Proc. Int. Conf. on Dublin Core and Metadata for e-Communities 2002: 229-234

Visual Representation and Contextualization of Search Results – List and Matrix Browser

Christoph Kunz, Veit Botsch
Fraunhofer IAO
Nobelstrasse 12
D-70569 Stuttgart
Tel +49-711-970-2334
{christoph.kunz, veit.botsch}@iao.fhg.de

Abstract

We present a new approach for the representation of search results in a graphical user interface that allows navigating and exploring these results. An interactive matrix display is used for showing the hyperlinks of a site search or other search queries in different hierarchical category systems. The results of a site search are not only shown as a list, but also classified in an ontology-based category system. So the user has the possibility to explore and navigate within the results of his query. The system offers a flexible way to refine the query by drilling down in the hierarchical structured categories. The user can explore the results in one category with the so called List Browser or in two categories at the same time with the so called Matrix Browser (Ziegler et al. 2002). A familiar and well known interactive tree widget is used for the presentation of the categories and located hyperlinks, so the handling of the system is very intuitive.
Keywords: *Search engine, meta data, ontology, matrix browser, list browser, topic map, information visualization, classification*

Introduction

Networked information structures are becoming increasingly important for exploring and navigating complex information spaces, such as Internet sites, knowledge repositories or engineering data. Information networks are also becoming important in the context of the Semantic Web (Berners-Lee et al. 2001) as metadata or ontologies for information indexing. Complex ontological information can, for instance, be expressed in formalisms such as Topic Maps (Biezunski et al. 1999) or DAML+OIL (Hendler 2001). Visualizing and exploring such network structures, however, still constitutes a major problem for user interface design, in terms of minimizing visual search, supporting user's understanding and providing efficient interaction for exploring the network.

The exponentially growing amount of information available for example on the internet, in an intranet or a file system increases the interest in the task of retrieving information of interest. Search engines usually return more than 1 500 results per query and the results are displayed in a plain list with only few meta information. In general, people have two ways to find the data they are looking for: they can search by entering keywords to retrieve documents that contain these keywords, or they can browse through a hierarchy of subjects until the area of interest has been reached. The two tasks of searching and browsing are separated in most of the search engines. The information located in the hierarchy of subjects is not used to classify and to display the search results.

The approach presented in this paper combines the two ways of searching and exploring information spaces in a new graphical user interface for search engines. Information units of the underlying information space are linked with the metadata layer as occurrences of the ontology topics (see Figure 1). Every information unit can be linked with different topics in different categories of the metadata structure. The ontology itself must be analyzed in a manner that structures with hierarchical or transitive properties can be automatically recognized and extracted (Ziegler et al. 2002), so the user gets the possibility to choose hierarchical parts of the whole ontology for representation of the search results. Choosing a specific domain by selecting a given hierarchical category system allows the user to refine his query and to get a structured result list. The list and matrix browser are used as front ends for the representation and navigation of search results, which are classified in an ontology-based, hierarchical category system. The results of a keyword search are prestructured by the system using the ontology-based metadata. So the user can navigate and explore the result set

