NEWSLETTER

Issue #4
October
2012

TOPICS
in this issue

Editorial

Allan Hanbury,
Vienna University of Technology, Austria
hanbury@ifs.tuwien.ac.at

The Khresmoi project has reached the end of the second year of its four year plan. This is an exciting time in the project as the first relatively complete and usable prototypes begin to appear, and the progress toward the Khresmoi vision becomes tangible. The Khresmoi system is constructed of a number of components. ezDL is the component that provides the search interface, and is described in more detail in this newsletter. Further components of the system include semantic search (presented in the previous edition of the newsletter), spelling correction, query completion and query translation. In such a complex system, it is important to understand the components' effects on each other and on the output of the complete system. For example, how does the query translation system handle badly spelled queries that have been erroneously corrected by the spelling checker? Doing such an evaluation in a quantitative way is challenging, and the second article in this newsletter describes the Khresmoi strategy for performing this evaluation. A further evaluation beginning soon is the user-centred evaluation, in which end users will carry out search tasks on the Khresmoi system and share their opinions and impressions of the system, which will guide the further development in the project. The Khresmoi webpage has also recently been redesigned to better organise the growing amount of information on the site, and all are invited to visit.

ezDL: Interactive Search Interface for Khresmoi

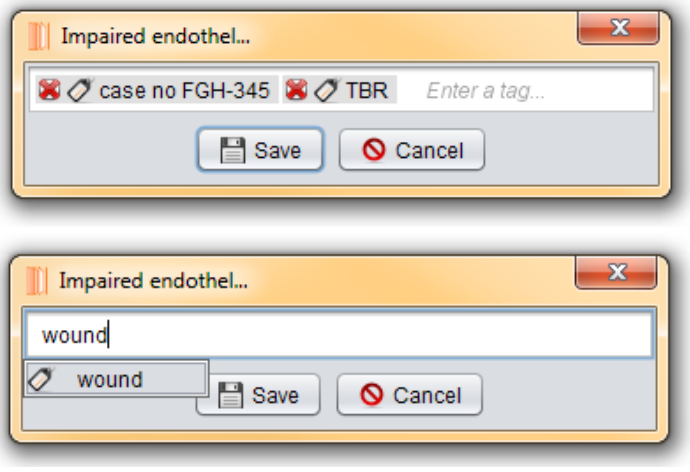
Sascha Kriewel
University of Duisburg-Essen, Germany
ikriewel@is.inf.uni-due.de

ezDL (short for "easy access to Digital Libraries") is the basis of the Khresmoi search interface. Originally designed for digital libraries, it has been extended to provide an interactive user interface to a large variety of search systems such as Khresmoi. The system allows for many different clients (such as a Java webstart application, an AJAX application, or an Android app) using common back-end services for user authorization, query conversion, or search suggestions. The version 1.7 of the ezDL Java client will be released in early October and will include many new features developed within the Khresmoi project.

The most noticeable improvements are the newly developed collaborative and organizing functionalities. The ezDL interface allows registered users to keep and organize results beyond a search session within a personal library. If you discover an interesting document during your search you can store it using the context menu, drag it to the open library, or save it with a button click from the document details. All documents that have been saved to the personal library will be marked with the personal library icon – so it is easy to see which documents have already been saved.

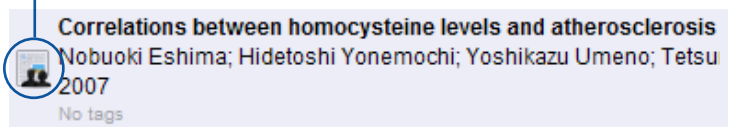
Saved results can be sorted or grouped by author, publication year, title and date of addition. In addition, it is possible to apply a filter to the personal library to quickly find a stored document – and of course the documents in the library can be exported and printed.

