

Exploratory Search in the Adaptive Social Semantic Web

Michal Tvarožek*

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

Abstract

Effective access to and sharing of information has become one of the most crucial needs of present day society directly affecting daily operation of many businesses and private individuals. To cope with issues such as information overload, unavailability of information, navigation problems and user diversity, and to facilitate the slow adoption of the Semantic Web, we devised an enhanced faceted semantic browser with support for multi-paradigm exploration, personalized recommendation and adaptive view generation. We employ facet and restriction selection, ordering and annotation to address information overload and user guidance, and adaptive view generation with incremental graph visualization to enable end-user grade exploration of semantic content. We present the highly promising results of several user studies performed with our browser prototypes in the job offers and digital image domains, which confirm the viability and practicality of our approach in terms of improved task times and user understanding of the explored information space. Based on our findings, we claim specific contribution to exploratory search and to the adoption of the Semantic Web.

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

*Recommended by thesis supervisor:
Prof. Mária Bieliková

Defended at Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava on February 10, 2011.

© Copyright 2011. All rights reserved. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from STU Press, Vazovova 5, 811 07 Bratislava, Slovakia.

Tvarožek, M. Exploratory Search in the Adaptive Social Semantic Web. Information Sciences and Technologies Bulletin of the ACM Slovakia, Vol. 3, No. 1 (2011) 42-51

Keywords

Semantic Web, exploratory search, faceted navigation, personalization, user modelling, user interface generation, graph visualization

1. Introduction

The information technology domain encompasses the access to, processing, organization and visualization of information together with the corresponding software and hardware infrastructure. Information technology has become a vital and indispensable part of daily life, shifting focus from the manufacturing of physical goods toward the creation, organization and sharing of information, thus giving rise to the Information age.

In this respect, the Web has become a global ubiquitous socioeconomic space providing information, services and facilitating both private and business communication with an estimated 1.8 billion users (survey by Miniwatts Marketing Group, 2010). Many use the Web daily as an integral part of their work, ultimately making the human society as we know it today dependent on the storage, availability and exchange of information on the Web. Due to the constant growth, complexity and size of the Web, several issues hamper online user experience:

- *information overload* (i.e., too much information being available),
- *the lack of information availability* (i.e., the required information is available somewhere in the Web, but unavailable to users who need it),
- *the navigation problem* (i.e., users losing track of their position in the information space effectively “getting lost in hyperspace”),
- *ignorance of user diversity* (i.e., the fact that sites are created to suit the “average user”).

These issues along with the evolution of the Web in terms of novel usage means and user expectations such as exploratory search, social networks, interactive applications or user created content together with the impact of the Web on the human society resulted in the emergence of Web Science as a new research field [5].

The constant improvement of web-based applications and information access and processing is also the focus of several web initiatives. The Semantic Web aims to provide

better search and browsing capabilities by enabling machine readability of information on the Web taking advantage of ontologies [13]. Adaptive Web approaches aim to learn about user preferences and adapt the user experience to the specific needs of individuals [3], while the Social Web aims to facilitate communication, collaboration, interaction and sharing on the Web by exploiting social wisdom to improve individual user experience [14]. The Exploratory search initiative aims to provide users with better tools for advanced information seeking tasks such as learning, investigation and analysis [8].

Although each of these initiatives addresses different problems and aspects of the Web's development, if successfully combined together they are likely to produce synergetic effects, which are already starting to surface. Ultimately, their combination would transform the current Web as we see it today into an entirely new and mature *Adaptive Social Semantic Web* of tomorrow.

Since there has been little cross-fertilization between these initiatives, our aim is to combine and extend their respective approaches into a highly interdisciplinary solution in order to facilitate Semantic Web adoption. We aim to devise an *end-user grade exploratory search approach* for seamless exploration of both semantic and legacy web content with specific focus on faceted browsing, user interface generation, advanced search and visualization.

2. Background and related work

To facilitate our goal of providing and improving end-user grade exploration of the Semantic Web, we need to address querying, visualization and exploration of Semantic Web resources via a combination of approaches from different fields. Typically, the information retrieval process consists of a search query, the successive navigation in the results and result exploration, with optional later revisitation of previously discovered results. We explore the faceted browsing paradigm [11], since it seamlessly combines these steps and has already been shown to be very promising and generally accepted amongst end-users. In order to better understand user behaviour in faceted browsers, Kules et al. performed a user study examining how searchers interact with faceted browsers. The study discovered that facets were an integral part of the exploration experience accounting for about one half of the time spent on actual search results [6]. We examine exploration capabilities and user support provided by faceted browsers and aspects of acquisition, search and visualization of history entries and metadata [9].

Wilson and schraefel performed a study comparing three prominent exploratory browsers – Flamenco, mSpace and RelationBrowser++ [22]. While Flamenco and RelationBrowser++ are more traditional faceted browsers, mSpace takes advantage of RDF data (native to Semantic Web) to provide users with a set of customizable filters that can be used to visualize a subspace of a high dimensional information space. The RelationBrowser++ is tailored to exploration of large statistical data and persistently displays all facets at the top unlike Flamenco, which hides exhausted facets [24].

VisGets is an advanced visualization and querying solution for legacy web data [4]. It crawls the Web and gathers news articles, and in turn enables users to interactively explore the data based on three dimensions - time, location

and topic. It does not however provide social recommendation support nor supports navigation or orientation after selecting a search result (i.e., once the user leaves the original search engine). Moreover as VisGets uses its own crawling and indexing engine it cannot be effectively used for general web search or Semantic Web exploration.

The BrowseRDF faceted browser provides elementary facet generation capability over simple RDF data [10]. It automatically identifies facets in source data based on several statistical measures, but offers only very limited interaction options and does not consider semantic metadata provided in the more expressive RDFS and OWL formats. Similarly, Tabulator enables users to browse Linked Data [2]. While Tabulator enables users to take advantage of different visualizations (e.g., map, calendar), it offers only very limited search support. Other Semantic Web browsers / query builders such as Disco Hyperdata browser or Zitgist Dataviewer offer even less user support and are thus useful only to experts.

