NEWSLETTER

Issue #6
August
2014

Editorial – Four years of Khresmoi

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

TOPICS
in this issue

Imagine that a radiologist is examining a computer tomography image and sees an anomaly that he has never come across before. A commonly adopted approach to identifying the anomaly is to page through books to try and find a similar-looking anomaly, or to ask a colleague if she knows what the anomaly is. With Khresmoi, he can do an automated search for similar-looking anomalies in cases in the hospital digital archives. He can then use anonymised reports written by his colleagues attached to the similar cases to guide and support his diagnosis. Based on a text analysis of the returned reports, Khresmoi also launches a search in the medical literature. Even if the returned radiology reports are in German, Khresmoi can access relevant publications in the much larger set of English medical publications. In effect, the radiologist gains immediate access to a huge amount of implicit medical knowledge of his colleagues and relevant information from the medical literature without entering a single search keyword.

Four years ago, such a scenario was not possible. Now that the Khresmoi project is nearing its end, the Khresmoi Radiology prototype returns all of the information listed above in less than 4 seconds after the radiologist has marked an anomalous region in an image and pressed the search button.

Khresmoi Radiology is the most exciting result to come out of the Khresmoi project, which will be ending on the 31st of August 2014 after four years of research and development in the area of medical information search and retrieval. Khresmoi Radiology integrates the largest number of technologies developed in Khresmoi into a single system. Khresmoi has developed a data-driven approach to visual similarity search in 3D medical images, taking advantage of the terabytes of medical images stored in hospital archives. As the approach is data-driven, it can be applied anywhere in the human body without the necessity for careful hand-crafting of techniques specific to various organs. Automated semantic annotation of the text of the radiology reports finds mentions of organs and medical conditions, which are used to propose a consensus diagnosis from a group of retrieved radiology reports and automatically construct a query for a search in the medical literature. A key characteristic has been the use of standard medical terminologies in both text and image analysis, allowing straightforward combination of the analysis results for both text and image modalities. Methods for large-scale search and classification of images in the medical literature have also been developed in Khresmoi.

The Khresmoi technology is also used to provide powerful access to medical information through text search. The semantic annotation linking documents to medical terminologies allows the user to easily locate relevant documents even if they only contain synonyms of the words used in the query. But even more powerful queries are possible, such as finding documents that mention medication used to treat diabetes, or documents mentioning diseases that have a dry cough as a symptom. The medicine-specific machine translation allows users to search English documents while entering a query in the language that they are most comfortable with. This currently works for French, German and Czech, but the statistical machine translation approach adopted makes the extension to further languages relatively straightforward. The prototypes specialised on text search have been optimised for two end user groups: Khresmoi Professional is designed for medical professionals, while Khresmoi for Everyone is easy to use by all. Khresmoi for Everyone puts particular emphasis on ensuring the trustworthiness of the websites presented in the search results.

A particular strength of Khresmoi has been the involvement of medical professionals and patients from the design phase to the testing phase, influencing all aspects of the project. In particular, all prototypes have been tested by the actual members of the target user group. Khresmoi Radiology has been evaluated by radiologists in hospitals in Austria, Germany, Switzerland and Greece; Khresmoi Professional has been tested by medical doctors while they attended symposia in Austria and Germany; while Khresmoi for Everyone has been tested by a diverse group of members of the general public in France, Switzerland and the Czech Republic, including patients in a hospital in France.

Around 50 people from 12 organisations worked together over four years to develop this innovative technology and produce new research results, while gaining invaluable experience in areas ranging from system integration to international cooperation. Young researchers have earned their PhD degrees, post-doctoral researchers have taken their first steps toward independent research, and more senior staff

have overcome the organisational challenges presented by such a large-scale multinational research and development project. But what happens now that the Khresmoi project is over? Are we going to turn off the servers and disappear? Not if we can prevent it. There are two initiatives to further develop Khresmoi technologies that are in too early a stage to write about in detail here. One has the target of making the Khresmoi medical text annotation, semantic search and machine translation available as commercial-grade web services and to adapt these technologies to patient record processing, while the other initiative deals with bringing the Khresmoi Radiology technology to the market. Keep watching the Khresmoi web page for updates on these initiatives

Khresmoi Prototypes

- Khresmoi for Everyone: http://everyone.khresmoi.eu
- Khresmoi Professional: http://professional.khresmoi.eu
- Khresmoi Radiology: http://radiology.khresmoi.eu

Evaluating Khresmoi Professional with medical doctors

Marlene Kritz
Society of Physicians in Vienna, Austria
marlene.kritz@billrothhaus.at

Veronika Stefanov
Vienna University of Technology, Austria
stefanov@ifs.tuwien.ac.at

During the development of a search system, it can be challenging to convince members of targeted user group to act as testers of an unfinished system. Especially for professional groups with a high workload, such as medical doctors, evaluation is typically done with substitute groups, e.g. medical students.

