Final Weeks to Register

CLICK HERE TO
Register Online!

Organized by Cambridge Healthtech Institute

15th ANNIVERSARY 2002-2017

Bio·IT World
CONFERENCE & EXPO '17

MAY 23-25, 2017 | BOSTON, MA | SEAPORT WORLD TRADE CENTER

FINAL AGENDA

3,300+ Global Attendees **13** Conference Tracks
14 Workshops **3** Industry Award Programs **190** Exhibitors

Building a global network for precision medicine by uniting the Bio-IT community

Plenary Keynote Speakers

Rommie E. Amaro, Ph.D.
Associate Professor, Chemistry and Biochemistry; Director, National Biomedical Computation Resource, University of California, San Diego

Rainer Fuchs, Ph.D.
CIO, Harvard Medical School

Edison T. Liu, M.D.
President and CEO,
The Jackson Laboratory

William Mayo
CIO, Broad Institute

Andrea T. Norris
Director, Center for Information Technology (CIT) and CIO, NIH

Aarti Shah, Ph.D.
Senior Vice President and CIO, Eli Lilly and Company

PLATINUM SPONSORS

aspera an IBM company · CYCLE COMPUTING · DASSAULT SYSTEMES · DDN STORAGE
DELL EMC · ELSEVIER · FLOWJO, LLC · Google Cloud
HGST a Western Digital brand · IBM · illumina · intel · SevenBridges

2017 Conference Tracks

1 Data & Storage Management
2 Data Computing
3 Networking Hardware - NEW!
4 Software Applications & Services
5 Cloud Computing
6 Bioinformatics
7 Next-Gen Sequencing Informatics
8 Clinical Research & Translational Informatics
9 Data Visualization & Exploration Tools
10 Pharmaceutical R&D Informatics
11 Clinical Genomics
12 Cancer Informatics
13 Data Security

BACK-TO-BACK WITH BIO-IT WORLD
5th Annual
Medical Informatics World Conference
May 22-23, 2017 | Boston, MA RENAISSANCE WATERFRONT HOTEL
MedicalInformaticsWorld.com

Final Weeks to Register

CLICK HERE TO
Register Online!

Organized by Cambridge Healthtech Institute

15th ANNIVERSARY 2002-2017
Bio-IT World CONFERENCE & EXPO '17

Welcome

to the 15th anniversary of the first Bio-IT World Conference & Expo

Much has changed since Bio-IT World's first conference in 2002. In March 2002, Blue Gene, IBM's new supercomputing effort, was not yet completed. Eric Lander, our first plenary speaker, was still leading the international Human Genome Project (HGP). The first conference had only 35 sessions.

Fifteen years later the event spans 13 tracks, 14 pre-conference workshops, and three industry awards. We are breaking new ground in data generation, data application, and ways (and reasons!) to share our insights. We've undergone dramatic shifts in how we think about research, health, medicine, and discovery.

Of course much has stayed the same. In his 2002 keynote, Lander observed: "We indeed have the keys to this remarkable library" of genomic information, but "we still don't know how to use it fully." We've made amazing progress since then, but his words still ring true.

And still, the Bio-IT World Conference remains the community home to industry veterans and newcomers alike who gather in Boston each year for expert commentary, industry insights, and the chance to grab a beer with the people driving the bio-IT industry forward.

Thank you for being part of our community. Thank you for joining us to celebrate our 15th anniversary.

Allison Proffitt
Editorial Director, Bio-IT World

Cindy Crowninshield
Senior Conference Director/Team Lead
Bio-IT World Conference & Expo

NEW FEATURE FOR BIO-IT WORLD!

A HACKATHON:

Making Data More Open, Fair, Collaborative, and Useful

We all know that life science data – sequences, screening results, and clinical information – is being generated at an increasingly rapid pace.

Far too much of that data is hidden away in silos – without good access or metadata – preventing past and potential collaborators from making better use of that data.

The purpose of this hackathon is to explore ideas to improve search and retrieval of data, to enhance sharing and collaboration, and to facilitate better conversion of that information into learnings and knowledge. Facilitated sessions will take place during the Bio-IT World event, with sharing of some key findings and results to be presented before the end of the conference.

Final Weeks to Register

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

#BioIT17

SCHEDULE-AT-A-GLANCE

TUESDAY, MAY 23, 2017

8:00 am - 4:00 pm	Pre-Conference Workshops
4:00 - 5:00 pm	Plenary Keynote Session
5:00 - 7:00 pm	Exhibit Hall Open
5:00 - 7:00 pm	Welcome Reception in the Exhibit Hall with Poster Viewing

WEDNESDAY, MAY 24, 2017

8:00 - 9:50 am	Plenary Keynote Session CIO Panel
9:50 am - 6:30 pm	Exhibit Hall Open
9:50 - 10:50 am	Coffee Break in the Exhibit Hall with Poster Viewing
10:50 am - 12:30 pm	Tracks 1-13
12:40 - 1:40 pm	Luncheon Presentations *(Sponsorship Opportunities Available)*
1:50 - 3:25 pm	Tracks 1-13
3:25 - 4:00 pm	Refreshment Break in the Exhibit Hall with Poster Viewing
4:00 - 5:30 pm	Tracks 1-13
5:30 - 6:30 pm	Best of Show Awards Reception in the Exhibit Hall with Poster Viewing

THURSDAY, MAY 25, 2017

7:00 - 7:50 am	Breakfast Presentation
8:00 - 9:45 am	Plenary Keynote Session, Benjamin Franklin and Best Practices Awards Programs
9:45 am - 1:55 pm	Exhibit Hall Open
9:45 - 10:30 am	Coffee Break in the Exhibit Hall and Poster Competition Winners Announced
10:30 am - 12:10 pm	Tracks 1-13
12:20 - 1:20 pm	Luncheon Presentations *(Sponsorship Opportunities Available)*
1:20 - 1:55 pm	Dessert Refreshment Break in the Exhibit Hall with Poster Viewing
1:55 - 4:00 pm	Tracks 1-13

THE BRIDGE BETWEEN HEALTHCARE AND LIFE SCIENCES STARTS HERE!

5th Annual
Medical Informatics World Conference

Renaissance Waterfront Hotel · Boston, MA

May 22-23, 2017
MedicalInformaticsWorld.com

Come Early to Attend Medical Informatics World Conference 2017

Medical Informatics World kicks off the week of informatics programs, bringing together senior-level healthcare providers, payers, technology vendors, biomedical scientists, academic researchers, informaticists and national health organizations to discuss emerging trends and collaborations in health IT for improved outcomes in the healthcare ecosystem.

SPECIAL 20% Registration Discount Available

See Registration Page for Details

Final Weeks
to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

Data & Storage Management
Data Computing
Networking Hardware
Software Applications & Services
Cloud Computing
Bioinformatics
Next-Gen Sequencing Informatics
Clinical Research & Translational Informatics
Data Visualization & Exploration Tools
Pharmaceutical R&D Informatics
Clinical Genomics
Cancer Informatics
Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

2017 Plenary Keynote Presenters

TUESDAY I 4:00 – 5:00PM

Keynote Introduction:

Kay Eron, General Manager, Intel Healthcare & Life Sciences, Data Center Group, Intel

Sponsored by intel

ROMMIE E. AMARO, PH.D.

Associate Professor, Chemistry and Biochemistry; Director, National Biomedical Computation Resource, University of California, San Diego

In 2012, Dr. Rommie Amaro opened her lab at UCSD in the Department of Chemistry and Biochemistry. The Amaro Lab is broadly concerned with the development and application of state-of-the-art computational methods to address outstanding questions in drug discovery and molecular-level biophysics. Her lab focuses mainly on targeting neglected diseases, Chlamydia, influenza, and cancer, and works closely with experimental collaborators to catalyze the discovery of new potential therapeutic agents. The Amaro Lab is also keenly interested in developing new multiscale simulation methods and novel modeling paradigms that scale from the level of atoms to whole cells, and beyond. Dr. Amaro earned her bachelor's and doctorate degrees in chemical engineering and chemistry, respectively, from the University of Illinois at Urbana-Champaign.

WEDNESDAY I 8:00 – 9:50AM

Keynote Introduction:

Jason Stowe, CEO, Cycle Computing

Sponsored by CYCLE COMPUTING

Plenary Keynote Session CIO Panel

RAINER FUCHS, PH.D., *CIO, Harvard Medical School*

Before joining Harvard Medical School in 2012, Dr. Fuchs spent almost 20 years in the biopharmaceutical industry, including 12 years with Biogen Idec, where he held the position of Vice President of R&D Information Technology. He also held senior leadership positions in life science informatics at Aventis, Ariad and Glaxo Wellcome. Fuchs was trained in Germany and holds a doctorate in biochemistry from the University of Darmstadt and a master's degree in microbiology from the University of Frankfurt.

WILLIAM MAYO, *CIO, Broad Institute*

Bill Mayo's background spans Consumer Products, Business Services, Life Sciences, Consulting and Non-Profit sectors. He has broad international experience, a deep understanding of technology and a strong strategic orientation. He is currently the CIO for the Broad Institute of MIT and Harvard in Cambridge. Bill also serves as the chairman of the governing board of Ascentria Care Alliance and as a Steering Committee member and mentor for Northeastern University's Venture Mentoring Network.

ANDREA T. NORRIS, *Director, Center for Information Technology (CIT) and CIO, NIH*

Since 2012, Andrea T. Norris has served as the Director, Center for Information Technology and Chief Information Officer at the National Institutes of Health. Prior to joining the NIH, she held leadership positions with the National Science Foundation (NSF) and National Aeronautics and Space Administration. Andrea was the 2008 recipient of the President's Meritorious Service Award. Andrea has a M.B.A. with a major in Information Systems Management from the George Washington University and a B.A. in Economics from the College of William and Mary.

AARTI SHAH, PH.D., *Senior Vice President and CIO, Eli Lilly and Company*

Aarti joined Lilly in January 1994 as a Senior Statistician after completing her Ph.D. in Applied Statistics from the University of California, Riverside. She was later promoted to Research Scientist in 1999 and since then has held many different technical and administrative leadership positions within the company. She was recently appointed in her current role in July 2016. Aarti has 20+ years of experience in the pharmaceutical industry and has worked across various phases of drug development, led various cross-functional and virtual organizations through development of a vision, strategy and successful execution of those strategies.

THURSDAY I 8:00 – 9:45AM

8:00 Organizer Remarks:

Allison Proffitt, Editorial Director, Bio-IT World

8:05 Awards Program Introduction:

Qi Li, M.D., Director, Product Innovation, InterSystems

Sponsored by InterSystems Health | Business | Government

8:10 Benjamin Franklin Awards and Laureate Presentation

Rafael Irizarry, Ph.D., Professor of Applied Statistics, Harvard and Dana-Farber Cancer Institute

8:35 Bio-IT World Best Practices Awards

8:50 Keynote Introduction:

George Vacek, Global Director, Life Sciences, Field Sales, DDN Storage

Sponsored by DDN STORAGE

9:00 EDISON T. LIU, M.D., *President and CEO, The Jackson Laboratory*

Edison T. Liu, M.D. is the president and CEO of The Jackson Laboratory. Previously, he was the founding executive director of the Genome Institute of Singapore (2001-2011), and was the president of the Human Genome Organization (HUGO) from 2007-2013. Between 1997 and 2001, he was the scientific director of the National Cancer Institute's Division of Clinical Sciences in Bethesda, MD, where he was in charge of the intramural clinical translational science programs.

Final Weeks to Register

CLICK HERE TO Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

2017 SPONSORS

PLATINUM SPONSORS

aspera an IBM company · CYCLE COMPUTING · DASSAULT SYSTEMES · DDN STORAGE · DELL EMC · ELSEVIER · FLOWJO, LLC · Google Cloud · HGST a Western Digital brand · IBM · illumina · intel · SevenBridges

GOLD SPONSORS

Abbott · ACD/Labs · AVERE · Benchling · CAMBRIDGE SEMANTICS · ChemAxon · Clarivate Analytics · CONGENICA GENOME BASED MEDICINE · Copyright Clearance Center · CORE INFORMATICS part of Thermo Fisher Scientific · CRAY · databiology · Deloitte. · dimension data · dotmatics knowledge solutions · ERT · Genedata solutions in silico · GENEFORMICS · Genestack · Hewlett Packard Enterprise · HCL · IBM Cloud Object Storage · idbs · INNOPLEXUS · INTERNET 2 · KPMG · Lab7 · LabAnswer LABORATORY INFORMATICS · Linguamatics · MEDABLE · Microsoft · OpenEye SCIENTIFIC · panasas · paradigm4 · PerkinElmer For the Better · PetaGene smaller, faster genomics data · PURESTORAGE Hello Possible · QIAGEN · Quantum · Qumulo · SCHRÖDINGER · SEAGATE · Signet Accel True interoperability realized in healthcare. · SILICON MECHANICS SERVERS • STORAGE • HPC · SINEQUA REAL-TIME BIG DATA SEARCH & ANALYTICS · SwiftStack · twigkit · WuXiNextCODE Leading Genomics

BRONZE SPONSORS

epam · EXOSTAR · the hyve · InterSystems Health | Business | Governme · LIAISON · Maven Wave · NODEUM SCALABLE DATA STORAGE MADE EASY · Roche · Tessella · WEKA.IO

OFFICIAL MEDIA PARTNER

Bio-IT World

Final Weeks
to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

Cambridge Healthtech Institute and Bio-IT World

2017 Awards Programs

Recognizing and celebrating leaders in innovation

Cambridge Healthtech Institute and Bio-IT World will again be recognizing and celebrating leaders in innovation through the following Awards Programs:

Bio-IT World Best Practices Awards

PROGRAM INTRODUCTION:

Qi Li, M.D., Director, Product Innovation, InterSystems

Bio-IT World has held the Best Practices awards since 2003, highlighting outstanding examples of technology innovation in the life sciences, from basic R&D to translational medicine. We particularly encourage vendors to nominate entries from valued academic and/or industry partners. Winners will be announced in a plenary session at Bio-IT World Conference & Expo, Boston, Thursday, May 25. Deadline for entry is February 3, 2017. Full details including previous winners and entry forms are available at Bio-ITWorld.com/BestPractices

Bio-IT World Best of Show Awards

PROGRAM INTRODUCTION:

Barbara Murphy, Vice President, Marketing, WekaIO

The Best of Show Awards offers exhibitors of the Bio-IT World Conference and Expo an exclusive opportunity to distinguish and highlight their esteemed products ranging from an innovative application, technology, tool, or solution from the competition. Judged by a team of leading industry experts and Bio-IT World editors, this awards program identifies exceptional innovation in technologies used by life science professionals today. Products considered are new products, or significant product upgrades, introduced between April 2016 and May 2017. Winners are judged based on the products' technical merit, functionality, innovation, and in-person presentations to the judges at the show. Judging and the announcement of winners is conducted live in the Exhibit Hall. Winners will be announced on Wednesday, May 24. To learn more about this program, contact Ryan Kirrane at 781-972-1354 or email rkirrane@healthtech.com. Please Note: Selection is not based upon level of sponsorship or exhibit participation.

