance and structured graph matching. The main idea is to use implicit semantic rather than explicit one. DSD-matchmaker [13] and iMatcher1 [29] are examples of non logic-based matching approaches.

The hybrid matching approaches use a combination of logic and non-logic mechanisms. OWLS-MX [16], WSMO-MX [11], SAWSDL-MX [15] are examples of hybrid matchmakers. The OWLS-MX matchmaker exploits both logic-based reasoning and content-based information retrieval techniques for OWL-S service profile I/O matching. The WSMO-MX matchmaker applies different matching filters to retrieve semantic web services. It computes logic-based and syntactic similarity-based matching degrees and returns a ranked set of services that are semantically relevant to a given user request. The SAWSDL-MX matchmaker is inspired from OWLS-MX and WSMO-MX. It performs hybrid matching for SAWSDL operations based on both subsumption reasoning (logic-based matching) and text retrieval technique (IR-based matching). It combines the results to provide a matching result for service interfaces with multiple operations.

The state of the art provides many works and formal tools in describing and matching web services, but is obviously deficient in specific tools for describing and matching multimedia web services (MaaS: Multimedia as a Service). Our work is located at the intersection of the IR (Information Retrieval) and Services communities. The present work aims at extending the existing works made in the literature to make further progress both in multimedia web services description and matching.

3 Multimedia Ontologies

The availability of huge amounts of multimedia objects implies the need for efficient information retrieval systems that facilitate storage, retrieval and browsing of not only textual but also image, audio and video objects. One potential approach can be based on the semantic annotation of the multimedia content to be semantically described and interpreted both by human agents (users) and technical agents (computers). Hence, there is a strong need of annotating multimedia contents to enhance the agents' interpretation and reasoning for an efficient search.

Expressing multimedia knowledge by means of ontologies increases the precision of multimedia retrieval information systems [31]. In addition, ontologies have the potential to improve the interoperability of different applications producing and consuming multimedia annotations. Hence, Ontologies play an important role in multimedia by exchanging the semantics of multimedia content between distributed information systems.

In last decade, significant research efforts have been made to build and implement multimedia ontologies. The authors of [31] compare well-known ontologies in the multimedia domain. The comparative study is done on 16 ontologies that are classified in four categories: (1) ontologies dedicated to describe multimedia objects in general; (2) ontologies describing images and shapes as visual elements for representing images; (3) ontologies for describing visual objects in general; and (4) music ontologies.

We are interested in the ontologies of the first category which can be considered to be generic for the multimedia domain. The most important of these ontologies are: COMM[a], an ontology with a modular design, which facilitates its extensibility and integration with other ontologies [3]; M3O[b], which is based on ontology design patterns and is targeted to multimedia presentations on the web [28]; and Media Resource Ontology[c], which provides a set of mappings with a great range of multimedia metadata [21].

In order to represent multimedia knowledge of MaaS services, we chose the Media Resource Ontology. Our choice is justified by the following: (1) it is W3C recommendation that is developed by W3C Media Annotation Working Group[d]; (2) it provides mappings with a variety of multimedia formats (Dublin Core, LOM 2.1, ID3, MPEG-7, EXIF, DIG35, etc.), which facilitates the interoperability; and (3) it is well documented, which benefits the ontology understanding. In addition, this ontology covers all the multimedia aspects, it is the most general for describing multimedia objects. The figure 3 presents an extract from Media Resource Ontology.

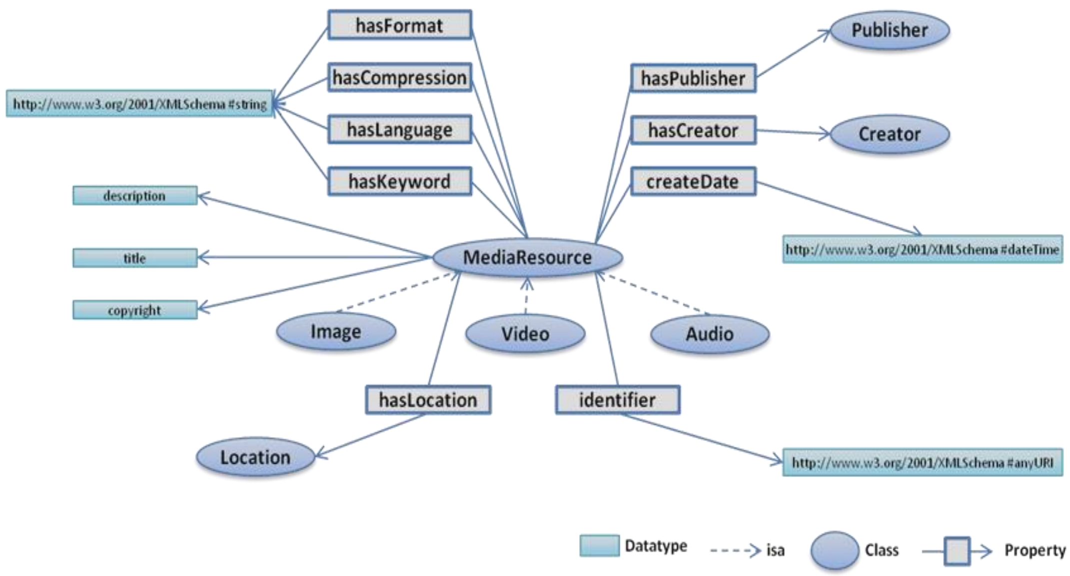


Figure 3 Media Resource Ontology.

