Chapter XIV
Semantic Visualization to Support Knowledge Discovery in Multi–Relational Service Communities

Nadeem Bhatti
Fraunhofer IGD, Germany

Stefan Hagen Weber
Siemens AG, Germany

ABSTRACT

Services provided through the Internet serve a dual purpose. They are used by consumers and by technical systems to access business functionality, which is provided remotely by business partners. The semantics of services, multi-relational networked data and knowledge discovery in multi-relational service communities (e.g., service providers, service consumers, and service brokers, etc.) become an area of increasing interest. The complex multi-dimensional semantic relationship between services demands innovative and intuitive visualization techniques to present knowledge in a personalized manner, where community members can interact with knowledge assets and navigate through the network of Semantic Web services. In this chapter, the authors introduce Semantic Visualization approach (SemaVis) to support knowledge discovery by using hybrid recommender system (HYRES). It makes use of the semantic descriptions of the Web services, and also exploits the dynamic evolving relationships between services, service providers and service consumers. The authors introduce a sample scenario from a research project TEXO, within the THESEUS research program initiated by the German Federal Ministry of Economy and Technology (BMWi). It aims to supply a service-oriented architecture for the integration of Web-based services in the next generation of Business Value Networks. The authors present as well the application of their approaches SemaVis and HYRES to support knowledge discovery in multi-relational service communities of future Business Value Networks.
INTRODUCTION
In recent years there has been an enormous increase in interest in novel collaborative web sites on the one hand and semantically annotated content on the other hand. The fusion of collaborative and semantic data leads to complex scenarios with multi-relational data that contain several entity classes and multiple interactions. As an example consider some of the recent developments in the Web 2.0 that have resulted in novel types of social network communities represented as multi-relational data. Similarly, it is expected that the emergent Semantic Web will produce an enormous amount of structured relational data. Another example is B2B and B2C transaction data that is collected and analyzed to better serve the customers. Yet another example is enterprise applications based on service oriented architectures (SOA), containing a network of semantic Web services. Multi-relational networked data are now easily accessible and knowledge discovery in multi-relational domains has become an area of increasing interest. As a sample scenario, a platform for dealing with semantic Web services will be described. Providers as well as consumers are part of a service community realizing future business value networks.

The success of Web applications and platforms increases proportional with the amount of high quality content provided and collected from all customers. Supporting customers as most valuable assets therefore is a crucial aspect of a flourishing web business. Additionally, exploiting all existing data with leading technology to the customer’s benefit is one of the most important tasks to address. The main driving force for a satisfied customer in his community is uniqueness. The customer has to be given the feeling of being perceived as an individual rather than let him drown in the pool of anonymity. This can only be achieved by understanding dynamic social network structures, the customers themselves as well as their relations. Existing approaches from social network analysis or machine learning concentrate on one single relation type that exists between two entity types. In practice, however, entity types are arranged in multi-relational networks and the involved relation types are usually highly correlated. Using all correlated relation types simultaneously will improve the performance of the relation prediction. Siemens AG developed HYRES (HYbrid REcommender System), an easy to apply, scalable and multi-relational matrix factorization model able to deal with any number of entity types and relation types.

Knowledge discovery in multi-relational service communities is a great challenge. However, the best technological background processes will be in vain if the knowledge presentation cannot be accomplished in an intuitive and user-friendly way. The complex multi-dimensional semantic relationship between knowledge assets (e.g. services and related recommended services) demands innovative visualization techniques to present knowledge in a personalized manner, where community members can interact with knowledge assets and navigate through the network of semantic Web services. The semantic visualization techniques offer a very promising solution to support knowledge discovery in service communities.

Fraunhofer IGD and Siemens AG have developed an approach of how semantic visualization can support knowledge discovery in multi-relational service communities. According to this approach, when a knowledge worker searches for a service he starts a query. As a result, the knowledge discovery system delivers a list of services which then are presented within knowledge spaces or clusters, e.g. services for automotive industry or services for the entertainment industry. In order to distinguish services and related recommended services, different graphical metaphors within knowledge spaces or clusters will be applied. These recommendations are based on HYRES that predicts a collaborative ranking of services according to the user’s profile and favoured ser-
services (preferred by other community members with similar behavioural patterns).

Innovative service bundles and their (semantic-) relationships will be visualized within clusters (knowledge spaces). Personalized graphical metaphors are exploited on the base of collaborative exchange of the data (knowledge sharing) to increase the service usage or generate new business models for customers and complete the value network. Furthermore, the semantic visualization allows knowledge workers to navigate through knowledge spaces or clusters to explore through automatic generated intelligent, dynamic hierarchical cluster of like-minded customers. This approach allows them to recommend suitable groups where the user’s participation is most valuable and makes the customers feel unique and perceived.

The following subchapters describe our approaches for multi-relational semantic services selection and semantic visualization. Furthermore the deployment of this approach in the project “TEXO” Business Webs in the Internet of Services (IoS) as case study will also be presented.

SEMANTIC VISUALIZATION

It is not surprising that the field of Human Computer Interaction (HCI) has been closely bound to Moore’s law. At its inception, HCI was concerned primarily with the only large group of people who had access to the technology: office workers. This type of user had very clearly defined tasks and goals that they tried to achieve on fairly limited hardware. As Moore’s law resulted in computer technology leaking out of corporations, HCI morphed to accommodate home usage and to look at how people set about completing less well-defined tasks. Currently, HCI has expanded to look at social and even whimsical application of technology (Marsden, G, 2008). In the same way, semantic technologies are not limited to scientific community or companies any more.

