Encore: A Robust and Easy to Use Radiology Information System and Dictation Query Tool

Ryan Ward, MD, Cleveland Clinic; David Piraino, MD, FSIIIM; James Wetzel; Michael Ciancibello

Background

The radiology information system (RIS) serves as a key component to the radiology department workflow, accumulating large amounts of patient scheduling, procedure billing, radiation dose, and exam interpretation information. Yet much of this information is textual, semi-structured, and embedded in systems inaccessible to the practicing radiologists. Few tools exist to easily query this information, and those that are available are costly to implement and limited in functionality. In this report, the authors describe the use of predominantly open source tools to create a “Google-like” search portal for the radiologist to query.

The potential uses of such a system are many:

Education

- Commercial comparative image databases are available but are expensive, often offer only a single slice “textbook” representation of the pathology in question, and have a limited number of case examples.
- Using this system, the radiology resident can find all comparable cases within the database, view them with multiplanar reformats, and see the pathology on multiple sequences. The process of finding similar cases can help solidify his/her knowledge of the disease process.
- This system also obviates the archaic practice of logging accession numbers in excel spreadsheet “case files” for future reference.

Research

- With a simple search, one can easily find every case dictated that mentions the pathology in question. As such, the researcher can find extraordinarily rare cases, often dictated years ago, that may have otherwise been impossible to find.

Productivity

- Radiology residency accrediting bodies stipulate minimum dictation quotas for various modalities that radiology residents must meet to become board certified. Manual tracking of this information is time consuming, can be inaccurate, and is not performed in real time.
- This tool provides easy filtering by dictating physician, by modality, and by day to see how many studies an individual dictates on a real-time basis.

Case Presentation

Using Lucene Solr, an open source enterprise search platform, we index approximately 48 million unique exams spanning over 20 years. Indexed data include all available radiology information system (RIS) fields as well as the results and impression sections from dictated reports. A summary of the data flow is provided below.

Figure 1
The database is queried via password protected web frontend which is accessed using enterprise credentials. Once logged in, the user enters his desired search term into the search box and results are immediately returned with the highest relevancy at the top of the page. Relevancy is determined by the proximity of the search terms to each other, the number of times each search term is mentioned within a result, and other variables. Solr has built-in functionality to fine tune search relevancy based on the use case.

**Figure 2**

Once the results are returned, additional filters can be applied to further refine the search by inpatient/outpatient status, imaging modality, department, etc. A user can click on a result to review the dictation results/impression and full set of RIS data elements for that accession. Images are immediately accessible via a “Get Images” button that spawns a mobile web viewing application with full-functioning image viewing capabilities (scroll, zoom, window, etc).

**Outcome**

The authors have successfully developed a fast and easily-queried multipurpose database. Use requires no special training, as the results are returned in a “Google-like” fashion. The current prototype database is hosted on decommissioned HP Z400 workstations with Intel Xeon W3520 Processors and 12Gb RAM. For data redundancy, we have implemented Solr in SolrCloud mode with four replicated nodes. The database consumes 70GB of space per node. To date, database deployment and front end development has utilized an estimated 300 hours of developer time.

**Discussion**

As we transition from beta to release, there remain a number of challenges that need to be overcome. First, we currently only use a snapshot of the RIS and have not yet experimented with live HL7 data streams. As an intermediate step, we anticipate merging daily or weekly snapshots to keep the database up to date. Second, we must make the search algorithm smarter. For example, in searching for “pneumoperitoneum”, it would be undesirable to have results containing “no pneumoperitoneum” be returned. Major improvement in this area will come from advances in machine learning and ranking algorithms.

**Conclusion**

The authors have created a robust and easy to use educational and research tool using predominantly open source software. While this implementation is still in beta stage, we hope to further expand its utility by integrating the EMR HL7 stream to pull in the problem list, ICD-10 codes, and other relevant medical record variables. Other features will be added soon including adding a tagging system to easily find cases and further expanding the user interface to add different groupings/filters.

**References**


**Keywords**

radiology information system, RIS, big data, search