4 MaaS services approach

The MaaS services are specific data web services that access to multimedia data. In a previous work [26], we have presented a full service approach to aggregate MaaS services for multimedia data retrieving. This approach is based on a four phases process: description; filtering; clustering; and restitution. In this work, which is a continuation of our previous contribution, we explain in detail how

[a] http://multimedia.semanticweb.org/COMM/
[b] http://www.uni-koblenz-landau.de/koblenz/fb4/AGStaab/Research/ontologies/m3o
[c] http://www.w3.org/TR/mediaont-10/
[d] http://www.w3.org/2008/WebVideo/Annotations/

MaaS services are described and matched with the user requests. The contribution of this paper is two-fold. First, we show how to extend SAWSDL standard for MaaS description. Second, we propose a MaaS matching process that is composed of domain and multimedia matching.

In the following, we present respectively the MaaS description and the MaaS matching phases.

4.1 MaaS description phase

As we said before, many languages and approaches have been developed with the goal to describe semantic web services. We distinguish two main classes of these approaches. Approaches of the first class are based on adding annotations, such as SAWSDL [20] and USDL [30]. Approaches of the second class are based on using of high-level ontology such as OWL-S [23] and WSMO [8], thus avoiding the problems of semantic heterogeneity that may occur. These last approaches use domain ontology to add semantic concepts in their description; they are a "closed approach": on the one hand, they manipulate a language ontology specification, e.g. OWL for OWL-S and WSML to WSMO [22]. On the other hand, they specify very limited set of concepts that are not easily extensible. However, SAWSDL remains an independent approach to language semantic representation. This independence is ensured by the separation between the mechanisms of semantic annotation and representation of semantic description. Without such a mechanism, developers do not have enough flexibility to select their favorite semantic representation of languages or to reuse their own ontology to annotate services [22]. In addition, SAWSDL is close to WSDL, it does not require more effort for developers familiarized with WSDL. This is considered an important advantage compared to other approaches. For all these reasons, we chose SAWSDL language to annotate semantically MaaS services.

The description and the annotation of all specificities and features of MaaS services with SAWSDL language is not enough, we need to take into account the multimedia aspects (e.g. format, location, creation, etc.). The main idea of this work is to extend SAWSDL for enhancing expressiveness of multimedia service description.

The use of *modelReference* attribute of SAWSDL to annotate MaaS services is not sufficient, this attribute allow to reference concepts describing a business domain of services. However, we need to reference separately the concepts defining the semantics of multimedia data of services. To achieve this goal, we propose in our approach to add a new attribute called *multimediaConcept*. This attribute allow to add a new level of description linked to multimedia aspects.

Figure 4 presents our extension of SAWSDL to describe MaaS services. This extension includes two types of ontologies: The first one is Domain Ontology containing concepts that covers a business domain (e.g. medical, tourism, etc.). The second type is MultiMedia Ontology containing concepts defining a set of annotation properties for describing multimedia content. These properties are URIs of multimedia ontology objects. This means that the annotation of a concept in our MaaS approach is a way to tie together this concept to a class that exists in ontology. The multimedia ontology used is Media Resource Ontology presented in the section above. The MaaS description was enriched by references to multimedia concepts such as: a type of media resource, a format, a location, a creation properties, etc.

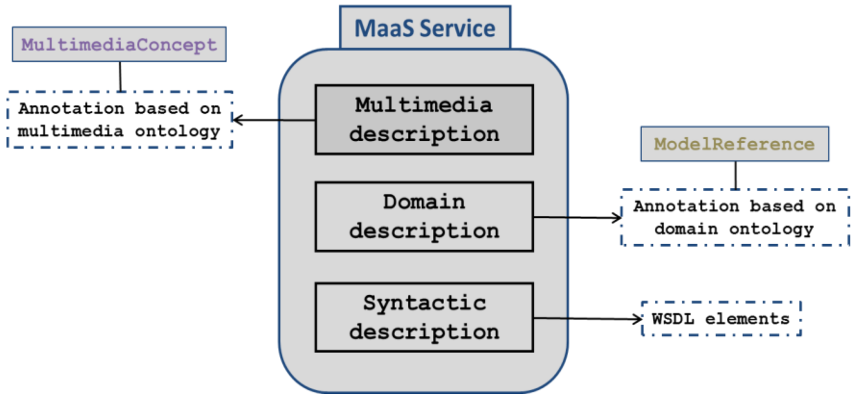


Figure 4 MaaS description approach.

Therefore, we believe that it is necessary to be able to differentiate the semantic annotation of services capabilities and the semantic annotation of data provided by services. Differentiation of semantic annotations for MaaS services aspects can be used to enhance discovery and multimedia data researching in our full service approach.

