Digital Pathology and Integrative Query System Subproject

Contents of this page

Overview - Digital Pathology Component

Digital pathology, unlike its more mature radiographic counterpart, has yet to standardize on a single storage and transport media facilitating an integrative approach. While DICOM has published a digital pathology standard, none of the major vendors in this space have adopted the standard to any real extent. The result of this lack of uniformly accepted standards is that outside a given laboratory of small collaborative groups, the integration of pathology data with radiographic, genomic, and proteomic data is all but impossible. Furthermore, because each pathology-imaging vendor produces its own image management systems, image analysis systems are also, by extension, proprietary and not standardized. The result is that images produced on different systems cannot be analyzed via the same mechanisms.

Project Description

This project proposes to leverage several open source and previously NCI-funded activities to provide an open source digital pathology image server that can host and serve digital pathology images for any of the major vendors without recoding, which often introduces additional compression artifacts. A single digital pathology server would allow NCI to include digital pathology images within TCIA / NBIA and provide a logical bridge from proprietary pathology formats to DICOM standards. Specifically, we are proposing to expand the functionality of the caMicroscope digital pathology platform to include support for some of the common formats adopted by whole slide vendors. This functionality would be made possible by use of the Openslide library. If time and funds allow, we would also expand the OpenSlide library to include the Olympus file format (the only format not currently supported).

The recently concluded FDA Imaging Pilot demonstrated the feasibility of data federation between NBIA and AIM, and the value of such data federation in streamlining the imaging review processes at the FDA. This data federation is made possible by the Bindaas middleware that is also used to build the backend infrastructure of caMicroscope. We propose to extend this backend infrastructure with a data federation capability that provides for query capability using multiple data points across TCIA and TCGA will be implemented.

Overview - Integration of Imaging & Molecular Data

All three research domains will clearly need an imaging archive that can be leveraged for integration across multiple data types and sources. For example, TCGA program has the goal of producing a comprehensive genomic characterization and analysis of 200 types of cancer and providing this information to the research community. TCIA and the underlying NBIA software stack were created to manage well-curated, publicly-available collections of medical image data, including diagnostic images associated with the tissue samples sequenced by TCGA. TCIA currently supports over 40 active research groups including researchers who are exploiting the existing linkages between TCGA and TCIA. TCIA has recently released an API — an Application Programmatic Interface that provides a REST API to TCIA metadata and image collections. This API is built using a middleware platform called Bindaas, an d this API is being designed to support federation of multiple information repositories using the concepts of a data mashups. This infrastructure can be expanded to include more data types and additional integration, and provide analytic and decision support, which will act as a foundation for a broader set of novel community research projects.

Subproject Team

Name	Role	Organization
Jose Galvez	Team Lead	NCI
John Freymann	SME	NCI - Leidos
Justin Kirby	SME	NCI - Leidos
Ashish Sharma	PI	Emory

Subproject Documentation

- Digital Pathology and Integrative Query System Documentation
- JIRA - Digital Pathology

- Scope
- Requirements