Newsbytes

Pathology lab uses homegrown tool to search AP report text

June 2022—Laboratories seeking a means to query their repository of archival anatomic pathology reports would do well to follow the advice of former tennis star Arthur Ashe: "Start where you are, use what you have, do what you can."

Adhering to the same guidance, pathologists at the University of Texas Southwestern Medical Center developed a tool to query the unstructured text in AP reports using the business intelligence software Tableau, for which the hospital already had a subscription. "From the get-go, the things we had on campus to build the tool ranked higher," says Ellen Araj, MD, a pathology informaticist at UT Southwestern.

Finding a solution to perform retrospective AP report queries became a priority in 2018, when the pathology lab transitioned from Sunquest's CoPath laboratory information system to Epic Beaker. Unlike CoPath, Beaker cannot perform retrospective full-text searches of AP reports. However, Dr. Araj and her colleagues were able to input data from Epic's Clarity clinical data warehouse directly into Tableau, which has the ability to conduct such text searches without using middleware.

The software tool can query UT Southwestern's repository of more than 333,000 AP reports. It functions much like the "control F" command in Microsoft Excel, Dr. Araj says, which allows users to search for characters, text, and phrases within a spreadsheet. The tool's ease of use and similarity to popular software programs are a boon for pathology department faculty, she adds. The platform is used by about one-third of the department, and these staff have employed it to perform a cumulative average of about 100 searches per month.

Dr. Arvisais-Anhalt

Simone Arvisais-Anhalt, MD, a former resident at UT Southwestern and coauthor (with Dr. Araj and others) of a *Journal of Pathology Informatics* article about the tool (doi.org/10.1016/j.jpi.2022.100014), says "this is one of the first times Tableau has been used effectively like a search engine." Many health care organizations use Tableau as a business intelligence tool, she adds, making it easy for other institutions to replicate the project. "If you connect with your enterprisewide team that is using Tableau, it's easy to extend the licensing and servers that are set up for Tableau to a project like this." UT Southwestern's quality and operational excellence group had set up a Tableau server for its own purposes, and it allowed the pathology department to allocate a small subsection for the search tool. "I say they built the ground we walked on," Dr. Araj jokes.

To program the tool, Dr. Araj used a custom SQL script to extract from the clinical data warehouse the unstructured text from AP reports and the structured data that can be included in end users' searches. She input both into Tableau's data model, which operates in much the same manner as a relational database, using the software's built-in extract, transform, and load, or ETL, functionality. She built the tool's search capabilities within Tableau using parameters for user input and nested calculated variables, as well as the software's regular expression functionality. The latter is a syntax built on pattern-matching strings of characters.

It took 80 hours, which included user feedback sessions, to develop the tool, plus an additional 40 hours to create data-governance documentation and a user tutorial, according to the *Journal of Pathology Informatics* article.

Maintenance typically takes one hour per month and primarily involves adding or removing users. The search dashboard itself requires little maintenance, as the tool automatically performs a monthly ETL data refresh.

Dr. Araj

Much more difficult than the programming component was ensuring the tool would comply with the medical center's data-governance standards. "It was something I underestimated," Dr. Araj says. The process, which involved getting buy-in from the hospital's security and HIPAA teams and chief medical information officer, took about four months. Only residents, fellows, and pathology department faculty can use the tool, and they must sign an agreement specifying how they will use it. Those who intend to use the tool for research must also provide an institutional review board number for a project that requires retrospective reviews of pathology reports and patient data.

Institutions interested in replicating the project should think about data governance early, Dr. Araj says. "Who's going to use it? Is it going to be only your department? Will it be researchers? How much data will you allow them to see? Will it be aggregated? Line level?" Tableau automatically authenticates users and tracks which data are being accessed and by whom, fulfilling HIPAA compliance, but a similar tool developed from scratch would have to address user authentication, she notes.

