Open Research Online

The Open University's repository of research publications
and other research outputs

Two-fold Semantic Web service matchmaking – applying ontology mapping for service discovery

Conference or Workshop Item

oro.open.ac.uk

Two-Fold Service Matchmaking – Applying Ontology Mapping for Semantic Web Service Discovery

Stefan Dietze[1], Neil Benn[1], John Domingue[1], Alex Conconi[2], Fabio Cattaneo[2]

[1] Knowledge Media Institute, The Open University, MK7 6AA, Milton Keynes, UK
{s.dietze, n.j.l.benn, j.b.domingue}@open.ac.uk

[2] TXT eSolutions, Via Frigia 27, 20126 Milano, Italy
{alex.conconi, fabio.cattaneo}@txt.it

Abstract. Semantic Web Services (SWS) aim at the automated discovery and orchestration of Web services on the basis of comprehensive, machine-interpretable semantic descriptions. Since SWS annotations usually are created by distinct SWS providers, semantic-level mediation, i.e. mediation between concurrent semantic representations, is a key requirement for SWS discovery. Since semantic-level mediation aims at enabling interoperability across heterogeneous semantic representations, it can be perceived as a particular instantiation of the ontology mapping problem. While recent SWS matchmakers usually rely on manual alignments or subscription to a common ontology, we propose a two-fold SWS matchmaking approach, consisting of (a) a general-purpose semantic-level mediator and (b) comparison and matchmaking of SWS capabilities. Our semantic-level mediation approach enables the implicit representation of similarities across distinct SWS by grounding service descriptions in so-called Mediation Spaces (MS). Given a set of SWS and their respective grounding, a SWS matchmaker automatically computes instance similarities across distinct SWS ontologies and matches the request to the most suitable SWS. A prototypical application illustrates our approach.

Keywords: Semantic Web Services, Matchmaking, Mediation, Vector Spaces.

1 Introduction

The increasing availability of a broad variety of Web services raises the need to automatically discover and orchestrate appropriate services for a given need. *Semantic Web Services (SWS)* [11] aim at addressing this challenge on the basis of comprehensive, machine-interpretable semantic descriptions. However, since Web services usually are provided by distinct and independent parties, the actual Web service interfaces as well as their semantic representations are highly heterogeneous. This strongly limits interoperability and raises the need of mediating between SWS descriptions as well as the actual Web services. However, despite the importance of mediation for widespread dissemination of SWS technologies, approaches to mediation are still limited and widely ignored by current SWS matchmakers [23].

In this paper, we propose a two-fold SWS matchmaking approach which implicitly tackels *semantic-level mediation* during SWS discovery. Semantic-level mediation refers to the resolution of heterogeneities between semantic representations of services – the actual SWS descriptions – as opposed to data-level mediation, i.e. mediation related to the structure, values or formats of input and output (I/O).

In our vision, semantic-level mediation can be perceived as a particular instantiation of the *ontology mapping* problem. In that, we argue that semantic-level mediation strongly relies on identifying *semantic similarities* between entities across different SWS ontologies [21][31]. However, semantic similarity is not an implicit notion within existing SWS representations (e.g. based on WSMO [30] and OWL-S [22]). Moreover, automatic similarity-detection as demanded by semantic mediation requires semantic meaningfulness. But the symbolic approach – i.e. describing symbols by using other symbols without a grounding in the real world - of established SWS representations does not fully entail semantic meaningfulness, since meaning requires both the definition of a terminology in terms of a logical structure (using symbols) and grounding of symbols [14]. Current approaches to mediation usually foresee the manual development of rather ad-hoc one-to-one mappings or the application of semi-automatic ontology mapping methodologies, mostly based on identifying (a) linguistic commonalities and/or (b) structural similarities between entities [20][5]. Since manually or semi-automatically defining similarity relationships is costly, current approaches are thus not capable to support SWS discovery on a web scale.

In our work, we investigate a mediation mechanism that is based on fuzzy similarity computations between instances as part of SWS ontologies in order to overcome the need for manual or semi-automatic mappings between distinct SWS representations. In this respect, we propose a general purpose matchmaking approach which implicitly addresses semantic-level mediation through (a) a representational approach allowing to implicitly represent similarities and (b) a general-purpose mediator exploiting similarities as represented through (a). In particular, we introduce the concept of *Mediation Spaces (MS)* to enable the implicit representation of semantic similarities across heterogeneous SWS representations through a grounding of SWS descriptions into vector spaces. We will demonstrate that refining heterogeneous SWS descriptions in multiple shared MS supports similarity-based mediation at the semantic level and implicitly facilitates SWS discovery. The provided general-purpose mediator – implemented as a dedicated mediation Web service – supports SWS discovery and is deployable for any semantic-level mediation scenario together with our proposed representational approach.

The remainder of the paper is organized as follows: Section 2 introduces the SWS matchmaking problem, while our two-fold matchmaking approach is proposed in Section 3. In Section 4, we a vector-based approach for semantic-level mediation and the implementation of a generic mediator is being presented in Section 5. Its deployment in a proof-of-concept application is proposed in Section 6 while we discuss and conclude our work in Section 7.

2 Semantic Web Services Mediation

Before formally introducing the SWS mediation problem, we report below the abstract definition of SWS as used throughout the remainder of the paper and a description of the SWS mediation problem, together with background information on current mediation approaches.