The semantic technologies are becoming part of daily life, that is why different user friendly and easy to use HCI techniques have to be introduced to improve acceptance of users.

According to (Gruber, T. R., 1993), an ontology is an explicit specification of a conceptualization. The term “Conceptualization” is defined as an abstract, simplified view of the world that needs to be represented for some purpose. It contains the objects concepts and other entities that rare presumed to exist in some area of interest and the relations that hold between them. The term “Ontology” is borrowed from philosophy, where ontology is a systematic account of existence. For knowledge-based system what “exists” is exactly that which can be (and has been) represented.

Therefore, as defined in Noy, N. F., & Mcguiness D. L. (2001), an ontology is a formal explicit description of concepts, or classes in a domain of discourse. Properties—or slots—of each class describe various features and attributes of the class, and restrictions on slots (called facets or role descriptions) state conditions that must always hold to guarantee the semantic integrity of the ontology. Each slot has a type and could have a restricted number of allowed values. Allowed classes for slots of type instance are often called arrange of slot. An ontology along with a set of individual instances of classes constitutes a knowledge base.

An ontology is more than just a hierarchy of concepts. It is enriched with role relations among concepts and each concept has various attributes related to it. Furthermore, every concept most probably has instances attached to it, which could range from one or two to thousands. Therefore, it is not simple to create a visualization that will display effectively all this information and allows the user to easily perform various operations on the ontology at the same time (Katifori, A. & Halatsis, C., 2007).

The following section describes the semantic visualization fundamentals (e.g. Human Visual Perception and graphical metaphor) which play
a key role for the semantic visualization approach.

**Human Visual Perception**

According to Ware, C. (2004) a simplified information-processing model of human visual perception consists of three stages. In Stage 1, information is processed in parallel to extract basic features from the environment. In Stage 2, active processes of pattern perception pull out structures and segment the visual scene into regions of different color, texture and motion pattern. In Stage 3, the information is reduced to only a few objects held in visual working memory by the active mechanism of attention to form the basis of visual thinking.

In Stage 1 processing, billions of neurons work in parallel, extracting features from every part of the visual field simultaneously. This parallel processing proceeds whether we like it or not, and it is largely independent of what we choose to attend to (although not of where we look). It is also extremely fast. If we want people to understand information quickly, we should present it in such a way that it could easily be detected by these large, fast computational systems in the brain. Important characteristics of Stage 1 processing include:

- Fast parallel processing
- Extraction of features, orientation, colour, texture, and movement patterns
- Transitory nature of information, which is briefly held in an iconic store
- Bottom-up, data-driven model of processing

At the second stage, rapid active processes divide the visual field into regions and simple patterns, such as continuous contours, regions of the same colour, and regions of the same textures. Patterns of motion are also extremely important, although the use of motion as an information code is relative neglected in visualization. The pattern-finding stage of visual processing is extremely flexible, influence both by the massive amount of information available from Stage 1 (parallel processing) and by the top-down action of attention driven by visual queries. Important characteristics of Stage 2 processing include:

- Slow serial processing
- Involvement of both working memory and long-term memory
- More emphasis on arbitrary aspects of symbols
- In state of flux, a combination of bottom-up feature processing and top-down attention mechanisms
- Different pathways for object recognition and visually guided motion

At the highest level of perception, the objects are held in visual working memory by the demands of active attention. In order to use an external visualization, we construct a sequence of visual queries that are answered through visual search strategies. At this level, only a few objects can be held at a time, they are constructed from the available patterns providing answers to the visual queries. For example, if we use a road map to look for route, the visual query will trigger a search for connected red contours (representing major highways) between two visual symbols (representing cities).

**Visual Information Seeking Mantra**

There are many visual design guidelines but the basic principle might be summarized as the Visual Information Seeking Mantra “Overview first, zoom and filter, then details-on-demand” (Shneiderman, B., 1996). Based on ontology visualization characteristics, this section attempts an analysis of tasks related to ontologies. The categorization of tasks is based on the tasks analysis proposed by Shneiderman (Shneiderman, B., 1996), who presents high-level tasks
that an information visualization application should support.

1. **Overview**: Gain an overview of the entire collection.
2. **Zoom**: Zoom in on items of interest
3. **Filter**: Filter out uninteresting items.
4. **Details-on-demand**: Select an item or group and get details when needed.
5. **Relate**: View relationships among items.
6. **History**: Keep a history of actions to support undo, replay, and progressive refinement.
7. **Extract**: Allow extraction of sub-collections and of the query parameters.

*Overview* strategies include zoomed out views of each data type to see the entire collection plus an adjoining detail view. The overview contains a movable field-of-view box to control the contents of the detail view, allowing zoom factors of 3 to 30. Replication of this strategy with intermediate views enables users to reach larger zoom factors.

Users typically have an interest in some portion of a collection, and they need tools to enable them to control the zoom focus and the zoom factor. Smooth zooming helps users preserve their sense of position and context. Zooming could be on one dimension at a time by moving the zoom bar controls or by adjusting the size of the field-of-view box. Dynamic queries applied to the items in the collection to *filter* out uninteresting items are one of the key ideas in information visualization. Filtering allows users to quickly focus on their interests by eliminating unwanted items. *Details-on-demand* allows users to get detailed information on a selected item or group of items. The usual approach is to simply click on an item to get a pop-up window with values of each of the attributes. Viewing *relationships* among items (e.g. “book written by an author”) as detail-on-demand allows users to understand different interaction patterns.

