How to create useful knowledge from pure data

April 15 2013, by Philipp Grätzel Von Grätz

Imagine a hospital where patient data from numerous sources is made accessible to ward physicians with the help of hyperlinks and intelligent indexing. Imagine a healthcare system that hands its patients – not an envelope or a CD-ROM – but an integrated dataset that allows them to truly understand their illness, and even use the Internet to obtain additional information. Imagine a radiologist who uses semantic technologies to navigate smoothly through the myriad imaging data. Welcome to the future of semantic technologies in health information retrieval.

Professor Alexander Cavallaro's vision of the educated lymphoma patient of the future is very different from today's patient, who carries the computed tomography (CT) images of his lungs and abdomen home on a CD or DVD after a routine radiological examination. In Cavallaro's vision of the future, the lymphoma patient would open his radiological report on a tablet-PC to find a document with the relevant hyperlinks – much like a webpage. The patient would learn, for example, that his spleen is enlarged. After clicking on "spleen," the corresponding radiological image would appear. It would show exactly where the spleen is located and what it looks like. The patient would also learn that, although his spleen is still larger than normal, it is in fact considerably smaller than it was at his last radiological examination: a sign that the chosen cancer therapy is working.

Another click would open a window listing hyperlinks to additional, context-specific patient information; for example, to the lymphoma pages of the Internet encyclopedia Wikipedia, to a patient self-help website or a drug database with patient information on the side effects and the importance of drug adherence.

A benefit not only for patients

Cavallaro is Senior Radiologist at the Imaging Science Institute (ISI) at Erlangen University Hospital, Germany. As such, he not only has a vision for patients, but also for the physicians responsible for their care. These doctors do not necessarily need hyperlinks to patient websites, but they do need links to previous images to make comparisons and assess the effectiveness of the treatment. They also need to integrate imaging and laboratory data, as well as information on clinical signs and symptoms, in order to make a diagnosis, or modify a treatment plan.

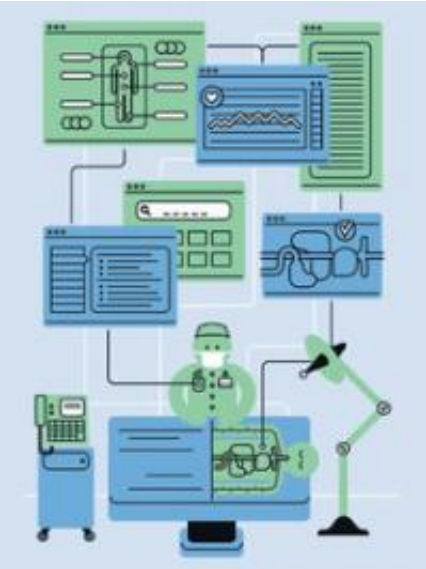
Equipping a medical report with this kind of context-specific and target group-specific information is not easy. It can only be done with the help

of semantic technologies that have been developed in artificial intelligence research. Semantic technologies structure information from different data sources with the help of an ontology: a structured dictionary. Thanks to semantic technologies, information from the Internet or another data collection can be filtered by content, so that a document search reveals only the most relevant and appropriate documents in a given context.

"Standard Internet searches provide millions of documents, most of which are useless," Cavallaro explains. "Semantic searches really stick to the point. This is what makes them so suitable for a highly specialized field like medicine."

Cavallaro is Senior Radiologist at the Imaging Science Institute (ISI) at Erlangen University Hospital, Germany. As such, he not only has a vision for patients, but also for the physicians responsible for their care. These doctors do not necessarily need hyperlinks to patient websites, but they do need links to previous images to make comparisons and assess the effectiveness of the treatment. They also need to integrate imaging and laboratory data, as well as information on clinical signs and symptoms, in order to make a diagnosis, or modify a treatment plan.

Equipping a medical report with this kind of context-specific and target group-specific information is not easy. It can only be done with the help of semantic technologies that have been developed in artificial intelligence research. Semantic technologies structure information from different data sources with the help of an ontology: a structured dictionary. Thanks to semantic technologies, information from the Internet or another data collection can be filtered by content, so that a document search reveals only the most relevant and appropriate documents in a given context.



"Standard Internet searches provide millions of documents, most of which are useless," Cavallaro explains. "Semantic searches really stick to the point. This is what makes them so suitable for a highly specialized field like medicine."