Benjamin Franklin Award

PROGRAM INTRODUCTION:

Qi Li, M.D., Director, Product Innovation, InterSystems

The Benjamin Franklin Award for Open Access in the Life Sciences is a humanitarian/bioethics award presented annually by the Bioinformatics Organization to an individual who has, in his or her practice, promoted free and open access to the materials and methods used in the life sciences. Nominations are now being accepted! The winner will be announced during the Bio-IT World Plenary Keynote and Awards Program on Thursday, May 25. Full details including previous laureates and entry forms are available at www.bioinformatics.org/franklin.

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

25 YEARS Organized by Cambridge Healthtech Institute

2017 Pre-Conference Workshops*

MORNING WORKSHOPS
Tuesday, May 23, 2017 | 8:00 - 11:30 am

W1. Data Management for Biologics: Registration and Beyond
Monica Wang, Ph.D., Principal Bioinformatics Architect, Takeda Oncology
Beth Basham, Ph.D., IT Director, Client Services, Biologics & Vaccines Discovery, Merck
Rudolf Kinder, Senior Scientist, Roche Innovation Center Penzberg
Sean Ward, CTO, Synthace
Timothy S. Gardner, Ph.D., Founder/CEO, Riffyn, Inc.
Peter Henstock, Ph.D., Senior Principal Scientist, Research Business Technologies, Pfizer

W2. An Intro to Blockchain in Life Sciences
Adrian Gropper, M.D., CTO, Patient Privacy Rights
Richard Shute, Ph.D., Consultant, Professional Services, Curlew Research
Maurizio Viviani, Ph.D., Lead Developer, IT, Encrypgen

W4. Data Visualization to Accelerate Biological Discovery
Nils Gehlenborg, Ph.D., Assistant Professor, Department of Biomedical Informatics, Harvard Medical School

W7. Introduction to Hadoop for Bioinformatics
Martin R. Gollery, CEO, Tahoe Informatics

W8: Digital Transformation for Life Sciences – Adopting Hybrid IT
Sponsored by dimension data
Rudy Gopaul, Senior Director, Business Development, NTT Innovation Institute
Casey Overman, Data Architecture and Analytics, Sequence Inc.
Kimberly Robasky, PhD, Senior Translational Scientist, Renaissance Computing Institute
Harsha Rajasimha, Ph.D., Global Director, Life Sciences R&D, NTT Innovation Institute
Peter Lutz, Global Lead Enterprise Architect – Pharmaceutical and Life Sciences, Dimension Data

AFTERNOON WORKSHOPS
Tuesday, May 23, 2017 | 12:30 - 4:00 pm

W9. Data Science Driving Better Informed Decisions
Wei-Yi Cheng, Ph.D., Data Scientist, Roche Innovation Center New York
Ahmed Abdeen Hamed, Ph.D., Scientific Knowledge Engineer, Scientific Information Management: Data Platform, Merck & Co.
Francesca Milletti, Ph.D., Senior Principal Scientist, Data Science, Roche Innovation Center New York
Venus So, Ph.D., Data Science, Pharma Research and Early Development Informatics, Roche Innovation Center New York

W10. Designing Storage Solutions for Life Sciences
Ari E. Berman, Ph.D., Vice President and General Manager of Consulting Services, BioTeam, Inc.
Aaron Gardner, Senior Scientific Consultant, BioTeam, Inc.

W11. Scientific Project Management
Gregg TeHennepe, Program Manager, Computational Scientist, The Jackson Laboratory

W12. Leveraging Cloud Technologies to Enable Large-Scale Integration of Human Genome and Clinical Outcomes Data
R. Mark Adams, Ph.D., CIO, Celmatix Inc.
Benjamin Breton, Biology, Senior Bioinformatics Engineer
Anna T. Fernandez, Ph.D., PMP, Senior Associate, Civil Health – Informatics, Booz Allen Hamilton
Andrew FigPope, Data Engineer, Celmatix Inc.
Michael Keller, Ph.D., Senior Associate, Civil Health, Booz Allen Hamilton
Christopher Matta, Systems Engineer, Confluent, Inc.

W13. Proteogenomics: Integration of Genomics and Proteomics Data
David Fenyő, Ph.D., Professor, Biochemistry and Molecular Pharmacology, Institute for Systems Genetics, NYU School of Medicine

W15. Virtual Pharma: Evolving toward True Data Exchange Highways beyond Corporate Firewalls
Ralph Haffner, Head Discovery Workflows Basel - pRED Informatics, Roche Pharma Research and Early Development, F. Hoffmann-La Roche Ltd
Michael Fowler, US Innovative Medicine and Early Development Lead, R&D Information, AstraZeneca
Dana E. Vanderwall, Director, Biology & Preclinical IT, Bristol-Myers Squibb
Michael H. Elliott, CEO, Atrium Research & Consulting LLC
Padmanabha Udupa, pRED Informatics, Roche Pharma Research and Early Development, F. Hoffmann-La Roche Ltd.
Laura Aguiar, Ph.D., Head, pREDi NY Operations, Data Science Decision Support Service Manager, Roche Innovation Center New York
Chris Morrow, Associate Director, Development Sciences IT Portfolio Strategy and Solutions, Genentech

** Separate registration required*

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

Organized by Cambridge Healthtech Institute

Data & Storage Management

INFRASTRUCTURE AND STORAGE SOLUTIONS TO ENABLE DISCOVERY AND EASE DATA BLOAT

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W3) Delivering Custom Mobile App Projects to the Cloud
(W7) Introduction to Hadoop for Bioinformatics

12:30 – 4:00 pm Afternoon Pre-Conference Workshops*
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing

Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing

Sponsored by InterSystems

DESIGNING DATA & STORAGE MANAGEMENT SOLUTIONS

10:50 Chairperson's Remarks
Chris Dwan, Senior Technologist and Independent Consultant

11:00 IT Design Patterns to Support Genomic Science in the Age of the Cloud: Challenges and Possibilities
Chris Dwan, Senior Technologist and Independent Consultant

11:30 Design Considerations for Genomic Data Archival
Saira Kazmi, Ph.D., Scientific Data Architect, Research Information Technology, The Jackson Laboratory
This talk focuses on issues related to managing and tiering storage for genomic sequence data. Architectural considerations for designing a solution for scalability, governance, and discoverability will be presented. The presentation will discuss some of the current hardware and software technologies and a solution using metadata indexing will be presented. The presentation will conclude with lessons learned and next steps.

12:00 pm Life Sciences at EXAScale: Applying a Novel IO System to Critical Workflows

Sponsored by DDN STORAGE

James Coomer, Technical Director, UK-Pre-Sales, Engineering, DDN Storage
Today, we generally assume that an IO interface and the filesystem choice are related and indeed this is usually true. We present an IOSystem based upon Flash which embeds within a filesystem, replacing the IO interface with one that removes common constraints – particularly for IO patterns seen in Life Sciences. DDN's Infinite Memory Engine radically changes the way IO is handled, providing new opportunities for complex life science workflows.

12:15 Novel Systems and Approaches for the Next Generation of Genomic Analysis and Data Management

Sponsored by Hewlett Packard Enterprise

Christopher Davidson, Life Science Solutions Manager, HPE
The pace and scale of genomics research is now less defined by the science itself than by the compute and storage architectures used to extract insight from the genomic data generated. This session focuses on how to enable genomic workflow acceleration and the democratization of data through flexible and scalable systems.

12:30 Session Break

12:40 Luncheon Presentation I: Broad Institute & Intel GATK 4.0 Optimization Overview

Sponsored by intel

Eric Banks, Director, Data Science and Data Engineering Group, Broad Institute
Geraldine Van der Auwera, Associate Director, Outreach and Communications, GATK, Broad Institute
Mark Bagley, Director, Center for Genomic Data Engineering, Intel
Paolo Narvaez, Senior Director, Engineering, Intel
Genomics research leader the Broad Institute of MIT and Harvard joins Intel to describe their collaboration to enhance the GATK environment and scale researchers' ability to analyze massive amounts of genomic data from diverse sources worldwide. Topics include performance best practices and the latest on Genomics DB and FireCloud.

1:10 Luncheon Presentation II: The New Era of Integrated Data Storage

Sponsored by Quantum

Mark Pastor, Director, Archive & Technical Workflow Solutions, Marketing, Quantum Corporation
Clinicians and researchers need high-performance and easy access to data — even as data repositories reach petabyte scale levels. Designing a future-proof, cost effective storage infrastructure to deliver performance access and protection for all your data, and can act as a gateway to all types of storage media including public cloud is now easier than ever before.

1:40 Session Break

1:50 Chairperson's Remarks
Asya Shklyar, Senior Scientific Consultant Infrastructure, BioTeam

1:55 Tools and Techniques for Making Data Less Scary and More Visible
Asya Shklyar, Senior Scientific Consultant Infrastructure, BioTeam
Data management has been and is an ongoing and rapidly escalating problem in the research and commercial world. The talk aims to summarize the tools

Data & Storage Management

INFRASTRUCTURE AND STORAGE SOLUTIONS TO ENABLE DISCOVERY AND EASE DATA BLOAT

that can be leveraged to wrangle the data across multiple heterogenous sources and make it more quantifiable, searchable, parsable, and actionable, including software and hardware, open source, and commercial.

2:25 Reducing the Size and Cost of NGS Data Storage and Transfer

Sponsored by PetaGene

*Dan Greenfield, Ph.D., CEO, PetaGene Ltd.**

PetaGene provides a software suite which allows up to 5x compression of NGS data and better utilizes your compute nodes. PetaGene's file system interface allows existing tools and pipelines to be used without modification, and this is provided as a free download so everyone everywhere can use the compressed files.

** Bio-IT World 2016 Best of Show winner*

2:55 Pushing the Limits of Discovery with Internet2

Sponsored by Internet2

Dan Taylor, Director, Business Development, Network Services, Internet2

3:10 Time to Results Matters: The Case for Performance Scale-Out NAS

Sponsored by Panasas

David Sallak, Vice President, Product Management & Industry Marketing, Panasas, Inc.

Acceleration. Decoding human genomes has gone from decades to mere days and hours. Your research infrastructure must keep ahead of the demands placed on it by the best scientists. Building the right storage foundation positions your organization for groundbreaking research insights. Learn how the Panasas accelerated scale-out NAS solution helped the Garvan Institute drive innovation faster and simplified their workflows.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

4:00 Evaluating Full Flash Scale-Out NAS Technologies for Some Bioinformatics Workloads

Youssef Ghorbal, Design and Technical Solutions Group Manager, Institut Pasteur

Scale-out NAS technology is an effective backend for bioinformatics workloads at Institut Pasteur. In our presentation, we will go through the shortcomings of the current setup for some identified use cases. We will also be giving feedback on how newly tested flash array technologies may overcome those limitations.

4:30 Integrating Data, Tools and Infrastructure to Enable Efficient Collaboration and Management in Large Scale Biomedical Studies

Sven Nahnsen, Ph.D., Head, Quantitative Biology Center (QBiC), University of Tübingen

We established an infrastructure that builds on multi-layer omics (genomics, transcriptomics and metabolomics), as well as imaging data from mice and from human material that is gained from clinical oncology studies. Furthermore, we developed a data and project management facility that facilitates the modeling of the experimental design, the interplay with the data acquisition facilities and the bioinformatics analysis.

5:00 Extreme Durability for Your Bioinformatics Data

Sponsored by HGST a Western Digital brand

Kent Ritchie, Solutions Architect, HGST, a Western Digital brand

HGST, a Western Digital brand, one of the largest storage companies in the world, can help with every stage of your workflow. Come visit us to hear about extreme durability for your bioinformatics data, focusing on high performance computing to archiving and analytics for your long-term discoveries. We are here to help you deliver possibilities at every stage.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM

Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

HADOOP

10:30 Chairperson's Remarks

Martin R. Gollery, CEO, Tahoe Informatics

10:40 Real World Data Platform and Analytics

Minnie Chou, Director, Information Systems, Amgen

The Real World Data (RWD) Platform is a game changer in Amgen's pursuit of serving patients by delivering innovative human therapeutic products faster. It provides a common high performance analytics ecosystem hosting large volumes of real world patient claims data and electronic medical records, enabling epidemiologists, analysts and scientists to deliver insights in a timely and cost effective manner. We used Agile approach to implement an enterprise RWD workbench on top of Hadoop based enterprise data lake to harmonize real world patient data assets, patient cohorts with diseases and/or receiving Amgen/competitor therapies to consistently address questions across the drug commercialization lifecycle.