The MaaS service description is structured in three layers:

1. Syntactic description based on WSDL standard.
2. Domain description represented by a set of annotations based on domain ontology. The SAWSDL *modelReference* attribute is used to add these annotations. We denote the domain description of any MaaS service S_i by $\mathbf{S_i.D}$.
3. Mulimedia description represented by a set of annotations based on multimedia ontology. The SA4MaaS *multimediaConcept* attribute is used to add these annotations. We denote the multimedia description of any MaaS service S_i by $\mathbf{S_i.M}$.

The MaaS description approach claims that introducing *multimediaConcept* attribute makes SAWSDL descriptions more expressive and significant for multimedia data retrieving. The search of a multimedia resource in our system is based either on annotations business domain, or on multimedia annotations or both.

4.2 MaaS matching phase

The purpose of this phase is to identify relevant MaaS services to meet user request, i.e. how to find MaaS services satisfying the query in the best way.

To achieve this goal, we need to identify a new matching mechanism able to find a similarity between the domain description and the multimedia description of both MaaS services and user queries. The comparison between syntactic description of MaaS service and user query is not necessary in our matching approach. We assume in this paper that the query is described in the same way as MaaS services.

Figure 5 explains our matching mechanism. At the MaaS description phase, business concepts are specified through the attribute "*modelReference*" and multimedia concepts are specified through the "*multimediaConcept*" attribute. The "*modelReference*" attribute specifies the concepts from the business domain ontology whereas the "*multimediaConcept*" attribute specifies the concepts from the multimedia ontology. Our proposed matching process is performed in two successive steps. The first step, which is called "Domain matching", consists to compare a domain description of MaaS services (S.D) with a domain description of query (Q.D). The second step, which is called "Multimedia matching", consists to compare a multimedia description of MaaS services (S.M) with a multimedia description of query (Q.M). This step is performed if and only if the previous step has succeeded. We explain in the following each of these two steps.

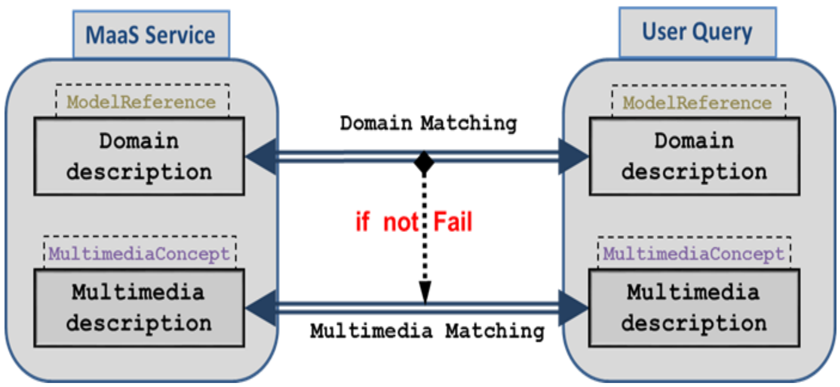


Figure 5 MaaS Matching process.

4.2.1 Domain Matching

This step focuses on the identification of relevant MaaS services to meet the user's request based on their domain descriptions. This refinement is done by applying our matching mechanism (described below) between the query and the candidate MaaS services. In this work, we assume that MaaS services and the query are annotated using the same ontology. The domain concepts of input and output MaaS services and the query are extracted from their SAWSDL files.

Our domain matching approach is based on an IO-matching, i.e. a matching process that considers only the inputs and outputs. The matching of these elements is summarized in a matching between the annotated concepts. We assume, for simplicity, that both input and output are annotated by a single concept. The similarity between two concepts (c_1, $c_2 \in C$) is evaluated by a matching degree. The different matching degrees used in our approach are:

- "Exact": if the first concept (c1) and the second concept (c2) are the same (or equivalent).
- "Subsumed": if the first concept (c1) is a sub-concept of the second concept (c2).
- "Subsumed-by": if the first concept (c2) is a sub-concept of the second concept (c1).
- "Has-Relation": if the two concepts (c1 and c2) are linked by a relation.
- "Has-same-Hierarchy": if the two concepts (c1 and c2) belong to the hierarchy of the same concept.
- "Unknown": if one of the concepts (c1 or c2) is not specified.

- "Fail": If no relationship can be determined between the two concepts.

We have extended the existing works of matching services by adding two degrees: "Has-Relation" and "Has-same-Hierarchy" that are specific to our approach.

We associate with these matching degrees the numeric values between [0, 1] (Table 2) representing similarity degrees, enabling the calculation of the similarity function *SIM* between a MaaS service and a query.

Matching Degree	Exact	Subsumes	Subsumed-by	Has-Relation	Has-same-Hierarchy	Unknown	Fail
Similarity Degree	1	4/5	4/5	3/5	2/5	1/5	0

Table 2 Similarity degrees.