Keep a *history* of actions to support undo, replay, and progressive refinement. It is rare that a single user action produces the desired outcome. Information exploration is inherently a process with many steps, so keeping the history of actions and allowing users to retrace their steps is important. Allow extraction of sub-collections and of the query parameters: Once users have obtained the item or set of items they desire, it would be useful to be able to *extract* that set and save it to a file in a format that would facilitate other uses, such as sending by email, printing, graphing, or insertion into a statistical or presentation package.

**Graphical Metaphors**

User interface design requires designing metaphors, the essential terms, concepts, and images representing data, functions, tasks, roles, organizations, and people. Advanced user interfaces require consideration of new metaphors and re-purposing of older ones. Awareness of semiotics principles, in particular the use of metaphors, can assist researchers and developers in achieving more efficient ways to communicate to more diverse user communities (Marsden, G., 2008).

Since Visicalc’s metaphorical ledger and the Xerox Star’s desktop metaphor, interface designers have been incorporating metaphors into user interfaces. User interface guidelines for most of the popular operating systems encourage the use of metaphors in interface design. They suggest that applications should build on the user’s real-world experience by exploiting concrete metaphors thereby making applications easier to use. Surprisingly little research supports the popular belief that metaphors in user interfaces facilitate performance, e.g. productivity.

The motivations for using metaphors in the design of user interfaces are similar to the reasons metaphors have long been popular in education. Many educators have observed that giving students comparisons can help them
learn. For example, an analogy commonly used in teaching about electricity is “Electricity is like water”. Imagine electricity flowing as water does. You can then imagine the wires as pipes carrying water (electrons). It follows that your wall plug can be thought of as a high-pressure source which can be tapped by inserting a plug. These types of comparisons are also used in teaching in the domain of human-computer interaction. For example, a physical metaphor for electronic storage is to think of “storage locations as buckets.” Experimental studies of the effectiveness of metaphor in teaching programming concepts have been conducted. Mayer showed that many programming constructs in BASIC (i.e., memory locations) could be learned more easily when they were presented in the context of a concrete metaphor. Thus, educators in many domains believe that students can import conceptual relations and operations from one domain to another.

**SEMAVIZ: A SEMANTIC VISUALIZATION TOOL**

“SemaViz” (Semantic Visualization), developed by Fraunhofer IGD, is a tool for ontology visualization, supporting different aspects, such as thematic co-occurrences, clusters, configurable domain-specific representations, and others. This section describes the semantic visualization approach of SemaViz, based on the semantic visualization fundamentals described in section 2. It also shows how the human visual perception and “visual information seeking mantra” were considered during the conceptualization of SemaViz and how they are supported by the use of various graphical metaphors.

Some suggested approaches by Colin Ware (Ware, C., 2004), which improve effectiveness of information visualization for users, are listed below:

- In any case where it is necessary to reveal fine detail, luminance contrast is essential.
- Low-level channels tell us about coding dimensions. It means that we can usefully consider colour, elements of form (orientation, size), position, simple motion, and stereoscopic depth as separate channels.
- For clearly related information, the visual structure should reflect relationships between data entities. Placing data glyphs (graphical object) in spatial proximity, linking them with lines. Or enclosing them within a contour will provide the necessary visual structure to make them seem related.
- In terms of seeing patterns in rather abstract data displays, perception of contours is likely to be especially important.

As described in section 2, ontology is more than just a hierarchy of concepts. The concepts contain role relationships, attributes, and most probably have instances attached to them. This means that instances are related to concepts and have relationships between them. According to the approaches suggested by Colin Ware, to use visual structure to reflect relationships, SemaViz has chosen circles as visual structure to represent concepts and graphical objects within circles to represent instances. The graphical objects enclosed within a circle provide a visual structure to show that an instance is related to concepts. The lines between instances represent the relationships between them. For example, Figure 1 shows three concepts (person, book and music) as circles, and instances as graphical objects within the circles. The relationships between instances are represented as directed edges. Furthermore, the colour coding represents different concepts and the size of circles represents the number of instance within one circle, respectively one concept.

According to the visual information seeking mantra, SemaViz is able to show all semantic
information as **Overview**, as shown in Figure 1. The semantic overview contains all concepts, instances and relationships. For example when users seek for a book “Information Visualization”, then SemaViz shows all concepts and instances directly (first order siblings) or indirectly (second order sibling) related to the book. Additionally, users have the possibility to navigate through the semantic network by clicking on instances. With this approach, users can start querying the system and explore through the semantic network representing the query results. The relationships between instances can also be coded with colour or thickness to represent different types or attributes of relationships. For example, the relationship `isSimilar` might have the attribute `similarityValue` indicating the degree of similarity between two instances. In Figure 1, Colin Ware is marked and an arrow clarifies that the person “Colin Ware” has written the book “Information Visualization”.

SemaViz offers zoom in and zoom out capabilities, enabling users to focus on specific semantic information, as shown in Figure 2. If a user is just interested in two concepts, e.g. book and person, then he can zoom in until the view is only focused on the desired semantic information. The control panel allows users to filter the visible information, and thus reduce the displayed semantic information, as shown in Figure 2. If a user wants to see just the books published in 2008, he can use a filter to do this, and to fade out all other books. Following another of Shneiderman’s suggestions, SemaViz allows users to get Details-on-demand by selecting (clicking) one instance, as shown in Figure 2. The popup window shows detailed information, e.g. description, picture or video.

