

A Thin Client Interface to a High Performance Multi-modal Image Analytics System

James W. Cooper
IBM T J Watson Research Center
jwcnmr@us.ibm.com

Shahram Ebadollahi
IBM T J Watson Research Center
ebad@us.ibm.com

Ellen Eide
IBM T J Watson Research Center
eeide@us.ibm.com

Abstract

We describe a platform for performing text and radiology analytics (TARA). We integrate commercially available hardware and middleware components to construct an environment which is well-suited for performing computationally intensive analyses of medical data.

The system, termed a “medical analytical platform,” adopts a client-server approach for the display and processing of radiology images, utilizing a standard web browser client connected to a web server. All of the computation is carried out on the web server and computationally intensive tasks are carried out on the attached IBM Cell processor. The system is used to store, fetch, align and display radiology images and the analysis of the associated text reports, and allows for real time processing of the images.

We describe novel algorithms for displaying slices of 3D images within a web browser, and for displaying and searching down the RadLex tree and an annotation system for discovering drug dosages in the accompanying clinical reports.

1. Introduction

The advent of computer networking coupled with facilities for massive storage and computation, enables medical data to be handled in a revolutionary way. Rather than recording images on film and clinical reports on paper, doctors now have the option of storing patient data electronically. This option enables a myriad of further possibilities, including instantaneous access to patient records and the merging of image and text reports in a single dossier for each patient. The medical field has taken a step in that direction with the introduction of picture and archiving system (PACS) for managing medical image data.

However, PACS systems have not dealt in detail with the computational requirements for image processing, nor for the additional information that can be provided by text analysis of the accompanying patient reports. Our system, TARA (Text and Radiology Application) extends traditional PACS

systems by enabling two key features, real-time image analytics and the integration of imaging information with other relevant data sources such as electronic health records, in one convenient system for use by physicians. In this paper we describe the software architecture of the analytics platform, as well as one analytic function, fast image registration using an affine transform, which showcases the power of the platform. We have previously reported in detail [1] on the hardware aspects of this system, and report here on the software architecture of this unique thin-client system for combined text and image analytics.

2. Background

Conventional PACS systems allow you to fetch and examine single 3D images, and with some manipulation perhaps two 3D images at once. However, such systems are cumbersome and do not provide any convenient way for 3 or more historical images to be viewed at once.

In the TARA system we describe here we attempt to construct a system that is easier to use and more flexible than available PACS systems and as well allows convenient use of the analytics platform.

HCI is formally defined as “a discipline concerned with design, evaluation and implementation of interactive computing systems for human use [2].” In this paper we describe how the system was designed and implemented and touch on how its usability can be evaluated.

In constructing this system we went through the “gearing up phase,” understanding the user’s requirements in displaying multiple 3D images and have constructed an Early Prototype system as described by Gould [3], in order to gain user feedback.

3. System Architecture

This Medical Analytics Platform (MAP) consists of a web server, a file server, a computational server and a web browser based client. The web server can be as simple as the Apache HTTPD server [4] and Apache

Tomcat [5] or as robust as IBM Websphere. In both cases both HTML and JavaServer pages (JSPs) are supported. The web server either hosts or is connected to a database such as IBM DB2 or Apache Derby [6], where tables of patient information, text analytic data and the location of actual images are stored.

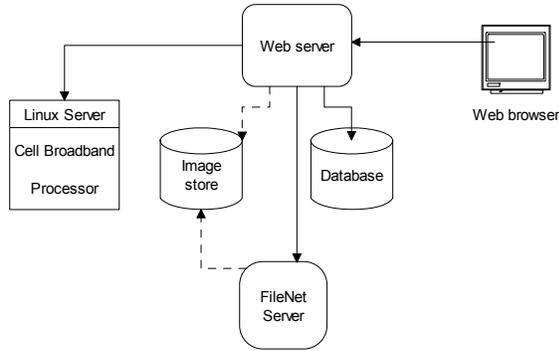


Figure 1 - The Medical Analytics Platform Architecture

While this basic system diagram implies a stand-alone system, the indicated connection to a database can be a connection to a standard hospital PACS, thus making TARA completely general and a useful extension to the standard approach for managing medical images.