Dr. Araj and her colleagues performed a vendor comparison and cross-benefit analysis before settling on Tableau. Though other solutions they considered had more powerful natural language processing functionality, they determined that Tableau's simple text-search functionality was sufficient for the pathology department's needs.

By building on Tableau, the developers of the tool were able to create an interface that allows users to enter up to three search concepts logically combined with the Boolean operator AND. For each of the search concepts, users can input up to four synonyms logically combined with the Boolean operator OR. Users can search a full pathology note or ICD-9 or ICD-10 diagnoses for each concept. The matching cases generated are displayed in the results bar, along with the patient's medical record number, pathology case number, time the order was placed, full text of the pathology note, and a section called "lines found," which displays the exact line of text within the report that matched the search concept provided. The "lines found" section, which arose from the initial feedback sessions, allows users to determine if a result is relevant without reading the entire report.

Tableau was also deemed a good fit for the project because it can be accessed through a Web browser, thereby avoiding the need to install or maintain software on end users' computers. Furthermore, the user interface can be modified quickly and easily, and viewer user licenses are affordable. Speed, too, is a factor. Though it isn't fast—Dr. Araj tells users to expect waits of about 30 seconds—it's comparable to CoPath. "I didn't expect it [Tableau] to scale at all," she says. "That was one of the biggest surprises."

Dr. Araj and her colleagues are in the early planning stages for the second version of the tool. The first step, she says, is to interview users about what they would like to see in version two. Dr. Araj, herself, hopes to incorporate into it additional discrete data elements from Epic's clinical data warehouse. Though users can include some structured data in their searches, such as the authorizing provider and test name, they cannot include structured data from new data fields in Epic that contain such information as whether a case is malignant. Incorporating these data would allow users to quickly find all malignant cases associated with their search by checking a box and exclude benign cases from the results, she explains. Dr. Araj also wants to create a method to search synoptic reporting.

For those interested in developing a similar tool, Dr. Araj advises beginning the process by interviewing potential stakeholders. "Obtain feedback as early in the process as possible," she says, and apply that to product versioning. "Version one should be the most slimmed down," with a minimal amount of functionality, she notes. The goal, she adds, is to get a version in front of end users as quickly as possible to observe how they use it and how it could be improved.

After Dr. Araj developed a demo of the tool, based on how she thought it should work, she solicited feedback from Justin Bishop, MD, chief of anatomic pathology, and pathologist Jyoti Balani, MD (coauthors of the *Journal of Pathology Informatics* article). "We went through multiple iterations, rapid fire, for about two weeks," she says, determining the bare minimum functionality needed for the tool to work.

Dr. Bishop

On Dr. Bishop's wish list for future versions is "more granularity in what I can search. Most helpful would be the addition of a NOT feature," he told CAP TODAY. "For example, when I search salivary gland oncocytomas [a rare condition], I end up getting a lot of renal oncocytomas. In this scenario, I'd like to be able to search 'oncocytoma' AND NOT 'renal' OR 'kidney.'"

Based on lessons learned during her time at UT Southwestern, Dr. Arvisais-Anhalt, who is now a clinical informatics fellow at the University of California San Francisco, is laying the groundwork to develop a similar tool at UCSF. Among the strengths of UT Southwestern's tool is that it is EHR and LIS agnostic, she says. "It's not tied to your vendor—it's just a matter of connecting Tableau to your clinical data warehouse." Yet gaining access to the Epic Clarity data warehouse can be tricky. "Access varies depending on how governance and security practices have been set up," Dr. Arvisais-Anhalt notes, a consideration for those interested in replicating UT Southwestern's project. UCSF offers access to Clarity only for clinical informatics fellows and designated clinical informatics personnel. And those staff members must first become certified in the use of Clarity, which differs from the requirements at UT Southwestern.

