

JPL Informatics Team Report

September 2009

<http://cancer.jpl.nasa.gov> (informatics center)

<http://edrn.nci.nih.gov/> (ops)

<http://nci-test-ws04-2-v.nci.nih.gov/> (test)

<http://edrn.jpl.nasa.gov/> (beta)

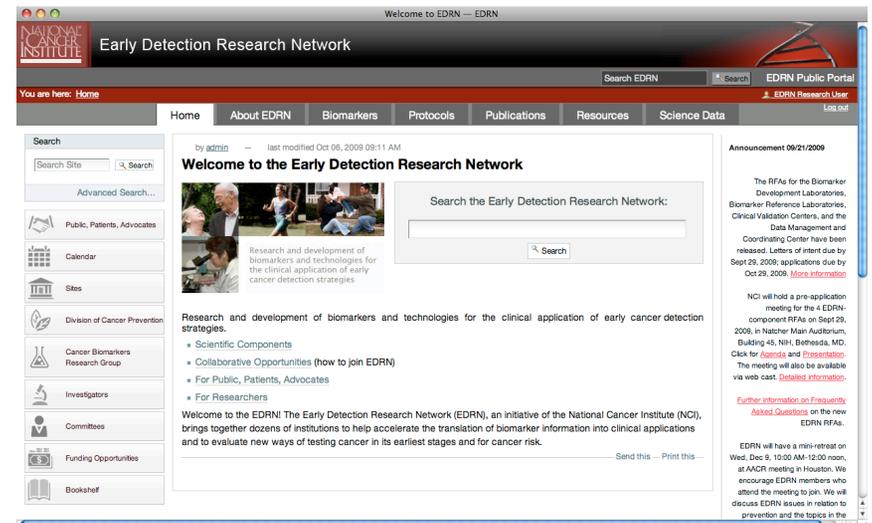


EDRN Knowledge Portal Accomplishments and Status

- New Plone 3 Version in internal testing at JPL. To be released soon to DMCC and NCI for further testing and feedback.

- New Features include:

- Complete new look and feel
- EDRN Knowledge Environment (EKE) has been integrated into the top level of the portal including:
 - Google-like search
 - New Tabs: Biomarkers, Protocols, Science Data



- High-level security has been implemented to allow groups of users to login. For Example, EDRN Researchers can now login and see curated EDRN biomarkers.
- All pages that were generated by DMCC have been replaced with data sharing between DMCC and JPL and are now generated by Plone 3
- Initial code to run new Plone 3 version of EDRN Portal sent to TerpSys

EKE Portal

Plans and Next Steps

- Fine-tune final features
- Define granular security plan for EDRN Public Portal
- Review granular security plan with DMCC and NCI
- Implement granular security plan
- Release to NCI for feedback and comments



Biomarker Database

- Accomplishments and Status
 - Supports real-time curation of growing number of biomarkers
 - Information about biomarkers, studies, publications, resources, sites, investigators and more can be linked together and annotated by a curator to provide a comprehensive picture of the current state of research for a given biomarker.
- Plans and Next Steps
 - Add relevant security groups to support the security model being implemented in the Portal.
 - Support the increasing richness and depth of the research data and scaling to meet the needs of a growing virtual organization as it incorporates research from additional sites and investigators.

eCAS

- Accomplishments and Status
 - Data loaded from 11 EDRN Protocols and 19 sites with a total of 1223 products ingested.
 - Additional metadata added to UI and curated
- Plans and Next Steps
 - Next software release:
 - PubMed data import similar to BMDB
 - Links to protocol data in the public portal
 - Breadcrumb links to data in the public portal
 - Curate data received from Alvin Liu
 - Curate data received from University of Michigan

ERNE

- Accomplishments and Status
 - Developing mock-up of Integration into EDRN Public Portal
 - Continued monitoring of data quality
 - Total of 10 sites connected
 - Resolved data quality issues at Duke University
 - Analyzing data quality at other sites
- Plans and Next Steps
 - Tightly integrate ERNE Specimen Searches into Public Portal
 - Develop plan to include Specimen Reference Sets into Portal search results
 - Automate QA

EDRN/Canary

Never Smokers Lung Cancer Project

Background

- Work with Genologics on a collaborative pilot project to collect biomarker discovery data from the EDRN/Canary Never Smokers Lung Cancer Project from 5 sites. Using the proposed joint architecture the researchers will deposit the data in a single virtual repository, and integrate the results to find the most promising biomarkers.