However, current research and our survey at the beginning of the project have shown that information needs and search behaviour varies amongst different subgroups of medical doctors. We therefore considered it unlikely that students could give us a valid representation of our target user group's needs.

For the development of Khresmoi Professional, we implemented three rounds of user evaluations with doctors.

After the completion of an initial prototype of the search system, members of the Society of Physicians in Vienna were invited to hour-long evaluation sessions, during which they were asked to complete pre-defined search tasks on the Khresmoi Professional prototype. The tasks were designed so that they were solvable with the functionality available at the time.

The results were useful for further development of the early system but we found that the majority of our participants belonged to two groups: either young doctors fresh from university or retired doctors. Most practicing doctors seemed to be too busy to attend an extensive system evaluation. From the little input we got from them, we could however see that their needs and interests were indeed quite different from less time-constrained physicians.

For the next round of user tests, we decided to go and meet the doctors where they gather – at conferences and educational events. CME (continuing medical education) is mandatory for practicing doctors. We designed shorter test sessions that could be run during break times at medical conferences.

We first implemented this approach at STAFAM, the largest Austrian conference for general practitioners in autumn 2013, in Graz, Austria. Khresmoi project members set up a stand with laptops in the exhibition area, handed out flyers, information folders, and – most importantly – free coffee. Testers were given pre-defined tasks but also asked to use the system for a "free task" where they could search for whatever they wanted. Khresmoi Professional had been developed to a point where free exploration was becoming feasible. A significant number of bug-fixes, feature requests, and another large increase in content coverage resulted from these tests. While the final version of the system was being developed, we designed the third and final round of user tests. These were again to be run at conferences and events, but this time gathering feedback on a purely "free" task.

The final user evaluations were started at a two-day CME symposium in Wiesbaden, Germany, and continued at a number of smaller half-day events in Vienna. In addition, a bilingual online version of the evaluation questionnaire was made available, used to collect further feedback and include doctors from other countries.

The analysis of these extensive results is currently underway, and the report will be on the Khresmoi website within the next two months.

The Khresmoi Professional evaluations at the STAFAM

How well do search systems find medical information? The CLEF eHealth campaign

Lorraine Goeuriot
Dublin City University, Ireland
lgoeuriot@computing.dcu.ie

This year, similar to last, some members of the Khresmoi consortium are organisers of the CLEF eHealth evaluation lab, with two of its members leading the 2014 edition (Liadh Kelly and Lorraine Goeuriot, from Dublin City University). The lab is held as part of the Conference and Labs of the Evaluation Forum (CLEF). The first edition of CLEF eHealth, in 2013, included three evaluation tasks:

(1) Named entity recognition and normalization of disorders;
(2) Normalization of acronyms/abbreviations; and
(3) Information retrieval to address questions patients may have when reading clinical reports.

Task 3 was managed by members of the Khresmoi consortium, in collaboration with the University of Turku (Finland), CSIRO and NICTA (Australia). The datasets included a document crawl provided by Khresmoi, queries manually built by the nursing group at the University of Turku, and relevance judgements provided by this group. 175 people registered their interest in the lab (64, 56 and 55 respectively for tasks 1, 2 and 3), and 53 teams participated (39, 5 and 9 respectively for tasks 1, 2 and 3). Teams participating included renowned groups from the clinical/medical natural language processing (NLP) and information retrieval (IR) domains. The task organization led to two official CLEF publications for the Khresmoi team, and several additional analysis publications. Moreover, through the official release of the 2013 task 3 dataset, more teams should be using it and investigating new approaches to improve medical IR. A one day workshop was dedicated to CLEF eHealth at the CLEF 2013 conference in Valencia, where one team per task was invited to give a presentation, and several teams presented posters.

The 2014 edition of the lab also included three evaluation tasks:

(1) Visual-Interactive Search and Exploration of eHealth Data;
(2) Information extraction from clinical text;
(3) User-centred health information retrieval.

Again, task 3 was managed by members of the Khresmoi consortium, and a cross language subtask was added. The dataset was created in a similar manner to 2013. 224 people registered their interest in the lab (50, 79 and 55 respectively for tasks 1, 2 and 3), and 53 teams participated (1, 10 and 13 respectively for tasks 1, 2 and 3). Teams participating again included renowned groups from the clinical/medical NLP and IR domains. The task organization led to four official CLEF publications. The organizers and participants will gather this year at CLEF 2014 in Sheffield to report results for each task and learn from participants' presentations and posters.

