Factic: Personalized Exploratory Search in the Semantic Web

Michal Tvarožek and Mária Bieliková

Institute of Informatics and Software Engineering, Faculty of Informatics
and Information Technologies, Slovak University of Technology
Ilkovičova 3, 842 16 Bratislava, Slovakia
{Name.Surname}@fiit.stuba.sk

Abstract. Effective access to information on the Web requires constant improvement in existing search, navigation and visualization approaches due to the size, complexity and dynamic nature of the web information space. We combine and extend personalization approaches, faceted browsers, graph-based visualization and tree-based history visualization in order to provide users with advanced information exploration and discovery capabilities. We present our personalized exploratory browser Factic as a unique client-side tool for effective Semantic Web exploration.

1 Introduction and Related Work

The Web has become an almost ubiquitous virtual information space where information is stored, published and shared in many different forms. To address problems associated with information access on the Web (e.g., the navigation problem, complex [semantic] query construction, information overload), several directions and initiatives are being pursued such as the Semantic Web [1], Adaptive web-based systems [2] or Exploratory search [3].

In our previous work, we have demonstrated the practicality of adaptation in a faceted browser by applying personalization principles to a faceted semantic browser thus improving overall user experience when working with an ontological repository [4]. While our previous approach provided good personalized navigation and query construction support, it offered limited support for visual information discovery, e.g., using different widgets to render topic-, time- and location-based information as demonstrated by VisGets [5].

Furthermore, it had to be configured for a specific domain model and thus was not seamlessly applicable across multiple (open) domains at the same time. Consequently, we focused on improving two aspects of our solution – *exploratory search enhancement* and *automatic user interface generation* based on domain metadata in order to improve *end user experience* when accessing information in the Semantic/Deep web, which is typically inaccessible via search engines.

2 Personalized Exploratory Search Browser

Our original personalized faceted browser *Factic* [4] performed faceted browsing and personalization to adapt facets, i.e. order, hide and annotate them while also

recommending restrictions based on user characteristics. To further improve user experience, we completely reworked *Factic* to facilitate two new features:

– *Multi-paradigm exploration* – combined view-based (faceted), keyword-based, content-based exploration with multiple adaptive result overviews and collaborative resource annotation (see Fig. 1), graph-based visualization and browsing of information (see Fig. 2) and custom/external content rendering.
– *User interface generation* – (semi)automated generation of the browser's user interface based on semantic metadata from the domain ontology to facilitate use in multiple or changing information domains such as the Web. We primarily focus on *facet generation*, i.e. what data correspond to facets, how data sources are queried, how facets are visualized and what interaction options users can employ; *result overview generation*, i.e. what attributes of results and how they are visualized; *graph view generation*, i.e. which nodes and how are visualized, what attributes to show and what layout to employ.
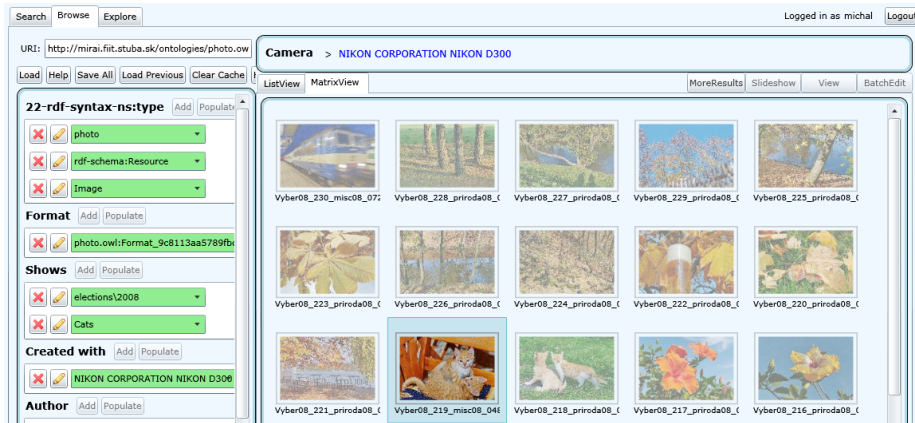
Fig. 1. Example of a faceted query result overview in a matrix (right) with one selected result whose attributes are edited in the annotation pane (left). The facets used for querying are hidden below the annotation pane, which is normally invisible.

We developed *Factic* as a client-side Silverlight application running inside a web browser (see Fig. 2). This simplifies deployment and enables it to process and store information on the client device. *Factic* is primarily centered around end-user specific functionality such as visualization, personalization, user modeling and profile management. We also process end-user specific data (e.g., activity logs and the derived user models) on the client thus reducing unwanted privacy exposure (with optional sharing with the server-side).

