

An Improved Matching Algorithm for Developing a Consistent Knowledge Model across Enterprises Using SRS and SWRL

Saravanan Muthaiyah
George Mason University
Department of Computer
Science
4400 University Dr
Fairfax, VA 22030, USA
smuthai@gmu.edu

Marcel Barbulescu
George Mason University
Department of Computer
Science
4400 University Dr
Fairfax, VA 22030, USA
mbarbulescu@gmu.edu

Larry Kerschberg
George Mason University
Department of Computer
Science
4400 University Dr
Fairfax, VA 22030, USA
kersch@gmu.edu

Abstract

This paper highlights an end-to-end framework and process methodology for developing a consistent knowledge model across enterprises. We demonstrate an improved matching algorithm i.e. the Semantic Relatedness Scores (SRS) and the Semantic Web Rule Language (SWRL) and how they can be coupled together to achieve better reliability and precision in matching heterogeneous data schemas. We introduce a process methodology support this. The goal here is to develop a consistent knowledge model across enterprises that are more precise and reliable. We have also implemented a multi-agent system (MAS) prototype based on the service oriented architecture (SOA) for proof-of-concept. Finally we demonstrate how our approach is represented in the Zachman Framework.

1. Introduction

Schema heterogeneity is a major barrier for enterprises to seamlessly share and reuse organizational data. Although the federated data approach and platform integration technologies provide much support for this area, with the advent of Web 3.0 technologies and beyond such as the Semantic Web, enterprises have to start rethinking as to how they can start sharing and reusing data in a seamless fashion. This is paramount in the services sector that involves government as well as private entities.

The Semantic Web platform enables software agents to sift, winnow and integrate data on-the-fly. However, technologies that make up the Semantic Web such as ontologies, Web Ontology language (OWL) and resource description framework (RDF) must be aligned to achieve this interoperability. Tim Berners-

Lee states that it is important for data to be made machine understandable and processable for Semantic Web to enable data integration across the Semantic Web. This also includes the provision for metadata and data interchange formats such as N3 (Notation 3), XML (Extensible Mark-up Language), Turtle (Terse RDF Triple Language) and N-Triples.

Ontologies, which are at the core of this technology, are meant to resolve schema heterogeneity. However, it is necessary for ontological commitments to be made by enterprises for a shared vocabulary before. As such, the vocabulary available in ontologies *per se* is not able to provide any solutions for a consistent information model. For example, various ontologies exist for the travel domain today. Although they describe the same travel information, the schema, structure and data type defined in them may not necessarily be the same. In other words semantically identical information for travel may be expressed differently by various ontologies. As such the need for integration becomes necessary for the extraction and reuse of information from such sources.

Let's look at an example to understand the problem better. We describe the data heterogeneity problem by providing two cases. Case 1 is a travel domain example and case 2 is an e-government example. Case 1, is illustrated by figure 1 and 2. Figure 1 shows a travel reservation ontology of enterprise A that has a super class called "Travel Reservation", below which, two subclasses are defined i.e. "Accommodation" and "Transportation". The accommodation class is made up of subclasses i.e. hotel, motel and others. The transportation class is made up of subclasses i.e. sea, air and land. "Airlines" appears as a subclass under "Air" and "Coach" appears as a subclass under "Land". The arrows here represent the "is-a" relation between classes for example hotel is a type of accommodation.

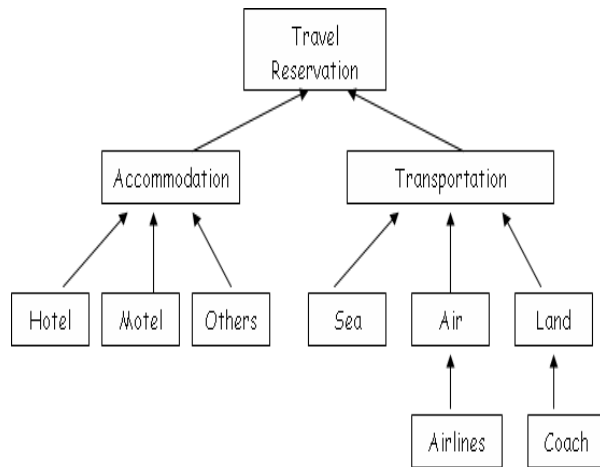


Figure 1. Travel Reservation Ontology A

Figure 2 also illustrates a travel reservation ontology for enterprise B which uses a slightly different naming convention. It has the same super class “Travel Reservation” but it’s subclasses have different names compared to figure 1, such as “Lodging” instead of “Accommodation” and “Transport” instead of “Transportation”. Also air transportation is defined with more granularity compared to the description in figure 1, where flight is described as domestic and international. “Flight” has also been used to describe “Airlines” in ontology B.