Link: http://clefehealth2014.dcu.ie

Retrieval tied into the clinical workflow

Georg Langs
Medical University of Vienna, Austria
georg.langs@meduniwien.ac.at

Dimitrios Markonis, Henning Müller
University of Applied Sciences and Arts Western Switzerland
dimitrios.markonis@hevs.ch, henning.mueller@hevs.ch

The integration of Khresmoi retrieval technology in a clinical setting made large strides after getting feedback in the initial user evaluation round. The project teams focusing on image computing, user interface design, content rights, and semantics worked closely together and came up with a second version that is a substantial advance from the initial proof of concept implementation. We could improve several aspects and add features that were on the top of the wish-list of radiologists involved in the feedback and testing of the round one prototypes.

Speed and accuracy

When a user indicates a region of interest in an imaging volume such as a CT, and starts the retrieval, results of similar regions across thousands of cases are now shown within 4 seconds. During this time the visual features of the query region are compared to millions of indexed regions, top regions are identified, and imaging volumes are ranked based on the configuration of those regions. To provide the user with most informative feedback when browsing the search results, image thumbnails that show the relevant portion of the image are rendered. Overall result accuracy is much improved, and the system now accurately identifies similar anomaly patterns across images and patients. The improved accuracy is due to advanced feature extraction and learning methods developed and incorporated into the prototype. The speed of retrieval is due to new indexing algorithms that make the visual information of many millions of image segments comparable within seconds. This is not trivial, since the necessary information cannot be held in the memory, and intelligent query strategies are necessary to ensure speed, and at the same time minimise deviation of distance estimates encoded in the index from the actual distance between examples.

Semantics and visual information

Aside from improving speed and accuracy, the integration of visual and semantic information proved to open very exciting possibilities for which we have even now only scratched the surface. The semantic information corresponding to the search results, such as radiology reports, is mined to generate a summary of the observations made in the majority of reports. The mining algorithms "understand" the meaning of words and their categories based on terminologies such as RadLex. The prototype engine identifies relationships, and is able to determine that a report states that a certain anomaly was observed in a certain anatomical region. This allows for summarising and analysing the entire set of reports retrieved together with the top ranked search results. The structuring of the search result list gains from the information known about individual data. For example, it relates the individual report of a case in the index to the other reports in the search result list. This allows the system to provide simple tagging of results into a consensus set, and a set that might be relevant for differential diagnosis. This also happens within the 4 seconds until the search result is displayed.

Exploring additional information

After the search results are displayed, the user can explore additional relevant information that is linked to the query case, or even the query region, and the search results obtained from the hospital imaging database. One of the prominent requests by users was the ability to directly access medical literature and educational material, or reference cases and examples that can help during diagnosis. This search is now directly triggered by the initial image based search. Based on the consensus diagnosis extracted semantically from the reports, we query databases such as PubMed and the educational literature, which can provide representative examples and explanations for observations that might correspond to the query case. This happens before the user has even entered a single keyword. Now the search can be tuned based on textual user input, and the radiologists are enabled to further explore the relevant context of their query.

Khresmoi Radiology evaluations with end users

The Radiology search prototype has been evaluated in practice, when used by real users. 26 radiologists in Austria, Germany, Greece and Switzerland were asked to carry out a set of radiology-related information retrieval tasks. Their actions were monitored and their comments were recorded. The system was found to be easy to use and radiologists felt quickly comfortable with it. More importantly, less satisfactory aspects and open challenges for future research were identified. Content-based image retrieval using regions of interest needs to be further improved, especially when the regions contain multiple pathologies. Image search using the semantic information can assist as well but in a non-visible layer

The Khresmoi prototype for clinical radiologists in action

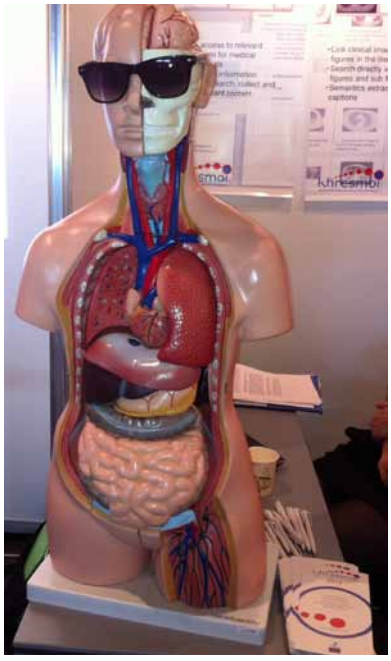
of the system as a crowded interface can sometimes be confusing to the user. In general the participants seemed to appreciate the added value that such a complete system would have in their work and most commented that they would use such a system.
Overall the second phase of prototype development brought a tight integration of components from image computing, semantic processing, and user interface design. It demonstrated how the Khresmoi approach of making components as easily useable across the project as possible enabled the experimentation with different approaches towards integrating retrieval in daily workflows, and also how it supported the development of rich data mining tools that cross borders between medical image analysis, semantic processing and search in a variety of relevant sources, beyond the clinical environment.