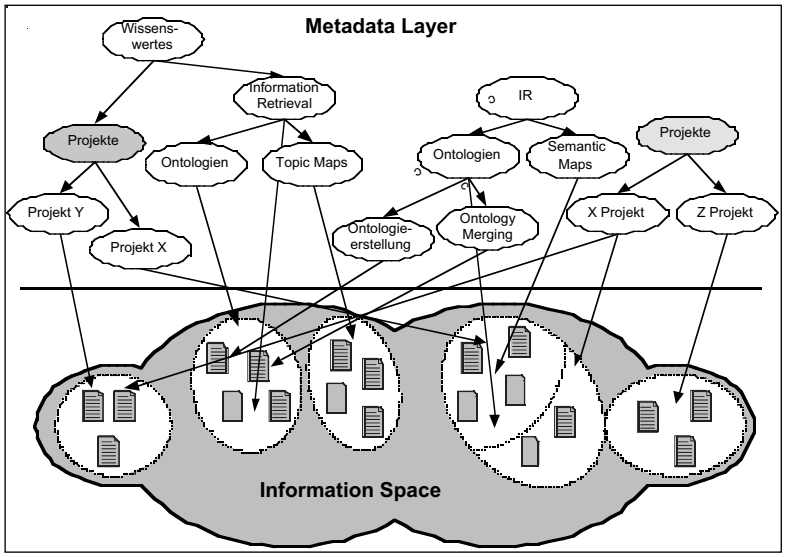

Figure 1. The Metadata Layer and the Information Space

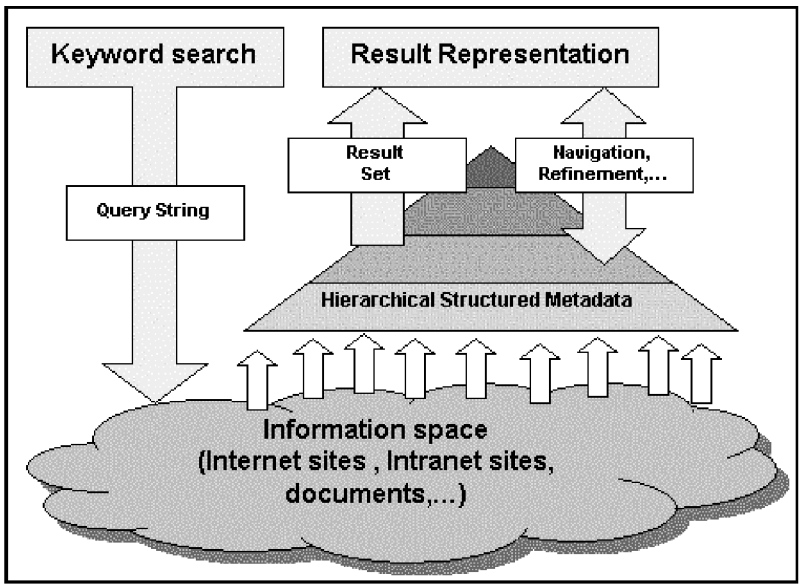

Figure 2. The search process

using the new graphical interfaces, without entering new queries (see Figure 2). The user can look at the returned hyperlinks with different, interactive "glasses", the chosen category system represents the glass.

The results of a keyword search are either displayed in one interactive category tree, the list browser or in an adjacency-like system with two interactive trees on the axis of a matrix, the matrix browser. In both cases the system provides a good review about the results and it offers the possibility to refine the query in a flexible manner. The first overview generated from the graphical user interface shows how many hits are found in the hierarchical ordered categories. If only one hierarchy is displayed, the number of hits is equivalent to the size of the displayed bar, if the user selected two hierarchies to structure the search results, the size of the squares inside the matrix encodes the number of hits belonging to the two categories. The user can explore the search result by expanding and collapsing the interactive, "window explorer" trees containing the hierarchical categories and the hyperlinks of the result set. If the user choose another hierarchical part of the category to visualize the search result, no re-querying is needed, because the result set is the same, only the representation has to be changed.

Related Work

Hierarchical ordered category systems are used in several search engines in the web (e.g. www.yahoo.com, www.google.de, www.dmoz.org, www.altavista.com) to allow the user not only to search the internet by keywords, but also present structured hyperlinks which can be explored by users. A new kind of search engine which presents the results in a graphical manner and which allows query expansion using a metadata system can be found at www.kartoo.com (see Figure 3).

The USU KnowledgeMiner is a modular software product for the structure of topics and for rendering access to information in heterogeneous IT environments uniform. The meta data extracted from existing data sources are semantically linked based on the topic map standards ISO 13250 and enable access to information from one central point. The structure thus established is displayed graphically (see Figure 4). The user quickly obtains an overview of topics and is led to the desired information via the navigation and retrieval components

A search engine which use a personalized fuzzy ontology to refine queries is proposed by Widyantoro (Widyantoro 2001). Gauch et al. (Gauch et al. 1999) use a personalized ontology for a re-ranking and filtering of query results. The search system incorporates users' interest into the search process to improve the results by creating user profiles. The