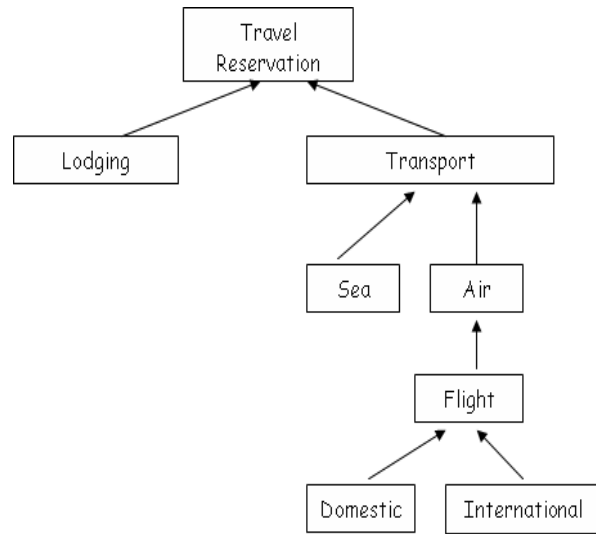


Figure 2. Travel Reservation Ontology B

In conclusion, enterprise A and enterprise B illustrate semantically identical information that is not defined explicitly in the same fashion. The ontology differences here result from mainly different naming conventions and the structure of the taxonomies in use. In terms of naming conventions the usage of “Lodging” instead of “Accommodation”, “Flight” instead of “Airlines” and “Transportation” instead of “Transport” illustrate the problem caused by different naming conventions. The exclusion of accommodation types in ontology B as well as land transportation that do not exist at all in ontology B, are the main causes for structural heterogeneity.

So it is clear from the examples above that ontologies are not the solution for mitigating machine processable data and resolving data heterogeneity. Ontologies alone cannot resolve data heterogeneity problems without commitment for shared vocabulary [1][2]. Ontologists, quite often do not agree on the same definitions, semantics and structure when developing their ontologies [6][8][15]. This causes two or more ontologies of similar contexts to be expressed differently as we have just discussed previously.

So, how can we share and reuse data in such a situation? To build a good travel web service we must be able to reuse existing travel definitions and schemas

without reinventing the wheel. Our solution this problem is to build a consistent information model across enterprise A and B where existing travel data definitions from their own local ontologies can be reused and shared. We approach this problem via ontology mediation and mapping. Ontology mediation is a process of finding a common ground for interoperability and ontology mapping is a process of producing concept and attribute matching. Now let’s take a look at case 2.

Currently, public service departments have implemented XML schemas with Web Service interfaces such as the Danish e-government project [21]. The effort is similar to that of maintaining a shared repository to ensure interoperability for all government systems by using the same schema language to avoid reusability problems of syntax specific definitions [3][4][7].

Schema interoperability guidelines are issued for this purpose and in most cases made mandatory to ensure all government agencies adhere to the same naming conventions. For example in England the schema has to be registered with UK GovTalk [22]. Examples of such schemas for addresses are CorrespondenceAddress, HomeAddress, BusinessAddress and ElectoralAddress. In some cases interoperability is only limited to central control such as the UN/CEFACT initiative [23].

This creates a barrier for inter-organizational services between public agencies of different domains outside that boundary. The lack of semantics causes data exchange to be impossible. For example figure 3 illustrates a data heterogeneity problem for inter-organizational services between public agencies where

a customer wants to renew his driver’s license online. The customer logs onto the DMV (Department of Motor Vehicle) portal and selects the type of service. Then he enters his full_name, DOB (10-10-1974), DMV customer ID (A11-03-5767) and address (3234 Hampton Rd Fairfax VA 22030). This data is passed onto a license renewal inspector to be verified. Also the payment mode is chosen and verified. Data collected is then sent to the DMV License Renewal server which validates and updates the renewal data. When payment is validated by the DMV License Renewal server, the records inspector receives this information with updates exchanged from DMV Records server. The bank is then notified for the charges and the customer’s account is debited.