Neither of these approaches can be effectively used for complex interactive exploration of Semantic Web content, which in addition to advanced (faceted) querying needs to support interactive information visualization, personalization and exploration of graphs (Semantic Web being a graph). Here, also graph visualization and interaction approaches must be considered as described in [12].

3. Framework for exploratory search

To address the issues outlined in section 1, we devised a comprehensive faceted exploration browser for the Semantic Web, which acts as an integrated tool for search when it acts like a client-side semantic search engine front-end, and for navigation when it supports navigation across a collection of “pure” information artifacts accessed via a semantic endpoint. We employ these principles [16]:

- *Semantic information space representation*
- *Personalized recommendation*
- *Multi-paradigm exploration*
- *Adaptive view generation*

We build upon the opening-midgame-endgame paradigm used in Flamenco [23] and extend it into comprehensive multi-paradigm exploration approach. We add user support for the individual stages of the information seeking process and populate them with additional complementary approaches to facilitate end-user grade exploration experience (see Figure 1, explained in following sections).

The opening stage is populated with several views that can be used to initiate and exploratory search session. Our *classical view* augments the traditional keyword-based search window with a tag-based overview of the information space content (tags correspond to information artefact types). The *faceted view* corresponds to our faceted browser interface without any selections. It is mostly used when we already have some information about user preferences and are thus able to provide a personalized set of initial facets for exploration. The *history view* supports information revisitation via the Semantic history map, which semantically organizes a user's search and browsing history.

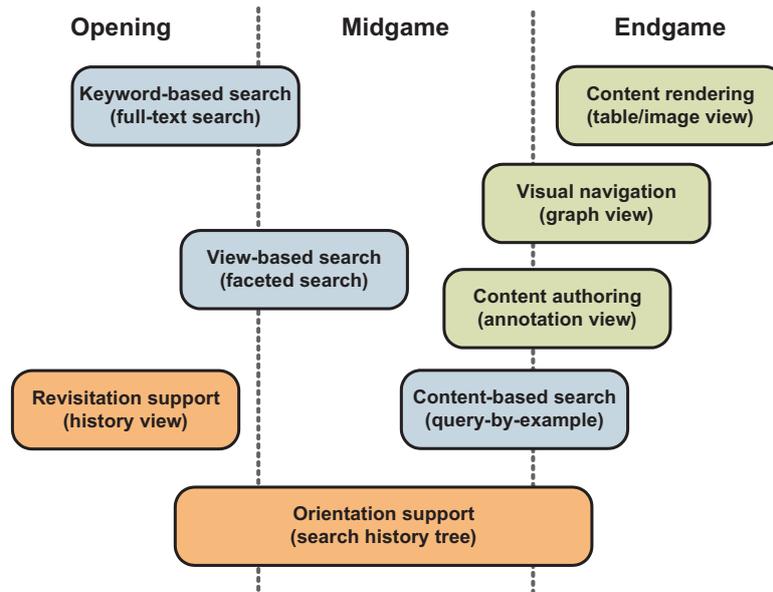


Figure 1: Overview of our multi-paradigm exploration approach showing the scope and applicability of individual sub-approaches to specific stages of the exploration process. Search approaches span primarily the opening and midgame (blue), content viewing, annotation and browsing approaches focus on the endgame (green), while support approaches span all stages (orange).

After the initial query, users proceed to the midgame stage where they continuously refine their query and explore the search results. This is done mainly via our enhanced faceted browser by extending the faceted view from the opening with *list and matrix result view*, the *search history tree* (provides orientation and history support), and the *graph view* (enables users to interactively explore the properties and relations between individual information artefacts).

Although typically user sessions end with the endgame, we see the opening-midgame-endgame as an iterative process which allows users to return to a previous stage. We thus populate the endgame with tools that enable the user to get a better understanding of the information space, to shape it or to simply view its contents in a natural way. This includes a *nested table view* (displays the properties of individual information artefacts), *specialized image view* (enables users to view photos similarly to popular web-based photo galleries), *graph view* (enables users to interactively view the information artefacts as a graph showing their relations and attributes), and *annotation view* (allows users to see a list of existing or optional properties of individual information artefacts and edit them).

4. Personalized recommendation

We extend the typical request handling of faceted browsers with additional steps that perform specific tasks. We extend search results processing with result recommendation which includes support for result annotation and adaptation. We employ external tools that evaluate the relevance of individual search results, e.g., by means of concept comparison with the user model. Subsequently, we reorder search results or annotate them with additional information. For example, in the domain of scientific publications, we can display the suitability of an article, based on its estimated relevance to the user’s research, as background colour or via emoticons.

Facet processing is extended with facet recommendation, which includes the adaptation, annotation and recommendation of facets and restrictions, which improve orientation and guidance support, reduce information overload and alleviate some disadvantages of faceted classification. If the set of available facets is insufficient, we use dynamic facet generation to add new facets at run-time on a per user basis thus allowing the user to refine the search query and improving support for open information spaces.

4.1 Model for relevance evaluation

Our browser logs events that occurred as results of user interaction with the current state of the browser via a specialized semantic logging service which preserves the semantics of events as opposed to traditional web server logs, which store them only implicitly in request URLs. The acquired events are processed by a separate user modelling back-end [1] and in turn retrieved as an updated user model, which drives our personalization engine. Each logged event uses our event ontology to specify the semantics of the respective user action and also references the domain and user ontologies as required.