This similarity degrees have not been arbitrarily chosen, they are based on the importance of the semantic link between the elements to match. Strong semantic links are close to 1 and weak semantic links are close to 0.A similarity degree equal to 1 means that the matching is correct or equivalent, whereas a similarity degree equal to 0 means a failure of matching. We consider the two matching degrees "Subsumes" and "Subsumed-by" having the same similarity degree that equal to 0.8.

Similarity (SIM) is calculated based equally on the similarity of inputs (SIM_I) and similarity of outputs (SIM_O).The similarity function is given by the following formula: $SIM = (SIM_I + SIM_O) / 2$.

Assuming we have a set of MaaS services $\{S_1, S_2... S_n\}$ and a query Q. In order to match a MaaS service with a user query, we adopt the following algorithm (figure 6) able to filter relevant MaaS services by comparing their two domain descriptions.

```
Algorithm 1: Domain Matching
Inputs: Query Q, set of MaaS services S {s1,s2, ..., sn}, Thresholdθ
Outputs: set of relevant MaaS services R {s1,s2, ..., sm}
/* the function Similarity returns the similarity degree between two concepts. */
/* the function SIM returns the global similarity degree. */
Begin
    R← ∅
    for each service si in S do
        SIM_IN ← Similarity(si.idc,Q.idc)
        SIM_OUT ← Similarity(si.odc,Q.odc)
        SIM ← (SIM_IN+SIM_OUT)/2
          if ( SIM >= θ) then
                R←R∪{si}
          end if
    end for
    return R
End
```

Figure 6 Domain matching algorithm.

In Algorithm 1, the SIM_IN and SIM_OUT terms denote respectively the input similarity and the output similarity. SIM_IN is calculated between the input domain concept of the first service S_1 (S_1.idc) and the input domain concept of the query Q (Q.idc) (line 11). SIM_OUT is calculated between the output domain concept of the first service S_1 (S_1.odc) and the output domain concept of the query Q (Q.idc) (line 12). The global similarity function is calculated, *SIM = (SIM_IN + SIM_OUT) /2* (line 13). The same process is repeated for all remaining MaaS services $\{S_2, S_3, ..., S_n\}$ and the results are sorted in descending order (line 16). At the end, only are retained the services that have a similarity measure *SIM* greater or equal than a threshold θ (θ is a numerical value chosen by the user), $\theta \in [0, 1]$.

For example, let's consider the MaaS service sample presented in figure 7, which was presented initially in [26]. The aim of this example is to show how a MaaS service can be annotated both with business domain concept and multimedia concept.

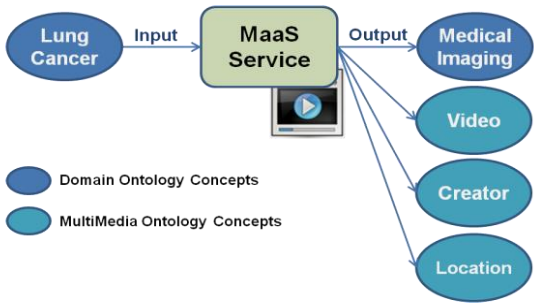


Figure 7 MaaS service sample.

The MaaS service sample returns a video showing the progress of a fiberoptic bronchoscope of a patient suffering from lung cancer, knowing that fiberoptic bronchoscope is an imaging test that can diagnose a patient case. So we annotate the input that is lung cancer by the concept "*Lung_cancer*" and the output that is fiberoptic bronchoscope by the concept "*Medical_Imaging*". We annotate also the service output with multimedia concepts such as *"Video"*, *"Creator"* and *"Location"*.

The semantic annotation of this MaaS service is ensured by two attributes. The first attribute *modelReference* contains a set of URI corresponding to business domain concepts: "*Lung_cancer*" (line 24) and "*Medical_Imaging*" (line 26). The second attribute *multimediaConcept* contains a SPARQL query (presented in the next section) corresponding to semantic multimedia annotation (line 27). We have used both *cancerOnto* ontology, presented in [26] to give semantic domain annotations, and Media Resource Ontology, defined in section 4 to add multimedia annotations.

