

Poster/Demo Presentations

Poster/Demo Number	Presentation Title
Research Highlights	
1 – Poster	Prediction of Putative Causal Variants and Genes at BMD GWAS Loci <i>Basel Al-Barghouthi</i>
2 – Poster	Understanding Lung Tissue Heterogeneity in Idiopathic Pulmonary Fibrosis <i>Panayiotis Benos</i>
3 – Poster	A Bayesian Causal Inference Method for Identifying Cancer Drivers of Individual Tumors <i>Chunhui Cai</i>
4 – Poster	Finding the Signal in the Noise: Social Media Early Hospital Notification of Mass Casualty Events <i>Rachael Callcut</i>
5 – Poster	A Need for Better Data Sharing Policies: A Review of Data Sharing Policies in Biomedical Journals <i>Robin Champieux</i>
6 – Poster	Analysis of RNA Editing in Cancer Epithelial-to-Mesenchymal Transition <i>Tracey Chan</i>
7 – Poster	Cataloguing and Curating BRCA1/2 Genetic Variation <i>Melissa Cline</i>
8 – Poster	Big Data Imaging Processing & Analysis (BigDIPA) <i>Michelle Digman</i>
9 – Poster	Predicting Adverse Cardiovascular Events for Tyrosine Kinase Inhibitors From Molecular Features <i>Anders Dohlman</i>
10 – Poster	Consumer Wearable Devices for Health Surveillance and Disease Monitoring <i>Jessilyn Dunn</i>
11 – Poster	IRRCMC - A Public Database of Macromolecular Diffraction Experiments <i>Marek Grabowski</i>
12 – Poster	Large-Scale Biological Text Mining: A Data-Driven Approach <i>Jiawei Han</i>
13 – Demo	Modeling Disease Progression From Sparsely Sampled Observations <i>Lukasz Kidzinski</i>
14 – Poster	Using Twitter to Study Autism <i>Denise McGinnis</i>
15 – Poster	Fine-Mapping of Obesogenic cis-Regulatory eQTL Variants Using High-Resolution Capture Hi-C <i>David Pan</i>

16 – Poster	Creating a Standard Programmatic Interface for Genomic Data With the GA4GH API <i>Benedict Paten</i>
17 – Poster	Systems Biology, Meet Evolution and Protein Structure for Characterizing Disease Variant Mechanisms <i>Jeremy Prokop</i>
18 – Poster	KnowEnG: Scalable Knowledge-Guided Analysis of Genomic Data Sets on the Cloud <i>Saurabh Sinha</i>
19 – Poster	QuBBD: SMART - Spatial-Nonspatial Multidimensional Adaptive Radiotherapy Treatment <i>David Vock</i>
20 – Poster	Extraction and Analysis of Signatures From the Gene Expression Omnibus by the Crowd <i>Zichen Wang</i>
Data Commons	
21 – Demo	Common Credit: Evaluating the Scalability of Open Source Applications Across Cloud-Based Providers <i>Paul Avillach</i>
22 – Poster	The MO-LD Project: Enhancing the FAIRness of Yeast and Other Model Organism Data <i>Michel Dumontier</i>
23 – Poster	The smartAPI Initiative: Making Web APIs FAIR <i>Michel Dumontier</i>
24 – Demo	Building The Commons: Interoperable Big Data Publication and Analytics <i>Ian Foster</i>
25 – Demo	BDbags and Minids: Tools for Managing Complex Big Data Sets <i>Carl Kesselman</i>
26 – Poster	FAIR Dataset Landing Pages, Digital Research Objects, and Software Tools for LINCS and BD2K <i>Amar Koleti</i>
27 – Demo	Cloud-Based Drag-and-Drop Scalable RNA Sequencing Pipeline <i>Alexander Lachmann</i>
28 – Demo	The Harmonizome: A Collection of Processed Datasets Gathered to Serve and Mine Knowledge About Genes <i>Avi Ma'ayan</i>
29 – Poster	Large-Scale, Cloud-Based Analysis of Cancer Data <i>Brian O'Connor</i>
30 – Poster	Automated Deployment of KnowEnG Portal Via Docker Containers in AWS Cloud <i>Pramod Rizal</i>
31 – Demo	Catalyzing Biomedical Research Through the NIH Commons Credits Cloud Computing Paradigm <i>David Tanenbaum</i>

