COMMUNITY DRIVEN REQUESTS FOR PROPOSALS
Applying Semantics to Match Customer Purchase Intents to Vendor Offers

Christophe Debruyne
Semantics Technology and Applications Research Lab, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

Davor Meersman
Digital Ecosystems and Business Intelligence Institute, Curtin University, GPO Box U1987, Perth, WA, 6845, Australia

Mathias Baert, Rami Hansenne
iChoosr BVBA, Bourlastraat 3 bus 14, B-2000 Antwerpen, Belgium

Keywords: Product and RFP Ontology, Request for proposals, Semantic matching, Data integration.

Abstract: This paper presents a platform for requests for proposals and describes how ontologies drive the different components: the creation of a proposal, the annotation of vendor data, the transformation of vendor data into other formats and the semantic matching of a proposal against annotated vendor data. The ontology construction started from DOGMA, a methodology with its grounding in the linguistic representation of knowledge that is suitable for community participation in the creation process. The ontologies were created in a modular way, with general product and meta-models that can be extended depending on the domain. In the case of the pilot, the product were holiday packages, more precisely winter sports holiday packages.

1 INTRODUCTION

When consumers want to buy a certain item on the World Wide Web today, they have to browse through literally hundreds of offerings and results and this number is expected to increase in the future. In this model, the vendors drive the process by publishing products and providing means to buy these online. For example, travel agencies in the Netherlands need to query many different tour operators to find holiday packages meeting their customers' requirements. They often have an API that facilitates this process, but the granularity of the specific search is often limited due to the heterogeneous nature of all vendor databases.

A solution to this problem would be to allow the consumers to specify their requirements and to match these to offers of different vendors. Such a specification is called a Request for Proposals (RFP). The COMDRIVE RFP project resulted in a platform enabling consumers to drive the requirements process by expressing their intent to buy a certain product in a tool and language they are comfortable with. This platform sends out the request to a distributed vendor infrastructure, which responds to the request with offers.

The RFP, the heterogeneous vendor data, the matching of both and an appropriate interface for the customer all need to be driven by a agreed upon conceptual model, called an ontology. An ontology is commonly defined as: *"a [formal,] explicit specification of a [shared] conceptualization"* (Gruber, 1995) and are now key to enable semantic interoperability between information systems and services on the Web (Guarino, 1998). In general, interoperability is defined as the ability of two or more information systems or their (computerized) components to exchange data, knowledge or resources and to interpret the information in them (De Leenheer and Mens, 2008), in this case the RFP platform and the different vendor applications.

This paper presents the platform and explains how ontologies and semantic technology are key in achieving success. The content of this paper is organized as follows: Section 2 provides more context about the project and the motivation why such a portal should

Debruyne C., Meersman D., Baert M. and Hansenne R..
COMMUNITY DRIVEN REQUESTS FOR PROPOSALS - Applying Semantics to Match Customer Purchase Intents to Vendor Offers.
DOI: 10.5220/0003330905250530
In *Proceedings of the 7th International Conference on Web Information Systems and Technologies* (WEBIST-2011), pages 525-530
ISBN: 978-989-8425-51-5

be driven by ontologies. Section 3 presents the ontology engineering methodology adopted and applied within the project as well as how the resulting ontologies were applied to achieve the platform's tasks. Section 4 provides some preliminary results during the pilot before concluding and discussion our work in Section 5. Throughout this paper, the examples used stem from the domain of winter holiday packages (including winter sports, accommodation, facilities).

2 CONTEXT AND PROBLEM

iChoosr BVBA[1] is an internet company who specializes in supporting communities in organizing *group buying events* (GBE) in spirit of Vendor Relationship Management (VRM). VRM means providing customers with tools for engaging with vendors in ways that work for both parties, as it is difficult for vendors to fully relate with customers. These GBEs gives customers significant volume discounts. In return, iChoosr supports vendors in finding the appropriate customers, who also diminish their costs in logistics. A GBE always consists of a community identified by a particular need for a product (e.g., heating oil) who is represented by a community leader.

Milq Media[2] is the publisher of the *wintersporters.nl/* platform and agreed to take part as pilot partner in this project. Wintersporters.nl's content is characterized by its actuality and a large amount of objective information. With an average reach of more than 600.000 visitors per month, it is the largest winter sports platform in The Netherlands. The community of wintersporters.nl, with Milq Media acting as the community leader, was therefore chosen to be the pilot partner within this project.