Tags can be used to organize the personal library according to a user's individual needs. For example, a user might want to tag documents that she still wants to read with "TBR" (to be read) or a physician could add tags corresponding to a specific case or patient that he is working on. Tags can be used to group the documents in the personal library.



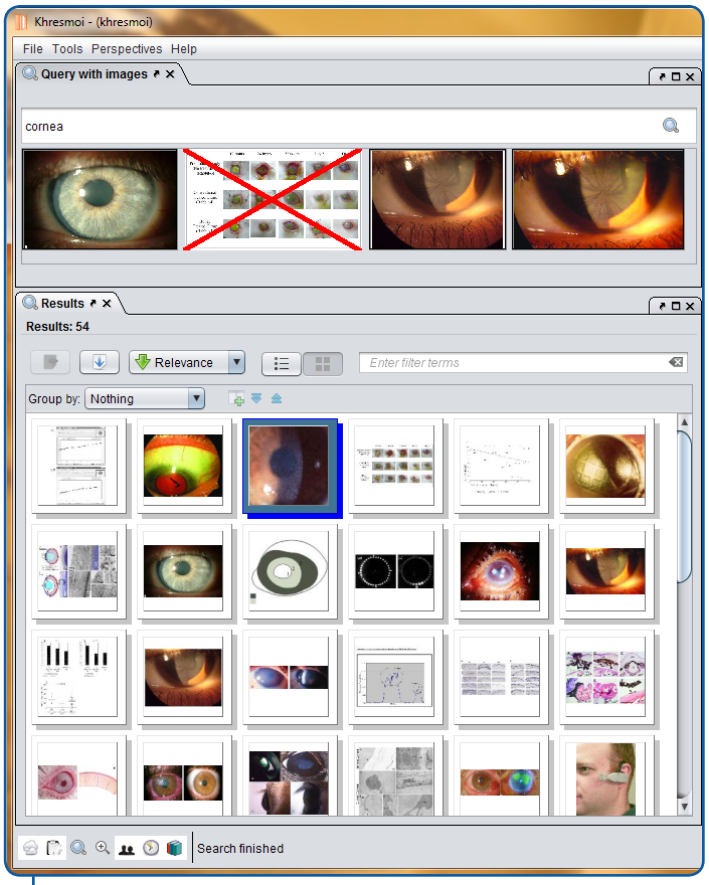
The personal library not only allows for personal information management, but can also be used to collaborate with other users. Documents that have been stored in the personal library (including new documents uploaded by the user) can be shared, and in future releases it will be possible for all users who share a particular

document to comment on it and discuss it with each other. Shared documents are marked with a special icon:



To facilitate collaboration and sharing, users can create a personal profile, which currently covers the user's name, country, preferred language, and a free-text description. The privacy setting allows users to control if they want to be found by other users. The search functionality of the interface can be used to search for users based on their name or the description used on their profile. To further support collaboration users can create their own personal contact and sharing lists, or public groups around a specific topic. These will allow for easier sharing of documents and discussions in future versions.

Another new functionality of ezDL is the support for image search by providing positive and negative examples. For search systems that allow similarity search, the "image search" perspective of ezDL can be used to collect example images as seen in the accompanying screen shot. The image examples can be used as positive or negative relevance feedback and allow for easy specification of queries that cannot be expressed through search terms. All previously found images can be used for searching, as can be images from a web browser, or even the local file system (provided the search system allows for upload of user images).



More information:

- ezDL on the web: http://ezdl.de
- Follow ezDL on twitter: https://twitter.com/ezdl

Global Empirical Evaluation Strategy (Investigating Khresmoi System Effectiveness)

Liadh Kelly, Lorraine Goeuriot, Gareth Jones
Dublin City University
lkelly@computing.dcu.ie

Achieving high levels of effectiveness in enabling users of the Khresmoi systems to find interesting and valuable information is a key motivator for the project research activities. Following a substantial evaluation design phase, project partners are currently engaged in the first of two large scale empirical and user-centred evaluation exercises. The overall aim of these evaluations is to enable an assessment of the efficiency and effectiveness of the component Khresmoi technologies and their efficacy in supporting the target users' medical information retrieval needs. The results of these evaluations will be used to guide further research and development in the project. In this, the first of a series of evaluation articles, we overview the empirical evaluation design, subsequent Newsletters will look at user-centered evaluation and findings from the evaluation exercises.

