

Homegrown search platform addresses the need for speed

May 2021—Laboratory information systems contain a wealth of diagnostic patient case data, but accessing it promptly can be difficult if not impossible—without a workaround, that is.

While LISs are searchable, they don't necessarily offer the type of instantaneous search functionality that would allow pathologists to quickly compare current cases to past cases, explains Scott Robertson, MD, PhD, a pathologist at Cleveland Clinic. "In my experience in the places I've worked, it has always been that way," he says. Therefore, Dr. Robertson built an inexpensive search platform for the Cleveland Clinic, which could be replicated by other institutions.



Dr. Robertson

The first key step in such a project, Dr. Robertson says, is to move a copy of relevant LIS data to a separate database that is not restricted by vendor-specific tools or protocols. "You need your data where you can access it in a programmatic way outside of the production database," he explains. "What I mean is, you have to be very careful on the database that is running your day-to-day business because there is a risk that the searches will stress the system. Once you have a copy of your data outside of that, you can do what you want with it and there's no risk to your clinical operations."

The search platform that Dr. Robertson constructed, called Pathtools, uses a standalone SQL server database. Cleveland Clinic allocated space for the platform on a SQL server hosted in its secure data center. Pathtools users at Cleveland Clinic can now search more than 4 million anatomic pathology cases dating back up to four decades and receive search results in seconds. (An automated process via Linux uploads new case information from the LIS into Pathtools every 24 hours.) In the past, Dr. Robertson says, a small team of analysts conducted all data searches, carefully managing the load on the production database to minimize negative effects on routine operations. "The analyst-centric workflow was slow," he adds, "and it took more than 24 hours to return results to the requester."

Dr. Robertson designed the Web-based graphical user interface for Pathtools and set up the database in his spare time over the course of two years. However, some medical institutions may have to hire a software developer on a project or part-time basis, he says. Because LIS data includes protected health information, the Web interface for a Pathtools-type platform would need to be built using a professional-grade Web framework and include extensive security features, Dr. Robertson explains. Therefore, he chose Django, a Web framework used by major corporations. Django is open source and based on the Python programming language, with which he is familiar.

For any project that affects how data is accessed, it is essential to collaborate with institutional oversight teams to ensure appropriate data protections are in place, Dr. Robertson says. He worked closely with Cleveland Clinic's institutional review board and the clinic's legal experts to ensure the platform met the medical center's data governance standards.

To safeguard data, Pathtools has a three-step registration process. Users must complete a registration form online and then click an email confirmation sent to their institutional email address. The Pathtools administrator, who in this case is Dr. Robertson, must then approve the registration form to grant access. Only select pathology

department staff, such as pathology residents and fellows, and clinic researchers who have been vetted are allowed to use the system. This last registration step allows Dr. Robertson to verify that only approved personnel gain access.

Security considerations also influenced how Pathtools' search functionality was structured, resulting in two search modes with different degrees of data access. "From a data governance perspective, you want to give people the least amount of information needed for their purpose," Dr. Robertson says. Therefore, the default, or preview, mode returns search results with case numbers and diagnostic data and no possibility of accessing protected health information. Preview mode searches also return only the most recent 100 cases, in reverse chronological order, and are intended for users who want to run quick searches during the workday. Cleveland Clinic, Dr. Robertson notes, is in the process of transitioning to a new LIS that takes more than five minutes to get results for searches that, on average, take less than a second using Pathtools.

Approximately 75 percent of searches in preview mode are conducted by pathology residents and fellows who want to compare past and present cases, Dr. Robertson says. The remaining 25 percent are primarily conducted by attending pathologists seeking case information for teaching, research, and presentation purposes.

Residents typically use Pathtools to look up factors considered in past cases when they preview current cases as part of an anatomic pathology training process that requires trainees to "preview" or write preliminary diagnoses that are reviewed and revised by attending pathologists. Before Pathtools, Dr. Robertson explains, there was no way to access this information, so writing preview diagnoses was more of a trial-and-error undertaking.

In a Pathtools study conducted by Dr. Robertson and published in the *Journal of Pathology Informatics* (doi.org/10.4103/jpi.jpi_43_20), 18 trainees were surveyed and all "strongly agreed" or "agreed" that Pathtools helped them write diagnoses that required less editing by attending pathologists.

For pathologists who want access to more detailed information, Dr. Robertson created the multifaceted research mode in Pathtools. The initial search results in research mode are similar to those in preview mode, except that in addition to returning 100 case results, this functionality provides a count of the total number of cases that meet the search criteria. "When you are starting a research project," Dr. Robertson says, "you want to be able to quickly get an idea of the scope of the project and its feasibility. For pathology projects, a lot of that centers around [determining] how many cases of diagnosis X we have at Cleveland Clinic and that exist in our archive."

Research mode also allows users to store their searches on a portal page and begin formal requests for more detailed data, including sensitive, protected health information. All formal data requests are routed to analysts who review the requests, perform the data downloads, and, for requests that are approved, send the data file to the user. These analysts also fulfill the requests for data from the LIS. However, unlike with LIS data requests that can take 24 hours to fulfill, Pathtools search requests are usually returned in an hour or two, Dr. Robertson says. Pathtools has eased analysts' workloads, he adds, because basic data searches can be run without their involvement and detailed searches for potentially sensitive information can be processed more quickly.

