Personalized View-Based Search and Visualization as a Means for Deep/Semantic Web Data Access

Michal Tvarožek
Institute of Informatics and Software Engineering
Slovak University of Technology
Ilkovičova 3, 842 16 Bratislava, Slovakia
tvarozek@fiit.stuba.sk

Mária Bieliková
Institute of Informatics and Software Engineering
Slovak University of Technology
Ilkovičova 3, 842 16 Bratislava, Slovakia
bielik@fiit.stuba.sk

ABSTRACT

Effective access to and navigation in information stored in deep Web ontological repositories or relational databases has yet to be realized due to issues with usability of user interfaces and the overall scope and complexity of information as well as the nature of exploratory user tasks. We propose the integration and adaptation of novel navigation and visualization approaches to faceted browsing such as visual depiction of facets and restrictions, visual navigation in (clusters of) search results and graph like exploration of individual search results' properties.

Categories and Subject Descriptors: H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval; H.5.2 [Information interfaces and presentation (e.g., HCI]: User Interfaces—*Graphical user interfaces (GUI)*; H.5.4 [Information interfaces and presentation (e.g., HCI]: Hypertext/ Hypermedia—*Navigation*

General Terms: Algorithms, Design, Experimentation

Keywords: navigation, visualization, view-based search, personalized faceted browsing, graph visualization

1. INTRODUCTION

Presently, effective access, navigation and user understanding of Semantic Web and deep Web resources has yet to be realized due to various issues with user interface complexity, query languages and usability among others. View-based search approaches have already been proposed as suitable means for integrated search and navigation in various application domains. Since both Semantic Web and deep Web repositories describe the structure of the respective information spaces via metadata, faceted browsers can be used as effective entry points providing users with powerful interfaces for querying, navigation and visualization of information.

We proposed the concept of an adaptive faceted browser to address issues concerning the size, complexity and user diversity of open information spaces in [3]. The faceted browser is personalized based on an automatically acquired user model using dynamic generation and adaptation of facets and restrictions according to user characteristics. However, our original approach was, as other existing faceted browsers, purely text-based, while many related approaches to searching and navigation focus on visual methods of navigation and visualization with promising results.

WWW 2008, April 21–25, 2008, Beijing, China.
ACM 978-1-60558-085-2/08/04.

Authors in [4] describe *CropCirles* – an approach to the visualization of OWL class hierarchies. The approach visualizes hierarchies as nested circles of different sizes and layouts in order to improve user understanding of the respective information domain. The ideas proposed by this approach can also be used to improve faceted browsing of deep faceted hierarchies, which are difficult to understand for many users.

TagSphere is an approach to visual presentation of search results originally developed for the digital image domain [1]. It presents search results visually in sets based on their associated tags. For each set, its similarity to the query is shown via distance and the overlap between a tag search and a classifier search based on low-level image properties. Applying a similar approach to faceted browser search results presents a great improvement in user understanding of the information space compared to the typical table-based list of search results displayed by faceted browsers.

2. VIEW-BASED SEARCH VISUALIZATION

We proposed an extension of our original adaptive faceted browser with support for advanced visualization techniques for facets, search results and instance details, and their respective personalization, taking into account the user's context (i.e., preferences, social networks or environment).

2.1 Facet and restriction visualization

We provide a search window for each facet, where users can type labels of the desired restrictions for quick access. Additionally, we employ three modes of facet visualization:

- *Text-based facet visualization* – all restrictions are shown via their respective textual labels. This is useful for enumerations, shallow hierarchies or if there are only few restrictions.
- *Graphical facet visualization* – all restrictions are shown graphically as nested circles/nodes. Size indicates the number of instances, color is used for recommendation, while their relative distance denotes similarity. Individual circles are annotated with labels (e.g., directly or via tooltips). Other shapes, icons or images can also be used to more closely describe facet contents (e.g., for multi-type queries). This mode is suitable for complex hierarchies or facets with many restrictions.
- *Hybrid textual and graphical visualization* – restrictions are shown both in textual and graphical form, which is suitable for low to medium number of restrictions yet forming a complex hierarchy.

2.2 Search results visualization

In addition to traditional search results tables and matrices with attributes, we employ a graphical overview of the returned search results based on a hierarchical clustering visualization (see Fig. 1). One central root cluster contains search results satisfying the current faceted query. One or more nested levels of clusters are shown corresponding to search results based either on a hierarchical attribute from the classification or on a custom clustering function. Individual clusters are annotated with short labels summarizing their contents.

Additionally, examples of instances are shown as tooltips, while the size, relative layout and color provide further information about instance counts, similarity, relevance (e.g., via a user's social network) and overall suitability (e.g., via user characteristics). A second layer of clusters can be optionally used to present "fuzzy" search results which do not satisfy the faceted query, yet are reasonably similar and thus possibly of interest to the user (similarly to [1]).