Research: A team effort

How semantic technologies can be applied to medicine was illustrated by the THESEUS MEDICO research project, which brought together radiologists from the University of Erlangen, experts from the German Research Center on Artificial Intelligence (DFKI), as well as researchers from Siemens, the Fraunhofer Society, and TUM (Technische Universität München). "MEDICO" was one of several cases put forward for the use of the THESEUS research program, which was initiated by the German Federal Ministry of Economics and Technology in 2007, in order to support technologies for an "Internet of services."

"When we discussed how medicine might benefit from semantic technologies, we soon realized that we really needed software solutions that could search and – to a certain degree – interpret images," Cavallaro recalls. Since images, and in particular radiological images, are an indispensable part of modern medicine, it simply did not make sense to talk about applying semantic technologies to medicine without considering images. The consequence was that, although THESEUS MEDICO was not fundamentally about computer-aided detection, it involved a considerable amount of computer-aided research. This led, among other results, to new algorithms for image analysis, which the Erlangen scientists reported in prestigious scientific journals.

An anatomical algorithm

The case chosen for the THESEUS MEDICO project was that of a lymphoma patient. "We chose lymphoma because we wanted to analyze the whole body, and we needed a disease with a certain radiological dynamic over time; so we could analyze, for example, treatment effects," Cavallaro explains. One of the first steps was to develop an algorithm that could identify various abdominal and thoracic organs on CT scans of the chest and abdomen automatically, within a reasonable time. It was a considerable success: "Our prototype solution can identify and visualize most relevant organs correctly within two minutes. This includes the esophagus and the pancreas: two organs that no other IT system seems ever to have been able to identify in such a short period of time," says Cavallaro.

Bone metastases, pathological lymph nodes, and alterations of blood vessels can also be detected automatically. This image information allows the automatic creation of hyperlinks between words in text documents, like a radiological report, that had been semantically analyzed by the system beforehand, and corresponding areas in the imaging dataset – provided that proper ontologies are used: in this case,

the radiological ontology RadLex and the anatomical ontology Foundational Model of Anatomy (FMA).

The ability to analyze imaging datasets automatically and to provide detailed links between radiological reports and radiological images is certainly impressive. But does it help in day-to-day medical scenarios? Cavallaro is convinced that it will do: "There are many, many ways in which the results of THESEUS MEDICO could become part of actual products in the future." For example, applying semantic technologies to radiological images could help physicians on the wards to better understand radiological images and reports.

This might also improve workflows, says Cavallaro: "Many radiological images are not self-explanatory. And the result is that we need many phone calls to explain images to our colleagues." By linking images with the pathological findings in reports, and by providing meaningful hyperlinks, the number of these explanatory phone calls could be reduced considerably.

Helping the physicians

Medical education, too, can benefit from a more semantic approach. Semantic technologies can be used to map anatomical drawings with radiological images of real patients, and at the same time, offer context-specific links to textbooks or scientific publications. "We have done this with three cases so far. They will be part of the anatomy course at Erlangen Medical School in the future," Cavallaro adds.

One problem remains, however: Even if the necessary meta-information is mainly accomplished by the system, the radiologist still has to validate or complement. A fact of which Cavallaro is very aware: "It is true that, in the beginning, it is more work for the radiologist, even though he is assisted by computer-aided capabilities. However, image analysis

methods are constantly getting better and one has to also look at the whole package. There are so many benefits of a more semantic approach to radiology. I am convinced that, in total, it will even out."

One of these benefits became apparent in the very early phase of THESEUS MEDICO. It had to do with the automatic analyses of organs and their locations. Once, for example, a pathological lymph node is described semantically, the system is able to "jump" to exactly that location in a follow-up examination. The radiologist no longer has to go through all the CT slices. With the help of semantic technologies, he is directed to the point of interest immediately. "This helps during the creation of a report. But it also makes radiological demonstrations far easier," Cavallaro explains. Time-consuming searches for follow-up images, while the rest of the department waits in the demonstration room, are eliminated. Each pathology that the radiologist wants to show and discuss with his colleagues is available instantly.

A step into the future

THESEUS MEDICO was a research project. It is not an off-the-shelf product at the moment. "But it certainly paves the way for the further development of the many IT solutions we use today – especially in the fields of radiologic reporting and data integration," says Cavallaro. The radiologist is convinced that semantic technologies are the way forward – for moving from the age of complex data collection toward the age of knowledge and interpretation of data. Using semantic technologies, it might be possible to conduct image-based searches of a patient database in order to identify patients with similar conditions. This could be of great help to a doctor who is confronted with a difficult or unusual case. "That is just one example. THESEUS MEDCIO has opened many windows," says Cavallaro. "At the Imaging Science institute, we are definitely planning to go on with this kind of research."

Provided by Siemens