An extract from the description of MaaS service sample is given in figure 8.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<wsdl:description
  xmlns="http://schemas.xmlsoap.org/wsdl/"
  targetNamespace="http://localhost/maas/LungCancerServices/"
  xmlns:wsdl="http://www.w3.org/ns/wsdl"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:sawsdl="http://www.w3.org/ns/sawsdl"
  xmlns:sa4Maas=" http://localhost/maas/ns/sa4Maas"
  xmlns:ma-ont="http://www.w3.org/ns/ma-ont"
  xmlns:tns="http://localhost/maas/LungCancerServices/"
  xmlns:mt="http://localhost/maas/MediaType"
  name="LungCancerDiagnosesWebService">
 <wsdl:types>
    <!-- Import the MetierType schema (CancerType schema) -->
    <xs:import namespace="http://localhost/maas/MetierType"
            schemaLocation="CancerType.xsd" />
    <!-- Import the MultimediaType schema -->
    <xs:import namespace="http://localhost/maas/MediaType"
            schemaLocation="MultimediaType.xsd" />
 </wsdl:types>
 <wsdl:interface name="LungCancerDiagnosesServiceInterface">
  <wsdl:operation name="getImagesDiagnosesLC">
      <wsdl:input element="tns:lungCancer"
        sawsdl:modelReference="http://localhost/canceronto#Lung_Cancer"/>
      <wsdl:output element="mt:videos"
        sawsdl:modelReference="http://localhost/canceronto#Medical_Imaging"
        sa4Maas:multimediaConcept="Select ..." />
  </wsdl:operation>
 </wsdl:interface>
 <wsdl:binding name="LungCancerDiagnosesServiceSOAPBinding"
  interface="tns:LungCancerDiagnosesServiceInterface"> ... </wsdl:binding>
 <wsdl:service name="LungCancerDiagnosesWebService"
  interface="tns:LungCancerDiagnosesServiceInterface">
  <wsdl:endpoint name="LungCancerDiagnosesServiceEndpoint"
    binding="tns:LungCancerDiagnosesServiceSOAPBinding"
    address="http://localhost/LungCancerDiagnosesWebService/" />
 </wsdl:service>
</wsdl:description>
```

SPARQL query presented in figure 9

Figure 8 SAWSDL file of MaaS service sample.

4.2.2 Multimedia Matching

As we said before, a multimedia concept is used to describe metadata of multimedia content. In our approach, we represent this metadata as SPARQL query defined over multimedia ontology. The domain matching mechanism based on IO-matching model cannot be used to represent the metadata of the MaaS services. This is explained by the fact that this model does not take into account the semantic relationships that may exist between the ontological concepts annotating the metadata part of a MaaS service. In this sense, we propose to annotate the metadata part of a MaaS service with a declarative semantics represented by a SPARQL query. In a SPARQL query, the semantic relationships between the ontological concepts is described by *ObjectProperties*.

```
PREFIX rdf:"http://www.w3.org/1999/02/22-rdf-syntax-ns"
PREFIX ma:"http://www.w3.org/ns/ma-ont"
SELECT  ?x, ?y
WHERE{
        ?M    rdf:type   ma:Video.
        ?M    ma:hasCreator  ?C.
        ?C    rdf:type   ma:Creator.
        ?C    ma:hasName    ?x.
        ?M    ma:hasLocation    ?L.
        ?L    rdf:type   ma:Location.
        ?L    ma:hasName    ?y.}
```

Figure 9 SPARQL query.

For example, if we consider a SAWSDL file of MaaS service (figure 8), the SPARQL query is represented in figure 9. It return the creator name (?x) and the location (?y) of video associated to MaaS service example. In this query, the semantic relationships between concepts annotating metadata of MaaS service are represented by the objectproperties: "*hasCreator*" and "*hasLocation*".

We use the principle of query containment [10] to enable the comparison between multimedia part of MaaS service (*S.M*) and multimedia part of query (*Q.M*). A *Q.M* is said to be contained in *S.M*, denoted by $Q.M \subseteq S.M$, if and only if the answer to *Q* is a subset of the answer to *S* for any knowledge base.

For this second step of matching, we have proposed an algorithm (figure 10) able to compare the multimedia description of MaaS service with the multimedia description of user query. In the first part of algorithm (lines 9-18), we compare each class node C_Q in Q to each class node C_{Si} in the service S_i.M and, if classes match, we continue the process. In the second part (lines 20-29), we check that all object properties in the query Q are covered by the metadata query of the service. The implementation of the functions *classNodeCovering()*, and *objectPropertyCovering()* is provided in the paper [4].

```
1  Algorithm 2: Multimedia Matching
2  Inputs: Query Q, a MaaS services S_i
3  Outputs: Boolean value isMatched (Match or no match)
4  Begin
5        classNode C_Q in Q;
6        objectProperty OP_Q in Q;
7        isMatched ← false;
8        C_Q.first();
9        do
10             classMatch ← false;
11             for each class node C_Si in  S_i.M do
12                   if (C_Q and C_Si have the same class type) then
13                         if (classNodeCovering(C_Q,C_Si)) then
14                               classMatch← true;
15                               break;
16             end for
17             C_Q.next();
18       while (C_Q.hasnext() and classMatch)
```