Semantic Web Services: a SWS description (either the description of the Web service or the description of the service request) is formally represented within a particular ontology that complies with a certain SWS reference model such as OWL-S [22] or WSMO [30]. Following the formalisation of [9][9], we define a populated *service ontology O* – as utilised by a particular SWS representation – as a tuple:

$$O = \{C, I, P, R, A\} \subset SWS$$

With *C* being a set of *n concepts* in *O w*here each concept C_i is described through *l(i) concept properties pc*. *I* represents all *m instances* where each instance I_{ij} represents a particular instance of a concept C_j and consists of *l(i) instantiated properties pi* instantiating the concept properties of C_j. The properties *P* of an ontology *O* represent the union of all concept properties *PC* and instantiated properties *PI* of *O*.

Given these definitions, we would like to point out that properties here exclusively refer to so-called data type properties. Hence, we define properties as being distinctive to relations *R*. The latter describe relations between concepts and instances. In addition, *A* represents a set of *axioms* which define constraints on the other introduced notions. Since certain parts of a SWS ontology describe certain aspects of the Web service, such as its capability *Cap*, interface *If* or non-functional properties *Nfp* [6], a SWS ontology can be perceived as a conjunction of ontological subsets:

$$Cap \cup If \cup Nfp = O \subset SWS$$

The capability description, as central element of a SWS description, consists of further subsets, describing the assumptions *As*, effects *Ef*, preconditions *Pre* and postconditions *Post*. However, for simplification reasons we prefer the exclusive consideration of assumptions/effects:

$$As \cup Ef = Cap \subset O \subset SWS$$

The SWS mediation problem: mediation aims at addressing heterogeneities among distinct SWS to support all stages that occur at SWS runtime, namely *discovery*, *orchestration* and *invocation*. In contrast to [23][6], we classify the mediation problem into (i) *semantic-level* and (ii) *data-level mediation*. Figure 1 illustrates the chronological order of different mediation tasks at SWS runtime. Whereas (i) refers to the resolution of heterogeneities between concurrent semantic representations of services – e.g. by aligning distinct SWS representations – (ii) refers to the mediation between mismatches related to the Web service implementations themselves, i.e. related to the structure, value or format of I/O messages. Hence, semantic-level mediation primarily supports the discovery stage, whereas data-level mediation occurs during orchestration and invocation. Please note that, for the sake of simplification, Figure 1 just depicts mediation between a SWS request and multiple SWS, while leaving aside mediation between different SWS or between different requests.

Several approaches, such as [1][2][3][19][25][28][31], aim at addressing the mediation issue partially by dealing either with (i) or (ii). For instance, [2] proposes a

semantic mediation framework for scientific workflows relying on the notion of semantic type and structural type, defined in a shared ontology. The semantic type gives a meaning to data, and the structural type is the data schema. As in [28] their work adapts data with a common semantic type but different structural types. In contrast, [31] provides an attempt to support similarity detection for mediation within SWS composition by exploiting syntactic similarities between SWS representations. However, it can be stated that all the above mentioned approaches rely on the definition of a priori mappings, the agreement of a shared ontology or the exploitation of semi-automatic ontology mapping approaches. Hence, providing a generic solution to mediation between heterogeneous SWS remains a central challenge to be solved by SWS matchmaking approaches.
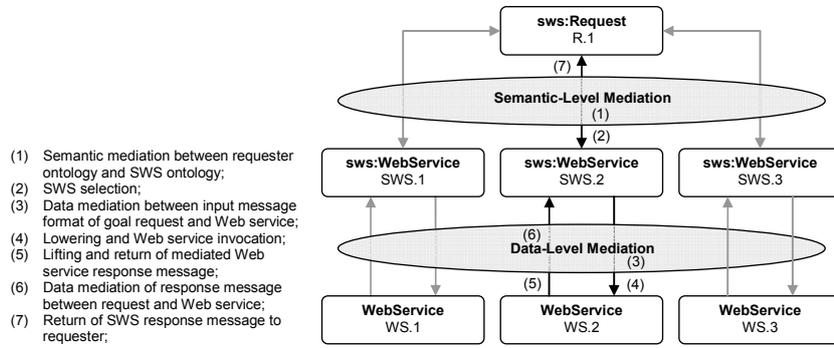

(1) Semantic mediation between requester ontology and SWS ontology;
(2) SWS selection;
(3) Data mediation between input message format of goal request and Web service;
(4) Lowering and Web service invocation;
(5) Lifting and return of mediated Web service response message;
(6) Data mediation of response message between request and Web service;
(7) Return of SWS response message to requester;

Fig. 1. Semantic-level and data-level mediation as part of SWS discovery, orchestration and invocation.

3 SWS Matchmaking as a Two-fold Process

In order to better understand the needs of semantic-level mediation, it is necessary to understand the requirements of the SWS discovery task to which semantic-level mediation is supposed to contribute. In order to identify whether a particular SWS S_1 is potentially relevant for a given request S_2, a SWS broker has to compare the capabilities of S_1 and S_2 , i.e. it has to identify whether the following holds true:

$$As_2 \subset As_1 \cup Ef_2 \subset Ef_1$$

However, in order to compare distinct capabilities of available SWS which each utilise a distinct vocabulary, these vocabularies have to be mapped. For instance, to compare whether an assumption expression $As_1 \equiv \neg I_1 \cup I_2$ of one particular SWS_1 is the same as $As_2 \equiv I_3 \cup \neg I_4$ of another SWS_2, where I_i represents a particular instance, matchmaking engines have to perform two steps:

S1. Semantic-level mediation: alignment of concepts/instances involved in distinct SWS representations;
S2. Matchmaking: evaluation whether the semantics of the SWS expressions match each other.