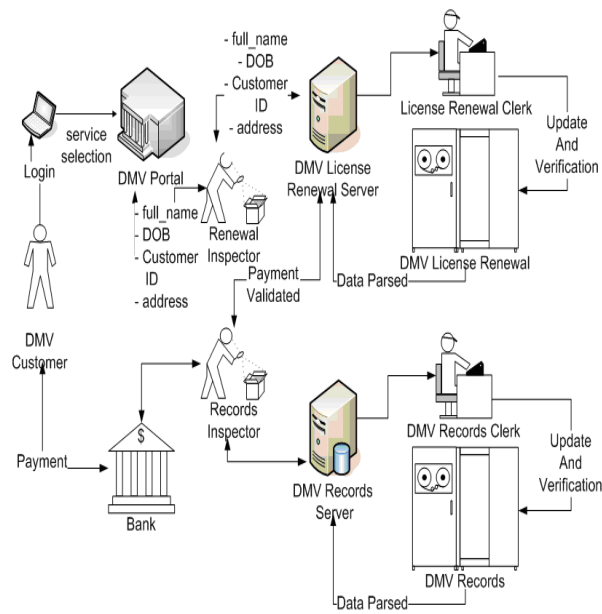


Figure 3. DMV License Renewal Process

The basis for data heterogeneity problems in the above example are clearly due to the different naming conventions by the two DMV agencies i.e. DMV Licence Renewal and DMV Records. For example, DMV Records maintains first_name, middle_name and last_name however DMV Licence Renewal maintains a complex string called full_name that literally combines first_name, middle_name and last_name. Also address in DMV Licence Renewal is treated as a complex string e.g. (“Hampton Rd Fairfax VA 22030”) and in DMV Records it is divided into street name, city, state and zipcode, e.g. (“Hampton Rd”, “Fairfax”, “VA”, “22030”). Since public agencies develop their own systems independently from each other, the granularity of how information is expressed can differ a great deal. As mentioned earlier having all agencies to adhere to one naming convention and making it

mandatory is not practical. This approach would be analogues to the federated schema approach (see figure 4) except that it does not have local schema mappings and a shared vocabulary. This approach is not practical as all local schemas have to be agreed upon and be integrated ahead of time. Sometimes they also have to be hard-coded. This does not allow local ontologies to be domain specific and maintain common data to exist at the higher level of the hierarchy. This approach also doesn’t produce semantically rich data because of the reasons stated earlier.

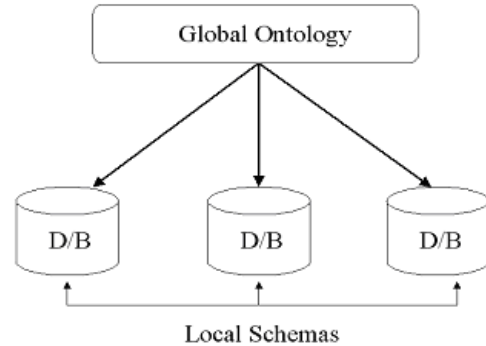


Figure 4. Domain Ontology Approach

Our approach is based on shared vocabulary and is quite similar to the federated approach but includes the source data in its own local ontology and shared vocabulary is achieved via inter-ontology mappings. A simple formalism is provided to capture only the domain knowledge pertaining to potential semantic conflicts (see figure 5). The advantage of our approach is that it is not only domain-specific but it also doesn’t lose semantic richness as it maintains local ontology definitions [14][15][16][17].

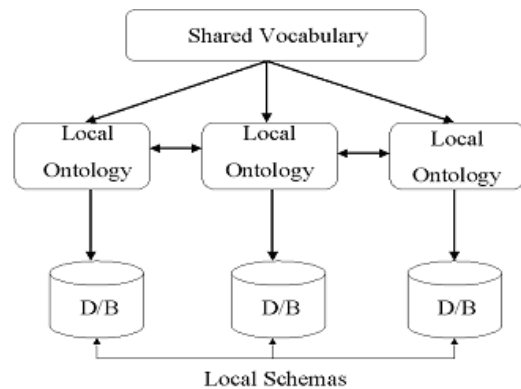


Figure 5. Hybrid Approach

This approach provides the semantic interoperability for inter-agency data exchanges and thus helps create the common information model that

is desired. To achieve this, we need to adopt a shared hierarchical information architecture.

2. Shared Hierarchical Structures

In the shared hierarchical structure [2] below, DMV License Renewal and DMV Records would maintain their own naming conventions for their customer records but at the same time borrow general concepts from an upper ontology. This helps facilitate knowledge reuse and allows domain experts to express their own specific definitions, even when they don't completely agree with each other (shown in figures 1 and 2 earlier). In this structure, knowledge is organized in different levels, each inheriting knowledge from upper level or parent ontologies. For multiple inheritance relationships it is important for knowledge inherited in local ontologies to be consistent with upper ontologies with no naming clashes.

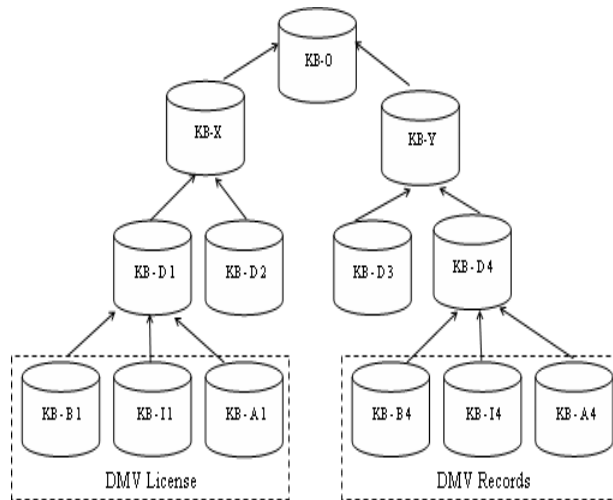


Figure 6. Shared Hierarchical Knowledge Repository

It will be impossible to find a perfect ontology to maintain rich definitions, as such a shared hierarchical structure is crucial [2][20]. Figure 6 illustrates a shared hierarchical knowledge repository that begins with knowledgebase O (KB-O) on the top. Two public agencies i.e. DMV License Renewal and DMV Records would have their own local knowledgebases that house their local ontologies e.g. KB-B1, KB-I1, KB-A1 KB-B4, KB-I4 and KB-A4. DMV license renewal contains KB-B1, KB-I1 and KB-A1 and would inherit knowledge from KB-D1 which inherits data from KB-X.