In addition, we introduced a connection to the FileNet Content Manager [7] system to illustrate how large quantities of images and their annotations can be managed from within the same system.

4. Using TARA

To use the TARA system, you log in and select a patient by name. When the server delivers the list of patient names to the web page shown in Figure 2 the patients' ages and sexes are cached in the client web page so that they change immediately in the display as the patient names are selected.

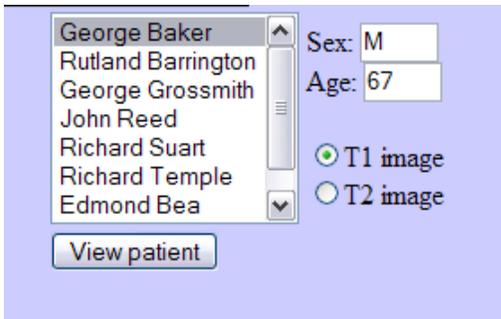


Figure 2 - Selecting a patient

TARA then displays two patient images, usually pre- and post-operative images and allows you to scan through them on your web browser in real time.

This display is unique to the TARA system, and allows you to grab the scrollbar with your mouse and scroll through image slices in real time (Figure 3).

This is accomplished by server-side software that converts the DICOM or ANALYZE-format images into a series of JPEG images that the embedding JavaScript code can rapidly reload. The image slices are cached in the browser when the page loads so that scrolling among is very smooth and rapid. Once an image has been expanded, this is recorded in the database so future users will not need to wait for the additional second or so this expansion requires. The date of the expansion and last access is also recorded in the database so that these expanded images can be purged if needed.

The Java server-side code for reading these image formats was adapted from code available in the open-source ImageJ package available from the NIH. [8]

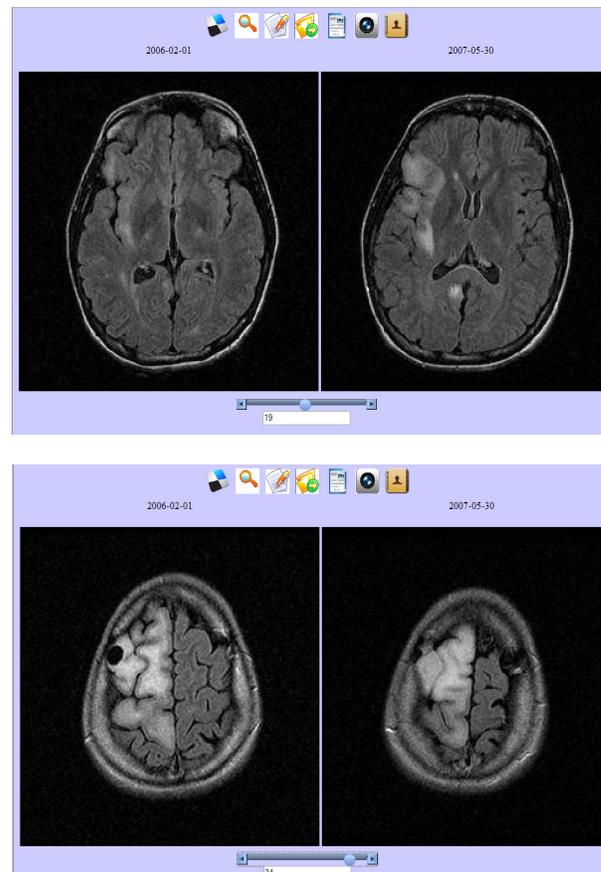


Figure 3 - Real time display of image slices in TARA. Upper: slice 17; Lower: slice 34 in both images.

The scrollbar is a JavaScript control from the Dojo library [9] as are the small floating icons above the images. The software is constructed so that a single scrollbar scrolls through both images simultaneously. This is an example of the *multi-window scrolling* described by Shneiderman [10].