Whereas current SWS execution environments exclusively focus on *S1*, SWS matchmaking also requires mediation between different SWS ontologies, as in *S1*.

3.1. Semantic-level mediation as an ontology mapping problem

Semantic-level mediation can be perceived as a particular instantiation of the *ontology mapping* problem [31]. With respect to [5] and [24], we define *ontology mapping* as the creation of structure-preserving relations between multiple ontologies. I.e. the goal is, to establish formal relations between a set of knowledge entities E_1 from an ontology O_1 – used to represent a particular SWS S_1 - with entities E_2 which represent the same or a similar semantic meaning in a distinct ontology O_2 [9] which is used to represent an additional SWS S_2. The term *set of entities* here refers to the union of all concepts *C*, instances *I*, relations *R* and axioms *A* defined in a particular SWS ontology. In that, semantic mediation strongly relies on identifying *semantic similarities* [1] between entities across different SWS ontologies. Hence, the identification of similarities is a necessary requirement to solve the mediation problem for multiple heterogeneous SWS representations [21][31]. However, in this respect, the following issues have to be taken into account:

Symbolic SWS representations lack meaningfulness and are ambiguous: similarity-detection across distinct SWS representations requires semantic expressions rich enough to inherently represent semantic similarity between represented entities. However, the symbolic approach, i.e. describing symbols by using other symbols, without a grounding in the real world, of established SWS representation standards, leads to ambiguity issues and does not fully entail semantic meaningfulness, since meaning requires both the definition of a terminology in terms of a logical structure (using symbols) and grounding of symbols to a conceptual level [14].

Lack of automated similarity-detection methodologies: Describing the complex notion of specific SWS capabilities in all their facets is a costly task and may never reach semantic completeness due to the issue described above. While capability representations across distinct SWS representations – even those representing the same real-world entities – hardly equal another, semantic similarity is not an implicit notion within SWS representations. But manually or semi-automatically defining similarity relationships is costly. Moreover, such relationships are hard to maintain in the longer term.

Given the lack of inherent similarity representation, current approaches to ontology mapping could be applied to facilitate SWS mediation. These approaches aim at semi-automatic similarity detection across ontologies mostly based on identifying linguistic commonalities and/or structural similarities between entities of distinct ontologies [20][5]. Work following a combination of such approaches in the field of ontology mapping is reported in [17][10][13][16][20][7]. However, it can be stated, that such approaches require manual intervention, are error-prone, and hence, similarity-computation remains as central challenge. In our vision, instead of semi-automatically formalising individual mappings, methodologies to automatically

compute or implicitly represent similarities across distinct SWS representations are better suited to facilitate SWS mediation.

3.2. Alternative approaches to similarity-computation

Distinct streams of research approach the automated computation of similarities through spatially oriented knowledge representations. *Conceptual Spaces (CS)* [12] follow a theory of describing entities in terms of their quality characteristics similar to natural human cognition in order to bridge between the neural and the symbolic world. [12] proposes the representation of concepts as multidimensional geometrical *Vector Spaces* which are defined through sets of quality dimensions. Instances are represented as vectors, i.e. particular points in a CS. For instance, a particular color may be defined as a point described by vectors measuring the quality dimensions hue, saturation, and brightness. Describing instances as points within vector spaces where each vector follows a specific metric enables the automatic calculation of their semantic similarity by means of distance metrics such as the Euclidean, Taxicab or Manhattan distance [16] or the Minkowsky Metric [28]. Hence, in contrast to the costly formalisation of such knowledge through symbolic representations, semantic similarity is implicit information carried within a CS representation. This is perceived as the major contribution of the CS theory. *Soft Ontologies (SO)* [15] follow a similar approach by representing a knowledge domain *D* through a multi-dimensional *ontospace A*, which is described by its so-called *ontodimensions*. An item *I*, i.e. an instance, is represented by scaling each dimension to express its impact, presence or probability in the case of *I*. In that, a SO can be perceived as a CS where dimensions are measured exclusively on a ratio-scale.

However, although CS and SO aim at solving SW(S)-related issues, several issues still have to be taken into account. For instance, similarity computation within CS requires the description of concepts through quantifiable metrics even in case of rather qualitative characteristics. Moreover, CS as well as SO do not provide any notion to represent any arbitrary relations [27], such as *part-of* relations which usually are represented within first-order logic (FOL) knowledge models. In this regard, it is even more obstructive that the scope of a dimension is not definable, i.e. a dimension always applies to the entire CS/SO [27].