11:10 Healthcare: Foundational Building Blocks: The Establishment of a Healthcare Data Ecosystem in a Hadoop Environment

Amy M. Andrade, MS, PMP, Assistant Vice President of Research, Meharry Medical College

Charles Boicey, MS, RN-BC, Associate Clinical Professor, Stony Brook Medicine

This talks presents insight and a working framework of how data and storage management, and clinical informatics in a Hadoop environment is plausible in months instead of years. A Data Science Center at a community-based academic health center focused on serving the underserved and minority populations has implemented a low cost HIPAA compliant cloud approach to

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

1 Data & Storage Management

INFRASTRUCTURE AND STORAGE SOLUTIONS TO ENABLE DISCOVERY AND EASE DATA BLOAT

"Big Data". Utilizing technologies new to healthcare, data from both within and outside of the healthcare environment was processed.

11:40 Kickstarting Breakthroughs in Life Sciences with Intelligent, Next-Generation Scale-Out Storage

Sponsored by Qumulo

Peter Godman, CTO & Co-Founder, Qumulo

Unprecedented storage and data management challenges resulting from advances in genomic IT are plaguing life sciences companies. How can companies stay competitive and handle the challenge of managing billions of small and large files? Discover how intelligent scale-out storage systems are providing enterprises with real-time answers into their data footprints at scale, providing breakthrough performance while balancing capacity and cost.

11:55 IBM Cloud Object Storage Solutions Enabling Better Patient Outcomes

Sponsored by IBM Cloud Object Storage

Piers Nash, Ph.D., Global Solutions Consultant, Genomics & Healthcare, IBM Cloud Object Storage, IBM

IBM and University of Chicago's Center for Data Intensive Science (CDIS) are accelerating medical discoveries. Utilizing IBM Cloud Object Storage, CDIS centrally stores and manages vast amounts of genomic and clinical data at web-scale. Discover how IBM Watson for Genomics can help researchers to collaborate via shared access to harmonized data sets, speeding discovery and enabling precision medicine.

12:10 pm Session Break

12:20 Luncheon Presentation: AI + RWD – Next Steps in your Big Data Journey

Sponsored by KPMG

Arun Ghosh, Principal Data & Analytics, KPMG LLP

Organizations' data repositories have evolved to accommodate real-world data. This moves the industry incrementally closer to the automated data management research scientists and epidemiologists need to accelerate improved patient outcomes. This presentation will illuminate how machine learning (as a service) and artificial intelligence can use data to accelerate value in small measures today while laying the foundation for what's next.

12:50 Luncheon Presentation: Optimized Scaling for NGS: Transfer, Storage and Archiving

Sponsored by GENEFORMICS

Rafael Feitelberg, CEO, Geneformics Inc.

As the volume of NGS data is increasing, so are the challenges of IT costs and infrastructure. In this session, we will cover solutions implemented by leading global organizations to reduce NGS footprint by up to 90%, with scalable Enterprise-Grade architectures that are lossless and transparent to bioinformatics applications.

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

FEATURED SESSION: BIOTEAM MICRO-SYMPOSIUM: 2017 BIO/IT TRENDS

1:55 Chairperson's Remarks

Chris Dwan, Senior Technologist and Independent Consultant

2:00 BioTeam Micro-Symposium: 2017 Bio/IT Trends

Chris Dwan, Senior Technologist and Independent Consultant (Moderator)

Ari E. Berman, Ph.D., Vice President and General Manager of Consulting Services, BioTeam, Inc.

Chris Dagdigian, Founding Partner & Director, Technology, BioTeam, Inc.

Aaron Gardner, Senior Scientific Consultant, BioTeam, Inc.

Adam Kraut, Director of Infrastructure and Cloud Architecture, BioTeam, Inc..

Asya Shklyar, Senior Scientific Consultant, Infrastructure, BioTeam, Inc.

Since 2010, the "Trends in the Trenches" presentation, given by Chris Dagdigian, has been one of the most popular annual traditions on the Bio-IT Program. The intent of the talk was to deliver a candid (and occasionally blunt) assessment of the best, the worthwhile, and the most overhyped information technologies (IT) for life sciences. The presentation tried to recap the prior year by discussing what has changed (or not) around infrastructure, storage, computing, and networks. This presentation has helped scientists, leadership, and IT professionals understand the basic topics involved in supporting data intensive science. For 2017, the "Trends in the Trenches" presentation will evolve and expand from 60-minutes to 120-minutes and feature more content, speakers, and interactive discussion. Short focused podium talks on current trends related to computing, storage/data transfer, networks, and cloud will be presented. A Q&A moderated discussion follows. Come prepared with your questions and commentary for this informative and lively session.

4:00 Conference Adjourns

CHI's INTRO-NET Networking at its Best

Maximize Your Experience Onsite

AT THE BIO-IT WORLD CONFERENCE & EXPO!

The Intro-Net offers you the opportunity to set up meetings with selected attendees before, during and after this conference, allowing you to connect with the key people that you want to meet. This online system was designed with your privacy in mind and is only available to registered session attendees of this event. Registered conference attendees will receive more information on how to access the Intro-Net in the weeks leading up to the event!

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

Organized by Cambridge Healthtech Institute

2 Data Computing

ADVANCES IN COMPUTING APPLICATIONS FOR BIG DATA

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W1) Data Management for Biologics: Registration and Beyond

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W11) Scientific Project Management
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing
Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing
Sponsored by InterSystems

NOVEL COMPUTING TECHNIQUES FOR BIG DATA PROCESSING

10:50 Chairperson's Remarks
Thomas Manganaro, District Manager, Pure Storage

11:00 Making Sense of Big Data
Chen Shen, Researcher, Computer Science, The George Washington University
Big data is a collection of large datasets that cannot be processed using traditional computing techniques. We suggest a different Big Data System that allows generating novel scientific hypotheses. The system is based on a novel method of clustering using Golay code transformations that are applied to diverse Boolean queries.

11:30 Using Ticket Tracking Software and Agile Techniques to Accelerate Science and Operations at The Broad
Bruce Kozuma, Principle System Analyst, Broad Information Technology Services, The Broad Institute of MIT and Harvard
Sadiya Akasha, Project Manager, Broad Information Technology Services, The Broad Institute of MIT and Harvard
The Broad has been using JIRA (often used for IT ticket tracking) and Agile techniques to speed science and streamline operations, e.g., handle the scale of genomic sample processing, visualize sample processing workflows, and track operational tasks such as data center moves. During this talk, we'll discuss the specific techniques and features used in those examples, such as JIRA Board of Boards.

12:00 pm Distributed Analytics at World Wide Scale and the Outbreak of Infectious Diseases
Sponsored by DELL EMC
Patricia Florissi, Ph.D., Global CTO, Dell EMC
In collaboration with Ben Gurion University, Dell EMC has prototyped a next-generation outbreak surveillance system based on metagenomic sequencing. Dell EMC uses bi-clustering to uncover common patterns of virulence factors among subsets of micro-biomes, unleashing the potential to not only identify early onset of outbreaks but to also uncover new combinations of virulence factors that may characterize new diseases.

12:15 How a Hosted Supercomputer Can Put a Little Spark in Your Variant Analysis
Sponsored by CRAY
Ted Slater, Global Head, Healthcare & Life Sciences, Cray, Inc.
NGS analysis workflows are big, and growing bigger every day as sequencing machines proliferate and analytics pipelines mature. Conventional compute architectures can be "good enough," but what are the hidden costs of waiting for results? Using a hosted supercomputer together with tools like Apache Spark™ will help you leverage data more effectively, recover those costs, and gain insight faster.

12:30 Session Break

12:40 Luncheon Presentation: Security & Regulatory Compliance in Google's Cloud
Sponsored by Google Cloud
Ben Lavallee, Customer Engineering Specialist, Team Manager, Google Cloud
Come learn how Google Cloud's innovations in Security, Networking and Machine Learning are helping Bio-IT Organizations move to the cloud and access new capabilities.

1:40 Session Break

DATA COMPUTING, AUTOMATION, INTEGRATION & WORKFLOWS: INNOVATIVE MODELS, CAPABILITIES, AND APPLICATIONS

1:50 Chairperson's Remarks
Brian Bissett, Senior Member, IEEE, Institute of Electrical and Electronic Engineers

1:55 Automation Proves Hard Work Pays Off Later More Than Laziness Pays Off Now

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

25 YEARS Organized by Cambridge Healthtech Institute

Brian Bissett, Senior Member, IEEE, Institute of Electrical and Electronic Engineers
Life is a zero sum game. Two places cannot utilize the same resources at the same time. The most valuable resource of any organization is its intellectual capital. When the best and brightest are occupied doing mundane tasks, an organization is unlikely to reach its full potential. The focal point of this presentation is how to leverage the right tools, techniques, and methods to automate the mundane tasks in an organization, so that its thinkers can be engaged in something that is worth thinking about.

2:25 Data Platform for a Community Ecosystem of Contextual Biological Information
Austin Huang, Ph.D., Associate Director, Biomedical Data Science Lead, Enterprise Science & Technology Operations, Pfizer
A key impediment to reproducible computational workflows in research is the treatment of dependencies on persistent data. We have implemented a data platform that can achieve the benefits of a more principled handling of data persistence with minimal analyst overhead. This is achieved by automating schema inference, metadata curation, versioning, and RESTful service production to reduce the engineering and administrative capacity typically required for data repositories.

2:55 Scalable Data Computing for Healthcare & Life Sciences Industry
Sponsored by IBM
Prashant Avashia, Senior Architect, Storage & Software Defined Solutions, IBM Systems, IBM
Fully leveraging all data, structured and unstructured, can enable more patient-centric care to help organizations achieve better outcomes. IBM software defined solutions provide a proven roadmap to transform fragmented data silos into a common, universally accessible platform for genomics, imaging and analytics. Clinicians can then apply IBM Watson technology to obtain cognitive insights from this platform for evidence-based decisions.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

DATA COMPUTING, AUTOMATION, INTEGRATION & WORKFLOWS: INNOVATIVE MODELS, CAPABILITIES, AND APPLICATIONS (CONT.)

4:00 Integrating Biomedical Devices into Information Systems
Jim McGinnis, Assistant Professor, Engineering Technology, University of Memphis
This presentation shares case studies that help formulate a framework for the integration of biomedical devices into information systems with an emphasis on information quality. The framework outlines the importance of data quality in projects to define a clear pathway integrating disparate systems together for better patient support.

4:30 AIDEAS 3.0: The New Generation Cheminformatics Platform
Rishi Gupta, Senior Research Scientist, Platform Informatics and Knowledge Management, AbbVie, Inc.
Vincent Le Guilloux, Platform Informatics and Knowledge Management, AbbVie, Inc.
AIDEAS is an integrated Cheminformatics solution that has brought together several scientific applications and methods under a single umbrella. This presentation will discuss the technology on which AIDEAS 3.0 is built and the scientific application built within AIDEAS 3.0. Examples will be presented to showcase the ability of AIDEAS to provide advanced scientific workflows within one of the best visual analytics frameworks that allows users to share information and define multiple analysis templates in a facile way.

5:00 Engineering for Insight, Building the Optimal Mix RWE Capabilities
Sponsored by DELL EMC
Michael Madden, Client Solutions Director, Dell EMC
At Dell EMC, we assist healthcare and life science organizations to find insights by connecting streams of data making it more viable for the entire development chain: from discovery to commercialization. We will highlight hybrid cloud capabilities that can lead to new insight, foresight and predictive findings on diseases, products, and patient populations.

5:15 Sponsored Presentation *(Opportunity Available)*

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM
Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

DATA COMMONS: DIGITAL ECOSYSTEMS FOR USING AND SHARING BIOMEDICAL DATA AT SCALE

10:30 Chairperson's Remarks

10:40 PANEL DISCUSSION: Data Commons
Matthew Trunnell, Vice President and CIO, Fred Hutchinson Cancer Center
Vivien Bonazzi, Ph.D., Senior Advisor for Data Science Technologies, National Institutes of Health (NIH)

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

2 Data Computing

ADVANCES IN COMPUTING APPLICATIONS FOR BIG DATA

Brian D. O'Connor, Ph.D., Technical Director, Genomics Institute – Computational Genomics Platform, University of California, Santa Cruz

Additional Speakers to be Announced

The Data Commons is an open science platform that allows producers and consumers of scientific data to connect, interact, exchange, create value and generate new discoveries, creating the basis for a digital ecosystem that can support scientific discovery in the era of biomedical big data. The computing platform is flexible and scalable, allowing scientists and researchers to transparently find and use services and tools they need, access large public data sets, and connect with other resources associated with scholarly research. During this 60-minute panel discussion, short focused podium talks will be presented on how this system has been adapting to the different evolving needs of research communities and technology innovations. A Q&A moderated discussion follows.

11:40 Storage Trends for Healthcare IT

Sponsored by PURESTORAGE

Vik Nagjee, Vice President & CTO, Healthcare and Life Sciences, Pure Storage

Healthcare organizations are requiring more of IT than ever before. IT is now a strategic element in the organization, tasked with facilitating innovation under new business models and driving improvements to the patient experience. In this session, Vik Nagjee will discuss current trends in healthcare IT, how modern IT systems impact outcomes, and storage trends for healthcare and biological research.

11:55 Sponsored Presentation *(Opportunity Available)*

12:10 pm Session Break

12:20 Luncheon Presentation *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

FEATURED SESSION: BIOTEAM MICRO-SYMPOSIUM: 2017 BIO-IT TRENDS

1:55 Chairperson's Remarks

Chris Dwan

2:00 BioTeam Micro-Symposium: 2017 Bio-IT Trends

Chris Dwan, Senior Technologist and Independent Consultant (Moderator)

Ari E. Berman, Ph.D., Vice President and General Manager of Consulting Services, BioTeam, Inc.

Chris Dagdigian, Founding Partner & Director, Technology, BioTeam, Inc.

Aaron Gardner, Senior Scientific Consultant, BioTeam, Inc.

Adam Kraut, Director of Infrastructure and Cloud Architecture, BioTeam, Inc..

Asya Shklyar, Senior Scientific Consultant, Infrastructure, BioTeam, Inc.