Graphical metaphors can assist researchers and developers to achieve more efficient ways to communicate to more diverse user communities, as described in section 2.3. SemaViz uses the cluster metaphor (circle) to represent the concepts.
of an ontology, graphical objects within clusters to represent instances, and arrows to represent relationships between instances. This metaphor should help users to easily understand complex ontology terms, such as concepts, instances, properties and relations. Furthermore, the use of icons for instances should help users distinguish different instances of different concepts. For example, abstract icons for books and persons make it easier for users to understand the semantic of an instance, as shown in Figure 1. The usage of individual book covers and person pictures, instead of abstract icons, will help users additionally to understand individual semantic relationships with respect to his/her environment, but also increases the amount of information in the overview.

**MULTI-RELATIONAL SEMANTIC SERVICE SELECTION**

This section focuses on the domain of personalized service selection, in an environment where users are offered a multitude of services, from which they must choose the best fitting ones. We also consider in which way services can be semantically described, and what relations exists between them. Furthermore, it is pointed out that service selection based solely on semantic annotation is usually not sufficient. The addition of an improved recommender system and the description of the additional relational concepts complete the service selection approach.

**Semantic Service Description**

The well established usage of Service Oriented Architectures (SOA) (Erl, T., 2005; Newcomer, E., & Lomow, G., 2005) in B2C and B2B applications is rapidly expanding the focus of Internet. These days the Internet provides not only textual and multimedia information, but also thousands of Web Services (Booth, D., Haas, H. & McCabe, F., 2004). A challenging task is the discovery and selection of yet unknown services in the huge space of the WWW, fitting to user requirements (Verma, K., Sivashanmugam, K., & Sheth, A., 2005; Sreenath, R. & Singh, M., 2004). Universal Description, Discovery and Integration (UDDI) (Bellwood, T., Capell, S., & Clement, L., 2004) repositories were a first and important step towards the registration and discovery of Web services. However several problems could not be solved by this approach: Web Services are described via the Web Service Description Language WSDL (Christensen, E., Curbera, F., & Meredith, G., 2001) only syntactically and not semantically, UDDI allows only keyword and taxonomy based searches, multiple categorizations and identification schemas are used and no search for unknown services is provided.
Hence the semantic description of Web Services has been the logical extension to the existing standards (Sheth, A., & Meersman, R., 2002). Domain independent ontologies for flexible modelling of the web service capability descriptions are defined by upper level ontologies, including OWL-S (David, M., Burststein, M., & Hobbs, J., 2004), WSDL-S (Akkiraju, R., Farrell, J., & Miller, J., 2005) and WSMO (De Bruijn, J., Bussler J., & Domingue, J., 2006). All models provide terms, concepts, and relationships to describe various service properties like inputs, outputs, preconditions and effects (IOPE).

Based on such models, semantic matching algorithms can be applied to discover fitting services (Smeaton, A., & Quigley I., 1996); Klein, M., & Bernstein, A. (2003); Rodriguez, M., & Egenhofer, M., 2003) and to rank the results (Cardoso, J., & Sheth A., 2003). Several ranking methods can be used to calculate the degree of match. The logic-based approach uses description logics and first order logic reasoning to match semantic capabilities (Paolucci, M., Kawamura, T., & Payne, T. R., 2002). Similarity-based methods apply information retrieval techniques like linguistic similarity or term frequency for match making (Wu, J., & Wu, Z., 2005). Graph matching calculates a structural match where two service descriptions match if they have the same structure and the corresponding nodes match (Bellur, U., & Kulkarni, R., 2007).

However the IOPE properties of services are not sufficient to carry out efficient service discovery. Non-functional properties define what the service is about. They describe the service in terms of its descriptive metadata, such as a reference to a classification type, or characteristics that are not directly related to the functional description of the service. O’Sullivan, J., Edmond, D., & ter Hofstede, A. H. M., (2005) gives a detailed formal description for non-functional service properties. They introduce complexity to the description of services but their inclusion is crucial to the automation of service discovery, comparison and substitution. The OWL Web service ontology offers “placeholders” for the description of non-functional service properties, along with a minimal number of specific non-functional properties. In the context of the OWL-S profile, the non-functional properties of services are considered to be almost entirely domain specific. The Web Services Modeling Ontology (WSMO) uses Dublin Core metadata as the core properties, and then extends these to include some Web service properties. The non-functional properties outlined thus far are more indicative of the categories of non-functional properties than a specific non-functional property set. The model is extensible and aims at domain-specific inclusions. The Web services Description Language (WSDL) presents an entirely functional view of services and was not intended to attempt the description of the non-functional properties of services. Non-functional properties in OWL-S and WSMO are an initial set that can be expanded.

Examples of non-functional properties resulting from the Dublin Core Metadata Ontology are: name, contributor, coverage, creator, description, date, format, identifier, language, publisher, relation, rights, source, subject, title or type. Examples of non-functional properties regarding Quality of Service (Toma, I., Foxvog, D., Jaeger, M., 2006; KangChan, L., JongHong, J., WonSeok, L., 2003; Maximilien, E. M., & Singh, M. P., 2004) include: accuracy, network related QoS, performance, reliability, robustness, scalability, security and trust. Other non-functional properties mentioned by O’Sullivan, J., Edmond, D., & ter Hofstede, A. H. M., (2005) are: availability, maintainability, portability, payment, price, discounts, obligations, penalties, quality, location of provisioning, version, delivery terms, benefits, supported languages, delivery unit, owner, type of match or category.