The user modelling back-end provides us with several sources of adaptation, which we employ with different weights depending on how closely related they are to the current user task:

1. *In-session user behaviour* – user navigation, facet and restriction selection during the current user session (i.e., user clicks). Frequent use of specific items indicates higher relevance to the current task and/or user interest in the corresponding domain concepts.
2. *Short/long term user model* – user characteristics acquired during multiple sessions described by their relevance to the user and the confidence in their estimation in the range $\langle 0,1 \rangle$.

3. *Similar/related user models* are assumed to belong to users with similar needs and are thus used for relevance evaluation if user specific data is unavailable or has low confidence. Social user context can be exploited by assigning custom weights to specific relations between users resulting in social recommendation.
4. *Global usage statistics* computed from the overall relevance and usage of individual domain concepts (e.g., facets, restrictions, target objects – be it images, publications or job offers) from all user models. The overall “popularity” of facets and restrictions increases the likelihood of their recommendation for a specific user, especially if his or her specific preferences are unknown or have low confidence.

4.2 Facet recommendation

With personalization, we empower users to make their own decisions more effectively via additional annotations while also providing sensible means of automatic adaptation. As opposed to most existing approaches, we perform personalization primarily on the client side (i.e., in the client browser), which has two benefits – personally sensitive data is kept entirely on the client side thus preserving privacy, server-side services need not have support for adaptation as it is performed by our browser on the client side, thus providing personalization for all information resources at no additional cost [17].

Facet recommendation distinguishes three types of facets adapted at run-time to the specific needs of individual users – active facets, inactive facets and disabled facets. The adaptation process first determines the relevance of individual facets and restrictions in our relevance model and then uses it in these steps (visualized in Figure 2):

1. *Active facet selection* – the total number of active facets is reduced to a relatively low number, e.g. 2 or 3 facets, since many facets are potentially available in complex information spaces. Active facets are selected based on relevance, recency and number of accesses. The rest of the facets is made inactive or left in disabled state.
2. *Facet and restriction ordering* – all facets are ordered in three groups (i.e., active, inactive, disabled) in descending order based on their relevance with the last used facet always being at the top. Restrictions are ordered alphabetically as alternative orderings based on relevance or the number of matching search results were not well accepted by users as they made it difficult to search for specific items.
3. *Facet and restriction annotation* – active facet restrictions are annotated with the number of matching instances, the relative number of matching instances by means of font size/type, or directly recommended (e.g., with background colour or the “traffic lights” metaphor) effectively providing shortcuts to deeply nested restrictions. Additional tooltips can describe individual facet/restriction meanings (e.g., the `rdfs:comment` annotation in ontologies).

Search result recommendation extends the processing of search results with support for personalized result ordering, annotation and view adaptation. We employ external

tools that evaluate the relevance of individual search results, e.g., by means of concept comparison with the user model or via the evaluation of (explicit) user feedback. Subsequently, we reorder the search results or annotate them with additional information.

5. Adaptive view generation

We generate user interfaces for each step of the exploratory search process (query construction, result browsing, resource exploration). We generate:

- Faceted browser interfaces for advanced query construction and modification.
- Result overviews for effective presentation of selected result attributes.
- Graph-based exploration views for incremental horizontal exploration of semantic resources and their relations with other resources.

5.1 Facet generation

During facet generation, we examine metadata describing the information space, identify object and literal facet templates, and select either an enumeration or hierarchical restriction template to use based on ontological metadata. The facet construction stage determines the interaction mode based on the overall number of potential restrictions; list mode is used for a small number of predefined values (e.g., days of the week), search mode is used for large numbers of values (e.g., all cities on Earth). If an ordering of values is defined in the ontology for object values, we also create restriction intervals to cover continuous values (e.g., real numbers or dates). The last facet mapping stage selects a suitable user interface widget to render the generated facet in the faceted browser, and maps the constructed facet and restriction values onto the widget. The widget provides facet visualization (see Figure 3) and handles user interaction, forwards events and facet metadata to the server back-end, which provides the corresponding querying services for the generated facet. Facet generation thus defines these facet properties [18]:

- A *facet template*, which corresponds to a pattern found in domain metadata and specifies the overall type and behaviour of the facet.
- A *restriction template*, which defines how the individual restrictions in the facet are constructed and mapped onto the domain ontology.
- A *query template*, which defines how the back-end query engine creates database queries and maps them onto facet restrictions.
- A *visualization and interaction template* (i.e., the corresponding widget type), which binds the facet to the graphical user interface and handles user input.

5.2 Result overview generation

We generate two result overviews – the *ListView* shows thumbnails and properties of individual results (see Figure 3, top left), while the *MatrixView* shows thumbnails, provides additional information in tooltips, and in addition offers a generated editing pane for modification of individual result attributes (see Figure 3, bottom right).

The screenshot shows a web application for job searching. At the top is a navigation bar with links: Home, Job Offers, Criteria Search, Company Registration, User Registration, About us, and Help. Below this, there are three main sections:

- Offered position:** A dropdown menu on the left shows 'Computing professionals (20)' as the selected category, with 'Physical mathematical and engineering science professionals (20)' as an alternative.
- Duty location:** A dropdown menu on the left shows '(4)' as the selected location, with various cities listed: Anaheim (1), Berkeley (2), Cypress (1), Glendale (1), Glendale (2), Irvine (2), Pasadena (2), Pasadena (4), San Diego (2), San Francisco (3), Santa Monica (2), US CA Irvine (2), US CA Pasadena (4), and US CA San Francisco (3).
- Current restrictions:** A red box at the top right shows 'Offered position: All > Professionals (24)' and 'Duty location: All > North America > United States > California (24)'. Below this, there are sorting options: 'Sort by: Name | Salary | Organization | Region' and 'Item per page: 10 | 15 | 25 | 50 | 100'.

The main content area is a table with 15 rows of job results. Each row includes a number, a job title with a link, salary, organization, region, a 'Rate It!' section with five stars, and an 'Actions' column with icons for edit, delete, and favorite.