DMV Records contains and manages ontologies in KB-B4, KB-I4 and KB-A4. These ontologies inherit

data from KB-D4, which then inherits from KB-Y. Both KB-X and KB-Y inherit data from KB-O (see figure 6). DMV License Renewal and DMV Records are aware that their inherited knowledge is common to KB-O when they collaborate. Knowledge reuse becomes easier in this way. Particularly when creating a new ontology, we can determine which parent ontology to inherit from and start adding new knowledge to what already exists.

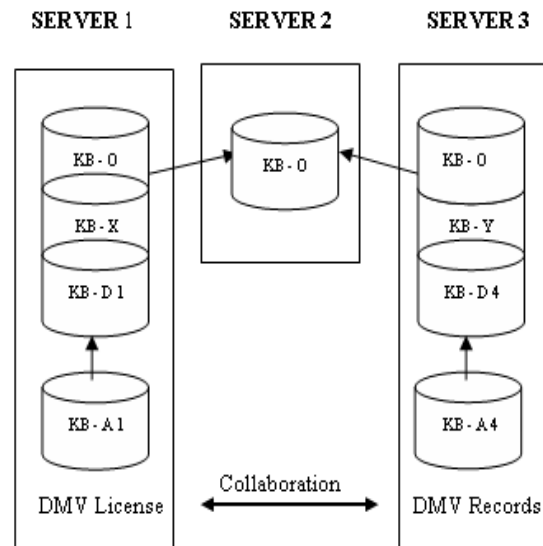


Figure 7. Distributed Shared Hierarchical Knowledge Repository

Although the hierarchical structure helps knowledge reuse, it will not be realistic to assume that all developed ontologies will be under one central control and available at all times [19]. As such, a distributed model of a hierarchical repository is more appropriate (see figure 7). There are three servers (i.e. server 1, server 2 and server 3) which are distributed. The servers could be maintained by other agencies within the DMV consortium or even other agencies. To solve the availability problem, when a new ontology inherits knowledge from another upper ontology, a copy of the inherited knowledge is made available locally. The reason for this is that the parent ontology does not have to be available online at all times to have its children ontologies to function properly.

3. Semantic Bridging Process Methodology

In this section we present our process methodology for semantically bridging the knowledgebases (see figure 8) [13][15]. The first step is called ontology development. It involves creating or selecting the source ontology (SO) and target ontology (TO). As

mentioned earlier DMV License Renewal could maintain its own set of definitions that is different from DMV Records. This is really a very important step for the ontologist as the domain specific local ontologies of public agencies e.g. DMV License Renewal and DMV Records have to be determined before mediation can be done. If DMV License Renewal is selected to be SO, then TO would be DMV Records.

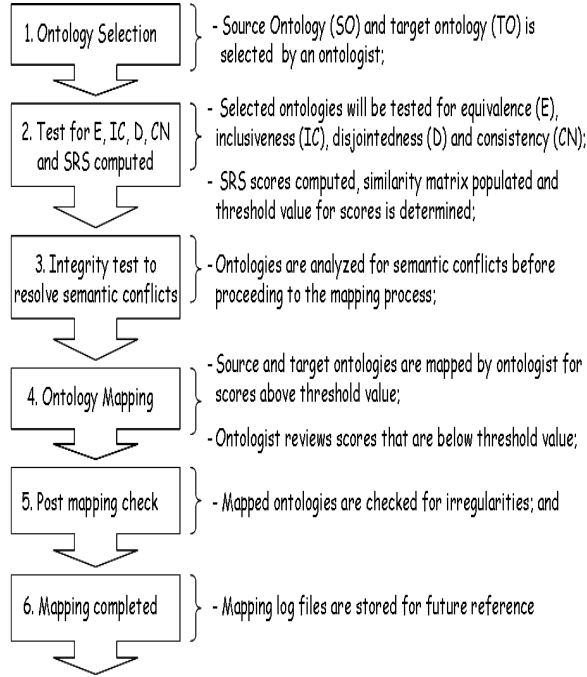


Figure 8. Semantic Bridging Process Methodology

In the second step, ontologies are checked for equality (E), inclusiveness (IC) and all disjoint (D) concepts are negated. We use the following symbols: 1) “C” for concept or class, 2) “c” for attributes or slots and 3) “O” for ontology, for simplicity [15]. The tests for E, IC, CN and D were based on definitions in [11]. The definitions are as follows:

- **Equality (E)**

Two classes (C) of DMV Licence Renewal and DMV Records are **equal** if, they: 1) have semantically equivalent data labels, 2) are synonyms or 3) have the same slots or attribute names. For example: 1) $C_1=Customer$ & $C_2=Customer$, 2) $C_1=Customer$ & $C_2=Client$ and 3) C_1 and C_2 have same slots (c) names, e.g. $c_1= \langle CustID, name, address, DOB \rangle$ and $c_2= \langle CustID, name, address, DOB \rangle$.