Since 2010, the "Trends in the Trenches" presentation, given by Chris Dagdigian, has been one of the most popular annual traditions on the Bio-IT Program. The intent of the talk was to deliver a candid (and occasionally blunt) assessment of the best, the worthwhile, and the most overhyped information technologies (IT) for life sciences. The presentation tried to recap the prior year by discussing what has changed (or not) around infrastructure, storage, computing, and networks. This presentation has helped scientists, leadership, and IT professionals understand the basic topics involved in supporting data intensive science. For 2017, the "Trends in the Trenches" presentation will evolve and expand from 60-minutes to 120-minutes and feature more content, speakers, and interactive discussion. Short focused podium talks on current trends related to computing, storage/data transfer, networks, and cloud will be presented. A Q&A moderated discussion follows. Come prepared with your questions and commentary for this informative and lively session.

4:00 Conference Adjourns

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

3 Networking Hardware - NEW!

EQUIPMENT, DEVICES, AND INTEGRATION TECHNOLOGIES REQUIRED FOR MEDIATING DATA

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W2) An Intro to Blockchain in Life Sciences

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W10) Designing Storage Solutions for Life Sciences
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing
Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing
Sponsored by InterSystems

NETWORKING & DATA SECURITY

10:50 Chairperson's Remarks
Tom Johnson, Senior Director, Healthcare and Life Sciences Solutions, Exostar

11:00 Science DMZ and Security around Moving Data between Sites
Eli Dart, Network Engineer on Science DMZ, ESNET
The Science DMZ model describes a performance-oriented approach to network architecture, network design, network security, and performance engineering. It has been the basis of many successful IT infrastructure projects which support data-intensive science. This talk will describe the
Science DMZ model, with an emphasis on security aspects.

11:30 End-To-End Asymmetric Encryption of Biomedical Data In-Transit and At-Rest
Ryan Harrison, Ph.D., Head of Engineering, BioBright
We present a workflow for the end-to-end asymmetric encryption of biomedical data, that goes well beyond the typical (not encrypted at all) and conscientious (HTTPS/SSL in-transit, AES-256 at-rest) laboratory use-cases. The pipeline, which is commercially available to selected customers, allows the option of customizable metadata extraction, allowing efficient lab-workflow-related search to coexist with at-rest encryption.

12:00 pm Networking and Data Transfer in the Modern Life Sciences and Healthcare Era
Ari E. Berman, Ph.D., Vice President and General Manager of Consulting Services, BioTeam, Inc.
Data generation throughout the life sciences research and healthcare domains has risen at a rate far beyond that predicted by Moore's Law. As a result, organizations are accumulating 10's to 100's of petabytes (PB) of data, spending millions on storage systems, and doing it all in a manner consistent with out of date IT practices and policies. These practices include little to no data management, ineffective or non-existent data lifecycle policies, no metadata standards, and a dependence on network infrastructure and services that are unrealistic and unsustainable. In this presentation, we will discuss the general scope of the data generation landscape in Life Sciences as well as review generalized networking and security solutions that have successfully supported research missions throughout the industry and enable the of movement and sharing large amounts of data effectively and sustainably.

12:30 Session Break

12:40 Luncheon Presentation *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:40 Session Break

NETWORKING, DEVICES, AND DATA TRANSFER

1:50 Chairperson's Remarks
Sanjay Joshi, CTO, Healthcare and Life Sciences, Dell EMC Emerging Technologies

1:55 Connected Health: A Collaboration and Integration Story
David Delaney, M.D., Chief Medical Officer, Healthcare Sector, SAP
Sanjay Joshi, CTO, Healthcare and Life Sciences, Dell EMC Emerging Technologies
Across the globe, the value of unlocking data is top of mind and progressing rapidly. However, for a variety of reasons, most health data continues to reside in silos. Creating a truly connected health ecosystem, with the patient's interest and outcome at the center, will ensure that the bridge from Research Insights to Clinical Utility will get shorter. One side of the fence is historical data, registries and image archives while the other side includes clinical databases in genomics, proteomics, biomarkers and clinical decision support. Handling this content relies on the entire gamut of infrastructure: HPC; Data Security and Governance, In-Memory Analytics, Apps and Databases; Near-Line distributed archives with fast networking; and WAN connectivity to Hybrid and Public Clouds. SAP and Dell EMC will present real use-cases that include the Hadoop and Analytics ecosystem with specialized databases, and user access in a collaborative ecosystem.

2:25 PANEL DISCUSSION: The Medical Device and IoT Ecosystem: Data Mediation and Understanding beyond Networks and Networked Hardware

3 Networking Hardware - NEW!

EQUIPMENT, DEVICES, AND INTEGRATION TECHNOLOGIES REQUIRED FOR MEDIATING DATA

Moderator:
Sanjay Joshi, CTO, Healthcare and Life Sciences, Dell EMC Emerging Technologies
Panelists:
Manoj Dadlani, MEng, CEO, CosmosID
Todd Gray, CEO, autonomous_ID
Dinesh Shah, MD, FACC, FSCAI, Cardiologist; CEO and Medical Director, Michigan Physicians Group; Co-Founder and Medical Director, Real 3D Polymer
Neel Nabar, M.D., Ph.D., Karolinska Fellow, Fauci Lab, National Institute of Allergy and Infectious Diseases (NIAID)
Stanley Yang, CEO, NeuroSky, Inc.
Just beyond the current buzz in genomics is the "continuous integration" challenge of various devices that link to the phenotype. From the microbiome analytics to metabolomics sensors to gait and posture based ID to brain and muscle waves to immunology, we have assembled a stellar panel to discuss the real-world issues of data wrangling, cleansing and use.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

GLOBAL COLLABORATIONS: CONNECTING CANCER TREATMENT AND RESEARCH CENTERS

4:00 Computational Approaches to Cancer: Cooperation between India and the United States
Kenneth Buetow, Ph.D., Director, Computational Sciences and Informatics, Complex Adaptive Systems Initiative (CASI), Arizona State University
Timothy Lance, Distinguished Service Professor Emeritus, University at Albany; President Emeritus and Chief Research Officer, NYSERNet
Amit Saxena, Senior Technical Officer, Bioinformatics Group, C-DAC
Anil Srivastava, President, Open Health Systems Laboratory (OHSL)
IUCKA: Indo-US Cancer Knowledge Alliance is being designed as an integrated biomedical informatics cyberinfrastructure for cancer treatment and research in India. It will be a true translational research platform from bench to bedside connecting cancer treatment and research centers across the country with access and connection to global centers of research, especially in the United States. IUCKA is being implemented as a PPP (public private partnership) and is bringing together technology products and service providers and cancer treatment and research centers in an ecosystem to directly benefit cancer patients in India and contribute to global research collaboration, especially between cancer centers in India. ICTBioMed is a group of life sciences supercomputing centers brought together by OSHL. ICTBioMed members have been working together for almost four years to create a shared global cyberinfrastructure as a seamless and friction-free platform for the researchers worldwide for their collaborative research in consistent with the tenets of team science. The backbone research and education network in India and the United States are now connected by a direct 5+5 gig optical fiber links between Mumbai and New York making it possible for cooperation in biomedical research leveraging computational biology in a big way. OHSL and its IUCKA and ICTBioMed initiatives are playing a big role in this area. This panel session will discuss the biomedical applications and hardware engineering components, as well as the status, plans and prospects for US-India collaboration.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM
Please see pages 3 & 4 for details.
8:05 Benjamin Franklin Awards and Laureate Presentation
8:35 Best Practices Awards Program
8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

SUPERCOMPUTING INDUSTRY PERSPECTIVES: CURRENT APPROACHES, CHALLENGES AND FUTURE OUTLOOKS

10:30 Chairperson's Remarks
Arvindh Lalam, Founder and CEO, Fusion Memory

10:40 Modern View of Supercomputing and the Convergence with Data Science
Thomas Schulthess, Ph.D., Director, Swiss National Computing Centre (CSCS)
Brian D. O'Connor, Ph.D., Technical Director, Genomics Institute - Computational Genomics Platform, University of California, Santa Cruz

11:10 Keeping Afloat with Data and Compute Intensive HPC Workloads
Scott Yockel, Ph.D., Senior Team Lead of HPC, FAS Research Computing, Harvard University
For years, scientists have been using the latest technology to create and analyze data. In the first era, numerically intensive computations from traditional HPC disciplines like chemistry and physics created large amounts of data. Once the Internet took off, researchers had the ability to download large data sets from external repositories and the dawn of data-intensive research began. In the third era, scientific instruments began to overwhelm the instrument-attached computer, thus a need arose for a cluster of computers to process data. This creates a type of data velocity that is both numerical and data intensive, and it creates a rate of research productivity that must be maintained from a storage and compute standpoint. In this talk, I'll present examples of these three eras and how Harvard Research Computing has been tackling each one and what the future holds.

11:40 Bringing Compute and Data Together: Challenges and Opportunities in the Data Era
Arvindh Lalam, Founder and CEO, Fusion Memory
In the data era, insights are the new currency. As we are inundated with data, extracting insights has become more and more challenging. In order to address

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

these challenges, the traditional IT infrastructure model is gradually giving way to Cloud Infrastructure and Edge Infrastructure. This talk will delve into what the emerging paradigms mean for the data generators and compute elements and highlight challenges and opportunities in this space.

12:10 pm **Session Break**

12:20 **Luncheon Presentation** *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:20 **Dessert Refreshment Break in the Exhibit Hall with Poster Viewing**

FEATURED SESSION: BIOTEAM MICRO-SYMPOSIUM: 2017 BIO-IT TRENDS

1:55 **Chairperson's Remarks**
Chris Dwan

2:00 **BioTeam Micro-Symposium: 2017 Bio-IT Trends**
Chris Dwan, Senior Technologist and Independent Consultant (Moderator)
Ari E. Berman, Ph.D., Vice President and General Manager of Consulting Services, BioTeam, Inc.
Chris Dagdigian, Founding Partner & Director, Technology, BioTeam, Inc.
Aaron Gardner, Senior Scientific Consultant, BioTeam, Inc.
Adam Kraut, Director of Infrastructure and Cloud Architecture, BioTeam, Inc..
Asya Shklyar, Senior Scientific Consultant, Infrastructure, BioTeam, Inc.
Since 2010, the "Trends in the Trenches" presentation, given by Chris Dagdigian, has been one of the most popular annual traditions on the Bio-IT Program. The intent of the talk was to deliver a candid (and occasionally blunt) assessment of the best, the worthwhile, and the most overhyped information technologies (IT) for life sciences. The presentation tried to recap the prior year by discussing what has changed (or not) around infrastructure, storage, computing, and networks. This presentation has helped scientists, leadership, and IT professionals understand the basic topics involved in supporting data intensive science. For 2017, the "Trends in the Trenches" presentation will evolve and expand from 60-minutes to 120-minutes and feature more content, speakers, and interactive discussion. Short focused podium talks on current trends related to computing, storage/data transfer, networks, and cloud will be presented. A Q&A moderated discussion follows. Come prepared with your questions and commentary for this informative and lively session.

4:00 **Conference Adjourns**

4 Software Applications & Services

TOOLS THAT BEST UTILIZE DATA TO DRIVE SCIENTIFIC DECISION MAKING

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W3) Delivering Custom Mobile App Projects to the Cloud

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W11) Scientific Project Management
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing

Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing

Sponsored by InterSystems

IMPROVING OPERATIONAL EFFICIENCIES IN THE LAB TO INCREASE DATA QUALITY THROUGH DIGITIZATION

10:50 Chairperson's Remarks
Paul Denny-Gouldson, Vice President, Strategic Solutions, IDBS

11:00 Internet-of-Things for Pharma R&D
Jim Bernhard, Scientist, Materials Discovery & Characterization, Vertex Pharmaceuticals

11:30 Challenges and Opportunities with Deploying Computerized Systems for Communication of Laboratory Critical Results: The View from the Brigham and Women's Hospital
Milenko Tanasijevic, M.D., MBA, Vice Chair for Clinical Pathology and Quality, Department of Pathology, Brigham and Women's Hospital and Dana-Farber Cancer Institute; Director of Clinical Laboratories, Brigham and Women's Hospital; Associate Professor of Pathology, Harvard Medical School

Our research group has previously developed a unique computerized system to identify and communicate critical laboratory values to the covering physicians. The in-house developed system consisted of an event monitor that identified the events of interest in the laboratory and drug administration database, identified the specific caregiver responsible for the patient's care contemporaneous with the triggering event, paged the caregivers and required an acknowledgement of the call on their part.

12:00 pm Managing Business Processes through Agile Project Methodology
Gurpreet Kanwar, MBA, PMP, Senior Project Manager, Information Management, NAV CANADA

12:30 Session Break

Sponsored by idbs

12:40 Luncheon Presentation: Lab of the Future, Internet of Things & Machine Learning: A Pragmatic Approach
Jarrod Medeiros, Product Manager, IDBS
While the current technology landscape can seem impressive and alluring, the typical state of instrumentation and technology in an R&D lab is an ad hoc collection that spans decades. In this talk we will discuss how a pragmatic approach focused on business value can pave the road to operational efficiency, enhanced communication, and increased innovation in your R&D laboratory.

1:40 Session Break

DATA INTEGRATION & MODELING WITH SEMANTIC TECHNOLOGIES

1:50 Chairperson's Remarks
Derek A. Debe, Ph.D., Senior Principal Scientist, Platform Informatics & Knowledge Management, Abbvie, Inc.

1:55 Semantic Searching and Linking of Related Information Sources in R&D with High Accuracy
Etzard Stolte, Ph.D., Global Head, Knowledge Management, Roche
Thousands of users leverage Hoffman-La Roche's semantic platform to, for example, federate related search results across document management systems, find internal experts, visually explore concept evolution, or enrich documents on the fly with related numerical and high-dimensional data. This talk presents our learnings and real world requirements to develop such a semantic search and integration platform, while maintaining speed and high relevance.