#	Name	Salary	Organization	Region	Rate It!	Actions
1	At-Tech, Orange, CA US	25000.0	At-Tech	California	☆☆☆☆☆	✎ ✕
2	6181 - Systems Programmer	85000.0	Genesis10	Pasadena	☆☆☆☆☆	✎ ✕
3	C# / ASP.NET Programmers	35000.0	At-Tech	Anaheim	☆☆☆☆☆	✎ ✕
4	Manpower Professional, San Diego, CA US	29.0	Manpower Professional	California	☆☆☆☆☆	✎ ✕
5	Systems Programmer	50.0	Genesis10	California	☆☆☆☆☆	✎ ✕
6	COBOL BASIC Programmers VMS DEC Alpha	35.0	Indotronic International Corp.	Glendale	☆☆☆☆☆	✎ ✕
7	Sr. Mainframe Database Programmer	75000	Genesis10	Passadena	☆☆☆☆☆	✎ ✕
8	Sr. Mainframe Database Programmer	85000	Genesis10	Passadena	☆☆☆☆☆	✎ ✕
9	Sr. Mainframe Database Programmer	75000	Genesis10	Passadena	☆☆☆☆☆	✎ ✕
10	6081 - Sr. Mainframe Database Programmer	75000.0	Kaiser Permanente	Passadena	☆☆☆☆☆	✎ ✕
11	Senior Programmer Analyst	50.0	Genesis10	Berkeley	☆☆☆☆☆	✎ ✕
12	Programmer Analyst II	40.0	Genesis10	Berkeley	☆☆☆☆☆	✎ ✕
13	Oracle 7-9i Programmer (PL/SQL, DataStage, ADX)		Strategi	Cypress	☆☆☆☆☆	✎ ✕
14	SAP_SD_WM_Consultant	20.0	MBA - Eindhoven	Glendale	☆☆☆☆☆	✎ ✕
15	Junior Programmer	35000.0		Glendale	☆☆☆☆☆	✎ ✕

Figure 2: Example of facet adaptation, annotation and restriction recommendation showing active, and inactive facets (left), also showing a list view of search results with attributes and additional operations (right).

ListView shows attributes of a specific result directly derived from the domain ontology visualized as label-value pairs. For multi-value properties such as Type in Figure 3, a column with all values is shown. We either show all result properties to maximize information or apply personalization to select only the most relevant properties.

5.3 Annotation view generation

Similarly to result overviews, we generate the *annotation view* (accessible from result overviews) separately for each specific result type. We identify all applicable properties from the domain ontology metadata, construct editing widgets based on property types (e.g., text boxes with language selection or auto-complete combo boxes, with single/multi-value support). Properties with existing values are shown first, while properties without values are shown at the bottom (see Figure 3, bottom left).

6. Multi-paradigm exploration

Multi-paradigm exploration is principal to approach as it integrates a set of search, navigation and visualization approaches into a comprehensive exploratory search solution as outlined in Figure 1.

6.1 Searching and browsing

The *classical view* is based on the initial screens of existing web search engines such as Google or Bing. In addition to the search box, we employ a tag cloud-based multi-purpose view that can correspond to:

- Information artefact types thus giving users an idea

of what kinds of information can be explored (e.g., photos, events, regions).

- Tags (topics) of recently added or modified information artefacts thus giving users an overview of what new information is available
- Popular information artefacts effectively providing social (either global or community based) recommendation and providing users with an overview of current trends.

Since prior work by Kules et al. has shown that users mostly use the facets and the result overview when working with faceted browsers, we focused mainly on their improvement [6]. Our *faceted view* (adaptation details described in sections 4 and 5) integrates these approaches [20]:

- Faceted browsing based on the traditional layout with facets on the left, query at the top and results in the centre.
- Adaptive search result overviews (list view, matrix view) providing users with quick and easy understanding of the current result set.
- Search history tree based on interactive graph visualization for orientation and history support.
- Query-by-example via search result rating or similarity search (performed via external tools) which improves querying capability of users and supports the exploration of similar information artefacts.

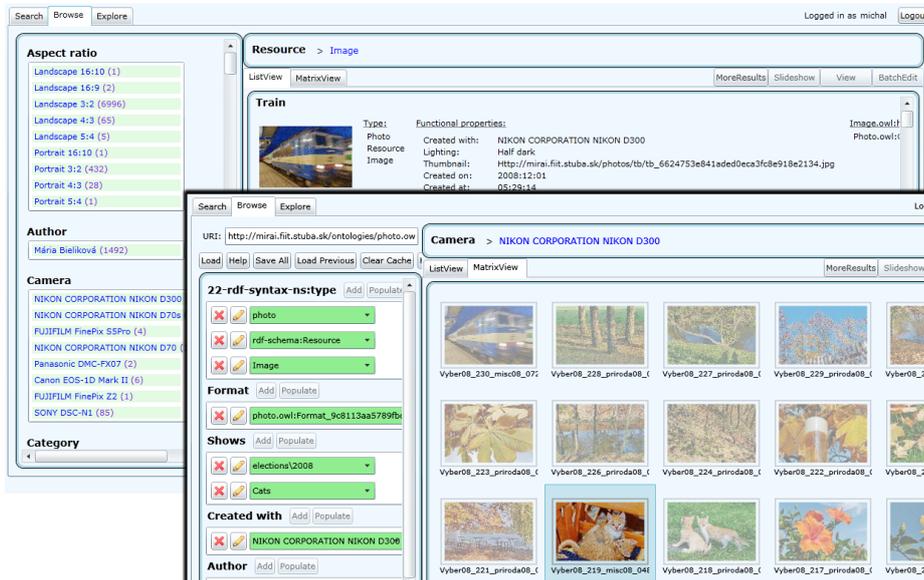


Figure 3: (A) Generated facets with a list-based result overview showing all result properties (top left). (B) A matrix result overview with image thumbnails and the correspondingly generated annotation pane for collaborative content creation (bottom right).