The pilot of the platform consists of a personal RFP for travel in cooperation with the publisher of online travel platforms Milq Media. During this pilot, members of the Milq Media user community can set out Personal RFPs with the system and get offers from interested travel vendors. The matching engine automatically found some of these offers; other offers were matched manually by an employee of a travel vendor. The user and the vendor can also communicate through a message system. This allows them to further clarify the request and the offer. The main components of this platform (see Figure 1) are:

- **Personal RFP Ontology.** A semantic and extensible conceptualization of RFP. An ontology that describes the domain-specific Personal RFPs, allowing communities to contribute product- and domain specific concepts through collaborative knowledge engineering.
- **An Automated Group Buying.** (AGB) module that allows community leaders to organize their group buying activities online with their community members with support for member group buying process initiation.
- **Semantic Matching Engine.** Matching of customer intent and vendor offering based on shared, personal customer purchasing profile and community profile for accurate offerings based on predefined and implicit criteria. Matching of semantic product data and flat vendor data annotated with the ontology.
- **Rapid Semantic Node Cloud Navigation.** Dynamic node cloud underpinned by semantic product data for rapid navigation through correlated product concepts.

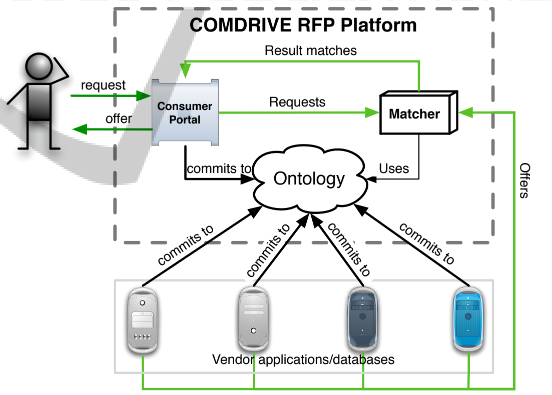

Figure 1: COMDRIVE RFP: The role of semantics in the annotation of vendor data and customer intent for semantic matching.

For this solution to be truly effective, buyers and vendors need to share a common vocabulary of the domain. More specifically, software agents need to interpret the information in RFPs to (semi-) automatically match this information with data in the vendors' product database based on their semantics. The semantic of the domain is provided by its conceptualization, which enables the alignment of the vendors' and buyers' perspectives. A conceptualization provides a shared agreement on the semantics of core concepts and the relationships between them imposing a structure on the domain that is readable by both humans and machines. In COMDRIVE RFP, the personal RFP ontology is thus key in the other modules' functioning.

[1] http://www.ichoosr.com/

[2] http://www.milq.nl/

3 APPROACH

The creation and application of ontologies within COMDRIVE RFP were constructed in an iterative manner. This section explains how the ontologies came to be and were then used to annotate vendor data, which were in turn used by the semantic matching to transform that data in a internal format suitable for the task.

3.1 Ontology Engineering Methodology

Many collaborative ontology engineering methodologies exists (Sure et al., 2009). We adopted DOGMA (Jarrar and Meersman, 2009), which stands out for its groundings in linguistics. DOGMA relies on the fact that its knowledge building blocks expressed in natural language are easily obtained and agreed upon (as inspired by database modeling methodologies such as NIAM (Wintraecken, 1990) and ORM (Halpin, 2008)). As a result, domain experts and knowledge engineers can use natural language to communicate and capture knowledge.

These building blocks - called *lexons* (Jarrar and Meersman, 2009) - only need in principle to express "plausible" facts (as perceived by the community of stakeholders) in order to be entered into the Lexon Base, a repository containing large sets of such lexons. A lexon is formally described as a 5-tuple ⟨γ, *headterm, role, co-role, tailterm* ⟩, where γ is an abstract *context identifier* pointing to a resource (e.g., a document on the Web). The context identifier is assumed to identify unambiguously (to human users at least) the concepts denoted by the term and role labels. Figure 3 shows some examples of lexons.

The *Commitment Layer* contains selections of lexons to annotate (types of) applications and constraints rules defining the use of the concepts in the ontology. In DOGMA, we distinguish two types of commitments: *ontologies* and *application commitments*. The first denotes a meaningful selection of lexons and rules that capture well the intended semantics of an application domain. The latter extends the community commitments mappings describing how the application symbols (e.g., fields in a table) of one individual application map to concepts in the ontology. The application commitment can furthermore contain additional lexons and constraints that describe how the application - as a whole - commits to the ontology. Figure 2 graphically depicts the different layers and Section 3.2 and 3.3 discusses respectively the construction of the first and the latter.