5. Software Architecture

TARA is written entirely in Java (server) and JavaScript (client) and is thus completely portable between browsers and operating systems and is not dependent on any version of Java or any other language system being installed on each client workstation.

Most actions you take on the client trigger calls to the host servlet and generate new client pages through JavaServer pages. On the server, a single servlet entry point detects the setting of a client page variable and uses it in a Factory pattern [11] to create an instance of the server side Java class to handle that particular command. Many of the commands then return the URL of a new client page which in turn obtains specific parameter settings from the server data stream and uses them to populate the information on the new client page.

Some pages utilize the AJAX [12] (Advanced JavaScript and XML) protocol to obtain data that updates part of a page without the need for the whole page to be reloaded.

6. Multi-modal Image Analytics

Diagnostic images are acquired using different imaging modalities, such as X-ray, CT-scan, MRI, etc., where each modality is suited for illustrating and capturing a different aspect of the patient’s anatomy. Radiologists use a variety of tools to perform information extraction, or reconstruction of multi-modal images to read them better and to make more informed decisions. Most of the image analytics routines are compute-intensive and are not suitable for use during image reading and diagnosis. For example, registering a brain MR image to a baseline for the purpose of assessing the progression or regression of a tumor, is a very time consuming process. The radiologist has to make a request for this computation to be carried out and wait for the results so he can assess the state of the patient’s health.

In the MAP, we provide these analytics on a high-performance computing platform, so that they can be applied to images at near real-time speed. The platform we employ for the high-performance image analytics is the Cell Broadband Engine processor.

7. The Cell Broadband Engine

The Cell Broadband Engine (CBE), is an asymmetric multi-core processor with high peak performance: it combines eight synergistic processing elements (SPEs) and a Power Processing Element (PPE), which is a general-purpose IBM PowerPC processor. Two of these PPEs can be connected via a high speed bus as shown in Figure 4, so 16 SPEs can run in parallel. Each SPE, furthermore, has a SIMD (Single Instruction, Multiple Data) unit, which can perform a floating or integer operation on four data elements at every clock cycle.

To accelerate an image analytic routine on this processor, you must optimize the program to exploit the parallelism at both task and SIMD-instruction levels and use the memory bandwidth efficiently. First, at a task level, you must partition the program into multiple tasks that fit in the local store on each SPE. Unlike conventional microprocessors, each SPE does not have hardware cache memory to manage the small on-chip local store. Thus, one can view this architecture as a distributed memory multiprocessor with a very small local memory attached to a large shared memory.

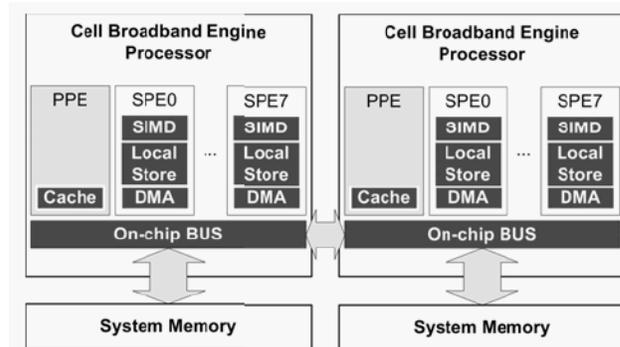


Figure 4 - Diagram of the IBM Cell Broadband Engine (CBE)

8. Real-time Image Registration

The TARA system provides near real-time image registration between two images, using the high performance computational capabilities of the Cell Broadband Engine (CBE). While experienced radiologists frequently make comparisons between images taken at two different time points for a given patient to assess changes, it has been reported that it is more accurate to look at the registered (aligned) images [14].

The algorithm takes two images from a given patient, one of which is identified as “fixed” and the

other as “moving,” and creates a new image corresponding to the 3D affine transformation of the moving image to best align it to the fixed image, and has been described previously. [1]

The specific registration algorithm we have implemented on CBE is based on Matte’s [15] mutual-information-based multi-resolution linear registration algorithm, as implemented in National Library of Medicine’s Insight Imaging Toolkit (ITK) [16] for the base implementation.