- Optional keyword-based full text search as a complementary approach to faceted- and content-based search. Note that in the semantic web environment, there normally is no full text to search so keyword based search is limited to the labels and comments of resources.

6.2 Result exploration

During the endgame of the exploration process, users need see and understand the properties (i.e., actual content) of the discovered information artefacts. We provide three types of exploration views:

- *Textual attribute exploration* via the nested table view for visualization of information properties and the annotation view for their modification.
- *Relation exploration* via the graph view for interactive exploration of relations between resources.
- *Content viewing* via the image view, which renders associated content (i.e., photos) in a native way.

Textual result exploration. The *table view* recursively renders the properties of individual search results in a nested table thus providing users with an exhaustive visualization of the details associated with a given resource. Compared to some other existing approaches, which only provide direct properties, the nested visualization improves user orientation by maintaining the context of resources,

as normally one would have to click a link to view properties of other resources thus losing the original context.

The *annotation view* supports collaborative content creation by allowing authorized users to create new information artefacts, modify or optionally delete existing ones via a generated form-based interface (see Figure 3, bottom left). The view is specific for each resource type and shows first properties with existing values and next properties corresponding to the given resource type without values. Users can either enter entirely new values or select pre-existing values from drop-down menus. Moreover, users can remove property values or entire resources, create new resources and even alter the schema of the repository (in ontologies schema and data are treated equally).

Content-sensitive result exploration. We employ domain specific visualizations based on resource type to support “natural” access to information artefacts. Since we also worked with an image collection, we devised a specialized image view, which enables users to view the photos similarly to popular web-based photo galleries. *Image view* supports image manipulation features such as zoom, rotate or slideshows, it shows image thumbnails and can also display basic image attributes.

Visual relation exploration. We provide visual result exploration support via a graph-based visualization of resource properties. The *graph exploration view* consists of the graph visualization window, predicate filtering win-



Figure 4: Example of our tree-based history visualization showing an initial keyword query (top left) and the successive faceted query refinements (left). The rest of the interface shows the list of available facets (centre) and the list of search results (right).

dows and an options toolbar (see Figure 5). Users can access the view either directly by typing in the URI of the node they wish to explore, or by exploring a result found in faceted view.

The graph view is generated directly from a domain ontology showing individual resources and their relations, also taking advantage of relevance evaluation from the personalization engine. Relations are intentionally visualized as separate nodes connecting resources to reduce information overload when one relation can have multiple values and to improve graph layout.

6.3 Revisitation and orientation support

The *history view* provides a tree-based visualization of search and browsing history that improves user orientation within complex navigation sessions and provides revisitation support for previously discovered (distributed) information during exploratory search sessions. We continually record user actions performed within our browser (e.g., facet selections, result exploration) and construct a tree of query modifications and result visits (see Figure 4). The tree is shown to users while they are browsing and also stored for future reference and processing [21]. We devised two integrated approaches to revisitation support:

- *Search History Tree* – an in-session tree-based history visualization,
- *Semantic History Map* – an interactive, semantically organized, graph-based visualization of longer-term browsing history that shows the original context of individual history entries.

Our method records user sessions (i.e., queries and visited web resources), identifies and separates individual user goals (i.e., coherent user sessions with similar terms), preserves their context by persistently storing history trees corresponding to relations between queries and visited web resources, and ultimately synthesizes navigable graphs from extracted terms, visited resources and user goals.

7. Evaluation and discussion

To validate our approach we performed experiments with two prototypes of our faceted semantic browser Factic. We used our initial prototype to evaluate the usefulness of our facet personalization approach in the job offers domain via a user study and to gather feedback on its design [15]. We then created a second, improved prototype, which extended the original functionality with support for facet generation, additional result visualization and the graph exploration view, while also addressing performance and usability issues of the first prototype [19]. Since analytical validation of user-centered approaches is difficult if at all possible, also considering the novelty of the exploratory search field and immaturity of methodologies for task design and browser evaluation [7], our evaluation goals focused on user studies and proof of concept validation of our individual approaches.

Our first study was in the job offers domain, where our approach proved to be particularly suitable, since it is a very complex information space with several deep hierarchical classifications (e.g., regions or positions) and intricate concept relations. Our evaluation showed that adaptive selection of active facets (i.e., fully rendered) and recommendations can significantly reduce information overload (i.e. the number of facets a user must examine) and thus total processing time which depends roughly linearly on the number of displayed facets. However, the number of clicks increased since the right facets were not always active and thus had to be manually enabled. This resulted in shorter refresh times and consequently shorter total task times [17]. Despite the very positive feedback and highly promising results, we encountered scalability issues with remote repositories due to repository querying limitations and network delays, which we addressed in our second prototype.

Our second study was in the digital image domain. We generated facets from the available data and examined how the browser behaved and whether the interface was still usable for its intended purpose in terms of usability