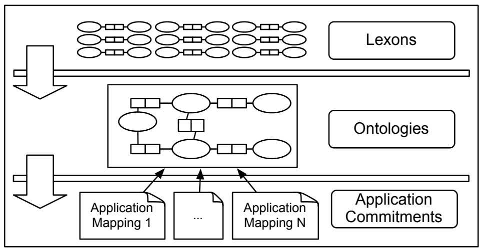

Figure 2: The different layers in DOGMA. Lexons in the Lexon Base are used to construct ontologies (with additional) constraints. Application commitments extends those ontologies with mappings of application symbols to the ontology.

3.2 Ontology Construction

Before building an ontology from scratch, one has to look around for existing meta-models. With meta-model, we mean any kind of conceptual model, not necessarily implemented with semantic technology (such as XML Schemas, or a code-based classification). (Hepp et al., 2007) analyzed and compared four important product meta-models: eCl@ss[3], UNSPSC[4], EOTD[5] and RosettaNet Technical Dictionary[6]. Both eCl@ss and UNSPSC are broad, but the first was were created by and driven by the German industry and thus a "de facto standard", whereas the latter was driven by the United Nations Development Programme. Both UNSPSC and eCl@ss provide very little detail for the traveling domain. The others were designed for more technical industries and did not fit the scope of this project.

Travel meta-models include Hi-Touch[7], OnTour[8], Harmonise (Dell'Erba et al., 2002) and the Open Travel Alliance (OTA) specification[9]. Harmonise focuses on accommodation and events (e.g., sports and conferences), but its main aim is to transfer data between tourism industry partners. Hi-Touch is a commercial thesaurus implemented in OWL to align different vendor databases. OnTour, a recent initiative, mainly covers accommodation and activities. OTA provides a structure for electronic messages, e.g., concerning flights, insurance, etc. Hi-touch and OnTour ontologies were developed based on international standards whereas OTA and Harmonise provide their own standards. We drew inspiration from these

[3]http://www.eclass-online.com/
[4]http://www.unspsc.org/
[5]http://www.eccma.org/
[6]http://www.rosettanet.org/
[7]http://www.mondeca.com/
[8]http://e-tourism.deri.at/ont/index.html
[9]http://www.opentravel.org/

meta-models to bootstrap an initial product ontology.

By our knowledge, no RFP meta-model exists. GoodRelations (Hepp, 2008) is a lightweight ontology to annotate most aspects of offers and their goods on the Web. Not exactly modeling an RFP, GoodRelations captures nicely what could be returned as a response or compared with a customer's intent. GoodRelation does provide a construct to describe what a business entity needs, but does not separate the concerns. It is thus possible for users to describe a product they are looking for and specify, for instance, a serial number.

We bootstrapped the product and RFP ontologies drawing inspiration from the existing meta-models and vendor applications, which were then refined and completed by several domain experts. We consulted such experts with different views on the domain, tour operators for the vendor perspective and a community of skiers for the buyer perspective. Figure 3 shows some of the lexons created for the purpose of this project. These lexons were built using Collibra Business Semantics Studio (BSS)[10].

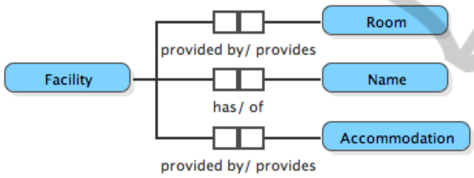

Figure 3: Some lexons describing domain knowledge developed during the project. These lexons describe that both rooms and accommodations can have facilities, which in turn have a name.

The ontology was developed in a modular way (see Figure 4). The *Upper Common Ontology* (UCO) contains the conceptualizations and semantic constraints that are common to and accepted by a general domain such as RFP and Product. For instance, the lexon ⟨γ, `RFP, has, of, Orderline` ⟩ states that an RFP has orderlines, is true for all applications of stakeholders within that domain and therefore belongs to that layer. The *Domain Application Ontology* (DAO) contains lexons specific to a certain application domain. In the case of COMDRIVE RFP, these lexons will contain the terms `Holiday Package` and `Accommodation`. The *Lower Common Ontology* (LCO) represents the interpretation of the domain from the perspective of an organization or community. For instance, the representation of a `Price` might change depending on the community: it is represented by a `Range` from a buyer's perspective, while it is represented as a `Value` from the Vendor's perspective. While the ontology evolves, this layer contains the information that is going to be refined by a core domain expert to be integrated in the UCO.

[10]http://www.collibra.com/products/

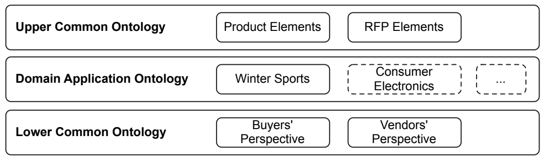

Figure 4: The modular structure of ontologies modeled with DOGMA within the COMDRIVE RFP platform.

