

# BD2K TARGETED SOFTWARE DEVELOPMENT PROJECTS

## DATA PRIVACY

Mark Gerstein, Yale University  
“Methods and Software to Enhance Genomic Privacy and Sharing of RNA-seq Data”

Haixu T. Tang, Indiana University Bloomington  
Xiaofeng Wang,  
“Encryption Methods and Software for Privacy-Preserving Analysis of Biomedical Data”

Darren Toh, Harvard Pilgrim Health Care, Inc.  
“Privacy-protecting distributed analysis of biomedical big data”

## DATA REPURPOSING

John Stelling, Brigham and Women's Hospital  
“Integrating The World’s Microbiology Laboratories into a Global Microbial Surveillance System”

Alexander Tropsha, University of North Carolina - Chapel Hill  
“Drug Repurposing for Cancer Therapy: From Man to Molecules to Man”

## APPLYING METADATA

Scott Campbell, University of Nebraska Medical Center  
“Deploying ONC National Standards in Support of Metadata for Big Data Research Warehouse Management of Repurposed Laboratory, Pathology & Patient Findings Data from the EHR”

Timothy W Clark, Massachusetts General Hospital  
“Argumentation and linked-metadata services for reproducible target validation”

Guoqian Jiang, Mayo Clinic Rochester  
“Tools for standardizing clinical research metadata using HL7 FHIR”

Christopher J Mungall, University of California - Lawrence Berkeley Laboratory  
“An Intelligent Concept Agent for Assisting with the Application of Metadata”

Cui Tao, University of Texas Health Science Center - Houston  
“Metadata Applications On Informed Content to Facilitate Biorepository Data Regulation and Sharing”

## DATA COMPRESSION & REDUCTION

ALI BILGIN, UNIVERSITY OF ARIZONA  
"Task-Specific Compression for Biomedical Big Data"

MARTIN FARACH-COLTON, RUTGERS, THE STATE UNIV OF N.J.  
"Meaningful Data Compression and Reduction of High-Throughput Sequencing Data"

MEHMET KOYUTURK, CASE WESTERN RESERVE UNIVERSITY  
"Theoretical Foundations and Software Infrastructure for Biological Network Databases"

OLGICA MILENKOVIC, UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN  
TSACHY WEISSMAN, STANFORD UNIVERSITY  
"Genomic Compression: From Information Theory to Parallel Algorithms"

PETER ROSE, UNIVERSITY OF CALIFORNIA SAN DIEGO  
"Compressive Structural Bioinformatics: High Efficiency 3D Structure Compression"

## DATA VISUALIZATION

HONG-WEI DONG, UNIVERSITY OF SOUTHERN CALIFORNIA  
"From Terabytes of Pixels to Intuitive Brain Networks"

JOHN NOVEMBRE, UNIVERSITY OF CHICAGO  
"Tools for Visualization of Geographic Structure in Population Genomic Data"

PETER J PARK, HARVARD MEDICAL SCHOOL  
"Visual Analysis of Genomic and Clinical Data from Large Patient Cohorts"

GRIFFIN M WEBER, HARVARD MEDICAL SCHOOL  
"Visualizing healthcare system dynamics in biomedical Big Data"

## DATA PROVENANCE

ZACHARY IVES, UNIVERSITY OF PENNSYLVANIA  
JUNHYONG KIM  
"Approximating and Reasoning about Data Provenance"

ERICH S. HUANG, DUKE UNIVERSITY  
"Flexible & Executable Provenance in Data-Intensive Biomedical Research: A Flexible Research Data Service"

SATYA SANKET SAHOO, CASE WESTERN RESERVE UNIVERSITY  
"A PROV standard-based data source agnostic provenance engine for Big Data analytics"

## DATA WRANGLING

SANDRA M. HARABAGIU, UNIVERSITY OF TEXAS - DALLAS

IYAD OBEID, TEMPLE UNIVERSITY

JOSEPH PICONE

“Automatic discovery and processing of EEG cohorts from clinical records”

WLADEK MINOR, UNIVERSITY OF VIRGINIA

“Integrated resource for reproducibility in macromolecular crystallography”

WEI WANG, UNIVERSITY OF CALIFORNIA LOS ANGELES

TYSON CONDIE,

SEAN YOUNG,

“Mining The Social Web to Monitor Public Health and HIV Risk Behaviors”

CHUNLEI WU, SCRIPPS RESEARCH INSTITUTE

“Community Platform for Data Wrangling of Gene and Genetic Variant Annotations”

JUN ZHU, ICAHN SCHOOL OF MEDICINE AT MOUNT SINAI

“Developing methods for curating multi-omics data”

### Reporter link to awards:

[https://projectreporter.nih.gov/Reporter\\_Viewsh.cfm?sl=12E8CA0C468ACED47598B8961CAA4A01A2FFCEB861BF](https://projectreporter.nih.gov/Reporter_Viewsh.cfm?sl=12E8CA0C468ACED47598B8961CAA4A01A2FFCEB861BF)