The CBE-based implementation of the registration algorithm was shown to register a pair of 256 x 256 x 34 3D MR images in one second, with better time improvements achieved compared to the implementation of the same algorithm on a regular high-end PC for larger volumes [17].

This is illustrated in Figure 5.

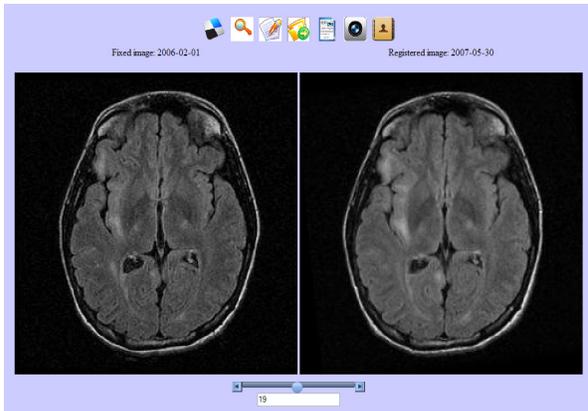


Figure 5 - Registration of the same images shown in Figure 3.

In the TARA system the Linux front end to the Cell processor has shared access to the same image storage space as the Web server does, making it easy to send a message to the Cell processor to carry out the real-time 3D registration. The resulting aligned image is written back to this shared disk space and expanded into individual JPEG images for each slice as described above.

9. Image History

It is not unusual for there to be a number of images available for a patient, accumulated over the course of his treatment, and TARA also provides a way to display and scroll through a number of images at the same time as shown in Figure 6.

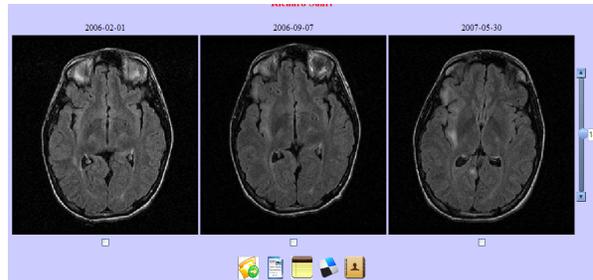


Figure 6 - A display of several images from a patient's history.

From this display you can also select a number of text analytic functions that summarize the findings of each clinical report. The buttons along the top of Figure 5 and the bottom of Figure 6 are created using the Dojo FisheyeList control, which means that they expand and show their captions as the mouse moves over them. This cross-platform JavaScript effect is much like the icons commonly found on the Mac-OSX platform.

10. Text Analytics

For each image there is an accompanying medical report that can be analyzed for a number of significant points. These example reports, which are a combination of clinical notes and radiological reports were written for us by Dr. Bradley Erickson of the Mayo Clinic, following the style he and his colleagues typically use in their practice. They contain the diagnosis, patient history, current medications, vital signs and physical examination results, a report and plan and a description of the MRI image.

Our group has developed a series of text recognizers, termed “annotators,” which recognize diagnoses, sites, drugs, Radlex [18] terms and various numeric concepts. These annotators operate in the Apache UIMA [19][20] framework where annotations can easily be combined to discover more complex concepts. These biomedical annotators have been collected into a package termed *MedTAS* (Medical text Analysis System). Central to this system is the ConceptMapper annotator that can recognize multi-word terms from various dictionary files even if the various words in the phrase are in a different order or contain intervening modifier terms. It operates with or without case folding, and can recognize a base term or any of its synonyms and relate them all to the canonical name for each concept. When sufficient diagnostic information is present in the report, the MedTAS annotators can fill in the slots in a *structured cancer model* [21] where the relationships between much of this information can be represented.

The documents, their annotations, annotation types and offsets are stored in the server database and can be

used to look up the data for any given document or to highlight particular terms as we illustrate in section 12 below.