Figure 3 depicts a semantically described Web service (OWL-S) with examples of non-functional properties.
Personalized Service Recommendation

The semantic discovery and ranking of services in the worst case still leads to an impersonalized list of hundreds of services. The personalized service selection is yet unaddressed. Web services can be traded as virtual products via the WWW. In the majority of cases product selection is done using a recommender system. Recommender systems are computer programs which attempt to predict items (movies, music, books, news, web pages or services) that a user may be interested in, given some information about the item or the user’s profile. Content based algorithms only use item information whereas collaborative filtering makes use of personal preferences and information of the user. Various recommender solutions can be applied to the service selection problem (Karta, K., 2005; Balke, W., & Wagner, M. 2003; Manikrao, U. S., & Prabhakar, T. V. 2005).

A pure content based recommender system takes the described functional and non-functional properties of a Web service as input to calculate a recommendation. Approaches using collaborative filtering make use of relations between the entities (Herlocker, J. L., Konstan, J. A., & Terveen, L. G., 2004). Hybrid solutions combine both methods resulting in improved prediction accuracy. The identified entities are users, services and actors. Obvious relations in a service ecosystem are listed in Table 1.

Service Ratings

The relation consumes could be used as an implicit feedback and interpreted as a user’s positive attitude towards the service. However, this behaviour

<table>
<thead>
<tr>
<th>Entity 1</th>
<th>Relation name</th>
<th>Entity 2</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>Consumes</td>
<td>Service</td>
<td>The user takes the role of a requester. A requester entity is a person or organization that wishes to make use of a provider entity’s Web service. It will use a requester agent to exchange messages with the provider entity’s provider agent.</td>
</tr>
<tr>
<td>User</td>
<td>Provides</td>
<td>Service</td>
<td>Here the user takes the role of a provider. The provider entity is the person or organization that provides an appropriate agent to implement a particular service.</td>
</tr>
<tr>
<td>Service</td>
<td>Orchestrates</td>
<td>Service</td>
<td>An orchestration defines the sequence and conditions in which one Web service invokes other Web services in order to realize some useful function.</td>
</tr>
<tr>
<td>Service</td>
<td>contactInfo contributor creator</td>
<td>User/Actor</td>
<td>Provides a mechanism of referring to individuals responsible for the service (or some aspect of the service). The range of this property is unspecified within OWL-S, but can be restricted to some other ontology, such as FOAF, VCard, or the now deprecated Actor class provided in previous versions of OWL-S</td>
</tr>
</tbody>
</table>
data does not necessarily accurately represent the user’s true opinion of a service. The implicit data collection does not demand any direct input of the user regarding his opinion. Instead, his actions are monitored and used as an indicator of his attitude towards the product. The implicit data collection can lead to a huge amount of data. Explicit data collection requires the user to rate content with an ordinal value. The users therefore are aware of the importance of their opinion which might lead to a more accurate feedback but less data available. Furthermore, the preferences can be divided into different aspects of the service to achieve a better understanding of the individual service consumers and their prioritization of certain aspects of the service itself. For example, one consumer might appreciate the performance of a certain service, while disapproving the availability and the security aspects. Since he puts more emphasis on performance, his overall opinion of the service still tends to be good. The above mentioned quality of service aspects are already described as non-functional properties of the service. The values of the non-functional properties only reflect the opinion of the service provider or an impartial partner. They cannot reflect the individual taste and rating of all the service consumers. Therefore, the preferences for these aspects have to be explicitly collected as ratings. Since there may be a large number of non-functional properties, a subset might be defined that is to be rated by the service consumers. In addition, it is not reasonable to rate all non-functional properties, e.g. for a service it makes no sense to rate Dublin Core properties like contributor, coverage, creator, date, format, identifier, language, publisher, source or subject. Note that the rating of all the different sub-aspects should be optional, as the user might be indifferent in his opinion towards some of them. Furthermore, the user should get the opportunity to describe his experience with a consumed service and his opinion about it in detail. This can be accomplished by offering a free text input field. Free text will not be further evaluated here since we focus rather on structured rating data than on unstructured data. Feature extraction algorithms, however, can be used to derive structured information from free text. Examples of additional relations resulting from non-functional properties are listed in Table 2.

Figure 4 depicts the described entities with their various interrelations. Each arrow symbolizes one relation type. To keep the figure layout clear, all the non-functional rating relations are subsumed in one relation called “X-Rating”. Dotted edges indicate attribute relations, dashed edges refer to rating relations and solid edges describe transactional relations.

Table 2. Examples for non-functional rating relations between user and services

<table>
<thead>
<tr>
<th>Entity 1</th>
<th>Relation name</th>
<th>Entity 2</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>overallRating</td>
<td>Service</td>
<td>overall impression of the service</td>
</tr>
<tr>
<td>User</td>
<td>accuracyRating</td>
<td>Service</td>
<td>precision and exactness of service results</td>
</tr>
<tr>
<td>User</td>
<td>performanceRating</td>
<td>Service</td>
<td>manner or quality of functioning</td>
</tr>
<tr>
<td>User</td>
<td>reliabilityRating</td>
<td>Service</td>
<td>worthy of reliance or trust</td>
</tr>
<tr>
<td>User</td>
<td>availabilityRating</td>
<td>Service</td>
<td>degree of being present and ready for use; at hand; accessible</td>
</tr>
<tr>
<td>User</td>
<td>priceRating</td>
<td>Service</td>
<td>cost/performance ratio</td>
</tr>
<tr>
<td>User</td>
<td>paymentRating</td>
<td>Service</td>
<td>quality and quantity of types of payment services offered</td>
</tr>
<tr>
<td>User</td>
<td>robustnessRating</td>
<td>Service</td>
<td>ability to operate despite abnormalities in input, calculations, etc.</td>
</tr>
<tr>
<td>User</td>
<td>scalabilityRating</td>
<td>Service</td>
<td>ability to either handle growing amounts of work in a graceful manner, or to be readily enlarged</td>
</tr>
</tbody>
</table>
Semantic Rating Description