4   A Vector-based Approach to Semantic-level Mediation

To overcome the issues introduced in Section 3.1, we propose a mediation approach which utilises a novel representation mechanism that extends the expressiveness of SWS representations with implicit similarity information.
In particular, we claim that basing service models on either SWS or CS is not sufficient and propose a representational approach which grounds a SWS representation into so-called *Mediation Spaces (MS)*. MS are inspired by CS and enable the implicit representation of semantic similarities across heterogeneous SWS representations provided by distinct agents. MS propose the representation of

concepts which are used as part of SWS descriptions as CS defined through sets of quality dimensions. Instances as part of SWS descriptions are represented as vectors (members) in a MS where similarity between two vectors is indicated by their spatial distance. Hence, refining heterogeneous SWS descriptions into multiple shared MS supports similarity based mediation at the semantic-level and consequently facilitates SWS selection.

Whereas CS allow the representation of semantic similarity as a notion implicit to a constructed knowledge model, it can be argued, that representing an entire SWS through a coherent MS might not be feasible, particularly when attempting to maintain the meaningfulness of the spatial distance as a similarity measure. Therefore, we claim that MS are a particularly promising model when being applied to individual concepts – as part of SWS descriptions – instead of representing an entire SWS ontology in a single MS. In that, we would like to highlight that we consider the representation of a set of n concepts C of a SWS ontology O through a set of n MS. Hence, instances of concepts are represented as members (i.e. vectors) in the respective MS. While still taking advantage of implicit similarity information within a MS, our hybrid approach – combining ontology-based SWS descriptions with multiple vector-based MS – allows to overcome CS-related issues, such as the lack of expressivity for arbitrary relations, by maintaining the advantages of ontology-based SWS representations. Please note that our approach relies on the agreement on a common set of MS for a given set of distinct SWS ontologies, instead of a common agreement on the used ontologies/vocabularies themselves. Thus, whereas in the latter case two agents have to agree on a common ontology at the concept and instance level, our approach requires just agreement at the schema level, since instance similarity becomes an implicit notion. Moreover, we assume that the agreement on ontologies at the schema level becomes an increasingly widespread case, due, on the one hand, to increasing use of upper-level ontologies such as DOLCE[1], SUMO[2] or OpenCyc[3] which support a certain degree of commonality between distinct ontologies, and on the other hand, to SWS ontologies often being provided within closed environments where a common agreement to a certain extent is ensured. In such cases, the derivation of a set of common MS is particularly applicable and straightforward.

In order to refine and represent SWS descriptions within a set of MS, we formalised the MS model into an ontology, currently being represented through OCML [18]. The ontology enables the instantiation of a set of MS to represent a given set of concepts as part of SWS descriptions. Referring to [26], we formalise a MS as a vector space defined through quality dimensions d_i of MS. Each dimension is associated with a certain metric scale, e.g. ratio, interval or ordinal scale. To reflect the impact of a specific quality dimension on the entire MS, we consider a prominence value p for each dimension [26]. Therefore, a MS is defined by

$$MS^n = \left\{ (p_1 d_1, p_2 d_2, ..., p_n d_n) \middle| d_i \in MS, p_i \in \Re \right\}.$$

Please note that we enable dimensions to be detailed further in terms of subspaces. Hence, a dimension within one MS may be defined through another MS by using

---

further dimensions. In such a case, the particular quality dimension d_j is described by a set of further quality dimensions. In this way, a MS may be composed of several subspaces and consequently, the description granularity can be refined gradually. Furthermore, dimensions may be correlated. Information about correlation is expressed through axioms related to a specific quality dimension instance.

A member M – representing a particular instance – of the MS is described through through a vector defined by the set of valued dimensions v_i:

$$M^n = \{(v_1, v_2, ..., v_n) | v_i \in M\}$$

With respect to [7], we define the semantic similarity between two members of a space as a function of the Euclidean distance between the points representing each of the members. However, we would like to point out that different distance metrics could be considered, dependent on the nature and purpose of the MS. Given a MS definition MS and two members v and u, defined by vectors v_0, v_1, ...,v_n and u_1, u_2,...,u_n within MS, the distance between v and u can be calculated as:

$$dist(u, v) = \sqrt{\sum_{i=1}^{n} p_i ((\frac{u_i - \bar{u}}{s_u}) - (\frac{v_i - \bar{v}}{s_v}))^2}$$

where $\bar{u}$ is the mean of all values of data set U and s_u is the standard deviation of U.

The formula above already considers the so-called Z-transformation or standardization which facilitates the standardization of distinct measurement scales utilised by different quality dimensions in order to enable the calculation of distances in a multi-dimensional and multi-metric space. Please refer to [8], for a detailed description on how distinct MS can be derived for arbitrary SWS, i.e. a methodology to represent SWS through MS.

5 Implementing Two-Fold SWS Matchmaking based on WSMO and IRS-III

The representational model described above had been implemented by and aligned to established SWS technologies based on WSMO [30] and the Internet Reasoning Service IRS-III [4]. However, please note that in principle the representational approach described above could be applied to any SWS reference model and is particularly well-suited to support rather light-weight approaches such as SAWSDL or WSMO Lite [29].
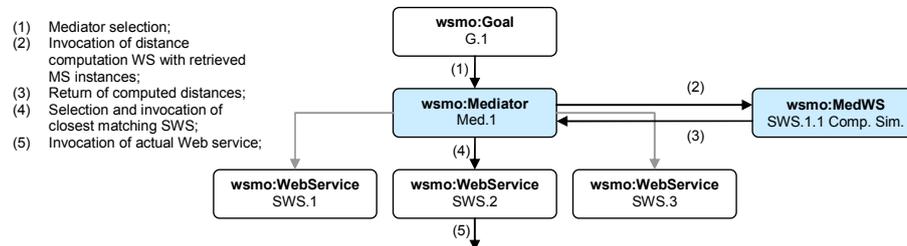

(1) Mediator selection;
(2) Invocation of distance computation WS with retrieved MS instances;
(3) Return of computed distances;
(4) Selection and invocation of closest matching SWS;
(5) Invocation of actual Web service;

Fig. 2. WSMO SWS matchmaking utilizing a similarity-based Mediator for semantic-level Mediation.