Accomplishments and Status

- Agreed on distributed informatics architecture between EDRN and Canary
- Completed conference calls with two of the participating sites (Sam Hanash and Ite Laird) to capture their data
- Held security requirements telecon with Genologics

Plans and Next Steps

- Complete data calls with additional sites (Adi Gazdar and Wan Lam)
- Gather test data from sites

Early

Detection
Research
Network

Biomarker Data Management (collaboration w/ Dartmouth)

- Accomplishments and Status
 - In collaboration with NCI (Christos Patriotis), worked with Dartmouth on Biomarker Database content management
 - Continue to curate biomarkers providing deep content such as publications, alternate names, etc.
 - Completed Review of Hanash CARET markers biomarker annotations and science data
 - Populating next set of markers for review (PLCO – Crammer and GI – Brenner)
 - Contacted several EDRN Investigators for Biomarker content (Dr. Crammer, Dr. Brenner, DMCC and NCI)
 - Presented Biomarker Database and Biomarker Review Process to Breast/ GYN Collaborative Group and GI Collaborative Group at EDRN September Steering Committee Meeting
 - Conducted the following ongoing conference calls:
 - Bi-weekly calls to discuss the Biomarker Database curation, system development and process development and evaluation of biomarkers with NCI and the DMCC
 - Bi-weekly calls to discuss eSIS curation, process development and data modeling for eSIS
 - Complete Protocol Records for Biomarker and eCAS content as it is curated



Biomarker Data Management Continued

- Plans and Next Steps
 - Continue curating PLCO – Crammer markers for review in December
 - Continue curating GI – Brenner markers for review in December
 - Create Biomarker Review Report and Check Lists to send to Pis for review in November
 - Continue researching and contacting investigators to provide deep content for existing markers (populating sensitivity and specificity, etc.)
 - Continue working with Dartmouth and DMCC to provide specifications on Biomarker Database and Biomarker Data Model
 - Continue updating documentation of biomarker database curation process
 - Continue working with EDRN sites to gather feedback on applicable metadata for eCAS
 - Continue contacting PIs to complete Protocol information
 - Continue all BMDB, eCAS, eSIS and informatics calls.



Publications, Presentations and Outreach

- Publications

- Chris Mattmann, John Tran, Heather Kincaid, Dan Crichton, Andrew Hart, Kristen Anton, Jackie Dahlgren, Mark Thornquist, Deanna Stelling, Suzanna Reid, Christos Patriotis, and Sudhir Srivastava. EDRN Catalog and Archive Service (eCAS) – The eCAS Model for Scientific Data Warehousing of Biomarker Data. Poster. Presented at the EDRN Scientific Workshop, August 2009. Bethesda MD.
- Heather Kincaid, Andrew Hart, Kristen Anton, Christos Patriotis, Mark Thornquist, Jackie Dahlgren, Chris Mattmann, Dan Crichton, and John Tran. EDRN Curation – Curation of EDRN Cancer Biomarker Research. Poster. Presented at the EDRN Scientific Workshop, August 2009. Bethesda MD.
- Dan Crichton, Mark Thornquist, Sean Kelly, Andrew Hart, Heather Kincaid, Chris Mattmann, Jackie Dahlgren, Kristen Anton, Deanna Stelling, Greg Warnick, Susanna Reid, Cim Edelstein, John Tran, Christos Patriotis, and Sudhir Srivastava. The EDRN Knowledge Environment (EKE) – Providing Integrated Access to Scientific Information and Knowledge in Cancer Biomarker Research. Poster. Presented at the EDRN Scientific Workshop, August 2009. Bethesda MD.

- Meetings

- Attend EDRN Scientific Workshop and Steering Committee Meeting – August/September 2009
- Genomics/Canary meeting on the informatics architecture – August 2009
- Meeting with University of British Columbia to discuss Biomarker Atlas – August 2009
- Meeting at the DMCC in Seattle to discuss and coordinate Informatics progress – August 2009

- Presentations

- Dan Crichton was invited (and sponsored) to present on “Archiving Science Data Systems in the 21st Century” at Indiana University in Bloomington, IN on Sep 28.

Early

Detection
Research
Network



EDRN Meetings/Interactions

- Informatics Planning Meetings (bi-weekly)
 - Thursdays, 11 AM PT
- ERNE QC teleconferences (bi-weekly)
 - Wednesdays, 2 PM PT
- ERNE Integration Meetings (bi-weekly)
 - Tuesdays, 9:00 AM PT
- Biomarker Curation Teleconference (bi-weekly)
 - Thursdays, 9:00 AM PT
- eSIS Coordination Teleconference (bi-weekly)
 - Tuesdays, 8:30 AM PT
- eCAS Coordination Teleconference (bi-weekly)
 - Tuesdays, 8:30 AM PT
- JPL Informatics Teleconference (weekly)
 - Tuesdays, 10 AM PT
- Canary/Lung Project calls (bi-weekly)
 - Thursdays, 9:30 AM PT
- Canary/Lung Technical calls (as needed)

Links, etc

- EDRN operational website
<http://edrn.nci.nih.gov/>
- EDRN beta website
<http://edrn.jpl.nasa.gov/>
- NCI test website
<http://nci-test-ws04-2-v.nci.nih.gov/>
- JPL EDRN Informatics Center Website
<http://cancer.jpl.nasa.gov/>
- eCAS Browser
<http://cancer.jpl.nasa.gov/ecas>
- eCAS Curator
<http://cancer.jpl.nasa.gov/ecas-curation/>
- BMDB curation system
<http://cancer.jpl.nasa.gov/bmdb>



Informatics Personnel

- Dan Crichton, PI
- Chris Mattmann, Development Lead
- Heather Kincaid, Operations Lead
- Sean Kelly, Technologist
- Steve Hughes, Information Architect
- Andrew Hart, Developer
- John Tran, Developer
- Thuy Tran, Developer
- Andrew Clark, NIH Bioinformatics Intern