Consequently, *Factic* works as an intelligent front-end to (multiple) search and/or information providers thus effectively delegating querying, indexing and crawling services to third-party providers (e.g., in the future Google, Sindice,
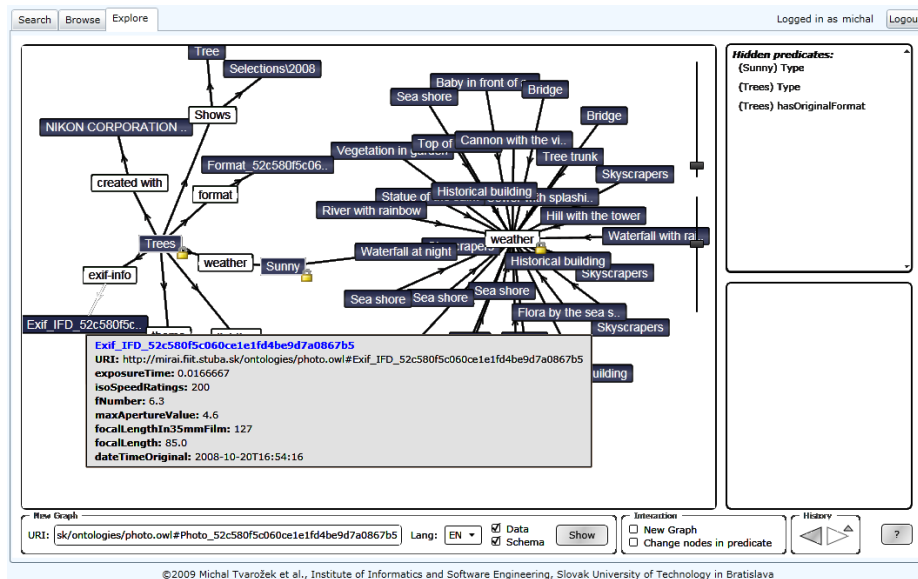
Fig. 2. The graph-based view shows different resources (e.g., photos, authors) as nodes. Users can expand nodes or navigate by centering the visible window on new nodes. Hovering over nodes shows additional annotations and associated thumbnails.

DBPedia). Our modular architecture allows us to easily incorporate new views and also gives us the flexibility to add new data sources provided that they contain enough semantic metadata to generate user interface widgets (see Fig. 3).
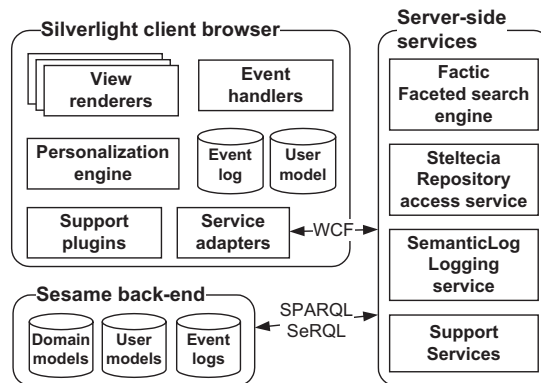
Fig. 3. The client renders the user interface and provides personalization support (top). The server includes web (WCF) services for faceted search (Factic), ontological repository access (Steltecia), and event logging (SemanticLog) for global statistics tracking (right). All services store their data in a common ontological repository in Sesame.

Our solution is primarily suitable for digital libraries or other (semi)structured domains which contain semantic metadata (we also experimented with publications and job offers). We evaluate our browser via user studies in the image domain where our data set contains roughly 8 000 manually and semi-automatically annotated images. Our results have shown the viability of adaptive interface generation in addition to user experience improvements in terms of better orientation, revisitation support and shorter task times.

3 Summary and Future Work

Our exploratory search browser *Factic* enables *adaptive Semantic Web exploration* by empowering end-users with access to semantic information spaces with:

- Adaptive view generation
- Interactive multi-paradigm exploratory search
- Personalized recommendation for individual users

Our client-side personalization and user modeling focus enables us to address privacy and identity issues as personally identifiable data never have to leave a user's system. Furthermore, we see the extension of our approach to accommodate legacy web resources as one possible direction of future work ultimately leading toward a *Next Generation Web Browser* for seamless exploration of semantic and legacy web content as outlined in [6].

References

1. Shadbolt, N., Berners-Lee, T., Hall, W.: The semantic web revisited. IEEE Intelligent Systems **21**(3) (May/June 2006) 96–101
2. Brusilovsky, P., Kobsa, A., Nejdl, W., eds.: The Adaptive Web: Methods and Strategies of Web Personalization. LNCS 4321. Springer, Berlin (June 2007)
3. Marchionini, G.: Exploratory search: from finding to understanding. Comm. of the ACM **49**(4) (2006) 41–46
4. Tvarožek, M., Bieliková, M.: Personalized faceted navigation in the semantic web. In Baresi, L., Fraternali, P., Houben, G.J., eds.: ICWE. Volume 4607 of Lecture Notes in Computer Science., Springer (2007) 511–515
5. Dörk, M., Carpendale, S., Collins, C., Williamson, C.: Visgets: Coordinated visualizations for web-based information exploration and discovery. IEEE Transactions on Visualization and Computer Graphics **14**(6) (2008) 1205–1212
6. Tvarožek, M., Bieliková, M.: Reinventing the web browser for the semantic web. In: WI-IAT '09: Proc. of the 2009 IEEE/WIC/ACM Int. Conf. on Web Intelligence and Intelligent Agent Technology, IEEE CS (2009) 113–116