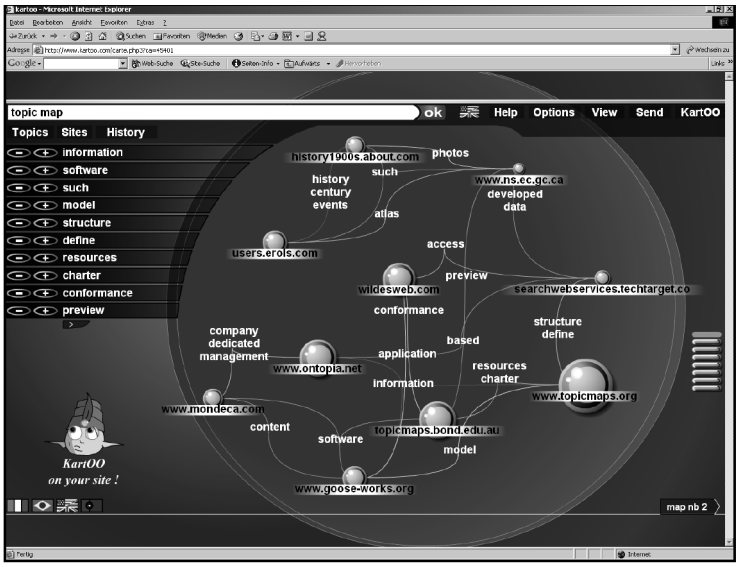

Figure 3. Search engine www.kartoo.com

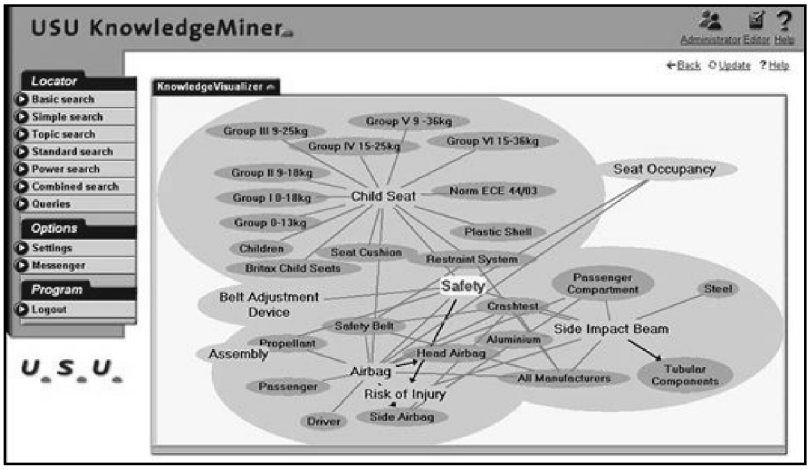

Figure 4. The USU KnowledgeMiner

user profiles are structured as a concept hierarchy of 4,400 nodes and are populated by 'watching over a user's shoulder' while he is surfing.

System Overview

The idea of the List and Matrix Browser visualization is based on different strategies, like usage of known metaphors, high interactivity and user-controlled diminution of the amount of information by filtering mechanisms. The user has different options to reduce the shown information and also to increase the interactivity of the representation. One possibility is to narrow or to widen the search query by using the well known keyword search. Another possibility is to select specific category hierarchies, which are used to refine the query and to classify the result set.

Hierarchies placed along one (List Browser) or two axes (Matrix Browser) can be explored directly by expanding and collapsing their interactive trees. This kind of exploration is familiar and effective and works both on the horizontal and the vertical axis. By expanding and collapsing the familiar interactive tree of the hierarchies, the user can increase or reduce the displayed amount of information and refine his query.

Requirements

The system needs an information space with a metadata network which is linked with the instances of the information space. Topics of the metadata structure contain occurrences (hyperlinks) of the information units. A site search engine with a category system is a possible domain for the new graphical user interface. Another potential usecase is a document management system which offers a structured hierarchical folder system and metadata belonging to specific documents. The results of searches in such systems with more than one hierarchical category system can be displayed either by the List or by the Matrix Browser. The information resource itself should be referred to in more than one category, other than in search engines like yahoo or google, where websites are only in one specific category.