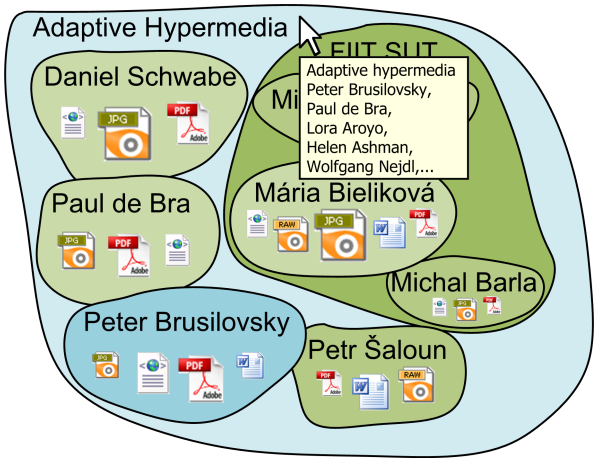

Figure 1: Search results visualization using hierarchical clusters.

2.3 Instance details visualization

To improve user understanding of individual search results, we proposed three visualizations of instance details:

- *Textual attribute visualization* – individual attributes of the viewed instance are displayed in a nested table with a predefined or adaptive nesting depth. The nesting is used for complex data types, e.g. object properties for OWL ontologies.

- *Graphical neighborhood visualization* – the context of a search result instance is displayed via a hierarchical cluster view of similar instances based on its properties (e.g., photos with the same author, job offers offered by the same company) or an aggregate clustering function. Size, colors and relative positions of clusters are used for further annotation along with tooltips similarly as with facets and restrictions.

- *Graphical attribute visualization* – the attributes of the search result instance are displayed as a graph, which can further be expanded to show attributes of associated instances. Simple data type attributes are shown directly in nodes, while object type attributes correspond to other nodes connected to each other via edges corresponding to properties (for OWL data). Furthermore, the center can be moved from the original instance to another instance effectively resulting in a moving window showing the information space. The neighborhood of individual attributes can be shown, similarly to the previous visualization based on instance types (e.g., similar instances of the same type can be shown in hierarchical clusters) (see Fig. 2).

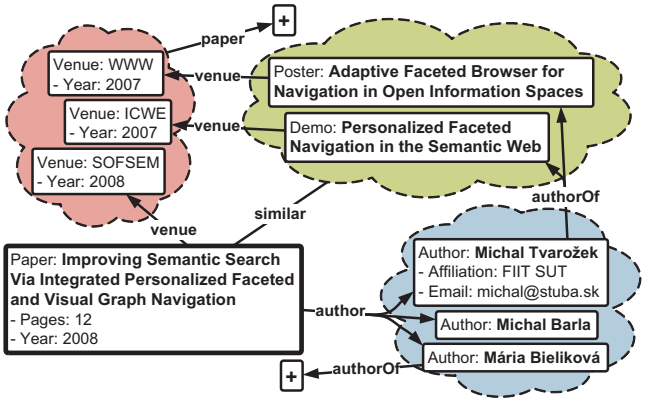

Figure 2: Graphical instance attribute visualization.

3. CONCLUSIONS

We have presented a novel method of navigation and visualization for faceted browsers, which are suitable for dealing with large complex information spaces such as Semantic Web metadata represented by OWL ontologies, or deep web data stored in large repositories. An important feature is the dynamic generation of facets and their visualization that enables access to these large information spaces without prior knowledge about their structure.

We extended the concept of faceted browsing with support for graphical depiction of facets, search results and instance attributes while also applying our personalization principles to the new visualization approach. Our extensions improve overall user experience and the understanding of the respective information space especially for open-ended exploratory tasks. We evaluate our extensions in three domains – job offer domain, digital libraries and photo galleries.

Acknowledgments. This work was supported by the Slovak Research and Development Agency, contract No. APVT-20-007104, the State programme of research and development, contract No. 1025/04 and the Scientific Grant Agency of Slovak Republic, grant No. VG1/3102/06.

4. REFERENCES

[1] M. Aurnhammer, P. Hanappe, and L. Steels. Augmenting navigation for collaborative tagging with emergent semantics. In I. F. Cruz et al. [2], pages 58–71.

[2] I. F. Cruz et al., editor. *Proc. of the 5th Int. Semantic Web Conference*, volume 4273 of *LNCS*. Springer, 2006.

[3] M. Tvarožek and M. Bieliková. Adaptive faceted browser for navigation in open information spaces. In C. L. Williamson et al., editor, *WWW*, pages 1311–1312. ACM, 2007.

[4] T. D. Wang and B. Parsia. Cropcircles: Topology sensitive visualization of owl class hierarchies. In I. F. Cruz et al. [2], pages 695–708.