11. A drug dosage annotator

As part of this project, we developed a text analytics technique for recognizing drugs and dosages so that they could be summarized to the clinician using the TARA interface. We developed a drug annotator based on the ConceptMapper annotator described above. It recognizes drug names based on a dictionary of base terms and synonyms, so that “aspirin,” for example is tagged so that the canonical (generic) name is “acetylsalicylic acid.”

We recognize that dictionary-based drug annotation may be insufficient in the general case, where drugs have differently spelled names and synonyms in different countries, but within a single domain, such as brain cancer, for example, it is quite effective.

To expand this system to allow greater generality, we would require a combination of a multi-lingual drug dictionary and some annotated documents where we can apply machine learning to recognize untabulated drug names from the context.

In addition we developed annotators to find the period (hour, day, week, p.r.n, [as needed] b.i.d [twice a day], t.i.d [three times a day], etc.) and the quantity (mg, ml, drop, tablet) and numerical values. We then developed a drug-dosage annotator that looks for all four of these concepts within a specific window: usually a sentence or phrase. If all four are found, or if one of the abbreviations which designate both period and frequency, a drug dosage annotation is created. A typical annotation is shown in Figure 7.

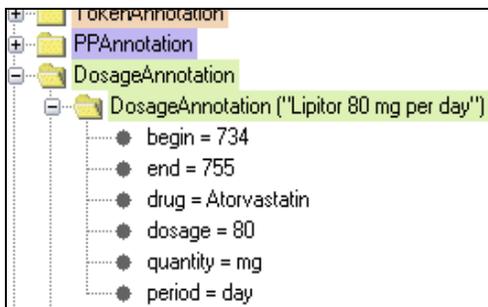


Figure 7 - A dosage annotation for Lipitor 80 mg per day

In this figure, the phrase “Lipitor 80 mg per day” is annotated as a drug dosage, and, as shown, the dosage, quantity and period are correctly recognized. Note that the generic drug name “atorvastatin” is entered as the canonical form for the drug name.

10.1 Preliminary Evaluation of the Dosage Annotator

We computed some preliminary accuracy numbers for the dosage annotator on a small series of reports. This is not an exhaustive study but merely an estimate of its utility.

The annotator was developed for a specific document format provided to us by scientists at the Mayo Clinic, where each drug/dosage phrase is on a separate line in the report. In reports in that format, the precision and recall of the drug dosage annotator is 100%.

The annotator does not do very well when several drugs and dosages are comma-separated in the same sentence, such as:

The patient is now taking Diovan 160 mg a day, Neurontin 300 mg t.i.d., Norvasc 2.5 mg a day, Bactrim three times a week, and Nexium 40 mg a day.

However, it is a simple matter to write a pre-annotator to separate these phrases. In that case, in a series of 10 oncology reports provided to us by another partner, we observed 100% precision and 86% recall, with all of the errors related to a the number annotator’s misrecognition of the number “2.5.”

12. Displaying text analysis in TARA

In the history display shown in Figure 6, you can select any of the images and then select one of the buttons to display the diagnosis, the drugs prescribed or the actual document. For example, if you select the Drugs button, the text in Figure 8 is added to the bottom of the display using Ajax:

Drug	Dosage	Units	Period
Lipitor	80	mg	day
Aspirin	81	mg	day
Vicodin tablet	-	-	day

Figure 8 - The drug-dosage analysis for this clinical note

In a similar fashion, you can display the entire report with site, diagnosis and drug dosage highlighted.

History of Present Illness Mr. BrainTumor2 is a 44-year-old gentleman who developed a right thalamic glioma in early 2002. He was treated with radiation and chemotherapy. He had a decrease in function after this which was felt to be secondary to radiation necrosis. He was treated supportively with steroids and improved. He remains off of steroids at the present time.

Since he has last been here, he has started working. He is working as a cashier at McDonald's 2-1/2 hours per day three days per week.