2:25 Allotrope Foundation: A Framework for Simplified Data Integrity, Reproducibility and Data Management through Improved Data Modeling with Semantic Technologies
Dana Vanderwall, Ph.D., Director, Biology & Preclinical IT, Discovery Information Technologies, R&D IT, Bristol-Myers Squibb
Allotrope Foundation has built a framework comprised of: an open, standards-based, extensible file format; ontologies and semantic models to provide

4 Software Applications & Services

TOOLS THAT BEST UTILIZE DATA TO DRIVE SCIENTIFIC DECISION MAKING

controlled vocabulary for R&D applications; APIs to facilitate consistent adoption of the standards. This talk covers current data integration and management challenges, the Framework architecture, examples of how the Framework is currently used in the Pharmaceutical industry, and the roadmap for ongoing development and releases.

2:55 Key Software Interconnections between Data Sources and Decision Makers

Sponsored by ACD/Labs

Graham McGibbon, Manager, Strategic Partnerships and Scientific Solutions, Advanced Chemistry Development (ACD/Labs)

Andrew Anderson, Vice President, Innovation & Informatics Strategy, Advanced Chemistry Development (ACD/Labs)

Reliable foresight depends upon knowledge, which supposes coherent assembly and presentation of information. People attempting to assess the value and risk of particular work products need software applications and interfaces that allow communication between discrete data sources because evaluations are often of comprehensive material characterization. Software can construct layers of data with interpreted meaning aiding determination of suitability for use.

3:10 Integration from the Ground Up: Transforming Biologics R&D Informatics with Benchling

Sponsored by Benchling

Sajith Wickramasekara, Founder and CEO, Benchling

Most R&D processes are scattered across disparate software. Benchling unifies experiment workflows, ensuring that cutting-edge science is never held back by obsolete software. In this talk, we will describe how we worked with scientists to empower them to streamline biologics R&D workflow on a single platform.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

ENTERPRISE APPLICATIONS

4:00 New Middleware Concepts for Enterprise Applications

Nathan McBride, Senior Vice President & CIO, Innovation Architects, AMAG Pharmaceuticals

4:30 Amgen Research Connected Data Landscape – Enabling Self-Service Knowledge Discovery

Filip Pattyn, Ph.D. Scientific Lead, Product Manager, ONTOFORCE

Wolfgang Hoeck, Ph.D., Senior Manager, Amgen, Inc.

This talk will explain how enhancing the findability of research data (and more) is done at Amgen using a linked data approach. The goal of this session is to explain and discuss the value of how a connected data landscape enables researchers in early drug discovery research.

5:00 Team Productivity Reimagined: Abbvie's Knowledge Notebook and Dash Collaboration Platforms

Derek A. Debe, Ph.D., Senior Principal Scientist, Platform Informatics & Knowledge Management, Abbvie, Inc.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM

Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

COLLABORATIVE SCIENCE: PLATFORMS & SOFTWARE TOOLS FOR DECISION MANAGEMENT & DATA SHARING

10:30 Chairperson's Remarks

Christopher Southan, Ph.D., Database Curator, IUPHAR/BPS Guide to PHARMACOLGY, University of Edinburgh

10:40 Software Facilitated Collaboration through Cognitive Networks: Connecting Data and Researchers for Big Discoveries

David King, Founder, Exaptive

Kristophe Diaz, Ph.D., Scientific Program Manager, Cohen Veterans Bioscience, Inc.

There is increasing agreement in the scientific community that the difficult problems we are attempting to solve today will not be solved by a single researcher or single lab, but will require coordinated collaborative effort across larger teams of increasing heterogeneity. This talk explores a new approach to data analytics called the "cognitive network", an important perspective to software architecture for facilitating innovation at a broader level.

11:10 AbbVie Unite: A Connected Enterprise of Ideas, Projects, People and Medicine

Abhik Seal, Senior Scientist, Platform Informatics and Knowledge Management, AbbVie, Inc.

Rishi Gupta, Senior Research Scientist, Platform Informatics and Knowledge Management, AbbVie, Inc.

The drug discovery landscape is evolving quite rapidly and dynamically. There is a growing demand to encourage individuals or functional groups in an organization to connect and communicate with peers, partners, and customers on the projects they are working on. Silo mentality has to be broken down for new discoveries, setting up effective collaboration and improving the overall efficiency of the organization. The UNITE platform is developed to address this current challenge of social collaboration and in order to connect people knowing their interest and scientific topics they may be working on. The platform is not only made for social collaboration but it has

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

4 Software Applications & Services

TOOLS THAT BEST UTILIZE DATA TO DRIVE SCIENTIFIC DECISION MAKING

been extended for a better understanding of a disease and adverse effects. In this presentation, we will focus on the R Shiny based web application and provide ideas on how AbbVie is attempting to address this challenge.

11:40 Research Informatics: *In silico* Approaches for Early Assessment of Immunogenicity

Sponsored by DASSAULT SYSTEMES

Kevin Merlo, BioSafety Development Engineer, Dassault Systemes, BIOVIA

Immunogenicity, an immune response, can possibly reduce biotherapeutic treatment efficacy and provoke adverse effects. Predicting immunogenicity is proving difficult because of the complexity of the underlying biological processes. We present here an informatics application based on modeling and simulation approaches that can help the pharmaceutical R&D to prioritize promising drugs with respect to the immunogenicity risk

11:55 Sponsored Presentation *(Opportunity Available)*

12:10 pm Session Break

12:20 Luncheon Presentation I: Making Big Data Work for You

Sponsored by twigkit

Stefan Olafsson, CEO, Twigkit

Big data technologies have profound implications for research, translational science and the ability to provide a 360° view across all your data. Unfortunately, many big data projects fail to deliver the promise; due to a lack of focus on the applications and the user experience meant to provide this single pane of glass. Find out how to easily change that.

12:50 Luncheon Presentation II *(Sponsorship Opportunity Available)*

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

COLLABORATIVE SCIENCE: PLATFORMS & SOFTWARE TOOLS FOR DECISION MANAGEMENT & DATA SHARING

1:55 Chairperson's Remarks

Kees van Bochove, CEO, The Hyve

2:00 Implementation of the pRED External Partner Tool for Decision Management of External Collaborations

Laura Aguiar, Ph.D., MBA, PMP, Head, pREDi NY Operations and Manager, Data Science Decision Support Service, Roche Innovation Center New York

Padmanabha Udupa, pRED Informatics, Roche Pharma Research and Early Development, F. Hoffmann-La Roche Ltd.

This talk describes attempts to eliminate duplication of efforts and costly oversights by streamlining access to data; increasing transparency across functions and DTAs; institutionalizing corporate memory on external activities and business decisions; and deploying a platform to manage external partner key entities/ relationships including institutions, contacts and opportunities.

2:30 Case Study: The Observational Health Data Sciences & Informatics (OHDSI) Collaborative

Alan Andryc, Technology Manager, Data Sciences, The Janssen Pharmaceutical Companies of Johnson & Johnson

Kristin Feeney, MPH, Senior Data Scientist I, ConvergeHEALTH by Deloitte

Kees van Bochove, CEO, The Hyve

Karthik Natarajan, Ph.D., Assistant Professor of Biomedical Informatics, Columbia University

The OHDSI program is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. This collaborate represents joined efforts of academia, pharmaceutical industry, and informatics companies to develop and leverage innovative methodologies and advanced analytics for real world observational health data and promote high quality evidence generation. All our solutions are open-source. OHDSI has established an international network of researchers and observational health databases with a central coordinating center housed at Columbia University. This talk presents our vision for high quality evidence generation through our commitment to open science and innovation leveraging open source tools, shared standards and technologies we are generating as a community. Come see how we are shifting the evidence paradigm in academia and industry. Join the journey.

4:00 Conference Adjourns

25 YEARS Organized by Cambridge Healthtech Institute

5 Cloud Computing

APPLYING CLOUD FOR EXPANDING APPLICATIONS

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W2) An Intro to Blockchain in Life Sciences
(W3) Delivering Custom Mobile App Projects to the Cloud

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W12) Leveraging Cloud Technologies to Enable Large-Scale Integration of Human Genome and Clinical Outcomes Data
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing

Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing

Sponsored by InterSystems

CASE STUDY CO-PRESENTATIONS: IMPLEMENTING CLOUD

10:50 Chairperson's Remarks
Ron Bianchini, CEO and President, Avere Systems

11:00 FEATURED PRESENTATION: Moving Our Corporate Data Center to the Cloud in 12 Months – >725 Servers, >250 Million Files and a Lot of Lessons Learned!
Martin Leach, Ph.D., Vice President, R&D IT, Enterprise Data Management & Analytics, Innovation, IT, Alexion Pharmaceuticals
ChoongTeik Ng, Associate Director, SAP Infrastructure, IT, Alexion Pharmaceuticals
During a recent move, we at Alexion challenged ourselves to move our corporate data center, servers, systems, applications and data to a cloud-based architecture. Our global IT team uncovered >725 physical and virtual servers, >350 business applications and >250 million files. With few exceptions we have successfully moved all this to a cloud based-environment. We now detail our approach, playbook, lessons learned, pitfalls and how we engaged our colleagues company-wide.

11:30 Neuroscience in the Cloud
Bridget Behringer, MBA, R&D IT Mergers & Acquisitions Engagement Leader, R&D Information, AstraZeneca
Immanuel Utomo, MSc, Senior Data Scientist, Informatics and Knowledge Utilization, Celgene
Neuroscience was a unique endeavor that adopted disruptive technologies to conduct drug discovery differently. Unlike the traditional model with internal labs, the full drug portfolio was conducted by a small internal team of key opinion leaders that leveraged capabilities, services and expertise of external partners and collaborators. Cloud computing enabled the group to be flexible, cost effective, and productive.

12:00 pm How Cloud has Changed Life Sciences
Sponsored by Cycle Computing
Jason Stowe, CEO, Cycle Computing
This talk discusses the impact cloud computing -- with its flexibility and scale -- has provided cost-effective and powerful access to life sciences researchers.

12:30 Session Break

12:40 Luncheon Presentation I: R&D Data Driven Organizations: Stop Treating Your Data Like Trash!
Sponsored by LabAnswer
John Conway, Global Director, Research & Development, Data Science and Informatics Strategy, LabAnswer
Without a coherent scientific data strategy, organizations fail to guard and cultivate one of their top assets, data. LabAnswer will walkthrough what a scientific data strategy entails, how it can help emulate a true data company and drive innovation/discovery. Additionally, LabAnswer will show how employing a scientific data strategy can be opportunistic and pragmatic as you develop your infrastructure/cloud strategy.

1:10 Luncheon Presentation II: From Humans to Mosquitos: Enabling Genomics Discovery in Microsoft Azure
Sponsored by Microsoft
Geralyn Miller, Director, Microsoft Genomics, Microsoft
This presentation will explore why genomics workloads are perfect for the cloud and how Microsoft Research is using Azure to accelerate genomics discovery.

1:40 Session Break

DESIGNING FLEXIBILITY WHILE MAINTAINING SECURITY

1:50 Chairperson's Remarks
Lance Smith, MBA, Associate Director, IT, Celgene

1:55 How BMS Uses cBioPortal for Cancer Genomics Research
Isaac Neuhaus, Ph.D., Director, Computational Genomics, Bristol-Myers Squibb
BMS has started using cBioPortal for visualizing cancer genomics datasets early

5 Cloud Computing

APPLYING CLOUD FOR EXPANDING APPLICATIONS

2016, supported by The Hyve, an open source bioinformatics company based in The Netherlands. The cBioPortal server runs on Amazon AWS and is tied to the company Active Directory for authentication. Currently loaded datasets are TCGA, CCLE and M2Gen and we will be loading clinical trial data later this year. Concurrently we are running a pilot for cBioPortal with mouse data.

2:25 Enabling Celgene's Innovation while Protecting the Enterprise in the Cloud

Lance Smith, MBA, Associate Director, IT, Celgene

Celgene uses Amazon Web Services (AWS) to run R&D workloads, enterprise systems, and collaboration environments securely and all controlled from a single pane. Topics covered include network design, HPC and job scheduling, workload isolation, storage, automated security, and environment auditing. We also discuss business use cases, challenges overcome (technical, legal, security, and organizational), and Turbot guardrails for management.

2:55 Automated High-Scale Computational Infrastructure

Sponsored by OpenEye Scientific

Jharrod LaFon, Chief Cloud Engineer, OpenEye Scientific

Large-scale computation is an important aspect of drug design, and increasingly is cloud-based. The true value of the cloud comes not just from computation but by enabling company-wide collaboration; the ability to build, modify, share, manage versions, and evaluate methods. OpenEye's cloud-native platform, Orion, and its workflow engine, Floe, provide the tools to create innovative approaches for computation and collaboration.

3:10 Research Informatics: Get Ready for the Cloud!

Sponsored by Dassault Systèmes

Ton van Daelen, ScienceCloud, Product Director, Dassault Systemes, BIOVIA

The life sciences industry is looking to the cloud to support outsourcing and external collaboration initiatives. At the same time the cloud can give research IT the 'agility' to innovate more rapidly and significantly reduce their IT footprint. BIOVIA is helping research organizations implement 'hybrid cloud' solutions with its ScienceCloud platform to facilitate this critical transition.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

4:00 5 Steps to Vet and Manage Cloud Service Providers

Dianne Pacheco, CISM, Information Security Officer, Information Technology, The Jackson Laboratory

The use of cloud services is essential, but for transmitting, processing or storing data that way, buyer beware – not all service providers are created equal. Many consumers do not realize they bear full responsibility for the security of the data held by service providers. This talk walks the audience through vetting and managing their service providers to minimize risk of exposing legally protected or sensitive research data.

4:30 Cloud Enabling Networks for Pediatric Care

Marcia M. Nizzari, MS, Consulting CTO and Software Architect, Ksandr Software Consulting

Cloud infrastructure enables new networks of pediatric hospitals, labs, clinicians, researchers, and pharma that can improve a child's care. The economies of scale, ease of data movement, shared tools, and communication enrichment are significant game changers in the world of pediatric rare disease. We cover how cloud technology and related new software architecture constructs, such as microservices and containers, can drive advancement in precision medicine and facilitate a better future for our patients.