Yet the Tableau search tool has advantages over an LIS with built-in text-search capability, Dr. Arvisais-Anhalt contends, because large queries can stall a lab system. In addition, with a built-in LIS tool, users may be limited in how far back they can search because of how the lab system's operational databases are organized. In contrast, she says, "when you hook directly into your clinical data warehouse—especially if the clinical data warehouse has archived a lot of data from previous years—your reach, in terms of how far back you can go, is much further."

Looking to the future, Dr. Arvisais-Anhalt says other hospital departments could benefit from using a similar search tool. It lends itself to any report-driven field that needs to query unstructured text, such as radiology or, on the clinical pathology side, microbiology and molecular pathology, she explains. "Today," she adds, "we're much more familiar with what it means to analyze structured data. Unstructured text is the next horizon."
—*Charna Albert*

Pramana launches platform to digitize pathology slides

Pramana has introduced a digital-pathology-as-a-service platform for digitizing standard-format glass slides and announced a commercial agreement with Mayo Clinic.

The company's automated solution digitizes archival pathology slides at scale to meet the needs of pathologists, clinicians, and researchers. As part of the service, Pramana will send technicians to customer sites to operate its

scanners. Clients are charged on a per-slide basis and have no equipment costs.

"It is often misconstrued that the challenges with going digital for tissue and body fluid slides end with the purchase of a scanner," said Prasanth Perugupalli, chief product officer for Pramana, in a company press statement. "We realized that the greater pain and costs lie in the human capital needed to operate the scanners, which includes making the correct parametric selections and qualifying each whole slide image for any errors after the scanning is completed."

Pramana's intelligent whole slide image-acquisition system places slides in position for scanning and can scan four slides simultaneously. It uses proprietary algorithms to assess the quality of slides during the digitization process and can rescan slides as needed. A single unit, consisting of one robot with four heads, can process more than 1,000 slides per day. A technician can oversee up to five units at once, generating more than 5,000 slides per day.

In its first commercial agreement for the service, Pramana signed a three-year deal with Mayo Clinic to digitize 5 million slides.

CDC establishes center to address public health threats

The Centers for Disease Control and Prevention recently launched the Center for Forecasting and Outbreak Analytics, with the goal of improving the ability to use data, models, and analytics in timely decision-making when responding to public health threats.

"CFA will produce models and forecasts to characterize the state of an outbreak and its course, inform public health decision-makers on potential consequences of deploying control measures, and support innovation to continuously improve the science of outbreak analytics and modeling," according to a CDC website posting.

The center has begun building an outbreak analytics team that will be made up of experts across several disciplines. The team will develop evidence to predict trends and guide decision-making during emergencies.

Proscia to add Visiopharm AI applications to Concentriq Dx

Proscia and Visiopharm have expanded their strategic partnership by striking a deal whereby Proscia will integrate Visiopharm's suite of CE-IVD artificial intelligence-based clinical applications, including breast IHC markers and solutions to detect breast and colon cancer metastases, with Proscia's CE-IVD Concentriq Dx platform for image-based data and workflow management.

Proscia's Concentriq Dx platform is available for in vitro diagnostic use in the United States during the COVID-19 pandemic under an FDA emergency use authorization. Visiopharm's clinical AI applications are available for research use only in the United States and are not covered under the EUA for Concentriq Dx.

Concentriq Dx will offer the full suite of Visiopharm AI applications to quantify the expression of biomarkers.

Proscia, 215-608-5411

PreciseMDX inks deal with Vanguard Laboratories

Salt Lake City-based Vanguard Laboratories has contracted with PreciseMDX for the PreciseMDX digital health platform. The platform will be integrated with Vanguard's laboratory information management system to enhance the molecular laboratory's PCR and diagnostic testing operations.

PreciseMDX, 866-462-4677

Dr. Aller practices clinical informatics in Southern California. He can be reached at raller@usc.edu. Dennis Winsten is founder of Dennis Winsten & Associates, Healthcare Systems Consultants. He can be reached at dwinsten.az@gmail.com.