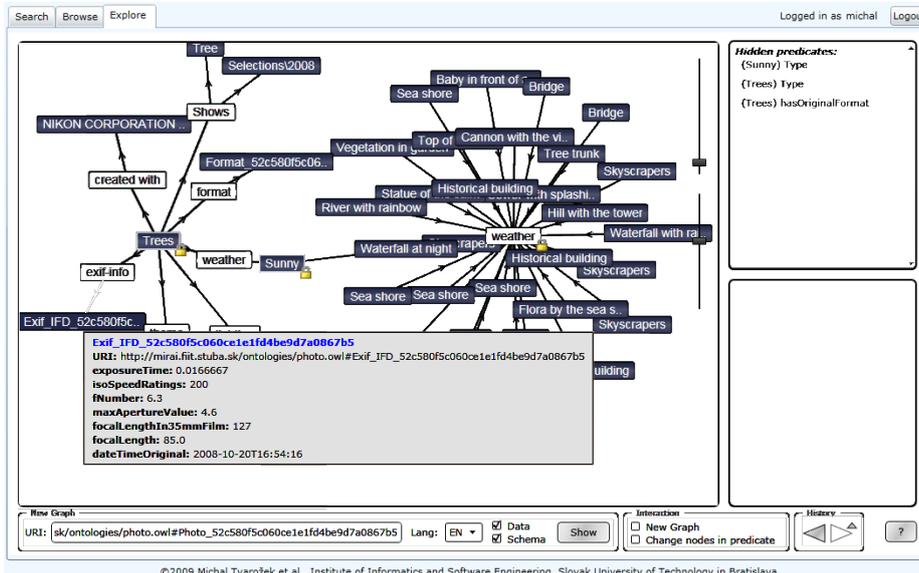


Figure 5: Example of our generated graph-view exploration interface. Dark nodes represent individual resources, white nodes correspond to relations (top). Hovering over nodes shows the attributes of a node (center); additional tools include zooming, spatial expansion, node hiding and history (right), with additional filtering options for languages and data/schema only visualization (bottom).

and performance. The experiments proved that the approach was viable for interface generation with minimal performance impact. The user study with the graph exploration interface was also successful as most users were able to accomplish their assigned tasks within a reasonable timeframe. Furthermore, users managed to answer 75% of the questions correctly leaving 25% false answers (this also includes answers that were close to the correct ones, but not exactly right). Based on these results, we concluded that graph-based exploration is viable for Semantic Web browsing as most users were able to accomplish the given tasks despite having no prior experience with a similar interface [18].

8. Conclusions and future work

Today, effective access to information has already become crucial to many aspects of daily life both in the corporate environment and in personal life. Our main focus was the development of novel methods for navigation and presentation, and the combination of approaches from the Semantic Web, Social Web and Adaptive Web initiatives. To improve and maintain our current information access capabilities, we devised a novel comprehensive approach to multi-paradigm faceted exploration of Semantic Web content with specific contribution to:

- *Multi-paradigm exploration* – integrating view-based, content-based and keyword-based search with advanced adaptive visualizations and incremental graph exploration of both content and browsing history.

- *Personalized recommendation* – devising a method of facet and restriction adaptation based on semantic logging of user action and continuous evaluation of the devised ontological user and relevance models.
- *Exploratory interface generation* – devising a method for facet identification in ontological metadata, its transformation into interface widgets and their mapping onto the ontological querying back-end (e.g., semantic search engines).

Our results in the job offers and digital image domains have shown the viability of the proposed approaches (personalization, faceted interface generation, graph exploration) for their intended purposes in terms of their practicality (i.e., it can be done) and improved user experience (i.e., improved task times, better understanding of the information space, efficient resource revisitation). Consequently, based on our findings we have improved upon the current state of the art in exploratory search in the Semantic Web by *empowering end-users with access to semantic information spaces* via an end-user grade exploratory browser for the Semantic Web with interfaces for effective query formulation, result overview browsing and individual result exploration.

We see the extension of our approach with support for legacy web content, i.e., for specific pages (e.g., personal browsing history) or for whole web sites (e.g., generating a faceted browsing interface for a typical corporate web site), and for interactive content exploration as a possible direction for future work.

Acknowledgements. This work was partially supported by the State programme of research and development “Establishing of Information Society” under the contract No. 1025/04, the Slovak Research and Development Agency under the contract No. APVT-20-007104, the Scientific Grant Agency of SR, grant No. VG1/0508/09 and it is a partial result of the Research & Development Operational Program for the project Support of Center of Excellence for Smart Technologies, Systems and Services II, ITMS 26240120029, co-funded by ERDF.