- **Inclusiveness (IC)**

Two classes (C) of DMV Licence Renewal and DMV Records are **inclusive** if, the attribute (c) of one is inclusive in the other. For example if $c_i = StreetAddress$ and $c_j = Address$, then c_i is a type of c_j . In other words *StreetAddress* is inclusive in *Address* $c_i (c_i \geq c_j)$. This is applicable to *hyponyms*.

- **Disjoint (D)**

Two classes (C) of DMV Licence Renewal and DMV Records are **disjoint** if, their attributes (c), c_i and c_j have nothing in common between them. For example *charge* and *name* are not equal or inclusive and have nothing in common which results in an empty set \emptyset .

Our Semantic Relatedness Score (SRS) is a hybrid matching technique that combines syntactic and semantic matching. SRS scores are populated into a similarity matrix which is then used as a basis to match the different schemas that the public agencies have defined in their ontologies. This is explained in detail in section 4. In the third step, the respective ontologies are tested for consistency. This is to ensure that the concepts that have been mapped are in fact consistent and that there are no conflicting concepts. We use a reasoning engine (i.e. RacerPro) [24] to check for these inconsistencies. If inconsistencies are discovered then they are resolved immediately. We define consistency as follows:

- **Consistency (CN)**

Two classes of DMV Licence Renewal and DMV Records are **consistent** if, all the attributes for $C_1 (O_1)$ i.e. *CustID*, *name*, *address* and *DOB* have nothing in common ($c_1 \sim CustID \neq c_2 \sim name \neq c_3 \sim address \neq c_4 \sim DOB$) s.t. $c_1 \cap c_2 = \{ \}$. All slots ($c_1 \sim CustID, c_2 \sim name, c_3 \sim address, c_4 \sim DOB$) must be subsets of $C_1 (O_1)$. This can be configured with RacerPro.

In step four, SO and TO are merged and integrated. At this point to ensure that schemas are perfectly matched data labels which have scores higher than the threshold score of 0.5 (e.g. $t > 0.5$) are matched first. Scores below the threshold are maintained in a log for the ontologist to refer to at a later point. SRS produces scores between 0 and 1. A score of 0 means that data labels have no match and 1 indicates a perfect match.

For instance if DMV License Renewal had the following schema, $\langle /Address \rangle$ and at the same time if DMV Records, had the following schema $\langle /Address \rangle$ it would result in a score of 1. We have developed a

detailed matching algorithm to illustrate and explain the process of arriving at the scores and also introduced a new measure for validation which is made up of the precision and relevance measure [15]. In step five, we use the same reasoning engine (i.e. RacerPro) [24] to check for post matching inconsistencies. This is to ensure that consistency is maintained even after data labels are matched. It also ensures integrity of matched data labels.

Lastly in step six, a log report is produced and published. This data is also annotated to document all changes that have been updated. We do this to provide other ontologists to trace the lineage of data that had been used to make any changes during the mapping process. In the next section we describe how SRS is determined.

4. Semantic Relatedness Scores (SRS)

We have researched thirteen well established linguistic and cognitive algorithms for finding the best measure for semantic correspondence. Some of the algorithms use pure syntactic (SYN) matching and others use semantic matching (SEM) only. Each algorithm was tested individually and then in combinations. The combination of these five algorithms i.e. Lin, Gloss Vector, WordNet, LSA (Latent Semantic Analysis) and SYN provided the highest the most accurate results [15]. SRS is a hybrid measure that comprises SYN and semantic matching SEM. Our experiments have proven that the combination of these five measures provide the highest reliability and precision [15].

Figure 9, shows that SRS had a higher precision and relevance score when matched against actual feedback received from human experts, i.e. human cognitive responses (HCR). An experiment was conducted with 30 word-pairs based on a study done at Princeton [6] to test the accuracy of our SRS. SRS had a 96.67% relevance score and a 40% precision score. These were higher compared to pure SYN algorithm scores which only had a 73.33% relevance score and a 16.67% precision score [15].