Clinical Diagnosis and History:

Current Medications
 Lipitor 80 mg per day. Atorvastatin
 Aspirin 81 mg per day.
 Vicodin approximately one per day.

Figure 9 - A portion of the clinical note for this patient visit.

In Figure 9, the diagnosis and site are highlighted in the top of the window, and the drugs and dosages near the bottom. Note, that if you hover over the drug name, the generic equivalent canonical name is displayed as shown.

13. Annotating and Displaying Radlex Terms

RadLex is a complete ontology [18] of terms commonly used in radiology, and is commonly represented as a tree of general terms and more specific terms as you move down the tree. The main top-level categories include treatment, imaging procedure, substance, imaging observation characteristic, modifier, imaging observation, anatomic entity, teaching attribute, relationship and image quality. Most of the observations we concern ourselves with in text analytics are found in the imaging observation and anatomic entity categories.

The RadLex ontology can be downloaded as an XML file that you can view and edit using the Protégé browser [23]. For the TARA server application, we converted this file into a database where keys in each row maintained the tree structure as well as the synonyms. In addition, we generated a dictionary for the ConceptMapper annotator so we could annotate all the Radlex terms in the Imaging Observation and Anatomical Entity categories.

This makes it possible to store the Radlex annotations we discover in the database of annotations, and using the tree structure in the RadLex term database find terms that are above or below a specific term in the RadLex tree. You can then use this information to make text searches more general.

In TARA, we annotated and indexed about 273 de-identified radiology reports that were originally

provided as part of the ImageCLEF competition [24]. This makes it possible to search for specific terms or their parents or children within any of the documents in the collection.

The TARA RadLex search screen is based on the Dojo ComboBox and fills the list of terms in the dropdown based on the characters you type. The ComboBox works using Ajax calls to the server to return a list of the terms that start with the characters you have just typed.

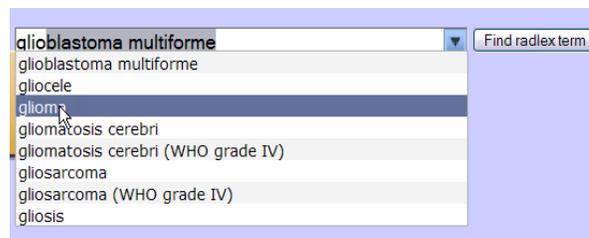


Figure 10 - Typing "glio" into the Radlex lookup ComboBox.

In Figure 10, the user has typed "glio" into the ComboBox, and all the RadLex terms beginning with that term are shown in the dropdown. You can then select any of these terms with the mouse cursor and search for them. In the figure, the umbrella term "glioma" is selected. Then, when you click on the "Find radlex term" button, the search term is sent to the server, and the related terms are returned in a Dojo Tree control.

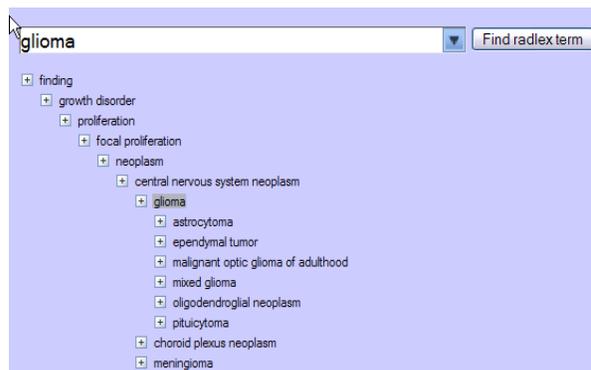


Figure 11 - The RadLex tree expanded below "glioma."

We designed the Dojo Tree display in a unique way to open only to the node that is to be displayed, *without* opening the outer nodes above it. This is a more effective use of screen space. However, any node that you click on, above or below the target node will open in the usual way.

Under the covers, this approach is accomplished using a pair of Ajax calls: one to fetch to nodes above the target node, without all their related nodes, and another to fetch the nodes that include the target term.