5:00 Aspera as a Data Aggregation and Distribution Hub for Global Tuberculosis Research Protocols

Sponsored by aspera an IBM company

Christopher Whalen, International Program Manager, Research Data & Communications ,National Institutes of Health (NIH), Office of CyberInfrastructure and Computational Biology , National Institute of Allergy and Infectious Diseases (NIAID)

Jay Migliaccio, Director, Cloud Products and Solutions, Aspera, an IBM Company

Michael Duvenhage, Global Clinical Data Support, Operations Manager, Research Data and Communications Technology, Office of CyberInfrastructure and Computational Biology National Institute of Allergy and Infectious Diseases, NIH

Clinical studies are leveraging new medical and informatics technologies. As a large multi-country study, the clinical data management program of the Tuberculosis Research Section of the NIAID requires robust methods to integrate many different data flows with advanced controls. This talk will focus on the technical issues that emanate from the study's ambitious goals and stringent security requirements and how the Aspera high speed transport software is used to solve these global data transfer challenges.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM

Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

APPLYING CLOUD FOR LARGE-SCALE GENOMIC RESEARCH

10:30 Chairperson's Remarks

Pamela Duffy, MS, Director, Core Clinical Solutions, IT, Pfizer

5 Cloud Computing

APPLYING CLOUD FOR EXPANDING APPLICATIONS

10:40 FEATURED PRESENTATION:
To the Cloud(s): Keeping Up with DNA Sequencing
Stacey Gabriel, Ph.D., Director of the Genomics Platform, The Broad Institute of MIT and Harvard
To keep up with sequencing at population scale, we require scalable, efficient, and cost-effective analysis approaches to eliminate computation as the gating factor in the progress of scientific discovery. We discuss challenges and findings from our experience, and an approach to empower other researchers to leverage our cloud-based best practice analysis pipelines capable of processing a whole genome every 4 minutes.

11:10 Multicloud in Action: A System for Variant Detection and Annotation Workflow in Multiple Clouds
Mohamed Abouelhoda, Ph.D., Head, Bioinformatics, Saudi Human Genome Project, Genetics Department, King Faisal Specialist Hospital and Research Centre (KFSHRC)
Exploiting competition in the cloud computing market is an effective means for reducing the cost of genomic data analysis. We present MC-GenomeKey, a novel package for running the variant analysis workflow in multiple cloud platforms, including Amazon, Google, and Azure. MC-GenomeKey can run the jobs in different clouds and it provides a reliable use of the spot instances of Amazon in combination with Google cloud for further cost reduction.

11:40 Re-Thinking the Cloud for Data-Intensive Science: The Lab7/IBM Genomic Cloud
Sponsored by Lab7
Christopher Mueller, Ph.D., President & CTO, Lab7 Systems
The Lab7/IBM Genomic Cloud presents a new approach to scientific cloud computing. By necessity, traditional cloud services cater to a wide range of use cases, limiting their utility for cost-effective high-performance computing. Understanding the needs of high-throughput genomic cores and limitations of commodity clouds, we designed a cloud service with a complete hardware, software, and support stack optimized for genomics.

11:55 Innovations in Large Scale Life Sciences Research
Sponsored by Google Cloud
Karan Bhatia, Ph.D., Cloud Specialist, Google Cloud
Google Cloud enables scientists to change the way the perform research and collaborate with one another. This presentation will highlight how Google Cloud is accelerating life sciences research and finding new ways to innovate.

12:10 pm Session Break

12:20 Luncheon Presentation I: Freeing Data: How to Win the War with Hybrid Clouds
Sponsored by AVERE
Ron Bianchini, CEO and President, Avere Systems
Adam Kraut, Director of Infrastructure and Cloud Architecture, BioTeam, Inc.
The public cloud is a necessary point of convergence for big science. Sustainable cloud infrastructures continue to leverage breakthroughs for life sciences. This presentation will focus on the factors leading to successful missions in public and hybrid cloud. Discuss the dichotomy of hybrid clouds and examples of real-world hybrid cloud use cases. Participants will gain valuable insights into how to free datasets to enable research and scientific progress with minimal infrastructure friction.

12:50 Luncheon Presentation II: Innovative Practices in the Cloud
Sponsored by Google Cloud
Evren Eryurek, Technical Director, CTO Office, Google Cloud
Bill Mayo, Chief Information Officer, Broad Institute
Google Cloud provides an opportunity to innovate traditional biomedical & research IT practices. In this session we will discuss how companies are managing workloads in the cloud, using big data to analyze large amounts of information, and leveraging machine learning to drive improvements from research methodologies to clinical practices.

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

FEATURED SESSION: BIOTEAM MICRO-SYMPOSIUM: 2017 BIO-IT TRENDS

1:55 Chairperson's Remarks
Chris Dwan

2:00 BioTeam Micro-Symposium: 2017 Bio-IT Trends
Chris Dwan, Senior Technologist and Independent Consultant (Moderator)
Ari E. Berman, Ph.D., Vice President and General Manager of Consulting Services, BioTeam, Inc.
Chris Dagdigian, Founding Partner & Director, Technology, BioTeam, Inc.
Aaron Gardner, Senior Scientific Consultant, BioTeam, Inc.
Adam Kraut, Director of Infrastructure and Cloud Architecture, BioTeam, Inc..
Asya Shklyar, Senior Scientific Consultant, Infrastructure, BioTeam, Inc.
Since 2010, the "Trends in the Trenches" presentation, given by Chris Dagdigian, has been one of the most popular annual traditions on the Bio-IT Program. The intent of the talk was to deliver a candid (and occasionally blunt) assessment of the best, the worthwhile, and the most overhyped information technologies (IT) for life sciences. The presentation tried to recap the prior year by discussing what has changed (or not) around infrastructure, storage, computing, and networks. This presentation has helped scientists, leadership, and IT professionals understand the basic topics involved in supporting data intensive science. For 2017, the "Trends in the Trenches" presentation will evolve and expand from 60-minutes to 120-minutes and feature more content, speakers, and interactive discussion. Short focused podium talks on current trends related to computing, storage/data transfer, networks, and cloud will be presented. A Q&A moderated discussion follows. Come prepared with your questions and commentary for this informative and lively session.

4:00 Conference Adjourns

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

25 YEARS Organized by Cambridge Healthtech Institute

6 Bioinformatics

COMPUTATIONAL RESOURCES AND TOOLS TO TURN BIG DATA INTO SMART DATA

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W4) Data Visualization to Accelerate Biological Discovery

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W13) Proteogenomics: Integration of Genomics and Proteomics Data
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing
Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing
Sponsored by InterSystems

APPLICATIONS & SOLUTIONS FOR DATA SHARING AND DECISION MAKING

10:50 Chairperson's Remarks
Kevin Merlo, BioSafety Development Engineer, Dassault Systemes, BIOVIA

11:00 Innovative Data Integration Applicable for Therapeutic Protein Development 2.0
Wolfgang Paul, Group Leader and Senior Scientist, Large Molecule Research, Roche
Therapeutic proteins are registered including sequence, structural and functional data and information. Millions of data points are captured during the development of Roche's innovative therapeutic proteins in data warehouse used by DAMAS (data acquisition, management and analyses system). Fast access and visualization of relevant process and analytical data drive scientific discussion and decision making. Analyzing the stored big data is key towards process development of therapeutic proteins 2.0.

11:30 Informatics – A Silver Bullet for Pharmaceutical Sciences?
William Loging, Ph.D., Associate Professor of Genomics & Head, Production Bioinformatics, Genetics and Genomics Sciences, Icahn School of Medicine at Mount Sinai
The Pharmaceutical Sciences field is in constant search for the next big innovative push that will increase the success rate of drug programs. The fields of computational chemistry, structural bioinformatics – just to name a few – have changed the way drug researchers look for and identify novel drug candidates. Utilizing more than 15 years of Pharmaceutical experience, and using real world examples of high provide drug projects, this talk will provide practical steps for the merger of informatics and the strategic approaches needed for drug discovery success.

12:00 pm Big Data-Driven Bioinformatics
Sponsored by IBM
Frank Lee, Ph.D., Healthcare Life Sciences Industry Leader, Software Defined Infrastructure, IBM Systems, IBM
IBM will discuss the IBM Reference Architecture for Genomics, its new features, and case studies: hybrid cloud with integrated workload and data management for high performance genomics analytics; container technologies for migrating and sharing application and data; and application portal and metadata engine for global access to and searching of distributed resources. A demo of a hybrid cloud-based bioinformatics solution will follow.

12:30 Session Break

12:40 Luncheon Presentation I: Towards the Use of Bioassays as Predictors of Adverse Events in Clinical Trials
Sponsored by ELSEVIER
Matthew Clark, Ph.D., Consultant Research & Development Solutions, Elsevier
To reduce clinical trial failures we have studied a method of using inexpensive bioassays to predict adverse event risks in clinical trials. We used data from FDA submissions and bioactivities from journals and patents to create 2x2 contingency tables for each target/event combination. The relationships were then studied with pathway analysis to understand the models.

1:10 Luncheon Presentation II: Informatics in the Fast Lane
Sponsored by Abbott
Shooki Grasiani, Senior Manager, Global Marketing and Product Management, Specialty Products, Product Marketing, Abbott Informatics
In order to survive in a rapidly evolving and competitive environment, organizations must accelerate innovation and adopt agile solutions which remove barriers to innovation. In this session, Shooki will cover strategies and tactics around the reality check of the laboratory informatics' promise, the shifts in the laboratory informatics macro-environment and technological trends that will reshape the laboratory informatics landscape.

1:40 Session Break

STANDARDS FOR CHEMICAL STRUCTURES

1:50 Chairperson's Remarks
Rich Lysakowski, Ph.D., Professor of Bioinformatics and Data Science; Principal

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

25 YEARS Organized by Cambridge Healthtech Institute

Software Engineer and IT Applications Specialist, Network Technology Academy Institute

1:55 PANEL DISCUSSION: Linking and Finding Information Using the IUPAC InChI Standard for Chemical Structures

Steve Heller, Ph.D., Project Director, InChI Trust; Scientific Information Consultant (Moderator)

Evan Bolton, Ph.D., Lead Scientist, National Center for Biotechnology Information (NCBI), National Library of Medicine (NLM), and National Institutes of Health (NIH)

Keith T. Taylor, BSc, Ph.D., MRSC, Principal, Ladera Consultancy

Tyler Peryea, Informatics Scientist, National Center for Advancing Translational Sciences (NCATS)

Lawrence Callahan, Ph.D., Chemist, Substance Registration System, Office of Critical Path Programs, Food and Drug Administration (FDA)

This session will highlight on-going efforts to strengthen and expand the non-proprietary IUPAC International Chemical Identifier (InChI) standard for chemical structures and its hashed-form, the InChIKey. Information standards are critical to enable effective communication of scientific content. Funding to maintain InChI comes from most major publishers and database providers as well as governmental agencies (NIH, FDA and NIST). The InChI is an open-source, widely adopted standard found in most chemical information containing databases, including those from Chemical Abstracts, Reaxys, ChEMBL, OpenPHACTS, PubChem, DrugBank, PDB, Sigma-Aldrich, and many others, such as internal Pharma corporate databases. InChI is an addition to a database, not a replacement. With the implementation of the ISO identification of medicinal products (IDMP) and the related ISO 11238 standards, adding and having an InChI will allow for an easier, effective, and more complete search for information on a particular drug.

2:55 Sponsored Presentation *(Opportunity Available)*

3:10 Integrated Informatics for Biologics Discovery

Sponsored by dotmatics

Robert Brown, Ph.D., Vice President, Product Marketing, Dotmatics

A presentation looking at the challenge of supporting biologics discovery and current solutions. Highlighting the value of an integrated informatics solution. It will also include an example biologics discovery workflow using Bio ELN, Biological Registration and Vortex Bioinformatics.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

MACHINE LEARNING AND SYSTEMS BIOLOGY TECHNIQUES AND APPLICATIONS TO PERFORM BIG DATA ANALYTICS ON –OMICS DATA

4:00 Building Disease Networks Using Text Mining and Machine Learning Techniques

Kamal Rawal, Ph.D., Assistant Professor, Biotech and Bioinformatics, Jaypee Institute of Information Technology

Obesity is a global epidemic affecting over 1.5 billion people and is one of the risk factors for several diseases such as type 2 diabetes mellitus and hypertension. We have constructed a comprehensive map of the molecules reported to be implicated in obesity. Using text mining & deep curation strategies combined with omics data, we have explained the therapeutics and side effects of several drugs (i.e., orlistat) at network level.

4:20 Big Data and Systems Biology: From Genome to Phenome (and Everything in Between)

Dan Jacobson, Ph.D., Computational Biologist, Oak Ridge National Laboratory

4:40 Novel Feature Selection Strategies for Enhanced Predictive Modeling and Deep Learning in the Biosciences

Tom Chittenden, Ph.D., D.Phil., Lecturer and Senior Biostatistics and Mathematical Biology Consultant, Harvard Medical School

We have built a robust AI approach that precisely assesses pathogenicity for all genomic missense variants. Coupled with our advanced deepCODE mathematical statistics feature selection strategy for constructing deep learning models, we are able to quantitatively integrate a priori pathway-based biological knowledge with multiple types of high-throughput omics data.