32 – Poster	RNA-seq Pipeline Tutorial With an Example of Reprocessing Data From a Recent Zika Virus Study <i>Zichen Wang</i>
33 – Poster	Deriving Signatures of Pharmacological Action via LINCS Signatures <i>Lixia Zhang</i>
34 – Demo	Big Data for Discovery Science (BDDS): Neuroimaging PheWAS <i>Lu Zhao</i>
Standards Development	
35 – Poster	Integrative Representation and Analysis of the LINCS Cell Lines Using the Cell Line Ontology <i>Caty Chung</i>
36 – Poster	An Urban Dictionary of Identifier Syntax for the Data Integration Jungle <i>Julie McMurry</i>
Training & Workforce Development	
37 – Demo	GUI Design and Big Data Visualization of BigDataU Website Development <i>Sumiko Abe</i>
38 – Demo	ERuDIte: The Educational Resource Discovery Index for Data Science Learning <i>Jose-Luis Ambite</i>
39 – Poster	Training Component Activities at the Center for Causal Discovery <i>Takis Benos</i>
40 – Poster	Preparing Underrepresented and First-Generation Students for Careers in Biomedical Big Data Science <i>Judith Canner</i>
41 – Demo	Biomedical Big Data Training for Novices: Initial Experience With a Short-Term Summer School <i>Brian Chapman</i>
42 – Demo	Decaying Relevance of Clinical Data When Predicting Future Decisions <i>Jonathan Chen</i>
43 – Poster	Training and Implementing Genomic Big Data Courses at Primarily Undergraduate Serving Institutions <i>Jeffrey Chuang</i>
44 – Poster	Big Data Research and Education Program in a Primarily Undergraduate Institution (PUI) <i>Math Cuajungco</i>
45 – Poster	Biomedical Research Data Management Open Online Education: Challenges and Lessons Learned <i>Julie Goldman</i>
46 – Poster	Demystifying Biomedical Big Data: A Free OnLine Course <i>Yuiry Gusev</i>
47 – Poster	Community Training and Outreach Activities of the BD2K-LINCS DCIC <i>Sherry Jenkins</i>

48 – Poster	Community Research Education and Engagement for Data Science <i>Patricia Kovatch</i>
49 – Poster	Get Real: A Synthetic Dataset Illustrating Clinical and Genetic Covariates <i>Ted Laderas</i>
50 – Poster	Getting Your Hands Dirty With Data <i>Ted Laderas</i>
51 – Demo	Engaging and Training Undergraduates in Big Data Analysis Through Genome Annotation <i>Wilson Leung</i>
52 – Demo	KnowEnG tools for Barrier-Free Learning of Genomic Data Clustering <i>Mohith Manjunath</i>
53 – Poster	Increasing Diversity in Interdisciplinary Big Data to Knowledge (IDI-BD2K) in Puerto Rico <i>Patricia Ordóñez</i>
54 – Poster	Data Science Educational Resources for Anyone, Anywhere <i>Nicole Vasilevsky</i>
55 – Poster	Pandem-Data: Using Big Data in High School <i>Chuck Wood</i>
56 – Poster	Big Data Training for Translational Omics Research <i>Min Zhang</i>
BioCADDIE & Resource Indexing	
57 – Demo	Aztec: A Cloud-Based Computational Platform to Integrate Biomedical Resources <i>Brian Bleakley</i>
58 – Demo	Reactome: A Curated Knowledge Base of Biomolecular Pathways <i>Antonio Fabregat</i>
59 – Demo	SATORI: A System for Ontology-Guided Visual Exploration of Biomedical Data Repositories <i>Nils Gehlenborg</i>
60 – Poster	A Framework for Metadata Management and Automated Discovery for Heterogeneous Data Integration <i>Ramkiran Gouripeddi</i>
61 – Poster	A Machine Learning Approach for Data Source and Type Identification to Support Metadata Discovery <i>Ramkiran Gouripeddi</i>
62 – Poster	A Scalable Dataset Indexing Infrastructure for the bioCADDIE Data Discovery System <i>Jeffrey Grethe</i>
63 – Poster	Deep Learning-Based Multi-Modal Indexing of Heterogeneous Clinical Data for Patient Cohort Retrieval <i>Sanda Harabagiu</i>
64 – Demo	The LINCS Data Portal and FAIR LINCS Dataset Landing Pages <i>Amar Koleti</i>