Each time you click on any node, a call is made back to the server to get all the expanded nodes and redraw them on the client.

12.1 Searching Documents for RadLex Terms

It is a simple matter to index all the medical documents in a collection using a search engine such as Lucene [25] and provide a way to search for strings within the documents. However, it is more useful to search *semantically* by selecting a RadLex term from the tree and searching both for the term *and any of its children*. We have implemented this search in the TARA system by providing a “Search Documents” button. If you select a RadLex term and click on this button, the server will return a list of all document which contain either that term or any of its children.

For example, if you search for “glioma,” a document is return which does not contain “glioma” at all, as shown in Figure 12.

Clinical Diagnosis and History:
History of **brain tumor** with recent syncopal episode
31-year old man with history of grade II **astrocytoma** of the **left frontal lobe**, status post resection four years previously and also external beam radiation therapy. The patient had done well until recently when he experienced a syncopal episode at work.

Figure 12 - Highlighted text from a document returned from the query “glioma.”

This document refers only to *astrocytoma*, but as you can see from Figure 11, this is a direct descendant of glioma. This approach allows the user to specify a general term and obtain documents containing more specific subsidiary terms where the general term is never mentioned. This greatly enhances the power of the search system.

14. Fetching Images from FileNet

For the image data associated with these ImageClef reports, we have integrated the FileNet content manager system as an image repository. FileNet provides the advantage of being able to handle and index very large quantities of data, and index any annotations you might want to store with either the images or the reports.

In TARA, we store only the images in FileNet. They are fetched when you click on any document in the list returned from the search, as shown in Figure 13.

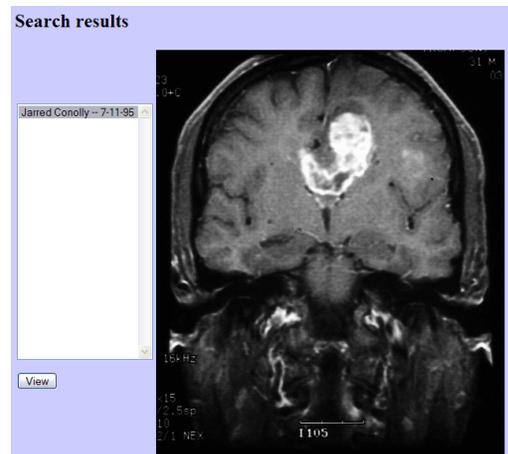


Figure 13 - An image returned from FileNet

15. Lessons Learned

The TARA system was constructed as a proof of concept system to show to scientists and gain more information about how they might use the Medical Analytics Platform in their clinical and research work. The system was first shown to radiologists and other scientists at the RSNA conference in November, 2007.

Comments at the conference included “you’ve just put me out of business,” from a radiologist who was used to carrying out the registration of images mentally while evaluating the images. This system was designed to illustrate new ideas and elicit feedback on its capabilities.

A number of users “played” with the demonstration system and for the most part had no trouble deducing how it was to be used. As noted recently by Greenberg and Buxton [26], there are times during the initial development of new ideas when an evaluation of the user interface is premature. We were also extremely gratified to learn that the system is sufficiently robust and light weight that we were able to demonstrate it in real time at the CARS Conference in Barcelona, Spain using a server running in Yorktown Heights, NY.

Most important, however, the TARA system led to a follow-on research project with another hospital which we are now carrying out. This project, part of which is based on the design of TARA, is undergoing a thorough usability evaluation now.

16. Summary and Conclusions

We report on the development and implementation of a thin web based client (TARA) for studying multi-modal text and image analytics of radiology images and their associated reports. Linear registration of images can be accomplished on the Medical Analytics Platform (MAP) using the Cell Broadband Engine to accelerate the computation. A novel algorithm is used to convert the 3D images to JPEG slices and for

displayed in an HTML web page. The images are cached in the browser so that you can scroll among them rapidly using a Dojo-based scrollbar.