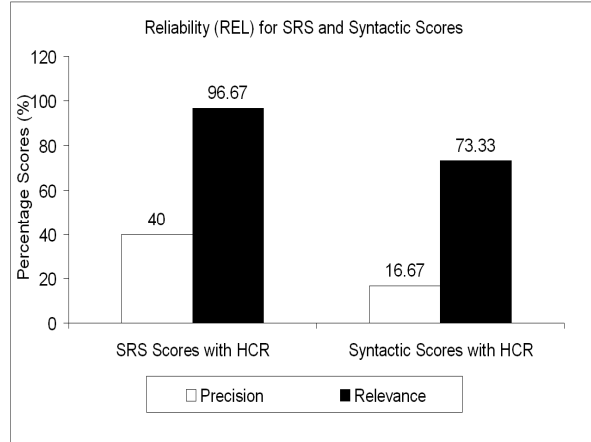


Figure 9. Semantic Bridging Process Methodology

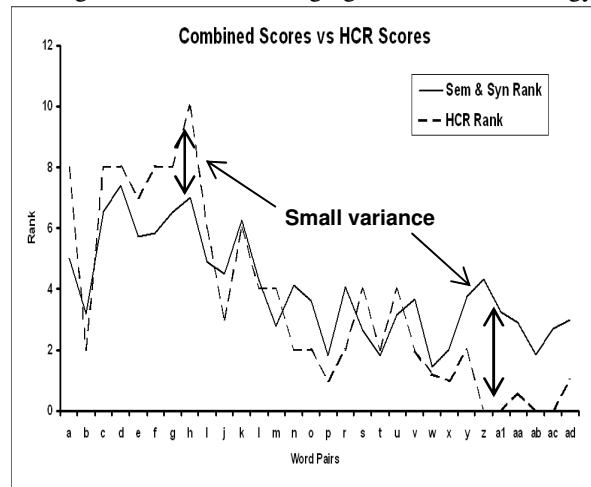


Figure 10. Variance for SRS scores and HCR scores

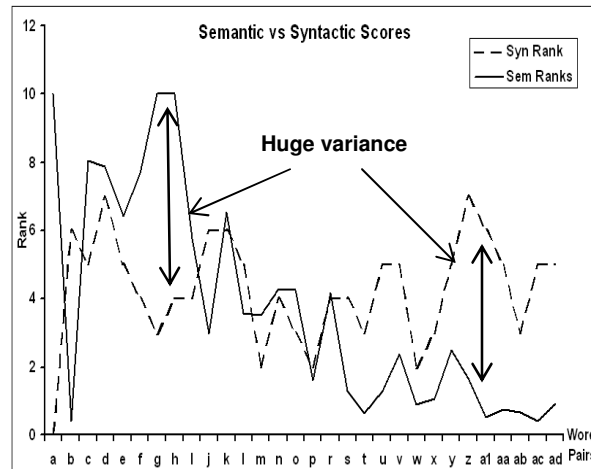


Figure 11. Variance for SYN and SEM scores

When tested for correlation, SRS had a positive correlation i.e. $r = 0.919$ (91.9%) with HCR scores and also had a smaller variance (see figure 10). SYN scores

when matched against HCR scores had a negative correlation and also had a higher variance (see figure 11). This validated that SRS scores were better than pure SYN scores for matching schemas. For the enterprise knowledge model to be efficient, SRS must be coupled with SWRL so that commonly used data schemas can be ascertained and matched ahead of time based rules without having to measure the scores every time that piece of information is used. This improves the efficiency and provides just-in-time data.

5. Bridging via SWRL

In this section we present how semantic bridging with SRS can be complemented with rules. SRS provides schemas that are likely to be matched with high reliability and precision [15]. Rules on the other hand are cardinality constraints that can be used for matching data labels, schemas and concepts. Rules can be predefined ahead of time so that frequently appearing schemas can be matched automatically [10].

To match schemas on names for instance we can write a rule that would match `</first_name>`, `</middle_name>` and `</last_name>` with `</full_name>`. If we had schemas for example `</street_name>`, `</city>`, `</state>` and `</zipcode>` to define an address in one domain ontology and defined as just `</address>` in another, a simple rule can be executed to match them on-the-fly. As mentioned earlier in section one and two, DMV Licence Renewal and DMV Records may have domain specific schema definitions that are different. Since establishing services between them will be an ongoing task rules could provide an automatic solution for creating homogeneity amongst heterogeneous schemas used by them [10] [18].

In view of the reasoning aspects that are possible in the Semantic Web, that supports web services, we propose an approach where rules can be used to match concepts. Reasoning is an approach when agents in a knowledge system perform tasks by inference [8][9]. Given the following statement “If X has a son Y, and X has a brother Z, given that X is a male” the agent is able to then infer that Y has an uncle, Z. The agent does not need to be explicitly told about the relationship between Y and Z. As long as uncle is defined earlier an agent will be able to infer this quickly. SWRL is based on OWL and RuleML (Rule Markup Language).

It enables OWL axioms to include Horn-logic that can be used to execute rules in a knowledgebase like the ones that public agencies will need to share (see section 2). SWRL rules show the implication between an antecedent (body) and consequent (head). In other words if the antecedent holds true, then the consequent

must hold true also. In our example earlier if antecedents “X has a son Y, and X has a brother Z”, X is a male” is all true then the consequent must also hold true which is “Y has an uncle, Z”.