To facilitate our MS-based approach, we provided a general-purpose matchmaking approach (Fig. 2) utilising a semantic-level mediator which implemented as a particular mediation service. Given the ontological refinement of SWS descriptions into MS as introduced above, the mediation service is reusable and can be deployed to solve all sorts of semantic-level mediation scenarios. Please note that our current Mediator assumes logical SWS capability expressions to be defined through simple conjunctions of instances. Arbitrary logical expressions will be considered within a revised implementation.

When attempting to achieve match a SWS request (*wsmo:Goal* in Figure 2), our mediator is provided with the actual SWS request SWS_i, named *base,* and the SWS descriptions of all *x* available services that are potentially relevant for the base – i.e. linked through a dedicated mediator:

$$SWS_i \cup \{SWS_1, SWS_2, ..., SWS_x\}$$

Each SWS contains a set of concepts $C=\{c_1..c_m\}$ and instances $I=\{i_1..i_n\}$. We first identify all members $M(SWS_i)$ – in the form of valued vectors $\{v_1..v_n\}$ refining the instance i_l of the base as proposed in Section 4. In addition, for each concept c within the base the corresponding conceptual space representations $MS=\{MS_1..MS_m\}$ are retrieved. Similarly, for each SWS_j related to the base, members $M(SWS_j)$ – which refine capabilities of SWS_j and are represented in one of the CS $CS_1..CS_m$ – are retrieved:

$$CS \cup M(SWS_i) \cup \{M(SWS_1), M(SWS_2), ..., M(SWS_x)\}$$

Based on the above ontological descriptions, for each member v_l within $M(SWS_i)$, the Euclidean distances to any member of all $M(SWS_j)$ which is represented in the same space MS_j as v_l are computed. In case one set of members $M(SWS_j)$ contains several members in the same MS – e.g. SWS_j targets several instances of the same kind – the algorithm just considers the closest distance since the closest match determines the appropriateness for a given goal. For example, if one SWS supports several different locations, just the one which is closest to the one required by SWS_i determines the appropriateness.

Consequently, a set of *x* sets of distances is computed as follows $Dist(SWS_i)=\{Dist(SWS_i, SWS_1), Dist(SWS_i, SWS_2) .. Dist(SWS_i, SWS_x)\}$ where each $Dist(SWS_i, SWS_j)$ contains a set of distances $\{dist_1..dist_n\}$ and any $dist_i$ represents the distance between one particular member v_i of SWS_i and one member refining one instance of the capabilities of SWS_j. Hence, the overall similarity between the base SWS_i and any SWS_j could be defined as being reciprocal to the mean value of the individual distances between all instances of their respective capability descriptions and hence, is calculated as follows:

$$Sim(SWS_i, SWS_j) = \left(\overline{Dist(SWS_i, SWS_j)}\right)^{-1} = \left(\frac{\sum_{k=1}^{n}(dist_k)}{n}\right)^{-1}$$

Finally, a set of *x* similarity values – computed as described above – which each indicates the similarity between the base SWS_i and one of the *x* target SWS is computed:

$$\{Sim(SWS_i, SWS_1), Sim(SWS_i, SWS_2), .., Sim(SWS_i, SWS_x)\}$$

As a result, the most similar SWS_j, i.e. the closest associated SWS, can be selected and invoked. In order to ensure a certain degree of overlap between the actual request

and the invoked functionality, we also defined a *threshold similarity value T* which determines the similarity threshold for any potential invocation.

Within our current implementation, we provided a new matchmaking function within IRS-III which automatically performs the similarity computation described above as part of the matchmaking procedure and hence, realizes our two-fold matchmaking approach.

6 Application – Similarity-based Selection of Video Retrieval Services

We provided a prototypical implementation which aims at similarity-based retrieval of public multimedia (MM) content exposed via Web services. Our prototypical application utilizes our approach to annotate (Web) services which operate on top of distributed MM metadata repositories. These services had been created in the context of the EC-funded project NoTube[4] and make use of the Youtube-API[5] as well as data feeds provided by BBC- Backstage[6] and Open Video[7]. The available services were annotated following the representational approach proposed in Section 4. We make use of standard SWS technology based on WSMO and IRS-III which had been extended with our two-fold matchmaking mechanism to tackle the semantic-level mediation problem.

6.1. Representing Video Retrieval Services through multiple MS

In fact, five different Web services had been provided, each able to retrieve content from distinct repositories through keyword-based searches. WS_1 is able to retrieve content from the Youtube channel of The Open University[8], while WS_2 provides Youtube content associated with the *entertainment* category following the Youtube vocabulary. WS_3 performs keyword-based searches on top of the Open Video repository, while WS_4 operates on top of the news metadata feeds provided by BBC Backstage. In addition, WS_4 provides Youtube content suitable for mobiles.