References

- [1] M. Barla and M. Tvarožek. Rule-based user characteristics acquisition from logs with semantics for personalized web-based systems. *Computing and Informatics*, 28:399–427, 2009.
- [2] T. Berners-lee, Y. Chen, L. Chilton, D. Connolly, R. Dhanaraj, J. Hollenbach, A. Lerer, and D. Sheets. Tabulator: Exploring and analyzing linked data on the semantic web. In *Proceedings of the 3rd International Semantic Web User Interaction Workshop*, 2006.
- [3] P. Brusilovsky, A. Kobsa, and W. Nejdl, editors. *The Adaptive Web, Methods and Strategies of Web Personalization*, volume 4321 of *Lecture Notes in Computer Science*. Springer, 2007.
- [4] M. Dörk, S. Carpendale, C. Collins, and C. Williamson. Visgets: Coordinated visualizations for web-based information exploration and discovery. *IEEE Transactions on Visualization and Computer Graphics*, 14:1205–1212, November 2008.
- [5] J. Hendler, N. Shadbolt, W. Hall, T. Berners-Lee, and D. Weitzner. Web science: An interdisciplinary approach to understanding the web. *Commun. ACM*, 51(7):60–69, 2008.
- [6] B. Kules, R. Capra, M. Banta, and T. Sierra. What do exploratory searchers look at in a faceted search interface? In *JCDL '09: Proc. of the Joint Conf. on Digital Libraries*, pages 313–322, New York, NY, USA, 2009. ACM.
- [7] B. Kules, R. Capra, M. Banta, and T. Sierra. What do exploratory searchers look at in a faceted search interface? In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital Libraries*, JCDL '09, pages 313–322, New York, NY, USA, 2009. ACM.
- [8] G. Marchionini. Exploratory search: from finding to understanding. *Commun. ACM*, 49:41–46, April 2006.
- [9] M. Mayer. Web history tools and revisitation support: A survey of existing approaches and directions. *Foundations and Trends in HCI*, 2(3):173–278, 2009.
- [10] E. Oren, R. Delbru, and S. Decker. Extending faceted navigation for rdf data. In I. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, and L. Aroyo, editors, *The Semantic Web - ISWC 2006*, volume 4273 of *Lecture Notes in Computer Science*, pages 559–572. Springer Berlin / Heidelberg, 2006.
- [11] G. M. Sacco and Y. Tzitzikas. *Dynamic Taxonomies and Faceted Search: Theory, Practice, and Experience*. Springer Publishing Company, Incorporated, 1st edition, 2009.
- [12] H.-J. Schulz and H. Schumann. Visualizing graphs - a generalized view. *Information Visualisation, International Conference on*, 0:166–173, 2006.
- [13] N. Shadbolt, T. Berners-Lee, and W. Hall. The semantic web revisited. *IEEE Intelligent Systems*, 21:96–101, May 2006.
- [14] S. Staab, P. Domingos, P. Mika, J. Golbeck, L. Ding, T. Finin, A. Joshi, A. Nowak, and R. R. Vallacher. Social networks applied. *IEEE Intelligent Systems*, 20:80–93, January 2005.
- [15] M. Tvarožek and M. Bieliková. Personalized faceted navigation in the semantic web. In *Proceedings of the 7th international conference on Web engineering*, ICWE'07, pages 511–515, Berlin, Heidelberg, 2007. Springer-Verlag.
- [16] M. Tvarožek and M. Bieliková. Collaborative multi-paradigm exploratory search. In *WebScience '08: Proc. of the Hypertext 2008 Workshop on Collaboration and collective intelligence*, pages 29–33, New York, NY, USA, 2008. ACM.
- [17] M. Tvarožek and M. Bieliková. Personalized faceted navigation in semantically enriched information spaces. In Marios C. Angelides et al., editor, *Advances in Semantic Media Adaption and Personalization*, volume 2, pages 181–201. CRC Press, 2009.
- [18] M. Tvarožek and M. Bieliková. Generating exploratory search interfaces for the semantic web. In P. Forbrig, F. Paternó, and A. Mark Pejtersen, editors, *Human-Computer Interaction*, volume 332 of *IFIP Advances in Information and Communication Technology*, pages 175–186. Springer Boston, 2010.
- [19] M. Tvarožek and M. Bieliková. Reinventing the web browser for the semantic web. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 03*, WI-IAT '09, pages 113–116, Washington, DC, USA, 2009. IEEE Computer Society.
- [20] M. Tvarožek and M. Bieliková. Bridging semantic and legacy web exploration: Orientation, revisitation and result exploration support. *Web Intelligence and Intelligent Agent Technology, IEEE/WIC/ACM International Conference on*, 3:326–329, 2010.
- [21] J. Šimko, M. Tvarožek, and M. Bieliková. Semantic history map: Graphs aiding web revisitation support. *Database and Expert Systems Applications, International Workshop on*, 0:206–210, 2010.
- [22] M. L. Wilson, M. C. schraefel, and R. W. White. Evaluating advanced search interfaces using established information-seeking models. *J. Am. Soc. Inf. Sci. Technol.*, 60:1407–1422, July 2009.
- [23] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *Proc. of the SIGCHI Conf. on Human factors in Computing Systems*, pages 401–408. ACM Press, 2003.
- [24] J. Zhang and G. Marchionini. Evaluation and evolution of a browse and search interface: relation browser. In L. M. L. Delcambre and G. Giuliano, editors, *DG.O*, volume 89 of *ACM International Conference Proceeding Series*, pages 179–188. Digital Government Research Center, 2005.

Selected Papers by the Author

- M. Barla, M. Tvarožek, M. Bieliková. Rule-Based User Characteristics Acquisition from Logs with Semantics for Personalized Web-Based Systems. *Computing and Informatics*, 28 (4): 399–427, 2009.
- M. Tvarožek. Personalized Navigation in the Semantic Web. In V. P. Wade, H. Ashman, and B. Smyth, editors, *AH 2006*, LNCS 4018, pages 467–472. Springer, 2006.
- M. Tvarožek, M. Bieliková. Generating Exploratory Search Interfaces for the Semantic Web. In P. Forbrig, F. Paternó, and A. M. Pejtersen, editors, *IFIP WCC: HCIS 2010*, IFIP 332, pages 175–186. Springer, 2010.
- M. Tvarožek, M. Bieliková. Visualization of Personalized Faceted Browsing. In P. Forbrig, F. Paternó, and A. M. Pejtersen, editors, *IFIP WCC: HCIS 2008*, IFIP 272, pages 213–218. Springer, 2008.
- M. Tvarožek, M. Bieliková. Adaptive faceted browser for navigation in open information spaces. In Proceeding of the 16th international conference on World Wide Web, pages 1311–1312. ACM Press, 2007.
- M. Tvarožek, M. Bieliková. Bridging Semantic and LegacyWeb Exploration: Orientation, Revisitation and Result Exploration Support. In Proceedings of IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies 2010, Workshop on Web Information Retrieval Support Systems, pages 326–329. IEEE CS, 2010.
- M. Tvarožek, M. Bieliková. Collaborative Multi-Paradigm Exploratory Search. In Proceedings of the Hypertext 2008 Workshop on Collaboration and Collective Intelligence, pages 29–33. ACM Press, 2008.
- M. Tvarožek, M. Bieliková. Factic: Personalized Exploratory Search in the Semantic Web. In B. Benatallah, editor, *ICWE 2010*, LNCS 6189, pages 528–531. Springer, 2010.
- M. Tvarožek, M. Bieliková. Personalized Faceted Browsing for Digital Libraries. In N. Furr, L. Kovacs, and C. Meghini, editors, *ECDL 2007*, LNCS 4675, pages 485–488. Springer, 2007.
- P. Bartalos, M. Barla, G. Frivolt, M. Tvarožek, A. Andrejko, M. Bieliková, P. Návrát. Building ontological base for

