

BioFortis

Empowering Researchers Through Innovative Software

A COLLABORATIVE TRANSLATIONAL INFORMATICS ENVIRONMENT AT THE NCI

Jason Levine¹, John Otridge², Steve Chen²

¹Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD

²BioFortis Inc., Columbia, MD

ABSTRACT

Translational research programs have been created to provide a better understanding of disease susceptibilities and pathway mechanisms. These programs require integration and analysis of multiple disparate data sets that come from different sources such as the hospital and pathology/molecular laboratories. At the NIH, a subset of the patients treated at the Clinical Center consent into various Institutes' research protocols. Here we describe how NCI investigators utilize a centralized database system that 1.) securely stores, annotates, curates, and tracks information such as patient data, clinical phenotypes, biospecimen & derivatives data, and experimental research data, 2.) communicates clinical and research information with the NCI Clinical Data Registry and BTRIS systems, 3.) enables collaborative exchange and sharing of information amongst research groups, and 4.) provides an intuitive environment for investigators to query and review their collected data with minimal need for direct IT support.

BACKGROUND

When patients are admitted at the NIH, their demographics information, orders, results, participating protocols, and other clinical care information are stored in the Clinical Center's central electronic medical record database, Clinical Research Information System (CRIS). NCI Center for Cancer Research (CCR) maintains its own separate Clinical Data Registry for patients who are enrolled in intramural NCI protocols. However, investigators traditionally have been "on their own" to capture, annotate, and store research information from consented, participating patients. As a result, many

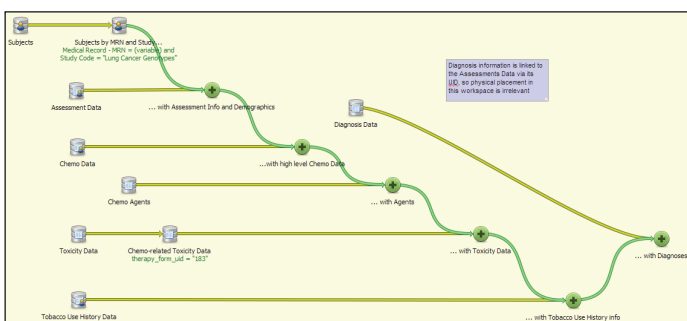
investigators managed clinical and research data in disparate information systems, often relying on spreadsheets and homegrown databases, which led to suboptimal retrieval and exchange of information that reduced the efficiency and effectiveness of translational medical research. Beginning in 2006, the Labmatrix software environment was introduced to CCR researchers with the following areas of emphasis and capabilities:

I. ACCESS & SECURITY

CCR's Labmatrix system is a clinical research data management platform that supports intramural research endeavors. Centrally-hosted and maintained on-site behind NIH's firewall, the system utilizes a web-browser thin-client/server model that allows for quick set up and management of users and groups based on laboratories, collaborations, research domains, and IRB protocols. While a set of central data resources are available to all Labmatrix users, a flexible security framework ensures private research data will only be accessible to designated group members and collaborators; furthermore, specific types of transactions against all data records (search, create, view, edit, delete) are configurable to ensure proper level of access. Individual's access to Labmatrix is initially facilitated by the existing user authentication authority at the NIH (Active Directory); upon signing in, the user's Labmatrix permissions are calculated via the individual's user group(s) assignment. To prevent misuse and interception of PHI, PII, and private research data, all user interactions with the system are encrypted while in transmission. Additionally, data values are logged in pre/post-state details for audit & security forensics.

III. INFORMATION RETRIEVAL

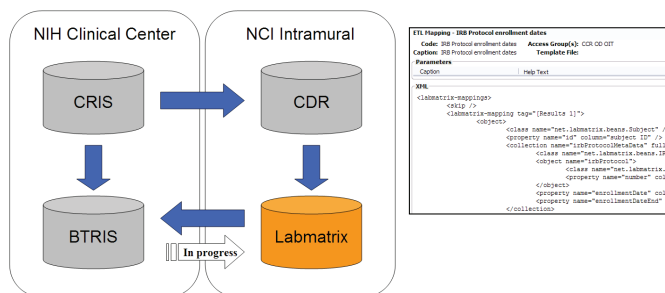
Traditionally, it has been difficult for investigators to query and obtain results for translational research because patient clinical information, biospecimen information, pathology information, molecular profile information, and other experimental data were kept in separate data repositories; even defining the correct linkage between data records from one repository to another can be very time-consuming (or even impossible) for research groups without dedicated IT support. With the deployment of Labmatrix, all types of data described above can be centrally federated/stored and accessed in a “one-stop shop” fashion. Utilizing the system’s graphical query user interface, investigators do not have to be trained in programmatic database languages in order to query across their data sets, hence receiving speedier data reports while freeing up IT resources. Query results can be exported in a variety of formats, while the queries themselves can be saved and shared. Furthermore, query parameters can be saved with static or dynamic values.



Graphical Query Interface

IV. PROGRAMMABILITY & MAINTENANCE

Labmatrix has a default web-browser user interface that presents different domain data in separate logical groups for ease of user consumption. This interface has been honed over multiple years of feedback from translational researchers. There are also alternative interfaces to access Labmatrix resources that adhere to the same system security rules. For example, Labmatrix’s Java API allows for programmatic access, so that designated IT personnel can develop and maintain custom projects, features, and user interfaces. Furthermore, with the “user scripting” and “extract-transform-load” (ETL) tools, advanced data workflows and complex data imports are possible. Exchange of information with other systems is supported by multiple connectivity options, working examples include NCI’s Clinical Data Repository and NIHCC’s BTRIS.



```

ETL Mapping: NCI Protocol enrollment dates
Code: NCI Protocol enrollment dates Access Group(s): CCR CD 007
Caption: NCI Protocol enrollment dates Template file
Parameters:
  - $DocID: //web/text
XML:
<labmatrix-mapping>
  <step />
  <labmatrix-mapping step="(Results 1)">
    <object>
      <class name="net.labmatrix.beans.Subject" />
      <property name="id" collection="subject" />
      <collection name="labProtocolMetaData" full="
      <class name="net.labmatrix.beans.IEIS
      <content name="labProtocol" />
      <class name="net.labmatrix.be
      <property name="number" column
      </object>
      <property name="enrollmentData" column
      <property name="enrollmentDate" co
    </collection>
  
```

Information Systems Data Flow (Left). Example ETL definitions interface (Right)

CONCLUSION

Labmatrix supports best practices in security, accountability, standardization, collaboration, efficiency, and data consumption for translational research. The implementation of CCR’s Labmatrix system provides a centralized, secure, integrated, and flexible informatics platform. NCI investigators utilize Labmatrix in their daily routines to store, review, manage, share, and explore clinical research data.