---

[4] http://projects.kmi.open.ac.uk/notube/

[5] http://code.google.com/intl/en/apis/youtube/

[6] http://backstage.bbc.co.uk/

[7] http://www.open-video.org/
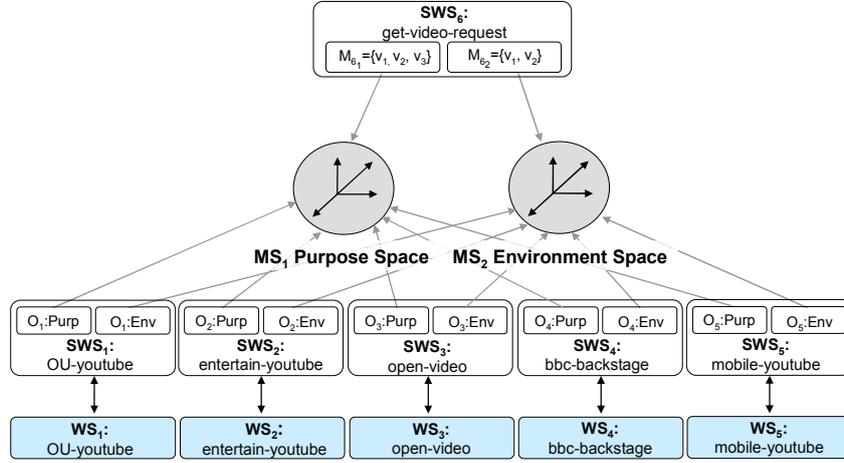
[8] http://www.youtube.com/ou

Fig. 3. MM service metadata refined in two distinct CS.

Based on the SWS reference model WSMO, we provided service annotations following the approach described above. Each service has distinct constraints, and thus distinct SWS metadata. In particular, we annotated the Web services in terms of the *purpose* they serve MM content for and the technical *environment* supported by the delivered content. In that, a simplified space (MS_1: *Purpose Space* in Figure 3) was defined to refine the notion of purpose by using three dimensions indicating the intended purpose of a particular piece of MM content: $\{((p_1*information), (p_2*education), (p_3*leisure))\} = MS_1$. The dimensions of MS_1 are measured on a ratio scale ranging from 0 to 100. For instance, a member P_1 in MS_1 described by vector $\{(0, 100, 0)\}$ would indicate a rather educational purpose. In addition, a second space (MS_1: *Environment Space* in Figure 3) was provided to represent technical environments in terms of dimensions measuring the available resolution and bandwidth $\{((p_1*resolution), (p_2*bandwidth))\} = MS_2$. For simplification, also the dimensions of MS_2 were ranked on a ratio scale. However, it is intended to refine the resolution dimension to apply an interval scale to both dimensions to be able to represent actual resolution and bandwidth measurements. Each dimension was ranked equally with a prominence of 1 in all cases.

By applying the representational approach proposed here, each concept of the involved heterogeneous SWS representations of the underlying services was refined as shared MS, while instances – used to define SWS and SWS requests – were defined as members, i.e. vectors. No explicit relations were formalised across SWS representations. Instead, similarities are computed by means of distance calculation following the algorithm proposed in Section 5. In that, assumptions (*Ass*) of available MM services had been described independently in terms of simple conjunctions of instances which were individually refined as vectors in shared MS as shown in Table 1. Each MM service was associated with a set of members (vectors) in MS_1 and MS_2 to represent its purpose and the targeted environment. For instance, SWS_3 which provides resources from the Open Video repository, which in fact are of rather educational or information nature, was associated with a corresponding purpose vector $\{(50, 50, 0)\}$. While SWS_5 represents a Web service dedicated to video content

suitable for mobiles, a vector *{(10,10)}* indicating low resolution and bandwidth values was associated with SWS_5.

Table 1. Assumptions of involved SWS (requests) described as vectors in MS_1 and MS_2.

	Assumption $Ass_{SWSi} = (P_{1SWSi} \cup P_{2SWSi} \cup .. \cup P_{nSWSi}) \cup (E_{1SWSi} \cup E_{2SWSi} \cup .. \cup E_{mSWSi})$	
	Members P_i in MS_1 (purpose)	Members E_j in MS_2 (environment)
SWS_1	$P_{1(SWS1)}$={(0, 100, 0)}	$E_{1(SWS1)}$={(100, 100)}
SWS_2	$P_{1(SWS2)}$={(0, 0, 100)}	$E_{1(SWS2)}$={(100, 100)}
SWS_3	$P_{1(SWS3)}$={(50, 50, 0)}	$E_{1(SWS3)}$={(100, 100)}
SWS_4	$P_{1(SWS4)}$={(100, 0, 0)}	$E_{1(SWS4)}$={(100, 100)}
SWS_5	$P_{1(SWS5)}$={(100, 0, 0)} $P_{2(SWS5)}$={(0, 100, 0)}	$E_{1(SWS5)}$={(10, 10)}

6.2. Similarity-based Matchmaking

An AJAX-based user interface (Fig. 4) was provided which allows users to define requests by providing measurements describing their context, i.e. the purpose and environment, and WS input parameters, i.e. a set of keywords. Fig. 4 depicts a screenshot of the Web interface after our mediator computed a ranking of most suitable SWS based on distances in MS.