3.3 Annotating Vendor Data

An application commitment represents an explicit interpretation of an ontology for an application or a family of applications. It consists of four things (Verheyden et al., 2004): (i) a selection of lexons from the ontology which are relevant for the application, (ii) constraints to specify how that selection can be used (mandatory and uniqueness constraints, for example), (iii) some scripting functionalities allowing database engineers/programmers to manipulate instances during the transformation process and (iv) a set of mappings between the application-specific symbols (e.g., XPath expressions) and the concepts used in the ontology. Figures 5 and 6 show how an application commitment of a vendor application constrains the use of lexons and the mapping to its application symbols. The transformation is performed by Collibra Business Semantic Enabler (included in BSS).

The effort needed to annotate the data of one vendor generally took one day when using BSS. The BSS commitment editor generates all occurring XPaths which the users need to annotate with lexons from the ontology by means of drag and drop. Constraints are created either via a graphical editor or a controlled natural language as shown in 5.

3.4 Creation of an RFP

A node cloud navigation is used as the user interface and lets users find the concepts from the ontology describing their wishes best. It is the result of an iterative design process involving several paper prototyping sessions and user tests during development. The interface (see Figure 7) consists of a list of starting points and a cloud of selected concepts, possibly with values assigned to them.

The starting points, chosen by the partner community since they know their members best, are there to get people started. After adding the concepts from the relevant starting points, the user can add more specific items using the search function. This is an instant

```
Holiday Package is identified by Name. Holiday Package has exactly 1 Name.
Holiday Package has at most 1 Description.
```

Figure 5: Constraining lexons.

```
map "/items/item" on Holiday Package. map "/items/item/title" on Name of Holiday Package.
map "/items/item/description" on Description of Holiday Package.
```

Figure 6: Mapping the application symbols to lexons. In this particular application, each "item" corresponded with one holiday package.

search-style search: while typing a keyword the result set is narrowed down. It usually takes only a few characters to find the concept one is looking for and also looks at possible values for these concepts (e.g., "France" for "Country") and synonyms (e.g., "Destination" for "Country"). Each time a user finds a relevant concept, he adds a node to the node cloud.

Figure 7: The user interface for creating an RFP driven by the ontology. The interface, driven by the ontology, aids the user in expressing their intent. In this example, the user is asked to give one (or more) possibilities for Ski Area ("Skigebied" in Dutch), Auto-completion relies on accessing the data through the application commitment.

3.5 Semantic Matching

The matching engine is an advanced search engine for analyzing structured, semi-structured and free text data. Contrary to classic searching and querying technology, matching functionality will take into account the semantic context of concepts to match upon and it will also return close matches if no exact matches can be found. The engine will score and rank the matches based on the degree in which they match, taking into account configuration parameters, e.g., weight, thresholds and optionality/requiredness of conditions.

The matching engine provides generic matching functionality and its interface is not specifically geared towards the notions of RFP or vendor offerings. Instead, from the match engine perspective, it simply matches *match queries* containing *match conditions* against a collection of *match objects*. A match property is a basic entity upon which can be matched. These match properties could be criteria such as price, age, distance, etc. A match object aggregates a number of match properties and assigns them specific values. Examples of match objects are: RFP, vendor offering, etc. A match condition represents the requested or "ideal" property. It is passed to the engine to compare with all known match properties that have the same id (and therefore same semantics) and to return all match objects which match the query conditions as close as possible.

A list of all properties known to the engine is defined in the match engines ontology mapping (Ω-RIDL file), which maps the relevant concepts in the ontology to the match engines internal format. The matching engine commits thus to the ontology in the same way as a vendor applications. The mapping used by the engine is quite straightforward and simply maps each "matchable" concept to a match property of the corresponding type, using the path in the ontology as ID. Figure 8 shows some mapping rules[11]. The transformation service mentioned in Section 3.3 takes care of transforming vendor offers into match objects.

4 PILOT AND OBSERVATIONS

The pilot ran in October 2010. It was agreed with Milq Media to encourage the wintersporters.nl community to test the interface to validate the ideas. Their forum provided feedback, which enabled the pilot to solve some of the initial bottlenecks (e.g., suggest a starting point when users don't know where to start). 38% of the RFP's where completed.