With rules in place we can easily automate matching of schemas on-the-fly, which otherwise would be very labor intensive. As such SWRL rules and SRS would be complementary efforts towards semantic bridging. For example (see figure 12), DMV Records maintains `first_name`, `middle_name` and `last_name` however DMV Licence Renewal (see figure 13), maintains a complex string called `full_name`. Also, the address in the DMV License Renewal is treated as a complex string and in DMV Records it is divided into `street_name`, `city`, `state` and `zipcode`.

```
<?xml version="1.0"?>
<rdf:RDF
.....
<owl:Ontology rdf:about=""/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="Customer"/>
<owl:Class rdf:ID="First_Name"/>
<owl:Class rdf:ID="Middle_Name"/>
<owl:Class rdf:ID="Last_Name"/>
</rdf:RDF>
<owl:Class rdf:ID="Address"/>
<owl:Class rdf:ID="Street_Name"/>
<owl:Class rdf:ID="Zipcode"/>
<owl:Class rdf:ID="City"/>
<owl:Class rdf:ID="State"/>
<owl:Class rdf:ID="Customer"/>
</rdf:RDF>
```

Figure 12. DMV Records schema definition

```
<?xml version="1.0"?>
<rdf:RDF
.....
<owl:Ontology rdf:about=""/>
<owl:Class rdf:ID="Client"/>
<owl:Class rdf:ID="Full_Name"/>
<rdfs:domain rdf:resource="#Address"/>
</rdf:RDF>
```

Figure 13. DMV License Renewal schema definition

5.1. Expressing Rules in SWRL

In this section we show how SWRL rules are written to accomplish the bridging task to execute schema matching based on data labels, schemas and concepts rules for DMV Records and DMV License

Renewal. The first rule would be to associate `</Street_Name>`, `</City>`, `</State>` and `</Zipcode>` from DMV Records to `</Address>` in DMV Licence Renewal.

The following is how the rule 1 is expressed:

Rule 1:

```
hasStreet_Name ^ hasCity ^ hasState ^ hasZipcode
→hasAddress
```

This implies that the (Antecedent (hasStreet_Name, hasCity, hasState and hasZipcode. Thus the consequent would be (hasAddress). The following is how rule 2 is expressed for matching two concepts that are similar using “sameAs” as part of the rule. In this case the schema client and customer are really the same so a simple rule can be used to match them on-the-fly:

Rule 2:

```
<?Client>owl:SameAs<?Customer>
```

Rule 3 determines the full name by concatenating first_name, middle_name and last_name and is expressed as:

Rule 3:

```
hasFirst_Name ^ hasMiddle_Name ^ hasLast_Name
→hasFull_Name
```

6. Mapping to the Zachman Framework

The Zachman Framework defines a logical structure for classifying and organizing the descriptive representations of an enterprise. It showcases a two-dimensional six-by-six matrix that is represented by players represented by the rows (i.e. Planner, Owner, Designer, Builder) and processes (i.e. Data, Function, Network, People, Time, Motivation); represented by the columns. In order to build a consistent model for information and knowledge across the enterprise, we have to address the columns and they can be described as follows:

- **Data (what)**

This column answers the question “what and enterprise is made of?” It is clear from our discussions in previous sections as to “what data” of the enterprise is significant to be semantically bridged. From case 1 and

2 we can deduce the entity relationship involved in each perspective of the enterprise.

- **Function (how)**

This column answers the question “how does it work?” We describe our hybrid approach based on SRS and SWRL on how the actual mediation process takes place transforming data and mapping them on-the-fly to achieve the goal of the enterprise, that is to achieve a common and consistent information model across the enterprise.

- **Network (where)**

This column answers the question “where does it work?” Our hybrid approach based transforms data, maps them and performs the actual mediation process in any platform that connects service providers, government departments, vendors and offices in any geographic location worldwide as its technology is based on Web 3.0 and beyond i.e. the Semantic Web.

- **People (who)**

This column answers the question “who is responsible?” Our approach suggests that the responsibility for the workflow has to come from the people involved in the organization such as involved in the workflow of DMV License Renewal and DMV Records.

- **Time (when)**

This column answers the question “when?” which an important business rule. We suggest using the workflow sequences in the DMV License Renewal and DMV Records example to determine the turn around time to minimize processing time and to provide just-in-time information to clients and service providers.

- **Motivation (why)**

This column answers the question “why?” and as we have explained earlier the goal is to prevent from “reinventing the wheel”. Data has to be shared and reused easily among all the service providers, government departments and vendors.

7. Conclusion

Our approach provides a two part solution for building a common information model. Firstly, we introduce a detailed process methodology for semantic bridging

via specific tests of equality (E), inclusiveness (IC) negating disjoint (D) data and maintaining data consistency (CN).

This is also supported by an empirical test of SRS, a hybrid measure that combines i.e. linguistic properties and syntactic properties. We have also shown validation of the success of SRS via our reliability and precision scores. Secondly, we have provided SWRL rules to increase the efficiency while agents perform matching on-the-fly.