```
        OP_Q.first();
        do
            objPropertyMatch← false;
            for each object property OP_Si in  S_i.M do
                if (OP_Q and OP_Si are the same) then
                    if (objectPropertyCovering(OP_Q,OP_Si)) then
                        objPropertyMatch← true;
                        break;
            end for
            OP_Q.next();
        while (OP_Q.hasnext() and objPropertyMatch)
        if (classMatch and objPropertyMatch) then
            isMatched← true;
        return isMatched;
End
```

Figure 10 Multimedia matching algorithm.

5 Experiments and solution validation

We chose to validate our approach in the tourism field. We assume that a tourist wants to find images about the different attractions of a given province. These images are published by "Tripadvisor" company in "jpeg" format. We have developed a set of MaaS services (Table 3) to validate our approach.

For the description of these MaaS services, we have used two ontologies: (1) the tourism-onto ontology (figure 11) developed in this work and used to add business domain concepts to the services; (2) the multimedia ontology described in section 3 and used to add multimedia concepts to the services. The third and fifth columns of Table 3 represent the ontological concepts annotating respectively inputs and outputs of services. However, the last column represents the ontological concepts annotating the multimedia aspects of services.

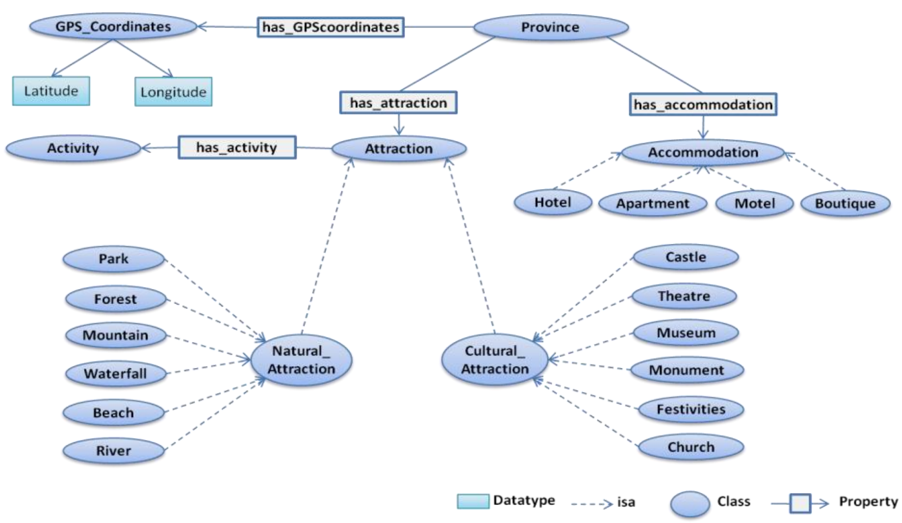


Figure 11 Tourism MaaS services.

Services	Service name	Service Input	SIM_I	Service Output	SIM_O	Service Multimedia Concept
S_1	mountains_province	Province	1	Mountains	0.4	Image, Publisher(tripadvisor), Format (jpeg)
S_2	mountains_province_v	Province	1	Mountains	0.4	Video, Publisher(tripadvisor), Format (png)
S_3	natural-attractions_province	Province	1	Natural_ Attractions	0.8	Audio, Publisher(youtube), Format (wav)
S_4	cultural-attractions_gps	GPSCoordinates	0.6	Cultural_ Attractions	0.8	Image, Publisher(tripadvisor), Format (jpeg)
S_5	hotels_gps	GPSCoordinates	0.6	Hotels	0	Image, Publisher(booking), Format (jpeg)
S_6	museums_gps	GPSCoordinates	0.6	Museums	0.4	Image, Publisher(tripadvisor), Format (jpeg)
S_7	waterfall_province	Province	1	Waterfalls	0.4	Image, Publisher(tripadvisor), Format (jpeg)
S_8	accommodations_ province	Province	1	Accommodations	0	Image, Publisher(tripadvisor), Format (jpeg)
S_9	motels_province	Province	1	Motels	0	Image, Publisher(booking), Format (jpeg)
S_{10}	cultural-attractions_province	Province	1	Cultural_ Attractions	0.8	Image, Publisher(tripadvisor), Format (jpeg)
S_{11}	activity_gps	GPSCoordinates	0.6	Activity	0.6	Image, Publisher(tripadvisor), Format (jpeg)

Table 3 Tourism MaaS services.

We annotate the user query by using the same principle. The result of this annotation is given in table 4.

	Query Input	Query Output	Query Multimedia Concept
Q	Province	Attraction	Image, Publisher(tripadvisor), Format (jpeg)

Table 4 Query Annotation.

To validate our proposals, we have made an evaluation that shows the impact of multimedia matching on the system performance. For that, the experimental evaluation focuses on comparing our proposed approach with and without multimedia matching step. The evaluation is based on calculating the well-known measures recall, precision and F-Measure.

— The recall R is the fraction between the pertinent results retrieved by the system (A) and the total pertinent results (B):

$$R = \frac{A}{B}$$

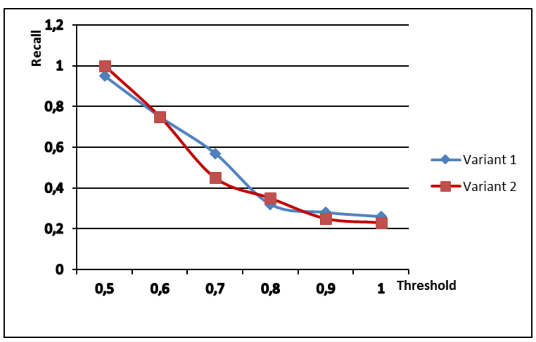
— The precision P is the fraction between the pertinent results retrieved by the system (A) and the total results (C):

$$P = \frac{A}{C}$$

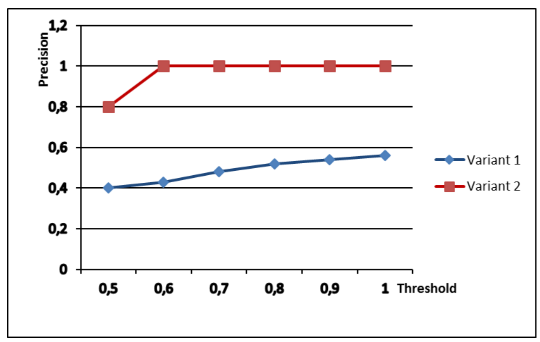
— The F-Measure gives a balanced score for testing the accuracy of the approach and is defined as:

$$F = \frac{2.R.P}{R+P}$$

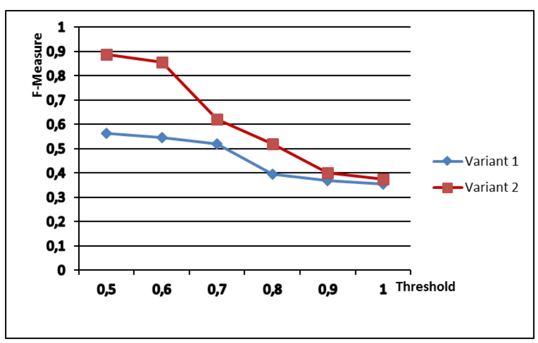
The following figures give the graphical representation of Recall, Precision, F-Measure values for the two variants of our proposed approach. The first variant (variant 1 in graphics) represents the evaluation results of our approach without multimedia matching. However, the second variant (variant 2 in graphics) represents the evaluation results of our approach with multimedia matching.



(a) Recall



(b) Precision



(c) F-Measure

Figure 12 The system performance.