Our evaluation strategy builds on the existing strategies for evaluation of the individual components and partial system integration of the technologies used in the Khresmoi system. In the Khresmoi project we aim to develop a complete integrated evaluation plan for the components, their role within the system and the system as a whole. In developing a holistic empirical evaluation plan, the target system, users and data, and how system components interact to support these users are key. The target users of Khresmoi technologies are: the general public, general practitioners, and radiologists. This represents two use cases and two systems for the Khresmoi technologies, specifically:

1) multilingual textual retrieval for the general public and general practitioners; and

2) multilingual and multimedia case retrieval for radiologists. Search systems such as Khresmoi are becoming increasingly complex combining multiple technology components in order to facilitate their specified functionality. The Khresmoi system must combine information retrieval techniques and some form of

language translation, and also integrates image analysis, knowledge management, user annotation and summarization components, along with tools designed to enhance user interaction such as spellchecking and query suggestion. In evaluating complex systems of this nature we need to investigate the performance of the individual components, but also, and importantly, be mindful of the fact that the components do not operate in isolation, but rather interact with each other in the quest to meet a user's information needs.

Generating Evaluation Resources

Evaluating the performance of components and components in interaction requires the collection, selection and development of the resources required for their evaluation in a holistic strategy. An important consideration is that the evaluation search tasks should be representative of the expected activities of our selected user groups. The target users of the Khresmoi system are diverse (speaking different languages, having different medical knowledge levels and differing levels of knowledge of the languages of the target documents). This means that input test data and the content searched and deemed relevant must be representative of the expected search tasks of these different classes of user. Suitable document sets have been identified and collected from the Web and professional organisations. Selection of this data must be realistic for the working environment of available data, and enable testing of multilingual technologies. Languages currently being supported within Khresmoi include English, French, German and Czech. Suitable sets of queries were sought in the first instance from existing query logs, for example from the Khresmoi partner Health on the Net and then manually translated in a culturally sensitive way into all languages of interest to Khresmoi.

Evaluating Components

Each of the Khresmoi system components (translation tool, spellchecker, etc) is evaluated in isolation using the use case driven generated resources. For example, the spellchecker is evaluated by measuring its ability to correct spelling errors; and the multilingual summarization functionality is evaluated in terms of the presence of required information in the summaries, translation accuracy and reading fluency.

The effectiveness of any of the system components may impact on system overall effectiveness, but this may also depend on how the components interact within the system. This means for example that beyond measuring the quality of the translation component in isolation, how the translation component operates when integrated into the overall system needs to be investigated. Thus in order to properly understand the effectiveness and impact of components within a complex search system, the individual components must be evaluated in a holistic way within the integrated system. Examples of the studies which will be carried out include: the impact of spelling mistakes in queries on the integrated search and the improvements available via, potentially erroneous, spelling correction; and the impact of query translation quality on the quality of retrieval.

Project News

Summer School on Image Processing 2012

Khresmoi partners Vienna University of Technology and Medical University of Vienna were the main organisers of the Summer School on Image Processing (SSIP) from July 4th to 13th in Vienna, Austria. The summer school was attended by 45 Students from 23 countries. In 2012, the theme of the summer school was Medical Image Analysis. Apart from lectures on this topic, students worked in international project groups to solve medical image analysis problems on provided medical image data. The prize for the best project work in 2012 was awarded to Miroslav Radojevic (Erasmus University Medical Center, The Netherlands), Roberto Gatta (University of Brescia, Italy), Jan Schier (Academy of Sciences of the Czech Republic) and Luka Šajn (University of Ljubljana, Slovenia) for their project on 3D Volume localization image retrieval using miniatures.