Despite its benefits, Dr. Robertson acknowledges that Pathtools isn't perfect. One problem is that it is very literal. Therefore, a database search for "Crohn's disease" would not capture "Crohns disease" if it were spelled without the apostrophe. Natural language technology could correct such an issue, he says, but it is not yet built into the system.

More importantly, although Pathtools contains approximately 4.3 million cases, it is not a complete copy of Cleveland Clinic's LIS, Dr. Robertson says. Some cases cannot be imported into the Pathtools SQL server because of details such as odd text characters. Between one and two percent of more current cases cannot be transferred to Pathtools, and the percentage of older cases that cannot be transferred is higher, he notes.

However, Pathtools' incomplete data set may soon be remedied by a data warehousing project underway at Cleveland Clinic. The medical center is organizing data from across the institution into a large centralized data repository that operates independent of the clinic's vendor-specific systems, Dr. Robertson says. Pathology

department data were expected to be loaded into the data warehouse by CAP TODAY press time.

“Once that has occurred,” Dr. Robertson says, “instead of searching my own SQL database that I built, I can point Pathtools to search the data warehouse, which will contain a much more comprehensive library of data.” □—*Renee Caruthers*

IICC announces publication of new LIVD specification

The IVD Industry Connectivity Consortium has released LOINC to Vendor IVD (LIVD) Specification V2.0.

LOINC for IVD (LIVD) is a format to facilitate the publication and exchange of LOINC codes for in vitro diagnostic test results linked to IVD test transmission codes or manual test identification. LIVD content is used by IVD software systems to automate the mapping between vendor IVD test transmission codes and LOINC codes.

Version 2.0 contains updates to the SARS-CoV-2 LIVD reference file maintained by the CDC. The updates are based on feedback from those using LIVD when reporting SARS-CoV-2 IVD test results. The latest specification also includes:

- the LOINC version ID as a LOINC data element to support mappings to multiple versions of LOINC in the same catalog.
- test kit and test kit-type fields to support test kit universal identifiers.
- the Health Level 7 International FHIR (Fast Healthcare Interoperability Resources) standards framework in lieu of the IICC-developed JSON (JavaScript Object Notation) open standard file format.

The LOINC for IVD standard can be implemented under a limited, nonexclusive, royalty-free public license. The LIVD V2.0 specification is available at <http://ivdconnectivity.org/livd/>.

Alliance to create sandbox to promote interoperability

The multi-organization collaboration InterOp.Community has announced plans to offer a virtual developer sandbox to promote open-source software and open platforms for health data exchange.

The community, a joint venture of the Interoperability Institute, Integrating the Healthcare Enterprise USA, Interopion, and Red Hat, will give developers the resources to collaborate on advancing interoperability. It will provide standards-driven interfaces for common components in EHR, laboratory, hospital, health information exchange, and insurer systems.

The sandbox is “the next-generation proving ground in the cloud, complete with tools to help build, launch, and test health apps as well as model real world health care scenarios with synthetic data,” according to an overview of the technology incubator project posted on the Interoperability Institute website.

The InterOp.Community was expected to launch the sandbox during the second quarter of this year.

Microsoft to purchase Nuance Communications

Microsoft has announced plans to acquire Nuance Communications as part of an effort to boost its portfolio of cloud-based products, which includes Microsoft Cloud for Healthcare.

Nuance is a provider of conversational artificial intelligence and cloud-based ambient clinical intelligence solutions, including the Dragon Medical One speech-recognition platform.

Microsoft and Nuance have collaborated since 2019 via a strategic development partnership.

The acquisition is expected to be finalized by the end of 2021.

[Nuance Communications](#), 781-565-5000

LigoLab Web portal supports school surveillance testing

LigoLab's TestDirectly direct-to-consumer platform supports COVID-19 school surveillance testing via test event scheduling and test collection, automated patient notification, and patient results reporting.

School districts can use the platform to organize and create testing events based on school location and patient type (student, teacher, or other staff member) by uploading patient demographics into the school's administrative portal via a CSV (comma-separated values) file or other method. School administrators, the CDC, and other government agencies can monitor test orders and results in real time.

TestDirectly can be used to prepare a batch testing event by generating all specimen labels simultaneously or automatically sending an email notification to each patient or parent of a minor inviting them to create a TestDirectly account.

TestDirectly can be integrated with laboratory information systems, revenue cycle management solutions, and EHR systems.

[LigoLab](#), 818-395-4659

Paige and Epredia announce global distribution partnership

The computational pathology company Paige has entered a commercial distribution agreement for its portfolio of diagnostic software solutions with Epredia, a precision cancer diagnostics firm.

Under terms of the agreement, Epredia will serve as a global distribution partner for Paige's imaging diagnostic products and hold exclusive distribution rights in Japan.

Epredia will commercialize Paige solutions alongside its products in the United States and select European countries beginning in the second quarter, according to an announcement from Epredia.

Paige's portfolio includes the Paige Platform imaging solution, which comprises a viewer, storage capabilities, and artificial intelligence-based diagnostic software and is compatible with most laboratory information systems.

[Epredia](#), 800-522-7270

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.