Fig. 4. Screenshot of AJAX interface depicting a suitability ranking of available services to match a given request.

For instance, a user provides a request *R* with the input parameter keyword "Aerospace" together with context measurements which correspond to the following vectors: *$P_1(R)$={(60, 55, 5)}* in MS_1 and *$P_2(R)$=(95, 90)}* in MS_2. These vectors indicate the need for content which serves the need for education or information and which supports a rather high resolution environment. Though no SWS matches these criteria exactly, at runtime similarities are calculated between *R* and the related SWS (*SWS_1-SWS_5*) through the similarity computation service described in Section 5.

This led to the calculation of the similarity values shown in Table 2. Given these similarities, our reasoning environment automatically selects the most similar MM service (SWS_3) and triggers its invocation.

Table 2. Automatically computed similarities between request R and available SWS.

	Similarities
SWS_1	0.023162405
SWS_2	0.014675636
SWS_3	0.08536871
SWS_4	0.02519804
SWS_5	0.01085659

Eventually, the most similar service is invoked and retrieves MM metadata records from the Open Video repository which match the requested search term "Aerospace". As illustrated above, our application utilises our two-fold matchmaking mechanism to support matchmaking of distributed SWS while tackling the semantic-level mediation problem.

7 Discussion and Conclusions

In order to further facilitate SWS interoperability we proposed a two-fold matchmaking approach which implicitly tackles the semantic-level mediation problem. Note, while our approach utilises a general-purpose mediation service which utilises SWS refinements in MS, different SWS alignment methodologies could be applied and combined to further optimise SWS alignment, i.e. semantic-level mediation. The introduced two-fold matchmaking approach supports implicit representation of similarities between instances across heterogeneous ontologies through dedicated representations in MS, and consequently, provides a means to facilitate SWS interoperability. To evaluate our approach, we deployed a prototypical application based on WSMO in a video metadata retrieval scenario.

The proposed approach has the potential to significantly reduce the effort required to mediate between distinct heterogeneous SWS ontologies and the extent to which two distinct parties have to share their conceptualisations. Whereas traditional matchmaking methodologies rely on either manual formalisation of one-to-one mappings or subscription to a common ontology, our approach supports automatic similarity-computation between instances though requiring a common agreement on a shared MS. However, even for the case of heterogeneous MS, traditional semi-automatic mapping methodologies could be applied to initially align distinct spaces. In addition, incomplete similarities are computable between partially overlapping MS. Given the nature of our approach - aiming at mediating between sets of concepts/instances which are used to annotate particular SWS - we argue that our solution is particularly applicable to SWS frameworks which are based on rather light-weight service semantics such as WSMO-Lite [29] or OWL-S [22]. Moreover, by representing SWS through vectors which are independent from the underlying representation language, we believe that our approach also has the potential to bridge between SWS across concurrent SWS reference models and modeling languages.

However, the authors are aware that our approach requires a considerable amount of additional effort to establish MS-based representations. Future work has to investigate on this effort in order to further evaluate the potential contribution of the proposed approach. Moreover, whereas defining instances, i.e. vectors, within a given MS appears to be a straightforward process of assigning specific quantitative values to quality dimensions, the definition of the MS itself is not trivial and dependent on individual perspectives and subjective appraisals. Furthermore, whereas the size and resolution of a MS is indefinite, defining a reasonable MS may become a challenging task. Nevertheless, distance calculation relies on the fact that resources are described in equivalent geometrical spaces. However, particularly with respect to the latter, traditional ontology and schema matching methods could be applied to align heterogeneous spaces. In addition, we would like to point out that the increasing usage of upper level ontologies, such as DOLCE or SUMO, and the progressive reuse of ontologies, particularly in loosely coupled organisational environments, leads to an increased sharing of ontologies at the concept level what also applies to SWS representations. As a result, our proposed hybrid representational model and mediation approach becomes increasingly applicable by further enabling similarity-computation at the instance-level towards the vision of interoperable ontologies.

8 References

[1] Bicer, V., Kilic, O., Dogac, A., Laleci, G.B. Archetype-based semantic interoperability of web service messages in the health care domain. International Journal of Semantic Web and Information Systems (IJSWIS), 1(4):1–23, October 2005.

[2] Bowers, S., Lud̈ascher, B. An ontology-driven framework for data transformation in scientific workflows. In E. Rahm, editor, DILS, volume 2994 of Lecture Notes in Computer Science, pages 1–16. Springer, 2004.

[3] Cabral, L., Domingue, J. Mediation of semantic web services in IRS-III. In First International Workshop on Mediation in Semantic Web Services (MEDIATE 2005) held in conjunction with the 3rd International Conference on Service Oriented Computing (ICSOC 2005), Amsterdam, The Netherlands., December 12th 2005.

[4] Cabral, L., Domingue, J., Galizia, S., Gugliotta, A., Norton, B., Tanasescu, V., Pedrinaci, C.: IRS-III: A Broker for Semantic Web Services based Applications. In proceedings of the 5th International Semantic Web Conference (ISWC 2006), Athens, USA (2006).