As seen in figure 12(a), we notice that both variants have similar performance in terms of recall. On the other hand, we notice in figure 12(b) a large superiority of the second variant relative to the first in terms of precision. We remark an increasing in the precision of the second variant, which will be equal to 1. This situation is caused by the elimination of impertinent results retrieved by the system when using both the domain and multimedia matching. Equal consideration of recall and precision using the F-Measure yields the results given in figure 12(c), which recapitulates the observations. In conclusion, the variant with multimedia matching offers better performance.

The results notice that the threshold range [0.7, 0.8] is a good compromise for the precision and recall. Our main challenge is then how to set the threshold when using a large test collection. We envisage in a near future to use some existing test collection in evaluating the proposed approach. However, up to our knowledge, there is no existing test collection adapted to our experiments. Hence, we need to prepare a big number of MaaS services fora deep evaluation of our approach.

6 Conclusions and Future Work

In this paper, we presented our recent work and experiments on multimedia data retrieving. We have proposed an extension of SAWSDL for MaaS services. This extension uses two types of ontologies: Domain Ontology and multimedia ontology. The Domain Ontology references business domain concepts of the web service. The Multimedia Ontology references multimedia concepts defining a set of annotation properties for describing multimedia content. In addition, we have presented how this extension is used to address the problem of matching between MaaS services and user needs. To achieve this goal, we have proposed a new matching mechanism for MaaS services. An experiment is conducted to validate the new proposed approach. Results indicate that the use of both domain and multimedia matching improve considerably the performance of multimedia data retrieving systems.

This proposal is part of an ongoing work for implementing the MaaS framework through the development of a general architecture for MaaS service description, discovery and invocation. Future work will concern the specification of a language query facility together with mechanisms for querying and searching a multimedia content of MaaS services. We recall that MaaS services can return one or more image, video, audio or text files. Another challenge, arising from this diversity, consists to give the user homogeneous and coherent results. Otherwise said, how to combine different types of MaaS services, such as an image MaaS service with a video MaaS service, to answer the user query. Accordingly, we believe that the proposed approach requires more reflection on both its theoretical and practical aspects. However, we need first to evaluate it in a large scale setting.

References

1. Akkiraju, Rama, et al. "Web service semantics-wsdl-s." (2005).
2. Aljoumaa, Kadan, Saïd Assar, and Carine Souveyet. "Publishing Intentional Services using extended semantic annotation." Research Challenges in Information Science (RCIS), 2011 Fifth International Conference on. IEEE, 2011.
3. Arndt, Richard, et al. COMM: designing a well-founded multimedia ontology for the web.6th International Semantic Web Conference ISWC'2007. Springer Berlin Heidelberg, 2007.
4. Barhamgi, Mahmoud, Djamal Benslimane, and Brahim Medjahed. "A query rewriting approach for web service composition." Services Computing, IEEE Transactions on 3.3 (2010): 206-222.

5. Brut, Mihaela, et al. "Adapting Indexation to the Content, Context and Queries Characteristics in Distributed Multimedia Systems." Signal-Image Technology and Internet-Based Systems (SITIS), 2011 Seventh International Conference on. IEEE, 2011.
6. Chabeb, Yassin, and Samir Tata. "Yet another semantic annotation for WSDL." IADIS International Conference WWW/Internet. 2008.
7. Chabeb, Yassin. Contributions à la description et la découverte de services web sémantiques. Diss. Institut National des Télécommunications, 2011.
8. ESSI WSMO working group: Web Service Modeling Ontology (WSMO), http://www.wsmo.org, 2004.
9. Fensel, Dieter, et al. "WSMO-Lite: Lightweight semantic descriptions for services on the web." W3C Member Submission 23 (2010).
10. Halevy, Alon Y. "Answering queries using views: A survey." The VLDB Journal 10.4 (2001): 270-294.