The rating relation is not binary, but ternary. In Semantic Web languages, e.g. RDF and OWL, we have only binary relations (properties) between individuals, such as a property $P$ between an individual $A$ and individual $B$, more precisely, $P$ is the property of $A$ with the value $B$. However, every rating from the user to a service also has a rating value attached or an additional timestamp. Noy, N., & Rector, A., (2004) describe a solution to represent properties of a relation, such as rating value as follows: $P$ becomes a relation among user, service, and rating. A common solution to representing $n$-ary relations such as these is to create an individual which stands for an instance of the relation and relates the entities that are involved in that instance of the relation. The original relation then becomes a class of all these relation instances.

OverallRating_1 is an individual instance of the OverallRating class representing a relation:

```
:OverallRating_1 a :OverallRating ;
 :user :Tom ;
 :service :Ecocalculator ;
 :ratingValue :4.
```

The following figure shows the corresponding classes and properties. Each rating has exactly one rating person, exactly one rated service and exactly one rating.

All other rating relations can be modelled as sub-classes of the OverallRating class. They inherit all the properties of the super class. Figure 7 shows the RDFS schema with three sample ranking instances: user Tom rates the service “EcoCalculator” in three different ways: the overall rating is 4, the performance rating is only 1, whereas the accuracy rating is 5.
Figure 5. N-ary rating relation

![Diagram of N-ary rating relation](image)

Figure 6. Classes and Properties of N-ary rating relation

![Diagram of classes and properties](image)

Figure 7. RDFS-Rating hierarchy with sample instances

![Diagram of RDFS-Rating hierarchy](image)
HYRES: A Multi-relational HYbrid REcommender System

Recommender Systems traditionally focus on a two dimensional model, usually recommending items to users or vice versa. However it has been argued that this approach based on singular rating values is not sufficient in the context of semantic Web services. Multidimensional recommender systems provide a solution to this problem.

Siemens implemented a hybrid recommender system (HYRES) that makes use of the semantic descriptions of the Web services, and also exploits the dynamic evolving relationships between services, service providers and service consumers. HYRES is based on a novel algorithm called multi-relational matrix factorization (MRMF) presented in (Lippert, C., Weber, S., Huang, Y., 2008). Established matrix factorization models are focused on a single relation class connecting two entity classes. MRMF can handle an arbitrary number of entity types and relation types in a given domain and exploits multiple relation types simultaneously. The approach picks up the idea from (Yu, K., Yu, S., & Tresp, V., 2005) of combining more than a single matrix factorization in a single optimization criterion and generalizes it to an arbitrary number of matrix factorizations. By applying a gradient descent algorithm similar to (Takacs, G., Pilaszy, I., & Nemeth, B., 2007), (Rennie J., Srebro, N., & Jaakkola, 2005) and (Funk, S., 2006) the approach can handle sparsely observed input matrices and is highly scalable.

SEMANTIC VISUALIZATION FOR MULTI-RELATIONAL SEMANTIC SERVICE SELECTION

HYRES exploits semantic features of the Web services together with the interactions between services, service providers and service consumers, as described in Section 4 The basic idea is to combine the power of semantics and recommender systems to support semantic service selection. This approach will help users easily identify useful services, as well as interesting sub-communities. The advantages of this approach can only be achieved if the results of a semantic search and personalized recommendations are presented in an intuitive and user friendly way.

Different and partially concurrent aspects and roles have to be visualized coherently. The ontology concept of a person can take one or more of the following roles. Each role has to be clearly visualized to the knowledge worker.

- **Initial service searching consumer:** The knowledge worker himself should play a central role in the visualization. He is connected through various relations to services and other users in their particular roles, e.g. he might have relations to the discovered services, to recommended services, to self-provided services or to a recommended sub-community of likeminded users to ease valuable intercommunication.
- **Service provider:** A person who provides his own services also has to be reflected in the visualization. His relations are e.g. to his service consuming customers, to recommended potential customers, to competitors or to recommended value chain completing services.
- **Recommended similar user:** this relation shows users with similar service usage behaviour - be it as provider or consumer or a combination of both.
- **User interested in service:** The interest can be derived from two different sources. First the user shows his interest by using the service directly and rating it. Second a collaborative recommendation to the user can identify a potential interest.

Accordingly, the roles of a service can be defined as services retrieved by a semantic search, by
a collaborative recommendation or by similarity to other services. Furthermore users and services might be distinguished by a category, e.g. their industry sector. Figure 8 shows an example of a user Eric who has relations to services retrieved by his query, to recommended services and to likeminded users. Furthermore, service providers, interested users and consumed services are shown.

The combination of all the different roles leads to complicated scenarios to be visualized in an acceptable manner. On the one hand, ontologies have to reflect the actual roles of all participants or the ontology has to be enhanced with these concepts during the visualization process. On the other hand, sophisticated relations have to be simplified without any information loss. One example is the visualization of ratings which are

Figure 8. Example of user and service roles
described in Section 4.3. Ratings as n-ary relations are modelled as ontology concepts themselves. It is not worthwhile and neither catchy to visualize ratings as entities. Users will still want to perceive ratings as relations notwithstanding the ontological model.