Lastly, we describe how our approach can be mapped to the Zachman Framework by describing in detail it's application in all six columns. We also imply that each column described is actually unique and that the columns have no order in terms of priority.

10. References

- [1] Ahmed-Macer, M., "Solving Interoperability Problems on a Federation of Software Process Systems, *Proceedings of the 6th International Conference on Enterprise Information Systems*, Porto, Portugal, April 14-17, 2004, pp. 591-594
- [2] Barbulescu M., Balan G., Boicu M., Tecuci G., "Rapid Development of Large Knowledge Bases" in *Proceedings of the 2003 IEEE International Conference on Systems, Man & Cybernetics*, Volume: 3, pp. 2169 - 2174, Washington, DC, October 5-8, 2003
- [3] Castillo, J. R., Silbescu, A., Caragea, D., Pathak, J., and Honavar, V.G., "Information Extraction and Integration from Heterogeneous, Distributed, Autonomous Information Sources- A Federated -Ontology-Driven Query-Centric Approach", *IEEE International Conference on Information Reuse and Integration*, 2003, pp. 183-191
- [4] Giunchiglia, F. and Zaihrayeu, I. "Making peer databases interact - a vision for an architecture supporting data coordination." *Proceedings of the Conference on Information Agents*, Madrid, September 2002
- [5] Gruber, T.R., *Toward Principles of Design Ontologies Used for Knowledge Sharing*, 1993
- [6] G.Charles, G. A. M. a. W. (1991). Contextual correlates of semantic similarity *Language and Cognitive Processes*, 6(1), 1-28.
- [7] H. Wache, T. Voegelé, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. Huebner. "Ontology-based integration of information - a survey of existing approaches". *Proceedings of IJCAI*, August 2001
- [8] H. Stuckenschmidt, U. Visser and H. Wache, ISWC-Tutorial, Sundial Resort, Sanibel Island, Florida, USA, October 20, 2003
- [9] Jerry, F., Brad, P., Marian, N., and Bruce, B., Agent-Based Semantic Interoperability in InfoSleuth, *SIGMOD Record* 28:1, March, 1999, pp. 60-67
- [10] Kurgan, L, Swiercz, W and Cios, K "Semantic Mapping of XML Tags using Inductive Machine Learning", Department of Computer Science and Engineering, University of Colorado at Denver.
- [11] Li, L., B. Wu, and Y. Yang. *Agent-Based Ontology Mapping Towards Ontology Interoperability*. in *Australian Joint Conference on Artificial Intelligence (AI'05)*. 2005. Sydney, Australia: LNAI 3809, Springer-Verlag.
- [12] Missikoff, M., Schiappelli, F., and Taglino, F., "A Controlled Language for Semantic Annotation and Interoperability in e-Business Applications", ", *Proceedings of the Second International Semantic Web Conference (ISWC-03)*, Sanibel Island, Florida, 2003, pp. 1-6
- [13] Muthaiyah, S and Kerschberg, L "Dynamic Integration and Semantic Security Policy Ontology Mapping for Semantic Web Services (SWS)", IEEE, *First International Conference on Digital Information Management (ICDIM'06)*, Bangalore, India, pp. 116-120
- [14] Muthaiyah, S and Kerschberg, L "Virtual Organization Security Policies: An Ontology-based Mapping and Integration Approach", *Information Systems Frontiers (ISF), A Journal of Research and Innovation*, Springer, USA, Special Issue on Secure Knowledge Management, 2007, pp. 505-515
- [15] Muthaiyah, S and Kerschberg, L "A Hybrid Ontology Mediation Approach for the Semantic Web", *Special Issue on Decision Technologies in E-Business, International Journal of E-Business Research (IJEER)*, Idea Group Publishing Inc, U.S.A, 2008. ISSN: 1548-1131, EISSN: 1548-114X
- [16] Noy, N.F.a.M.A.M. *PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment*. in *Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*. 2000. Austin, Texas: The MIT Press
- [17] Obrst, L., "Meditation and Data Sharing: Ontologies for Semantically Interoperable Systems", *Proceedings of the Twelfth International Conference on Information and Knowledge Management*, November, 2003, pp. 366-369
- [18] Park, J. and Ram, S., "Information Systems Interoperability: What Lies Beneath?" *ACM Transactions on Information Systems*, Vol. 22, No. 4, October 2004, pp. 595-632
- [19] Peter Haase, F.H., Zhisheng Huang, Heiner Stuckenschmidt and York Sure, *A Framework for Handling Inconsistency in Changing Ontologies* Springer-Verlag Berlin Heidelberg, 2005: p. 353-367
- [20] Tecuci G., *Disciple: A Theory, Methodology and System for Learning Expert Knowledge*, Thèse de Docteur en Science, University of Paris-South, France, 1988
- [21] <http://standarder.oio.dk/English/Guidelines/>
- [22] <http://www.govtalk.gov.uk/>
- [23] <http://www.unece.org/cefact/>
- [24] <http://www.racer-systems.com/>