
A Distributed Biomarker Atlas for Lung Research aiding the Discovery and Early Detection of Cancer Biomarkers

Wilbur Franklin¹, Dan Crichton², Mary Reid³, Chris Mattmann², Andrew Hart²,
Dayi Deng¹, Brian Logue¹, John Hayes¹, Deanna Stelling⁴

¹University of Colorado Health Sciences Campus, Aurora, CO 80045, USA

²Jet Propulsion Laboratory, California Institute of Technology Pasadena, CA 91109, USA

³Roswell Park Cancer Institute, Buffalo, New York 14263, USA

⁴Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

The EDRN Biomarker Atlas Working Group recently has developed a Biomarker Atlas software system that allows a researcher to correlate lung cancer patients with similar characteristics around regions of the lung in which their sample specimens including multiple types of bronchoscopies (e.g., white light, fluorescent), slide images of cancerous lesions, and other data products were collected. Researchers can observe characteristics of collected specimens over time and use the Biomarker Atlas to identify patient trends based on family characteristics, epidemiology, and the site at which the specimen were collected. Patient specimen data is annotated with a common set of data elements that allow the Biomarker Atlas system to properly filter through and co-locate the available data in the system.

Recent technological advances in location-based data discovery and AJAX-based user interfaces, e.g., Google Maps, have served user communities well, allowing even novice users to find businesses, restaurants and sites of interest using a few simple keywords and the selection of a "region" on the map. Not only is the learning curve for these user interfaces significantly smaller, but also the navigation and general interaction patterns between user and application are becoming pervasive in modern search engine technology.

Our "Biomarker Atlas" capability is predicated upon the above recent technical advances and provides a means for lung cancer researchers to browse lung cancer specimens collected at sites participating in the distributed specimen sharing network using a location-based map of a human lung. The backend distributed data system of the Biomarker Atlas is built upon the data grid middleware framework called OODT, the principal enabling technology of the EDRN Resource Network Exchange (ERNE).

The initial pilot sites for the Biomarker Atlas include University of Colorado Health Sciences Center (UCHSC) and Roswell Park Cancer Institute (RPCI). Future work entails collaborating with more EDRN sites to include their specimens in the distributed Biomarker Atlas system, and increase the changes of cancer biomarker discovery and observation.

Acknowledgement: This material is based upon work supported by the Jet Propulsion Laboratory, managed by the California Institute of Technology. Effort also supported by the National Cancer Institute's Early Detection Research Network (EDRN) program.