List Browser

If the information structure only contains one hierarchy of categories or if the user chooses only one hierarchy of categories, the results of a keyword search are displayed in an interactive representation of the categories, the List Browser (seeFigure 5). The categories, all subcategories and the hyperlinks to the sites of the result set are listed in a "Windows Explorer"-like tree widget, which allows the user to expand and collapse the categories and subcategories by clicking on the interactive symbols in front of the bar and the category name. Not all categories of the metadata representation are listed in the List Browser, empty categories, where no hits of the search are located, are not visible in the representa-

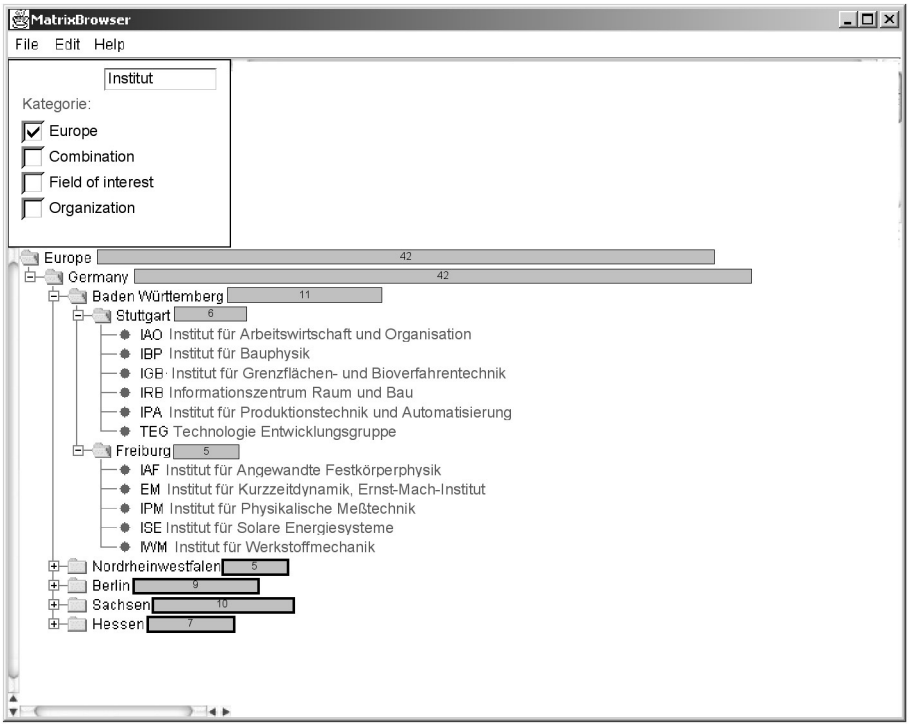

Figure 5. The List Browser

tion of the search result. The size of the bar represents the consolidated amount of search hits in the category and all subcategories. So the user gets an overview in which categories the search results are located and he can explore the results in a flexible manner.

Matrix Browser

The Matrix Browser (see Figure 6) allows the user to display the result list in two different categories. Choosing two categories also refines the query itself, because only links in the chosen categories are shown. The chosen categories are displayed as interactive, "Windows Explorer"-like trees on the two axis of an adjacency matrix. The user has the possibility to navigate within these trees and to explore the structured search result without re-querying the information space. The interactive matrix gives a good review about the search result and the reference of the found sites to the metadata category system. Different kinds of interactive symbols inside the matrix visualize on the one hand how many sites are found in two categories (the size of the circles) and on the other hand the site reference itself (the squares). Like in the List Browser the hyperlinks to the located sites are listed in the interactive category trees. If the hyperlink is listed in both trees, the horizontal and the vertical, this hyperlinks are connected visually with a square if both hyperlinks are in the focus. If one of the hyperlinks is in a collapsed state the square contains a plus and works interactively, so the user can also click on the square to expand the appropriate category. If a site is located in both category systems, but both categories are in a collapsed state, then a circle is shown in the matrix. The size of the circle represents how many located links are in the hierarchy under the referred nodes. Some hyperlinks are only listed in one hierarchy, they don't have a link to the other category system. In Figure 6 the listed institute CLS has no link to the category "Europe", because the institute is located in America.