11. Kaufer, Frank, and Matthias Klusch. "Wsmo-mx: A logic programming based hybrid service matchmaker." Web Services, 2006. ECOWS'06. 4th European Conference on. IEEE, 2006.
12. Keller, Uwe, et al. "Automatic location of services." The Semantic Web: Research and Applications. Springer Berlin Heidelberg, 2005. 1-16.
13. Klein, Michael, and Birgitta König-Ries. "Coupled signature and specification matching for automatic service binding." Web Services. Springer Berlin Heidelberg, 2004. 183-197.
14. Klein, Michael, Birgitta Konig-Ries, and Michael Mussig. "What is needed for semantic service descriptions? A proposal for suitable language constructs." International Journal of Web and Grid Services 1.3-4 (2005): 328-364.
15. Klusch, Matthias, and Patrick Kapahnke. "Semantic web service selection with SAWSDL-MX." The 7th International Semantic Web Conference. 2008.
16. Klusch, Matthias, Benedikt Fries, and Katia Sycara. "OWLS-MX: A hybrid Semantic Web service matchmaker for OWL-S services." Web Semantics: Science, Services and Agents on the World Wide Web 7.2 (2009): 121-133.
17. Klusch, Matthine. "Semantic web service coordination." CASCOM: Intelligent Service Coordination in the Semantic Web. Birkhäuser Basel, 2008. 59-104.
18. Küster, Ulrich, and Birgitta König-Ries. "Evaluating semantic web service matchmaking effectiveness based on graded relevance." The 7th International Semantic Web Conference. 2008.
19. Kuster, Ulrich, and Birgitta Konig-Ries. "Semantic Mediation between Business Partners-A SWS-Challenge Solution Using DIANE Service Descriptions." Proceedings of the 2007 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Workshops. IEEE Computer Society, 2007.
20. Lausen, Holger, and Joel Farrell. "Semantic annotations for WSDL and XML schema." W3C recommendation, W3C (2007).
21. Lee, WonSuk, et al. Ontology for Media Resources, W3C recommendation, 2012. http://www.w3.org/TR/mediaont-10/
22. M'BARECK, Nomane OULD AHMED. Une approche sémantique pour la description, l'abstraction et l'interconnexion de workflows dans un contexte inter-organisationnel. Diss. Institut national des télécommunications, 2008.
23. Martin, David, et al. "OWL-S: Semantic markup for web services." W3C member submission 22 (2004): 2007-04.

24. Martin, David, Massimo Paolucci, and Matthias Wagner. "Towards semantic annotations of web services: Owl-s from the sawsdl perspective." Proceedings of the OWL-S Experiences and Future Developments Workshop at ESWC. 2007.
25. Martin, David, Massimo Paolucci, and Matthias Wagner. Bringing semantic annotations to web services: Owl-s from the sawsdl perspective. Springer Berlin Heidelberg, 2007.
26. Midouni, Sid Ahmed Djallal, Amghar, Youssef, and Chikh, Azeddine. A Full Service Approach for Multimedia Content Retrieval. In: Model and Data Engineering. Springer International Publishing, 2014. p. 125-137.
27. Paolucci, Massimo, et al. "Semantic matching of web services capabilities." The Semantic Web - ISWC 2002. Springer Berlin Heidelberg, 2002. 333-347.
28. Saathoff, Carsten, and Ansgar Scherp. "Unlocking the semantics of multimedia presentations in the web with the multimedia metadata ontology." Proceedings of the 19th international conference on World Wide Web. ACM, 2010.
29. Schumacher M., Helin H., and Schuldt H.: Semantic Web Service Coordination. Chapter 4, CASCOM: Intelligent Service Coordination in the Semantic Web, Birkhäuser Basel, 2008.
30. Simon, Luke, et al. "A universal service description language." IEEE 20th International Conference on Web Services, 2013.
31. Suárez-Figueroa, Mari Carmen, Ghislain Auguste Atemezing, and Oscar Corcho. "The landscape of multimedia ontologies in the last decade." Multimedia tools and applications 62.2 (2013): 377-399.
32. Vitvar, Tomas, et al. "Wsmo-lite: Lightweight semantic descriptions for services on the web." Web Services, 2007. ECOWS'07. Fifth European Conference on. IEEE, 2007.
33. Wang, Hai H., et al. "A Formal Model of Semantic Web Service Ontology (WSMO) Execution." Engineering of Complex Computer Systems, 2008. ICECCS 2008. 13th IEEE International Conference on. IEEE, 2008.
34. Wu, Zhongbo, Zhumei Dou, and Chaonan Song. "Web Service Matching for RESTful Web Services Based on Parameter Semantic Network." 3rd International Conference on Computer Science and Service System. Atlantis Press, 2014.