Users generally welcomed the idea, though the results were often not accurate enough. This was due to two reasons. First because of the heterogeneous nature of the vendor data streams; data that was present

[11] The complete Ω-RIDL mapping for the matching engine is available at http://starlab.vub.ac.be/website/files/matchobjects.ridl_.txt. Vendor data mappings within this project are not publicly available, but parts of it can be requested by contacting the authors.

```
    map "/matchObjects/matchObject/numericalMatchProperty[@id='Value of Price of Holiday Package']
/value" on Value of Price of Holiday Package.
    map "/matchObjects/matchObject/stringMatchProperty[@id='Currency of Price of Holiday Package']
/value" on Currency of Price of Holiday Package.
```

Figure 8: Mapping the application symbols to lexons.

with one vendor, was sometimes not present with another. Second, and what we considered the most important issue we have experienced, was a serious discrepancy between what concepts and level of granularity the user thinks are important in defining an intent and what is offered by vendors (e.g., "distance to the lift"). We have seen that merely building ontologies on top of the data that is provided by vendors does not solve the problem of finding products that match user needs. When given the liberty of defining ones wish, users demonstrate the desire to involve information that pertains to the product application domain. Although some of these concepts were present in the ontology, there was no data to work with.

5 CONCLUSIONS

This paper presented the COMDRIVE RFP platform that enables customers to define purchase intent, called an RFP, which can then be matched against annotated vendor data. All modules (RFP creation, vendor data annotation and transformation, semantic matching) are driven by an ontology with different layers of abstraction (a general RFP and product ontology and extensions for particular domains). The method for ontology creation was based on linguistics, which facilitated the dialogue with the different domain experts.

Future work includes the application of the upper common ontologies (RFP, and product) in other domains for validation and refinement as well as the application of the domain application ontology to other types of holiday packages (e.g., camping). In the future, a more prominent role for community aspects will be investigated. These can be useful to provide even better matching results, as different communities behave differently (e.g., younger people tend to put more emphasis on budget constraints).

ACKNOWLEDGEMENTS

The results of this research were partially funded by IWT[12] 090214 SME innovation project "COMDRIVE RFP". We also would like to thank Quentin Reul, Maarten Smolders, Tanguy Coenen for the fruitful discussions.

[12]http://www.iwt.be/

REFERENCES

De Leenheer, P. and Mens, T. (2008). Ontology evolution. In Hepp, M., De Leenheer, P., de Moor, A., and Sure, Y., editors, *Ontology Management*, volume 7 of *Semantic Web And Beyond Computing for Human Experience*, pages 131–176. Springer.

Dell'Erba, M., Fodor, O., Ricci, F., and Werthner, H. (2002). Harmonise: A solution for data interoperability. In Monteiro, J. L., Swatman, P. M. C., and Valadares Tavares, L., editors, *I3E*, volume 233 of *IFIP Conference Proceedings*, pages 433–445. Kluwer.

Gruber, T. R. (1995). Toward principles of the design of ontologies used for knowledge sharing. *International Journal of Human and Computer Studies*, 43:907–928.

Guarino, N. (1998). Formal ontology and information systems. In *Int. Conference On Formal Ontology In Information Systems FOIS'98*, pages 3–15, Trento, Italy. Amsterdam, IOS Press.

Halpin, T. (2008). *Information Modeling and Relational Databases*. Morgan Kaufmann, San Francisco, CA, USA.

Hepp, M. (2008). GoodRelations: An ontology for describing products and services offers on the web. In Gangemi, A. and Euzenat, J., editors, *EKAW*, volume 5268 of *LNCS*, pages 329–346. Springer.

Hepp, M., Leukel, J., and Schmitz, V. (2007). A quantitative analysis of product categorization standards: content, coverage, and maintenance of eCl@ss, UNSPSC, eOTD, and the RosettaNet Technical Dictionary. *Knowledge Information Systems*, 13(1):77–114.

Jarrar, M. and Meersman, R. (2009). Ontology engineering – the DOGMA approach. In Dillon, T., Chang, E., Meersman, R., and Sycara, K., editors, *Advances in Web Semantics I*, volume 4891 of *LNCS*, pages 7–34. Springer Berlin / Heidelberg.

Sure, Y., Staab, S., and Studer, R. (2009). Ontology engineering methodology. In Bernus, P., Blazewicz, J., Schmidt, Günter, J., Shaw, M. J., Staab, S., and Studer, R., editors, *Handbook on Ontologies*, International Handbooks on Information Systems, pages 135–152. Springer Berlin Heidelberg.

Verheyden, P., De Bo, J., and Meersman, R. (2004). Semantically unlocking database content through ontology-based mediation. In Bussler, C., Tannen, V., and Fundulaki, I., editors, *SWDB*, volume 3372, pages 109–126.

Wintraecken, J. (1990). *The NIAM Information Analysis Method, Theory and Practice.* Kluwer Academic Publishers.