Project News

Workshop on Medical Information Retrieval

This year at the SIGIR conference (Special Interest Group On Information Retrieval), several members of the Khresmoi consortium (Lorraine Goeuriot, Liadh Kelly, Gareth Jones and Henning Müller), together with Justin Zobel (University of Melbourne), organised the first workshop on Medical Information Retrieval (MedIR). The objective of this workshop is to provide a forum to enable the progress of research in medical information retrieval to provide enhanced search services for all users with interests in medical information search. Participation from researchers in all fields related to medical information search including mainstream information retrieval, but also natural language processing, multilingual text processing, and medical image analysis was encouraged. The workshop programme committee gathers 15 experts from the domain. 20 papers were submitted to the workshop, of which 8 were accepted. Four of the accepted papers were written by Australian groups, three from European ones and one from a Singaporean group. Papers cover a broad range of topics, including information retrieval evaluation, natural language processing and information retrieval, and knowledge discovery.

Link: http://medir.dcu.ie



LREC workshop and booth

Khresmoi had a booth at the Language Resources and Evaluation Conference (LREC) conference in Reykjavik, Iceland, from the 28th to the 30th of May, where the three prototypes were demonstrated to visitors. LREC is the major event on Language Resources and Evaluation for Language Technologies. The LREC conference covers Language Resources and their applications, evaluation methodologies and tools, industrial uses and needs, and requirements coming from the e-society, both with respect to policy issues and to technological and organisational ones. Khresmoi also sponsored the Fourth Workshop on Building and Evaluating Resources for Health and Biomedical Text Processing (BioTxtM 2014), held after the LREC. The workshop featured an invited talk by Georgios Paliouras, coordinator of the BioASQ project, which organises challenges on large-scale biomedical semantic indexing and question answering. After presentations of papers and posters accepted for the workshop, a discussion session was held on identifying missing resources for health and biomedical text processing, and proposing ways in which these resources could be created.

The Khresmoi torso at the LREC Conference.

CeBIT 2014

In 2014 Khresmoi participated with a booth at CeBIT for the second time. For the first time the booth was exclusively for Khresmoi. One goal of this participation was to present clearly the prototypes to a larger public and get feedback on the prototypes for the preparation of the final Khresmoi prototypes tested this spring and summer. A second objective was to get commercial contacts and get linked to partners for the Khresmoi technology. Particularly on the first two days many people visited the booth, tested the technology and gave feedback. Many discussions with companies also led to technology exchange and several propositions to distribute the Khresmoi technology if products

become available. The Khresmoi torso also helped to clearly brand Khresmoi as a medical project and this attracted interest of many passing persons. CeBIT is the biggest computer fair in the world with a large and extremely varied participation from the entire world but in an important part from Germany.

The Khresmoi booth
ready to go for a day of the CeBIT.

eHealth day

On June 6, 2014 the second eHealth day took place in Sierre organised by the foundation theArk and the University of Applied Sciences and Arts Western Switzerland. Khresmoi was one of the focus parts of the event with a plenary presentation and with all prototypes being shown in the main hall where coffee breaks and the lunch took place. Many discussions happened around the Khresmoi posters and the Khresmoi demonstrations. In total over 160 registered participants were at the event, in large part from the French part of Switzerland but also in part from the German speaking part and France. As in the past edition, the speakers came from a varied background of physicians and technical persons and from medical institutions and small companies around the world including the US, Brazil and Sweden.

Dimitrios Markonis demonstrates the Khresmoi Radiology prototype.

Presentation of Khresmoi to Commissioner Kroes

The European Data Forum 2014 (EDF2014) took place from March 19-20, 2014 in Athens, Greece. EDF is the annual meeting-point for data practitioners from industry, research, the public-sector and community initiatives, to discuss the opportunities and challenges of the emerging Data Economy in Europe. The third edition took place in 2014. The Khresmoi project had a booth at the European Data Forum, where the three prototypes were demonstrated. We were also honoured to be able to present the Khresmoi results to Commissioner Neelie Kroes, Vice President of the European Commission responsible for the Digital Agenda for Europe.

Allan Hanbury presents Khresmoi for Everyone to Commissioner Kroes.

Read more on the Khresmoi webpage: http://khresmoi.eu

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

Join the Khresmoi group on LinkedIn: http://www.linkedin.com/groups/KHRESMOI-project-3745929

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

SEVENTH FRAMEWORK PROGRAMME