Open Interface for Connection of Different Search Engines to d.3

Today we recognize that the mere availability of information is not of great use. Only the rapid access to relevant information allows us to extend our knowledge and to quickly make the right decisions. With d.3 search, you get an intuitive access to unstructured information.

As an open interface, d.3 search allows to connect different full text search engines to d.3 archives in order to provide extensive, comfortable search options to the user. d.3 search constitutes an intermediary component between the d.3 server and the according search engines. On account of the search engines specialized in the search task, new possibilities for document searching (phonetic, semantic search, document ranking according to relevance, free text search, search for similar documents etc.) arise. With the integrated d.3 search engine and the infolution engine, which is available as an additional component, d.3 offers two alternative and partly complementary options for the full text search that is by default integrated in d.3 and based on relational databases.

Search engines such as Google are distinctive specialists. They make use of own data structures optimized for the specific task and do not resort to the general relational data model.

Therefore, these programs are capable of solving the search task faster and with additional functionality.

Instead of saving full text information itself, the d.3 server passes, through a uniform d.3 search API, these texts and up-to-date document attributes over to the d.3 search server for the purpose of key word generation. The d.3 search server can be connected through TCP/IP and the familiar client/server communication.

Special link-up modules convert these general inquiries for the particular search engine.

When searching documents, d.3 search provides a list of document IDs with the respective relevance to the d.3 server that in turn executes the authorization verification. The inquiry syntax is as well decoded by the link-up modules in order to confront the user only with the general d.3 inquiry language.

Search Options with d.3 search pro

Phonetic/fuzzy search: Different notations (e.g. of names) prevent that documents can be found if the notation is unknown. To a certain extent, every employee has his or her own idea of spelling. With phonetic search, even unisonous or similarly spelled words are supposed to be found. Especially in case of similar notations, the editing distance (Levenstein distance) serves as the yardstick. The more single steps (deleting, pasting of letters) are necessary in order to convert one word into another, the greater this distance is.

Semantic/associative search denotes the search for words that are similar as regards content. Thus, the search for "pantywaist" finds even documents containing "wimp". This search is the specialty of the infolution engine.

The free text search identifies relevant documents by means of a text extract. In case the user comes across a circumstance in a text, this text extract can be co-

d.3 search

Features

- optional extension of d.3 for the purpose of filing of full text information of the documents
- integrated search engine
- d.3 search engine
- utilization of search engines of different suppliers, e.g. infolution engine
- speed optimization at document import
- extended, clear-cut search options (phonetic, semantic/associative)

- search by means of semantic clouds on the basis of neural networks allows e.g. knowledge workers to quickly get an overview of huge unstructured information quantities
- automatic generation of abstracts accelerates the recognizing and judging of a hit

d.velop

pied to the text search field via clipboard. Then, associated documents are searched for in the d.3 archive.

In case of the search for similar documents the full text information of a d.3 document is interpreted as the search inquiry in order to find related documents as regards content. This function can be accessed through the option "search for similar documents" in the context menu.

The identification of search suggestions serves the purpose of restricting an extensive hit volume to a particular subtopic. These suggestions appear in the title row of the search hits in the d.3 explorer.

Through the recognition of phrases the search result can be considerably improved. The algorithm recognizes phrases in a self-learning way and provides the phrases "The Washington Times", "George Washington", Washington Redskins" etc. for the search term "Washington".

In addition, d.3 search provides the opportunity to generate summaries of documents. These summaries are independent of the search and variable as concerns their length. In this way, the user can evaluate the hit volume significantly faster and is ideally supported in his or her work.

Modular Composition

A general search inquiry is accomplished the way known from Google and is entered into the full text field in the d.3 ex-
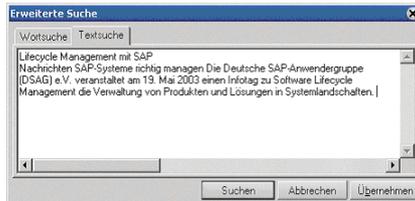

Fig. 1: Thanks to free text search, users can intuitively search for similar content.

plorer. Space characters combine words with a logical AND; the term OR stands for the OR combination.

Negative words are marked by a prefixed minus and *,? serve as wildcard symbols. ~ denotes the fuzzy search for the subse-

quent term. Inverted commas around several search terms define the search for the exact phrase.
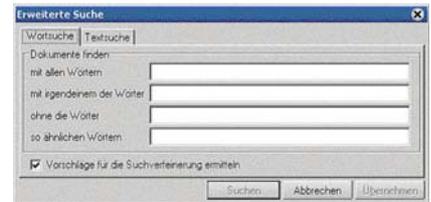

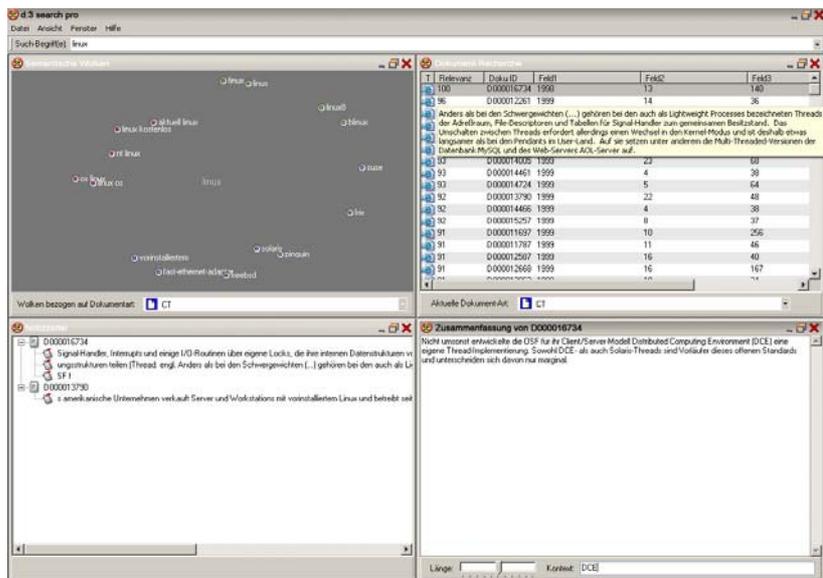Fig. 3: For the user who does not want to learn this syntax, a mask is available in the d.3 explorer

Fig. 2: d.3 search pro offers, besides full text search, the opportunity to search through associative combinations, through phrases as well as through fuzzy terms. Additionally, d.3 search provides the possibility to generate summaries of documents.

d.3 search

Anforderungen

Software: Windows NT 4.0, 2000, XP

Hardware: PIII, 650 Mhz, 256 MB RAM, 10 GB available space

Contact

Do you still have questions?
Please don't hesitate to contact us:

d.velop AG
digital business solutions

Schildarpstraße 6–8
GER - 48712 Gescher

phone: +49 (0) 2542 – 9307-0
fax: +49 (0) 2542 – 9307-20

email: info@d.velop.de