Any text attribute which can be recognized by an annotator is stored in a database and can be highlighted when displaying the document, or shown using an Ajax call in a designated area of the web page.

A unique algorithm has been developed for loading the Dojo tree-control in segments so that the entire expanded tree need not be displayed down to the point where a term is found.

A novel algorithm is described for recognizing drug dosages, by finding a mention of a drug, a quantity, an amount and a time period.

17. Acknowledgements

We gratefully acknowledge the cooperation of Dr. Bradley Erickson of the Mayo Clinic in providing us with hypothetical sample clinic reports that we used in this study. We also acknowledge the help of Anni Code, Michael Tanenblatt and Igor Sominsky of IBM Research who are authors of a substantial part of the MedTAS annotator system.

18. References

- [1] Cooper, J. W., Ebadollahi, S., Eide, E. Neti, C. and Iyengar, G. "A Medical Imaging Informatics Appliance," CARS Conference, Barcelona, Spain, June, 2008.
- [2] Hewett, Baeker, Card, Carey, Gasen, Mantei, Perlman, Strong and Verplank. *ACM SIGCHI Curricula for Human-Computer Interaction*. (1996)
- [3] Gould, John D, "How to Design Usable Systems," excerpted in Baecker, R., Grudin, J., Buxton, W. and Greenberg, S. *Human Computer Interaction: Toward the Year 2000, 2nd edition*, Morgan-Kaufman, 1995.
- [4] <http://httpd.apache.org/>
- [5] <http://tomcat.apache.org/>
- [6] <http://db.apache.org/derby/>
- [7] www.filenetinfo.com
- [8] <http://rsb.info.nih.gov/ij/>
- [9] www.dojotoolkit.org
- [10] Schneiderman, Ben, *Designing the User Interface*, Addison-Wesley-Longman, 1998. (459).
- [11] Cooper, J. W., *Java Design Patterns, A Tutorial*, Addison-Wesley, New York, 2000.
- [12] Aleson, R. and Schutta, N., *Foundations of Ajax*, Apress-Springler-Verlag, New York, 2006.
- [13] A US Patent application 12/098707 has been filed on the details of this process.
- [14] Erickson, B.J. et al. "Effect of Automated Image Registration on Radiologist Interpretation." *Journal of Digital Imaging*. June 2007, **20** (2) 105-113.
- [15] Mattes, D. Haynor, D.R., Vessele, H., Lewellen, T.K. and W. Eubank, "PET-CT Image Registration in the Chest Using Free-form Deformations," *IEEE Transactions on Medical Imaging*, **22** (1) (Jan 2003) 120-128.
- [16] Insight Segmentation and Registration Toolkit (ITK), <http://www.itk.org/index.htm>
- [17] Ohara, Moriyoshi; Yeo, Hangu; Savino, Frank; Iyengar, Giridharan; Gong, Leiguang; Inoue, Hiroshi; Komatsu, Hideaki; Sheinin, Vadim; Daijavad, Shahrokh, 2007 IEEE International Conference on Multimedia and Expo, July 2007, 272 – 275
- [18] www.Radlex.org
- [19] Ferrucci D., and Lally, A. "The UIMA System Architecture," *IBM Systems Journal*, March, 2004.
- [20] Mack, Robert L. et. al. "BioTeks: Text Analytics for the Life Sciences." *IBM Systems Journal*, March, 2004
- [21] Coden A, Savova, G, Sominsky, I., Tanenblatt, M., Masanz, J, Schuler, K., Cooper, J., Guan W., de Groen, P. "A Automatically extracting cancer disease characteristics from pathology reports into a novel disease knowledge model," submitted to *Journal of Biomedical Informatics*
- [22] www.incubator.apache.org/uima
- [23] http://protegewiki.stanford.edu/index.php/Protege_Web_Browser
- [24] www.imageclef.org
- [25] www.lucene.apache.org
- [26] Greenberg, S. and Buxton, B., "Usability Evaluation Considered Harmful (Some of the Time)," *Proceedings of CHI 2008*, Florence, Italy.