65 – Poster	Augmenting the the Capabilities for Semantic Search of the Medical Literature <i>Ani Nenkova</i>
66 – Demo	Reactome: New Services and Widgets to Ease Third-Party Integration <i>Konstantinos Sidiropoulos</i>
67 – Poster	Indexing Clinical Research Datasets Using HL7 FHIR and Schema.org <i>Harold Solbrig</i>
68 – Poster	Aztec I: Building a Technology Platform to Integrate Biomedical Resources <i>Justin Wood</i>
69 – Demo	Development of DataMed, a Data Discovery Index Prototype by bioCADDIE <i>Hua Xu</i>
70 – Poster	Metadata Mapping in bioCADDIE: Challenging Cases <i>Nansu Zong</i>
Software, Analysis, & Methods Development	
71 – Poster	Histology-Validated Neural Networks Enable Optical Coherence Tomography Virtual Histology <i>Vikram Baruah</i>
72 – Poster	REproducible by Design, A Docker-Based Tool for Workflow Design, Data Linkage, and Workflow Execution <i>Tyler Bath</i>
73 – Poster	The Georgetown Database of Cancer (G-DOC): A Web-Based Data Sharing Platform for Precision Medicine <i>Krithika Bhuvaneshwar</i>
74 – Poster	Flexible Bootstrapping Approaches Toward the Clustering of Complex Medical Data <i>Rachael Blair</i>
75 – Demo	KnowEnG: Cloud-Based Environment for Scalable Analyses of Genomic Signatures <i>Charles Blatti</i>
76 – Demo	Interactive Web Application for Visualization of Brain Connectivity <i>David Caldwell</i>
77 – Poster	GRcalculator: An Online Tool for Calculating and Mining Drug Response Data <i>Nicholas Clark</i>
78 – Poster	Formal Evidence Networks for Reproducibility in Biomedical Translation <i>Tim Clark</i>
79 – Poster	Augmenting Metadata With Models of Experimental Methods <i>Scott Colby</i>
80 – Poster	Old Medicines, New Uses: Upcycling Drugs Using Social Media and Cheminformatics <i>Nabarun Dasgupta</i>
81 – Poster	A Software Suite for Causal Modeling and Discovery <i>Jeremy Espino</i>

82 – Demo	Clustergrammer: Interactive Visualization and Analysis Tool for High-Dimensional Biological Data <i>Nicolas Fernandez</i>
83 – Poster	Weak Supervision: Biomedical Entity Extraction Without Labeled Data <i>Jason Fries</i>
84 – Demo	Reproducible Exploratory Data Analysis With Vistories <i>Nils Gehlenborg</i>
85 – Demo	TruenoDB: A Network Database System for Managing, Analyzing, and Querying Large Biological Networks <i>Ananth Grama</i>
86 – Poster	Increasing NCBO BioPortal and CEDAR Synergy for BD2K <i>John Graybeal</i>
87 – Poster	Establishing Context: Geospatial Health Context Cube <i>Timothy Haithcoat</i>
88 – Poster	Active Deep Learning-Based Annotation of Electroencephalography Reports for Cohort Identification <i>Sanda Harabagiu</i>
89 – Demo	BDDS Tools to Enable Transcriptional Regulatory Network Analysis <i>Ben Heavner</i>
90 – Demo	Biobank to Digibank: High-Frequency Mobile Sensor Data Collection for Long-Lasting Research Utility <i>Timothy Hnat</i>
91 – Poster	The Duke Data Service: Building an Infrastructure for Data and Provenance Microservices <i>Erich Huang</i>
92 – Poster	A Computational Framework for Identifying New Treatment Options in Glioblastoma <i>Haruka Itakura</i>
93 – Demo	SigNetA: Web Tool for Network Analysis of Gene Expression Signatures <i>Rashid Karim</i>
94 – Poster	Building Entity Matching Management Systems for Data Science Problems in Biomedicine <i>Pradap Konda</i>
95 – Poster	Dashboard Visualization and Tool Integration for Enrichr <i>Maxim Kuleshov</i>
96 – Poster	Gene Wiki Knowledgebase and Tool Development for Molecular Signatures of Cardiovascular Phenotypes <i>Jessica Lee</i>
97 – Demo	Pathfinder: Visual Analysis of Paths in Biological Networks <i>Alexander Lex</i>
98 – Demo	GEN3VA: Aggregation and Analysis of Gene Expression Signatures From Related Studies <i>Avi Ma'ayan</i>