For more information, see
http://www.cir.meduniwien.ac.at/ssip2012

Khresmoi at the LREC 2012

Khresmoi had a booth in the EU Projects Village of the Language Resources and Evaluation Conference (LREC) in Istanbul, Turkey from May 21st to 27th 2012. This activity resulted in good contacts being made with EU projects working in the language technology area. We were honoured to be able to present Khresmoi to Deputy Minister Prof. Dr. Davut Kavranoğlu of the Ministry of Science, Industry and Technology of Turkey during his visit to the booth. Allan Hanbury also presented an overview of the Khresmoi project in the special session on EU projects. In the photo Allan Hanbury explains Khresmoi to Thierry Declerck of the DFKI, Germany at the Khresmoi LREC booth.

Khresmoi at the MIE 2012

Khresmoi was very present at the Medical Informatics Europe (MIE) Conference in Pisa, Italy from August 26th to 29th 2012. The project had a stand in the Village of the Future, and a presentation was given in the Village of the Future session on People and Expectations. In this session, the scenario of Little Sam was considered. Sam is diagnosed with Cystic Fibrosis (CF) at an early age, and makes use of internet search engines to get information about the disease, and social networks and blogs to get into contact with fellow CF patients. The importance of access to trustable online medical information and the key role that search technology plays in this access was underlined in this session. Furthermore, Henning Müller and Allan Hanbury presented a tutorial on Searching Text and Images in the Medical Domain. Papers on Khresmoi results were also presented in the conference sessions.

Prestigious Award for Khresmoi Researcher

Norbert Fuhr, Professor at the University of Duisburg-Essen, a partner in the Khresmoi project, was presented with the Gerard Salton Award at the Association of Computing Machinery (ACM) Special Interest Group on Information Retrieval (SIGIR) 2012 Conference. The award honours his pioneering, sustained, and continuing contributions to the theoretical foundations of information retrieval and database systems. In Khresmoi, Norbert Fuhr's group is developing the ezDL interactive search interface described in this newsletter.

Upcoming events

- A demonstration of the Khresmoi system for medical information search by physicians and members of the general public will be given at the Health 2.0 Europe Conference in Berlin, Germany on November 6th to 7th 2012.
- The Khresmoi clinical radiology prototype will be demonstrated at the annual meeting of the Radiological Society of North America (RSNA) in Chicago, USA from November 25th to 30th 2012.
- Khresmoi technology will be demonstrated at the stand for EU language technology and big data projects at the CeBIT 2013 in Hannover, Germany from March 5th to 9th 2013.

Follow Khresmoi on Twitter: https://twitter.com/khresmoi

To order the newsletter and receive regular updates:
http://www.khresmoi.eu/contact-us/keep-me-updated/

At a glance

Project Coordinator
Henning Müller
University of Applied Sciences
Western Switzerland
Tel: +41 27 606 90 36
Fax: +41 27 606 90 00
henning.mueller[at]hevs.ch

Scientific Coordinator
Allan Hanbury
Vienna University of Technology
Tel: +43 1 58801 188310
allan.hanbury[at]tuwien.ac.at

Partners
University of Applied Sciences Western Switzerland (CH),
Vienna University of Technology (AT),
Atos Origin (ES),
ELDA (FR),
Ontotext (BG),
Dublin City University (IE),
University of Duisburg-Essen (DE),
Charles University in Prague (CZ),
The University of Sheffield (UK),
Health on the Net (CH),
Medical University of Vienna (AT),
Society of Physicians in Vienna (AT)

Duration:
09/2010 - 08/2014

Funding scheme: IP

Total Cost:
€ 10.534 m

EC Contribution:
€ 8.036 m

GA number: 257528