Further Developments

Hyperlinks between sites of the search results could also be shown in the Matrix Browser, using another symbol for this kind of link. This additional information shows the level of networking between the located sites. We also think about including the consolidated values of located sites for the categories, already shown in the List Browser, in the Matrix Browser. An important factor for the backend system is the automatic generation of hierarchical structured metadata, which is linked with the information units.

First user evaluations of the new graphical user interface have to take place to improve further developments of the visual front end for search engines. The task of searching and browsing with the List and Matrix Browser has to be compared with other keyword-based search engines and category systems.

Conclusions

This paper described a new graphical user interface for ontology-based search engines, that allow the user to navigate and explore the results in a familiar

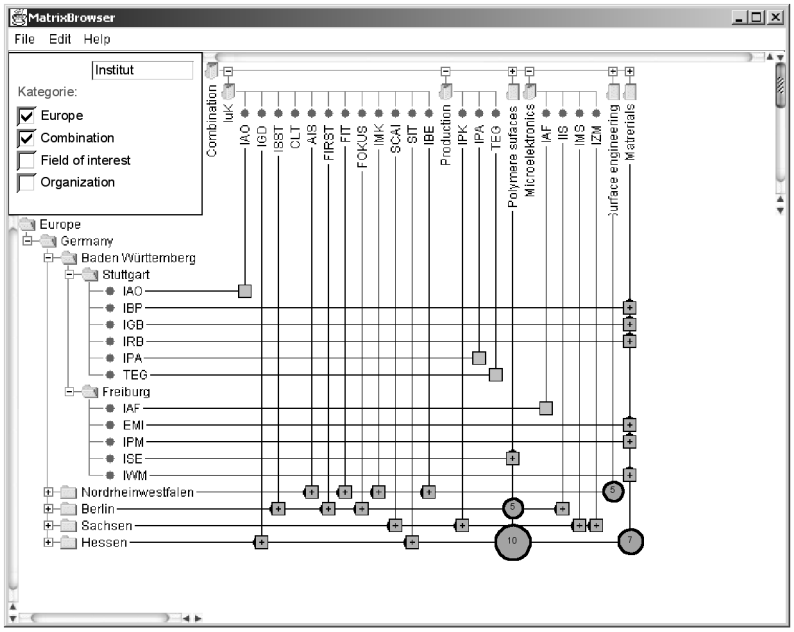

Figure 6. The Matrix Browser

and intuitive manner. Hierarchical category systems are displayed in interactive tree widgets, so the user can increase or reduce the displayed amount of information and refine his query without input of any text data. The two tasks of searching and browsing are combined in one graphical interface. Using the Matrix Browser provides an opportunity to visualize more details of the metadata structure together with the located sites. The result set of a keyword search is shown in a part of the metadata structure, so the user can chose different "glasses" (parts of the metadata structure) to look at the results.

References

Berners-Lee, T., Hendler, J., Lassila, O. 2001. The Semantic Web, Scientifc American, Vol. 284, p 34-43.

Biezunski, M., Bryan, M., Newcomb, S. R. 1999. ISO/IEC 13250 Topic Maps: Information Technology — Document Description and Markup Language. http://www.ornl.gov/sgml/sc34/document/0058.htm.

Gauch, Susan; Pretschner, Alexander 1999. Ontology Based Personalized Search. In: Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence, Chicago, 1999, pp. 391-398.

Hendler, J. 2001. DAML: The DARPA Agent Markup Language Homepage. http://www.daml.org.

Widyantot, D.H., Yen, J. 2001. Using Fuzzy Ontology for Query Refinement in a Personalized Abstract Search Engine. In: Proceedings of Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vancouver, Canada, 2001.

Ziegler, Jürgen; Kunz, Christoph; Botsch, Veit 2002. Matrix Browser - Visualizing and Exploring Large Networked Information Spaces, In: Proceedings of the International Conference on Computer Human Interaction SIGCHI 2002, ACM Press, Minneapolis, 2002.

Ziegler, Jürgen; Kunz, Christoph; Botsch, Veit; Schneeberger, Josef 2002. Visualizing and Exploring Large Networked Information Spaces with Matrix Browser. In: Proceedings of 6th International Conference on Information Visualization, London, England, 2002.