Semantic visualization concepts, as described in chapter 2, have to be extended to meet the needs of recommender systems like HYRES. After a semantic search the results of the query will be presented according to the semantic visualization approach. For example, industry sectors are concepts of the ontology in use and services within an industry sector are ontology’s instances. In a second step SemaViz visualizes recommendations of HYRES (Multirelational HYbrid REcommender System) as additional information. This can be one or a combination of the following aspects:

- Historical ratings about used services of the current user can be shown. The relation can be labelled as “rates” and includes the rating value as an attribute of the relation. The ratings could further be split in two relations. One reflects the ratings with very high values and therefore describes the relation likes. The opposite relation dislikes consists of all the ratings with very low values. Mostly only the relation likes will be of interest to the user.
- Personalized recommended services are predicted by HYRES based on the historical service usage of the user. The recommendation is a prediction of the rating value the user would give to a service if he would consume it. This relation can be labelled as recommendedService. Again this relation could be split in “highly recommended” for high predicted rating values and “avoid” for low predicted rating values. We focus only on recommended services with high predicted rating values.
- The reverse relation to recommendedService is the relation potentialUser which is of interest for service providers. This relation is also predicted by HYRES.
- Similar services are predicted by HYRES based on the historical service usage of the services themselves, and on their semantic description. The recommendation is a prediction of the similarity between services. The relation is labelled similarService.
- Similar users are predicted by HYRES based on their historical service usage and their semantic description. The recommendation is a prediction of the similarity between users. The relation is labelled similarUser.

These new recommender system specific relations have to be added to SemaViz. The following table represents all additional relations.

To visualize the HYRES specific relations, a colour coding as shown in Table 3 is used. Since the knowledge worker is only interested in a subset of all available information, there has to be a way to dynamically arrange the desired visualisation. Depending on the user’s objectives the visualization has to be interactively adapted: relations and roles are shown or (partially) hidden and the level of abstraction can be zoomed as described in section 3.

CASE STUDY: TEXO, A SERVICE BROWSER FOR SERVICE COMMUNITIES IN INTERNET OF SERVICES (IOS)

TEXO³ is a research project, within the THESEUS research program initiated by the Federal Ministry of Economy and Technology (BMWi). It aims to supply a service-oriented architecture for the integration of web-based services in the next generation of Business Value Networks. The research focus of TEXO addresses the full lifecycle (“from innovation to consumption”) of business
services via intuitive interfaces and technical systems, and aims to provide a new Internet-based infrastructure to support the development, the use, the retrieval, and the access of services, and to improve their value of knowledge.

Services provided through the Internet serve a dual purpose: they are used by consumers and by technical systems to access business functionality which is provided remotely by business partners. The goal of TEXO is to provide a platform which makes services tradable on the internet, composable into value-added services, and allows the integration of customized services into the environment of service consumers. An example for such a service is described in the next chapter: the certified calculation of a product’s eco-value to guarantee compliance with new laws.

TEXO takes advantage of semantics and web 2.0 technologies like Community and Mashups to support the search and the use of services in the web. In collaboration with Siemens AG, Fraunhofer Institute for Computer Graphics (IGD) has developed the semantic visualization tool SemaViz for the service discovery and selection. The following section describes a TEXO scenario and application of SemaViz in the TEXO project.

**TEXO Sample Scenario: Eco Calculator**

The European Union has recently introduced a voluntary scheme to encourage businesses to market products and services that are environmentally-friendly and for European consumers - including public and private purchasers - to easily identify them with an Eco-label. The EU Eco-label has a clear objective of encouraging business to market greener products. With growing concern for the environment, consumers are becoming more aware of the need to consume eco-friendly products. This market dynamics puts pressure on manufacturers and producers. The EU has published a list of products that have eco-friendly labels. It initially started with simple products such as laundry detergent. However, recent additions include personal computers, televisions, and washing machines. So far, there is no mention of any automotive product. While the EU Eco-label is a voluntary scheme, increasing market pressure from consumers with environmental concerns has made many manufacturers and producers realize that eco-labelling their products makes good business sense.

The use case plays in the automotive manufacturing industry. A seat manufacturer has been asked by a car manufacturer, to offer and adapt one of their seat models for the new Eco-label. This sets prescriptive limits for e.g. recyclable material and environmentally friendly manufacturing processes. The seat manufacturer is already aware of the advantages of TEXO Business Web and has registered their company profile.
Service providers also notice the recent increase of queries for Eco-label services. First basic versions of this new service are offered by several providers who compose and orchestrate several existing services. Some even add newly developed own services. The resulting new service only fulfils the basic demand for calculating an Eco value and issue a certificate.

The seat manufacturer figures out they have to calculate and proof that the eco balance for the seat is above 85%. Since his software environment does not offer such calculations he drags and drops the whole government requirement text on the TEXO icon in the taskbar. The visual service discovery SemaViz is started as front end to all background processes that are retrieving the results to the user query. Figure 9 depicts the visualization of the query results. A service “Eco it” is found. Additional semantical information is displayed indicating that the service is provided by “York Sure” and that the service itself is composed of two other services “EcoCalculatorDocumentation” and “ProfessionalCalculator”.

Their providers are also displayed, e.g. the service “EcoCalculatorDocumentation” is provided by “Nadeem Bhatti”.