5:00 Integration of Predictive Analytics and CBDD for Indication Prioritization

Sponsored by Clarivate Analytics

Marina Bessarabova, Ph.D., Senior Director, Discovery and Translational Science, Life Sciences Professional Services, Clarivate Analytics (Formerly the IP & Science Business of Thomson Reuters)

Elia Stupka, Ph.D., Director, Genomics and Computational Biology, Boehringer Ingelheim

Systems biology is a powerful approach to drug development. CBDD is a precompetitive consortium between Clarivate Analytics, Novartis, Pfizer, Sanofi, Janssen, Regeneron, UCB, Roche, Takeda, Biogen, Boehringer Ingelheim, Bristol-Myers Squibb, Merck, and Eli Lilly, focused on implementation of advanced systems biology methods. Clarivate Analytics will present CBDD scope and Boehringer Ingelheim will present application of CBDD developments for indication prioritization.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM
Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

Data & Storage Management
Data Computing
Networking Hardware
Software Applications & Services
Cloud Computing
Bioinformatics
Next-Gen Sequencing Informatics
Clinical Research & Translational Informatics
Data Visualization & Exploration Tools
Pharmaceutical R&D Informatics
Clinical Genomics
Cancer Informatics
Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

25 YEARS Organized by Cambridge Healthtech Institute

6 Bioinformatics

COMPUTATIONAL RESOURCES AND TOOLS TO TURN BIG DATA INTO SMART DATA

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

DATA COMPUTING AND BIOINFORMATICS IN AGRO CHEMICALS AND BIOTECHNOLOGY: CHALLENGES AND OPPORTUNITIES

10:30 Chairperson's Remarks

Bino John, Ph.D., Computational Biology Group Leader, Dow AgroSciences LLC

10:40 How Biotech and Big Data Are Changing Agro Industry

Bino John, Ph.D., Computational Biology Group Leader, Dow AgroSciences LLC

More than 70% of the increase in food production in the next 50 years is expected to come from technological advances. Indeed, recent advances in genomics and phenomics are beginning to transform the Agro-industry, whereby creating new opportunities for informatics disciplines. While informatics needs in managing, analyzing, and visualizing big data share commonalties between Agro and the biomedical communities, Agro companies face unprecedented challenges in big biological data, generally larger than their peers in the biomedical community.

11:10 Offering Outcomes: How Digital Farming Data Is Enabling New Business Models

Chris Paterson, Lead, Digital Farming North America, Bayer Crop Science

11:40 Building the Next-Generation R&D IT Infrastructure for Small Molecule Discovery

Paimun Amini, Chemistry IT Lead, R&D IT, Monsanto Company

Barrett Foat, Ph.D., Data Science Team Lead, Agricultural Productivity Innovations, Monsanto

The Pharma boom in the 90s & 2000s led to the emergence of a rich ecosystem of software companies focused on delivering the IT needs for small molecule discovery. Today, cloud data storage, IoT, and the growth of predictive analytics present new opportunities for the evolution of the R&D pipeline. New technologies allow for integrated software and hardware solutions that optimize productivity while removing the risk of technical debt.

12:10 pm Session Break

Sponsored by

12:20 Luncheon Presentation: CLC Genomics Cloud Engine: Enterprise NGS Made Easy

Jacob Nikolajsen, Cloud Architect, QIAGEN

Reaping the benefits of cloud computing takes more than just moving servers and data online. Ensuring secure solutions and smooth transition to the cloud is far from trivial, particularly if results are expected to be identical to an on-premise solution. QIAGEN CLC product line is the leading solution for NGS analysis, proven to minimize cost per analysis on Intel CPUs.

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

LOOKING BEYOND THE GENOME OF THE PATIENT: DATA, ANALYSIS AND TOOLS TO IMPROVE BETTER DISEASE UNDERSTANDING FOR CURRENT TREATMENTS AND DRUG DEVELOPMENT

1:55 Chairperson's Remarks

Michael N. Liebman, Ph.D., Managing Director, IPQ Analytics, LLC and Strategic Medicine, Inc.

2:00 Distinguishing between Precision Medicine and Accurate Medicine: Application to Heart Failure Patients and Clinical Practice

Michael N. Liebman, Ph.D., IPQ Analytics, LLC and Strategic Medicine, Inc.

Increasingly, patient stratification based on genomic analysis is being considered in disease management. Critically, the need to understand real world medical practice and real world patient complexities extends far beyond the genome of the patient. We have shown examples of this complexity in heart disease and how this impacts development of clinical guidelines, trial design, and development of new patient management approaches.

2:30 CARPEDIEM - Comorbidity and Risk Profiles Evaluation in Diabetes and Heart Morbidities

Sabrina Molinaro, Psy.D., Ph.D., Head, Department of Epidemiology and Health Services, Institute of Clinical Physiology, National Research Council of Italy

Our project uniquely develops a patient record that includes clinical and individual factors (EHR-driven phenotyping) that will be validated through the comparison of existing standards for building new risk algorithms. An understanding of the current limitations and biases of risk profiling in heart disease and diabetes and how an extended, integrated database and automatic rule-based classification system can be used to improve patient management.

3:00 PANEL DISCUSSION: Precision Medicine vs. Accurate Medicine: The Need to Understand Real World Medicine and Real World Patients

Michael N. Liebman, Ph.D., IPQ Analytics, LLC and Strategic Medicine, Inc. (Moderator)

Charles Barr, M.D., MPH, Group Medical Director and Head, Evidence Science and Innovation, Genentech

Jonathan Morris, M.D., Vice President, Provider Solutions and Chief Medical Informatics Officer, Real World Insights, QuintilesIMS

Krishnan Nandabalan, Ph.D., President, CSO and Co-Founder, BioXcel

Hal Wolf, Director, National Leader of Information and Digital Health Strategy, The Chartis Group

4:00 Conference Adjourns

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

Organized by Cambridge Healthtech Institute

7 Next-Gen Sequencing Informatics

ADVANCES IN LARGE-SCALE COMPUTING

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W6) Digging Bioactive Chemistry Out of Patents Using Open Resources

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W13) Proteogenomics: Integration of Genomics and Proteomics Data
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing
Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing
Sponsored by InterSystems

BIG DATA ANALYTICS ON CANCER GENOMIC DATA

10:50 Chairperson's Remarks
Bhanu Bahl, Ph.D., Director, Clinical and Translational Science Centre, Harvard Medical School

11:00 Deep Analytics of Cancer Genomics Data
Nicholas Camarda, Bioinformatics Analyst I, Carter Group, Dana-Farber Cancer Institute
The very large collection of cancer genomic data within Foundation Medicine allows us to see not only breadth of mutational variation, but re-occurring and significant patterns in many different cancers. These different mutational modes potentially have great relevance to more actionable diagnoses, as well as providing deeper insights for cancer research and drug development. We are exploring a new method to map each case into a statistically-defined multidimensional space along with functional and response associations.

11:30 Building a Harmonized Clinical Data Resource for Cancer Research and Translational Medicine: Challenges, Goals and Progress
John Methot, Director, Health Informatics Architecture, Dana-Farber Cancer Institute
Oncology's pioneering role in precision medicine has led to routine tumor sequencing. However, the incompleteness and inconsistency of data from EMRs and other sources presents significant challenges to building the patient-specific diagnosis, treatment and outcome "stories" that are required as correlates to genomic data to inform outcomes and clinical quality research. At DFCI, we are striving to improve the quality and coverage of clinical data for research and translational medicine, including decision support in the clinic. We are leveraging this data resource, using modern software architectures, to accelerate research and improve patient outcomes.

12:00 pm Graph Genome Tools for Precision Medicine
Sponsored by SevenBridges
Jack DiGiovanna, Senior Scientist, Director of Program Management, Seven Bridges
Constructing a genome graph captures variations within a population as branches which diverge from the common reference but later rejoin it. This approach is fundamentally different than alignment and variant calling against a single linear reference, but does it matter for precision medicine? Here, we explore two key features of graph technology - improved variant calling and capturing distributions of variations in a population. We project these features onto the current state of precision medicine.

12:30 Session Break

12:40 Luncheon Presentation I: NGS in the Fast Lane: How a Collaborative Approach between the Mayo Clinic and Illumina is Advancing Genetic Testing
Sponsored by illumina
Michael Ball, Vice President, EIBU Commercial, Illumina
Despite progress in sequencing technologies, obstacles still prevent a systematic method for deriving actionable information from an individual's genome. The BaseSpace® Suite from Illumina was designed to eliminate obstacles and enable users to easily store, manage, and aanalyze genomic data. In this session, learn how the Mayo Clinic is using BaseSpace Suite to expedite the delivery of its genomic expertise, and how Illumina is applying feedback from the Mayo Clinic to make the BaseSpace Suite even more efficient.

1:10 Making TCGA Tumor Datasets Accessible for Real-Time Multiomics Analysis and Visualization
Sponsored by WuXi NextCODE Leading Genomics
Jim Lund, Ph.D., Director, Tumor Product Development, WuXi NextCODE
The collection of cancer datasets in The Cancer Genome Atlas (TCGA) dataset has provided a wealth of insight into cancer biology; however, analyzing such a complex dataset is challenging for most investigators. This talk will demonstrate the Wuxi NextCODE approach using our proprietary genomically ordered relational (gor) database architecture to rapidly and simultaneously explore the diverse data types in TCGA.

1:40 Session Break

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

Organized by Cambridge Healthtech Institute

LARGE SCALE COMPUTING & DATA ANALYSIS

1:50 Chairperson's Remarks
Shanrong Zhao, Ph.D., Director, Computational Biology, Early Clinical Research, Pfizer Worldwide Research and Development

1:55 QuickMIRSeq: A Strand-Aware Pipeline for Quick and Accurate Quantification of Known miRNAs and isomiRs from Large-Scale Small RNA Sequencing
Shanrong Zhao, Ph.D., Director, Computational Biology, Early Clinical Research, Pfizer Worldwide Research and Development
Genome-wide miRNA expression data can be used to study miRNA dysregulation comprehensively. We designed and implemented a pipeline called QuickMIRSeq (https://sourceforge.net/projects/QuickMIRSeq/) for quick and accurate quantification of known miRNAs and miRNA isoforms (isomiRs) from large-scale sequencing experiments. QuickMIRSeq considers the unique nature of miRNAs and combines many important features into its implementation for the sake of computational efficiency and quantification accuracy.

2:25 Creating and Executing Large-Scale, Reproducible Genomics Analysis Pipelines Using Globus Genomics
Ravi K. Madduri, Ph.D., Scientist, Mathematics and Computer Science, Argonne National Laboratory; Senior Fellow, Computation Institute, University of Chicago

2:55 Transforming Data Storage for Life Sciences Research *Sponsored by* DDN STORAGE
Paul Jeffreys, Professor, Digital Operations, Programme Lead, The Institute of Cancer Research; Keble College Fellow, University of Oxford Digital Operations, The Institute of Cancer Research
Jon Lockley, Head of Scientific Computing,The Institute of Cancer Research
In the context of a leading cancer research institute working in close partnership with a world-leading cancer hospital, we will present an innovative approach to storing and managing research data through its life-cycle.

3:10 Sponsored Presentation *(Opportunity Available)*

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

4:00 Democratize Public Genomic Data Analysis through WebMeV
Yaoyu Wang, Ph.D., Associate Director, Center for Cancer Computational Biology, Dana-Farber Cancer Institute
WebMeV is a robust, open-source cloud based scalable data analysis software tool developed at the Dana-Farber Cancer Institute that uses intuitive visual interfaces to provide users with access to advanced data analysis methods. It will allow researchers and biotechnology companies considering tools for large scale genomic data analysis an additional option to all the proprietary software.

4:30 An Analysis of Cancer Genome Large-Scale Variation
Jeffrey Rosenfeld, Ph.D., Assistant Professor, Pathology and Laboratory Medicine, Rutgers Cancer Institute of NJ
Large-scale sequencing has been done on cancer genomes through projects such as TCGA, but there is still a lot to learn about genomics variation in cancer. Most of the current studies have looked at SNPs and other base-level variants in the genome. We have used advanced techniques and novel sequencing platforms to explore structural variation in cancer both at a DNA and a RNA level.

5:00 Democratizing NGS Analytics with HPE Purpose-Designed Solutions *Sponsored by* intel
Michael J. McManus, Ph.D. Senior Health & Life Sciences Solution Architect Intel
Haruna Cofer, Ph.D., Senior Applications,Engineer, Domain Expertise Group ,HPE
Genomics is the key to unlocking the power of precision medicine. As the demand for genome sequencing grows in areas like pediatric medicine and oncology, so does the amount of data that must be processed, stored and managed.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM
Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

APPLICATIONS OF NGS TO CANCER, IMMUNOLOGY, DIAGNOSTICS, THERAPEUTIC DEVELOPMENT, & EMERGING TECHNOLOGIES

10:30 Chairperson's Remarks
Jeffrey Rosenfeld, Ph.D., Assistant Professor, Pathology and Laboratory Medicine, Rutgers Cancer Institute of NJ

10:40 KEYNOTE PRESENTATION: Integrating Multiple High Content Technologies for Precision Immunology
Mario Roederer, Ph.D., Senior Investigator, ImmunoTechnology Section, Vaccine Research Center, NIAID, NIH

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

Organized by Cambridge Healthtech Institute

7 Next-Gen Sequencing Informatics

ADVANCES IN LARGE-SCALE COMPUTING

11:10 Subtyping Cancer Patients Using Immune Gene Expression

Jurriaan Brouwer, Ph.D., Data Science and Oncology, Roche Innovation Center, NY

We will present a novel method of identifying cancer patients that may respond to cancer immunotherapy using immune-related gene expression. By establishing the balance between immune activating and inhibiting signatures, we can suggest more personalized treatment options than classical subtyping methods.

11:40 Analysis and Exploration of n-Dimensional Single Cell-Data from Cytometry and Single Cell Sequencing

Sponsored by FLOWJO, LLC

Michael Stadnisky, Ph.D., CEO, FlowJo, LLC

The single cell is the basic unit of disease, but emerging technologies in cytometry and single cell sequencing are held back by time-consuming, sequential manual steps or computationally expensive, non-deterministic data reduction. Herein, we describe a powerful nonlinear, deterministic data reduction technique which we use to drive insight in single cell phenotyping and single cell sequencing of whole melanoma metastases.