- experimental evaluation of semantic web applications. In J. van Leeuwen, G.F. Italiano, W. van der Hoek, C. Meinel, H. Sack, and F. Plášil, editors, *SOFSEM 2007*, LNCS 4362, pages 682–692. Springer, 2007.
- M. Tvarožek, M. Barla, M. Bieliková. Personalized Presentation in Web-Based Information Systems. In J. van Leeuwen, G.F. Italiano, W. van der Hoek, C. Meinel, H. Sack, and F. Plášil, editors, *SOFSEM 2007*, LNCS 4362, pages 796–807. Springer, 2007.
- M. Tvarožek, M. Barla, G. Frivolt, M. Tomša, M. Bieliková. Improving Semantic Search Via Integrated Personalized Faceted and Visual Graph Navigation. In V. Geffert, J. Karhumäki, A. Bertoni, B. Preneel, P. Návrát, and M. Bieliková, editors, *SOFSEM 2008*, LNCS 4910, pages 778–789. Springer, 2008.
- M. Tvarožek, M. Bieliková. Personalized Faceted Navigation in the Semantic Web. In L. Baresi, P. Fraternali, and G.-J. Houben, editors, *ICWE 2007*, LNCS 4607, pages 511–515. Springer, 2007.
- M. Tvarožek, M. Bieliková. Bridging Semantic and Legacy Web Exploration: Orientation, Revisitation and Result Exploration Support. In Proceedings of IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies 2010, Workshop on Web Information Retrieval Support Systems, pages 113–116. IEEE CS, 2009.
- M. Tvarožek, M. Bieliková. Personalized View-Based Search and Visualization as a Means for Deep/Semantic Web Data Access. In Proceeding of the 17th international conference on World Wide Web, pages 1023–1024. ACM Press, 2008.
- M. Tvarožek, M. Bieliková. Personalized Faceted Navigation for Multimedia Collections. In P. Mylonas, M. Wallace, and M. Angelides, editors, *Proceedings of SMAP 2007 – 2rd International Workshop on Semantic Media Adaptation and Personalization*, pages 104–109. IEEE CS, 2007.
- M. Tvarožek, M. Bieliková. Personalized Faceted Navigation in Semantically Enriched Information Spaces. In M. C. Angelides, P. Mylonas, and M. Wallace, editors, *Advances in Semantic Media Adaptation and Personalization*, vol. 2, pages 181–201. CRC Press, 2009.
- R. Filkorn, M. Barla, P. Bartalos, P. Sivák, K. Szobi, M. Tvarožek. Ontology as an Information Base for a Family of Domain Oriented Portal Solutions. In G. Knapp, G. Wojtkowski, J. Zupancic, and S. Wrycza, editors, *Advances in Information Systems Development: New Methods and Practice for the Networked Society* volume 1, pages 423–433. Springer, 2007.
- J. Šimko, M. Tvarožek, M. Bieliková. Semantic History Map: Graphs Aiding Web Revisitation Support. In Proceedings of DEXA/WebS 2010: 9th International Workshop on Web Semantics, pages 206–210. IEEE CS Press, 2010.
- A. Andrejko, M. Barla, M. Bieliková, M. Tvarožek. User Characteristics Acquisition from Logs with Semantics. In A. Kelemenová, D. Kolář, A. Meduna, and J. Zendulka, editors, *ISIM 2007: Proceedings of the 10th International Conference on Information System Implementation and Modeling*, pages 103–110. 2007.
- M. Barla, P. Bartalos, M. Bieliková, R. Filkorn, M. Tvarožek. Adaptive Portal Framework for Semantic Web Applications. In M. Brambilla, and E. Mendes, editors, *AEWSE'07: Proceedings of the Second International Workshop on Adaptation and Evolution in Web Systems Engineering*, pages 87–9). 2007.
- M. Tvarožek, M. Adam, M. Barla, P. Sivák, M. Bieliková. Spot-it: Going Beyond the Vision Loss Boundaries. In *ISIE2006: Proceedings of International Symposium on Intelligent Environments - Improving the quality of life in a changing world*, pages 67–76. Microsoft Research Ltd, 2006.
- M. Tvarožek. Software quality attributes, Case study design (in Slovak). In M. Bieliková, P. Návrát, M. Barla, P. Bartalos, M. Ciglan, J. Hamar, et al., editors, *Selected studies on software and information systems*, Vol. 3, pages 14-24,58-70. STU Press, 2007.
- A. Andrejko, M. Barla, M. Bieliková, M. Tvarožek. Software tools for user characteristics acquisition. In P. Vojtáš, and T. Skopal, editors, *Proceedings of DATAKON '06*, pages 139-148. 2006.
- M. Tvarožek. Exploratory Search in the Adaptive Social Semantic Web. In M. Bieliková, and P. Návrát, editors, *Workshop on the Web – Science, Technologies and Engineering*, pages 21-22. STU Press, 2010.
- M. Tvarožek. Semantic Based Navigation in Open Spaces. In M. Bieliková, and P. Návrát, editors, *Selected studies on software and information systems*, Vol. 4, pages 273-302. STU Press, 2009.
- M. Tvarožek, M. Bieliková. Adaptive navigation, presentation and search in the large information space of the Semantic Web (in Slovak). In F. Babič, and J. Paralič, editors, *WIKT 2007: Proceedings of the 2nd Workshop on Intelligent and Knowledge oriented Technologies*, pages 76-80. 2008.
- M. Tvarožek, M. Bieliková. Integrating Web and Semantic Web browsing via a next generation browser (in Slovak). In P. Návrát, and V. Vranič, editors, *WIKT 2008: Proceedings of the 3rd Workshop on Intelligent and Knowledge oriented Technologies*, pages 1-4. STU Press, 2009.
- M. Tvarožek, M. Bieliková. Personalized navigation for support of search and understanding of information in the Semantic Web (in Slovak). In M. Laclavík, I. Budinská, and L. Hluchý, editors, *WIKT 2006: Proceedings of the 1st Workshop on Intelligent and Knowledge oriented Technologies*, pages 37-40. 2007.