[5] Choi, N., Song, I., and Han, H. (2006), A survey on ontology mapping, SIGMOD Rec., Vol. 35, No. 3. (September 2006), pp. 34-41.

[6] Cimpian, E., Mocan, A., Stollberg, M. (2006) Mediation Enabled Semantic Web Services Usage. Proceedings of the 1st Asian Semantic Web Conference (ASWC 2006), September 6, 2006, Beijing, China, 2006 - Springer

[7] Dietze, S., Gugliotta, A., Domingue, J., (2008) Conceptual Situation Spaces for Situation-Driven Processes. 5th European Semantic Web Conference (ESWC), Tenerife, Spain.

[8] Dietze, S., Gugliotta, A., and Domingue, J. (2009) Exploiting Metrics for Similarity-based Semantic Web Service Discovery, IEEE 7th International Conference on Web Services (ICWS 2009), Los Angeles, CA, USA.

[9] Ehrig, M, Sure, Y. (2004), Ontology Mapping - An Integrated Approach, in Proceedings of ESWS, 2004.

[10] Euzenat, J., Guegan, P., and Valtchev, P. OLA in the OAEI 2005 Alignment Contest. K-Cap 2005 Workshop on Integrating Ontologies 2005, 97-102.

[11] Fensel, D., Lausen, H., Polleres, A., de Bruijn, J., Stollberg, M., Roman, D., Domingue, J. (2006): Enabling Semantic Web Services – The Web service Modelling Ontology, Springer 2006.

[12] Gärdenfors, P. (2000), Conceptual Spaces - The Geometry of Thought. MIT Press, 2000.

[13] Giunchiglia, F., Shvaiko, P., and Yatskevich, M. S-Match: An Algorithm and an Implementation of Semantic Matching. ESWS 2004, 61-75.

[14] Harnad, S. (1999) The Symbol Grounding Problem CoRR cs.AI/9906002, (1999).

[15] Kaipainen, M., Normak, P., Niglas, K., Kippar, J., Laanpere, M., Soft Ontologies, spatial Representations and multi-perspective Explorability, Expert Systems, November 2008, Vol. 25, No. 5.

[16] Krause, E. F. (1987). Taxicab Geometry. Dover.

[17] Mitra, P., Noy, F. N., Jaiswals, A. (2005), OMEN: A Probabilistic Ontology Mapping Tool, International Semantic Web Conference 2005.

[18] Motta, E. (1998). An Overview of the OCML Modelling Language, the 8th Workshop on Methods and Languages, 1998.

[19] Mrissa, M., Ghedira, C., Benslimane, D., Maamar, Z., Rosenberg, F., Dustdar, S. A context-based mediation approach to compose semantic web services. ACM Trans. Internet Techn., 8(1), 2007.

[20] Noy, N.F., and Musen, M.A. The PROMPT Suite: Interactive Tools for Ontology Merging and Mapping. International Journal of Human-Computer Studies 59 (2003), 983-1024.

[21] Qu, Y., Hu, W., Cheng, G. (2006) Constructing Virtual Documents for Ontology Matching, *WWW 2006*, May 23–26, 2006, Edinburgh, Scotland. ACM 1595933239/06/0005.

[22] OWL-S 1.0 Release. http://www.daml.org/services/owl-s/1.0/.

[23] Paolucci, M., Srinivasan, N., Sycara, K. Expressing WSMO Mediators in OWL-S, in proceedings of Semantic Web Services: in 3rd International Semantic Web Conference (ISWC 2004).

[24] Pease, A., Niles, I., Li, J.(2002), The suggested upper merged ontology: A large ontology for the semanticweb and its applications. In: AAAI-2002Workshop on Ontologies and the Semantic Web. Working Notes (2002).

[25] Radetzki, U., Cremers, A. B. Iris: A framework for mediator-based composition of service-oriented software. In ICWS, pages 752–755. IEEE Computer Society, 2004.

[26] Raubal, M. (2004). Formalizing Conceptual Spaces. in: A. Varzi and L. Vieu (Eds.), Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004).Frontiers in Artificial Intelligence and Applications 114, pp. 153-164, IOS Press, Amsterdam, NL.

[27] Schwering, A. (2005). Hybrid Model for Semantic Similarity Measurement, in R. Meersman and Z. Tari (Eds.): CoopIS/DOA/ODBASE 2005, LNCS 3761, pp. 1449 – 1465, 2005..

[28] Spencer, B., Liu, S. Inferring data transformation rules to integrate semantic web services. In S. A. McIlraith, D. Plexousakis, and F. van Harmelen, editors, Int'l Semantic Web Conference, volume 3298 of Lecture Notes in Computer Science, pages 456–470. Springer, 2004.

[29] Vitvar, T., Kopecký, J., Viskova, J., Fensel, D. WSMO-Lite Annotations for Web Services. ESWC 2008: 674-689.

[30] WSMO Working Group, D2v1.0: Web service Modeling Ontology (WSMO). WSMO Working Draft, (2004). (http://www.wsmo.org/2004/d2/v1.0/).

[31] Wu, Z., Ranabahu, A., Gomadam, K., Sheth, A.P., Miller, J.A., Automatic Composition of Semantic Web Services using Process Mediation, Proceedings of the 9th International Conference on Enterprise Information Systems (ICEIS'07), Funchal, Portugal Jun. 2007. pages 453-461.