99 – Poster	A MACE2K Text Mining Tool to Extract the Impact of Genomic Anomalies on Drug Response <i>A.S.M. Ashique Mahmood</i>
100 – Demo	Faster and Better Metadata Authoring Using CEDAR's Value Recommendations <i>Marcos Martinez-Romero</i>
101 – Poster	New Algorithms for RNA-seq and ChIP-seq Data Compression <i>Olgica Milenkovic</i>
102 – Poster	kBOOM! Intelligent Merging of Different Disease Terminologies <i>Chris Mungall</i>
103 – Demo	Using Crowds to Design Biological Network Visualizations <i>T.M. Murali</i>
104 – Poster	ADAM Enables Distributed Analyses Across Large-Scale Genomic Datasets <i>Frank Nothaft</i>
105 – Poster	GTRAC: Fast Retrieval From Compressed Collections of Genomic Variants <i>Idoia Ochoa</i>
106 – Poster	Mining Electronic Health Records for Possible Drug Repositioning Opportunities <i>David Page</i>
107 – Poster	Dynamic Control Models for Strategic Interaction <i>John Pearson</i>
108 – Demo	Automatic Discovery and Processing of EEG Cohorts From Clinical Records <i>Joseph Picone</i>
109 – Demo	EEG Event Detection Using Deep Learning <i>Joseph Picone</i>
110 – Poster	Scalable EEG Interpretation Using Deep Learning and Schema Descriptors <i>Joseph Picone</i>
111 – Demo	Integrative LINCS (iLincs): Connecting Diseases, Drugs, and Mechanisms of Actions <i>Marcin Pilarczyk</i>
112 – Poster	NetLINCS: Correlation of Chemical Perturbagen and Signatures to Identify Biological Targets <i>John Reichard</i>
113 – Demo	Compressive Structural Bioinformatics: Large-Scale Analysis and Visualization of the PDB Archive <i>Peter Rose</i>
114 – Poster	Scientific Reproducibility Using the Provenance for Clinical and Healthcare Research Framework <i>Satya Sahoo</i>
115 – Poster	Computational Tools and Resources for LINCS Proteomics Data <i>Behrouz Shamsaei</i>
116 – Poster	SAP – A CEDAR-Based Pipeline for Semantic Annotation of Biomedical Metadata <i>Ravi Shankar</i>

117 – Poster	Using Eye Tracking to Enhance Usability of Big Data in Cancer Precision Medicine <i>Vishakha Sharma</i>
118 – Poster	Big Data Contrast Mining for Genetic Distinctions Between Disease Subtypes <i>Matt Spencer</i>
119 – Poster	Enabling Privacy-Preserving Biomedical Data Analytics in the Cloud and Across Institutions <i>Haixu Tang</i>
120 – Demo	Multitask Deep Neural Net Kinase Activity Profiler <i>John Turner</i>
121 – Poster	Patient Linkage Across Research Datasets in a Patient Information Commons <i>Griffin Weber</i>
122 – Demo	Visualizing Healthcare System Dynamics in Biomedical Big Data <i>Griffin Weber</i>
123 – Demo	KEGGlincs Design and Application: An R Package for Exploring Relationships in Biological Pathways <i>Shana White</i>
124 – Demo	CEDAR: Easing Authoring of Metadata to Make Biomedical Datasets More Findable and Reusable <i>Debra Willrett</i>
125 – Demo	BioThings APIs: Linked High-Performance APIs for Biological Entities <i>Chunlei Wu</i>
126 – Poster	Global Detection of Epistasis <i>Sihai Zhao</i>
127 – Poster	Curate Patient-Centric Multi-Omics Data for Precision Medicine <i>Jun Zhu</i>
Collaborative Presentations	
128 – Poster	Aztec and CEDAR: Extraction of Digital Object Metadata From Free Text <i>Brian Bleakley</i>
129 – Poster	Leveraging the CEDAR Workbench for Ontology-Linked Submission of AIRR Data to the NCBI-SRA <i>Syed Ahmad Chan Bukhari</i>
130 – Poster	Cloud-Based Integration of Causal Modeling and Discovery Tools With a Unified Patient Research Database <i>Jeremy Espino</i>
131 – Poster	Machine Learning in Textual Data of Cardiovascular Disease via Phrase Mining and Network Embedding <i>Vincent Kyi</i>
132 – Demo	ELIXIR: A European Distributed Infrastructure for Life-Science Information <i>Pablo Roman-Garcia</i>

133 – Poster	FAIR LINCS Data and Metadata Powered by the CEDAR Framework <i>Raymond Terryn</i>
Sustainability	
134 – Poster	Combining Protein and Genome Annotation for Interpretation of Genomic Variants <i>Peter McGarvey</i>
135 – Poster	Interoperability of NURSA, PharmGKB, dkNET, and DataMed <i>Neil McKenna</i>