In order to better visualize the presented results, services can be ranked by several methods: by level of ability to be automatically integrated in the software and the system environment of the seat manufacturer, the purchasing conditions, the pricing model and by a personalized ranking factor indicating how well the new service would fit to the historical service usage behaviour pattern of the seat manufacturer. The last ranking method, calculated by HYRES, is one of the most important since it covers the actual hidden preferences of the company. For example, if several Romanian colleagues work in the software-lab, support from a Romanian speaking helpdesk is generally preferred. This information is not explicitly modelled anywhere. However there exists a fraction of likeminded users from the TEXO environment that already made their decision and experiences. They prefer a certain service that now gets a higher ranking for the seat manufacturer as a
consequence. Since in the above example only one service was found by the semantic query, the seat manufacturer would like to have a more diversified choice of services. One way to expand the list of services is to display all services that are used in the same way as the “ECO-it”-service. HYRES recommends some services that feature similar behavioural patterns as the “ECO-it”-service and therefore are highly correlated to it.

Now the seat manufacturer has a better overview and decides to use one of the recommended services. To use it, the user accepts the displayed conditions and the software is automatically integrated in his working environment; in this case, extending the parts list table with a column for the “eco value”, displaying a button “eco calculator” in the navigation bar and a button “calculate eco balance” in the product & design tool. After running the calculator, the parts list table shows that the overall eco value is shortly under the required value of 85%, barely missing the norm.

At this point the manufacturer would like to know some more details about the root causes of the bad “eco value”. Therefore he needs some more services that complete the Eco Calculator service in its complexity. Initially he has no further keywords to query for. Again, likeminded users might have solved this problem for him. So the seat manufactured uses SemaViz to explore various semantic relationships. First he requests a set of top ten recommended services. Furthermore SemaViz displays all other consumers of the Eco Calculator service along with their used services. Immediately, the manufacturer notices that two of the recommended services are also broadly used by other consumers using the Eco Calculator. One delivers information about the amount of carbon dioxide released during the production process of polymethane cellular plastics used in the seats. Another service gets the detailed chemical composition of the used type of polymethane cellular plastics and the environmental compatibility of each component. Now the seat manufacturer has all information in his hands to start the improvement of his product. In the end the certificate

Figure 10. Similar services calculated by HYRES
can be downloaded and the service usage can be stopped via TEXO Business Web.

It may be that at some point the seat manufacturer still has some open questions regarding the new law concerning the eco-value regulations. There is no service that will answer his demand. However, a dynamic and agile community is available and ready to answer. HYRES finds a sub-cluster of service community members fitting to the seat manufacturer’s service usage profile and recommends the found users the manufacturer. Now he has all information on his fingertips.

The service provider at the other side is aware that the Eco Calculator offered by him is completed by other services on the customer side. Now he has several alternatives to explore service usage relations. First he requests a list of all recommended services for him as service provider. Second he shows all service consumers of these recommended services together with the services they are already consuming from his offered services. This combination clearly shows up what services really have the potential to complete his value chain of offered services. He updates his offer range and now can cover a broader bandwidth of his customer’s demands. To increase his customer base with his now completed value chain he decides for letting TEXO search for potential customers. Again HYRES predicts a list of service consumers that would need the offered services with a high probability. He can also get an overview about what his competitors offer. Other service providers however might only be displayed anonymized. Figure 11 shows a possible display of potential Users. The provider can now make a dedicated offer to e.g. “Stefan Weber”.

SemaViz harmonizes all these interactive service browsing activities in one easy to use graphical interface. It allowed users also to filter the services according to their system ranking and user rating with the filter panel as shown Figure 11. Users can also decide which concepts and relationships have to be visualized to reduce information and cognitive overload. Furthermore, details on demand about services will be offered

Figure 11. Recommended Customers are displayed for campaign management
by double click on services or by using different levels of zoom.

**CONCLUSION AND FUTURE WORK**

A number of important challenges for visualization and service selection research have been answered. Given an increasing quantity of available relational information, it is essential to use powerful tools for discovering and visualizing accurate information from semantic service communities. We have designed and implemented a graphical tool for the visualization of semantic metadata for the use of multi-relational service selection. SemaViz is capable of visualizing multi-relational data in an interactive and user-friendly way. The interactivity and appropriate visualization greatly helps the user to analyze relationships in service communities and browse complex semantic descriptions. The recommendation system HYRES enriches knowledge discovery in service communities by personalized recommendations of services and sub-communities. Recommendations based on machine learning methods help selecting services and open up business opportunities.

Other challenges still have to be approached. Among them is the inclusion of additional relations and other views like cushion trees and geographical visualization (e.g. via Google maps). Although scalability has been proven for the recommendation system (Lippert, C., Weber, S., Huang, Y., 2008) it has to be shown that the visualization approach is scalable for thousands of instances. Furthermore, similarities of services/users might be displayed using graphical means by displaying more similar entities nearer to each other. This will be a significant improvement of visual cognition.

Extensions to SemaViz to enhance the ease of use include the dynamic personalized view adaption, e.g. the interactive selection of relation types and the possibility to define dynamic fading in and out of relations over a defined number of edges. This can be done by more than one instance (e.g. two users simultaneously).

**ACKNOWLEDGMENT**

The project was funded by means of the German Federal Ministry of Economy and Technology under the promotional reference “01MQ07012”. The authors take the responsibility for the contents.

**REFERENCES**


Semantic Visualization to Support Knowledge Discovery in Multi-Relational Service Communities


ENDNOTES

1 These bundles consist of various services of different kind that are orchestrated together in an innovative way.
2 http://www.intel.com/technology/moore-slaw/
3 http://www.baddesigns.com/mswebenf.htm
4 http://dublincore.org/
5 URL: http://theseus-programm.de/scenarios/en/tezo