12:10 pm Session Break

12:20 Luncheon Presentation *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

1:55 Chairperson's Remarks

Yuval Itan, Ph.D., MRes, Research Assistant Professor, Human Genetics of Infectious Diseases, The Rockefeller University

2:00 Management, Analysis, and Interpretation of DNA and RNA Sequencing To Aid in the Interpretation of Variants of Unknown Significance for Studies of Rare Inherited Disease and in Oncology

David Wesley Craig, Ph.D., Professor of Translational Genomics, Keck School of Medicine, University of Southern California; Co-Director, USC Institute of Translational Genomics

With the emergence of DNA and RNA sequencing (RNA-seq) technologies, RNA-based biomolecules hold expanded promise for their diagnostic, prognostic and therapeutic applicability in various diseases, particularly in the context of variants of unknown significance. In this talk we describe ongoing efforts towards the establishment of benchmark standards, assay optimization for clinical conditions and demonstration of assay reproducibility are required to expand the clinical utility of RNA-seq.

2:30 Estimating Genotypic Heterogeneity Underlying Human Disease

Yuval Itan, Ph.D., MRes, Research Assistant Professor, Human Genetics of Infectious Diseases, The Rockefeller University

3:00 Challenges and Implications of RNA-Seq Quantification at the Transcript Level

Chi Zhang, Ph.D., Post Doc, Computational Biology, Early Clinical Research, Pfizer Worldwide Research and Development

3:30 NGS Informatics for Mouse Cancer Avatars

Govindarajan Kunde Ramamoorthy, Associate Director, Computational Sciences – Scientific Computing, Jackson Lab for Genomic Medicine

Vinod Yadav, Application Computational Scientist, The Jackson Laboratory

Mouse cancer avatars provide a power platform for rapidly investigating therapies for treatment resistant cancers. Patient tumors are engrafted into mice, and various treatment approaches can be evaluated in the context of their genomic and transcriptomic profiles. JAX has developed cutting edge approaches to assist preclinical oncology in the effort to understand tumor response or resistance to treatment, including sequencing and analysis for genomic, transcriptomic, and expression data. This talk will cover the bioinformatics associated with detecting a variety of molecular alterations in patient tumors as well as the unique challenges associated with working with animal avatars.

4:00 Conference Adjourns

8 Clinical Research & Translational Informatics

TRANSFORMING BIOLOGICAL DATA TO CLINICAL DEVELOPMENT

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W5) Mobile Health, Virtual Research, Wearables, and Sensors; How to Accelerate Their Use and Adoption in Your Company

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W12) Leveraging Cloud Technologies to Enable Large-Scale Integration of Human Genome and Clinical Outcomes Data
(W14) 10 Common Analytical Mistakes Made by Biomedical Data Scientists and How to Avoid Them
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing
Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing
Sponsored by InterSystems

RESEARCH PLATFORMS, DATABASES AND ANALYTICS FOR IMPROVING TRANSLATIONAL RESEARCH

10:50 Chairperson's Remarks
Jay Bergeron, Director, Research and Early Clinical Development BT, Pfizer

11:00 eTRIKS: Successes, Surprises and Lessons on the Verge of Completing the Largest General Translational Informatics Public Private Partnership to Date
Jay Bergeron, Director, Research and Early Clinical Development BT, Pfizer
The European Translational Research Information and Knowledge Management Services (eTRIKS), Innovative Medicines Initiative project, is finishing a five year, 17 partner public private partnership to establish products, services and best practices for translational informatics. eTRIKS has delivered technical platforms, data curation and analytics services, data standards/security/ethical guidelines, confidentiality and materials transfer master agreements, focused marketing platforms to encourage patients to share their health information for translation research and a post-consortium sustainability plan.

11:30 Brain-CODE: A Large-Scale Neuroinformatics Platform Offering Full Research Integration
Brendan Behan, Ph.D., Program Lead, Informatics and Analytics, Ontario Brain Institute
Brain-CODE is an extensible large-scale platform that manages the acquisition, storage, processing and analytics of multidimensional data collected from various brain disorders. Brain-CODE offers full research integration in terms of ensuring received data meets particular standards, offering data curation processes and facilitating data sharing through best practice governance, privacy and security guidelines. Brain-CODE also offers the computational infrastructure for researchers to perform analytics on their data.

12:00 pm Text Mining in Translational Research: Bench to Bedside and Back Again
Sponsored by Linguamatics
David Milward, Ph.D., CTO, Linguamatics
Critical translational research data is locked in textual format, such as scientific literature, clinical trial reports or electronic health records. Text mining can extract and connect relevant clinical and scientific data, providing evidence for clinical decision making and helping select patients for clinical trials. We will finish with a specific customer use case: mining clinical trial data for drug repurposing.

12:15 Buzzwords Aside: Delivering on the Promise of Big Data in Precision Medicine
Sponsored by Signet Accel
Philip R.O. Payne, Ph.D., FACMI, CSO, Signet Accel
Precision medicine and big data are oft-used phrases in BioIT, but we'll transcend the hype to truly know their impact. Payne strips away industry buzzwords to share why both must be scaled to population levels—and what it takes to get there: Unfettered access to vast quantities of disparate data; advanced data science and machine learning; and new approaches to true interoperability.

12:30 Session Break

12:40 Luncheon Presentation I: Data, Ecosystems and Outcomes using Evidence from Bench to Bedside and Back
Sponsored by Deloitte.
Brett Davis, Principal, Deloitte Consulting, LLP
Whitney Kent, Director, Global Program Management, Takeda Pharmaceuticals
Arun Nayar, Executive Director, R&D Informatics, Amgen
Patrick Loerch, Senior Director, Data Scientist, Celgene
The health care ecosystem is in a period of unprecedented change. The combination of value-based reimbursement and precision medicine is directing treatment toward personalized care—and requiring a reorientation around insights. Our panel of industry experts will discuss their approaches to leveraging data, evidence, and knowledge assets to deliver relevant information to researchers for evidence-driven decision making.

8 Clinical Research & Translational Informatics

TRANSFORMING BIOLOGICAL DATA TO CLINICAL DEVELOPMENT

1:10 Luncheon Presentation II: Before and After: Improving Laboratory Fitness with BaseSpace® Clarity LIMS

Sponsored by illumina

Melanie Febrer, Product Manager, Illumina

BaseSpace Clarity LIMS is a laboratory information management system that helps laboratories track samples and optimize procedures and workflows. Attend this informative session in which Q2 Solutions, a leading global clinical trials laboratory services organization, discusses their lab before and after BaseSpace Clarity LIMS. Presenters will share lessons learned on the implementation, as well as how the LIMS is helping them streamline and optimize their procedures to provide even better service to their customers.

1:40 Session Break

NEW TECHNOLOGY SUPPORTING TRANSLATIONAL RESEARCH

1:50 Chairperson's Remarks

John Reites, Chief Product Officer and Partner, THREAD

1:55 Briefing on Pharma's Use of Mobile Health, Virtual Research, Wearables and Sensors

John Reites, Chief Product Officer and Partner, THREAD
Joe Dustin, Director, Mobile Health, Medidata
Georgia Mitsi, Senior Director, Search Evaluation, Digital Healthcare, Sunovion

This will be a briefing on how mobile health, virtual research, wearables and sensors are being used in Pharma-focused programs with insights covering lessons learned, data standards, privacy, platforms in use, adoption strategies and much more.

2:55 Artificial Intelligence Powered Data and Insights

Sponsored by INNOPLEXUS

Gunjan Bhardwaj, Ph.D., CEO, Innoplexus AG

Innoplexus is creating a new paradigm of continuous decision support in clinical research and translational informatics. Instead of mining data silos, with a novel meta ontology based AI approach, we are bringing in relevant data from all corners of the clinical research space. In addition to this, we are building continuous analytics applications for decision support.

3:10 Democratizing Mobile and Machine Learning in Translational Research

Sponsored by MEDABLE

Michelle Longmire, MD, CEO, Co-founder, Medable

Investigators want to leverage mobile for continuous and connected participant data collection. Dr. Longmire will address how Medable is enabling investigators to deploy mobile apps into research studies without any development and then harness the power of machine learning with cloud systems to derive insights from app enabled big data.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

NEW TECHNOLOGY SUPPORTING TRANSLATIONAL RESEARCH

4:00 Co-Presentation: Building JAX-CKB: A Public Resource Supporting Tumor Profiling and Complex Queries in Cancer Genomics

Daniel Durkin, Manager, Software Development Computational Sciences, The Jackson Laboratory for Genomic Medicine
Susan Mockus, Ph.D., Manager, Clinical Analytics & Curation, The Jackson Laboratory for Genomic Medicine

The Jackson Laboratory Clinical Knowledgebase (CKB) is a semi-automated/manually curated database of gene/variant annotations, therapy knowledge, diagnostic/prognostic information, and clinical trials related to oncology. Variants are connected to targeted therapies through an efficacy evidence annotation, which has been optimized to capture and query complex molecular tumor profiles. Public CKB is accessible at https://ckb.jax.org. Design considerations, data capture transparency, and the technological platform will be discussed.

4:30 Data-Driven Patient Recruitment with Real World Data at Roche pRED

Liping Jin, Data-Driven Recruitment Lead, Pharmaceutical Research & Early Development, Roche Innovation Center New York

With the increasing use of Real World Data (RWD) in the pharma industry, Data-Driven Recruitment (DDR) team at Roche Pharm Research & Early Development (pRED) would like to share our experience of integrating RWD with trial metrics data to optimize study protocol design and target patient recruitment strategy. While the team has received positive feedback from our business partners, we would like also to share the challenges to expanding the effort in broader US and international settings.

5:00 A High-Performance Solution for Clinical Genomics and Translational Medicine Research

Sponsored by IBM

Janis Landry-Lane, Worldwide SD and Flash Solutions, Life Sciences Industry, IBM Systems, IBM
Kathy Tzeng, Ph.D., STSM, Life Sciences & NGS Solution Enablement IBM Systems, IBM

Clinical genomics and translational medicine research is driving the need to process raw data from a sequencer, and integrate and analyze genomics with complex data sources in a timely manner. IBM's high performance, secure and scalable integrated NGS and analytics solutions can address the challenges, while minimizing IT resources. IBM will discuss the approach of building these integrated solutions and collaboration with customers and partners.

5:15 Democratizing Data Science: Balancing Flexibility and Usability for Scientific Applications

Sponsored by PerkinElmer For the Better

Jens Hoefkens, Director, Strategic Marketing, PerkinElmer Informatics, Inc

We will present a novel data analysis pipelining tool which enables even non-expert users to create complex scientific analysis pipelines. Having been developed entirely inside TIBCO Spotfire, the system leverages the full interactivity, connectivity, extensibility, and ease of use of the world's leading research data analysis platform.

8 Clinical Research & Translational Informatics

TRANSFORMING BIOLOGICAL DATA TO CLINICAL DEVELOPMENT

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM
Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

LEVERAGING REAL WORLD DATA FOR TRANSLATIONAL RESEARCH

10:30 Chairperson's Remarks

10:40 Transforming Early Pharmaceutical R&D Strategies with Real World Evidence

Cliona Molony, Ph.D., Senior Director & Head of Computational Biomedicine, Computational Biology & Biomedicine, Pfizer

Pfizer is capitalizing on maturing healthcare informatics, expanding sources of real world evidence, and deeply characterized populations to adopt value-driven and precision medicine R&D strategies. This presentation will discuss our approach and the informatics capabilities that are driving real world data into Pfizer's early drug discovery pipeline research efforts.

11:10 Co-Presentation: rHEALTH (Real World Evidence Health Analytics Hub): Transition Janssen from Opportunistic to Systematic Strategic Use of Real World Evidence

Xiaoying Wu, M.D., MS, Director, RWE IT CoE & Medical Informatics, Data Sciences, Janssen IT, Johnson & Johnson

Asha Mahesh, Senior Manager, Emerging Technologies, Janssen R&D IT, Johnson & Johnson

rHEALTH is a Janssen developed enterprise Real World Data (RWD) and analytics platform to address challenges that the Janssen RWE community is facing. The ecosystem consists of four components: a global RWE knowledge portal; a Smart Catalog; a data platform that ingests data from disparate data sources and transforms the data to a Common Data Model, and a tailored analytical environment. With the new enterprise RWE-generating platform rHEALTH, we reduced the time for data access and analysis from months to days, enabling business partners to deliver valuable research insights faster than previously possible.

11:40 Accelerating Research through Full Text Semantic Enrichment and Data Integration

Sponsored by Copyright Clearance Center

Mike Iarrobino, Product Management, Copyright Clearance Center

Anna Lyubetskaya, Data Scientist, Copyright Clearance Center

R&D-focused organizations are processing scientific content at scale to extract relationships between biological features. You will gain insights from diverse cases for applying full-text semantic enrichment integrated with other data sets at your organization. In this session, you'll learn how to: use a network environment where text, raw data, and structured data are easily connected to help you make informed decisions, save time and increase efficiency by uncovering and integrating expert knowledge computationally to address potential challenges that result from an increased volume of information and novel facts present in unstructured text and experimental output.

12:10 pm Session Break

12:20 Luncheon Presentation I: Addressing the Pain Points of Bioinformatics R&D in the Post-Genomic era: Lessons Learnt from the Pharmaceutical Industry

Sponsored by Genestack

Misha Kapushesky, Ph.D., CEO, Genestack

A key challenge in pharmaceutical R&D is to leverage high-throughput multi-omics data, combined with clinical data to inform the drug discovery process. This requires data federation, data governance, and scalable storage and compute, enabled by a modular environment which would fast-track genomics data analysis. Here, we present case studies illustrating how Genestack has helped its customers achieve these goals.

12:50 Luncheon Presentation II *(Sponsorship Opportunity Available)*

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

DATA VISUALIZATION FOR CLINICAL AND TRANSLATIONAL RESEARCH

1:55 Chairperson's Remarks

Christel Chehoud, Ph.D., Scientist, Data Sciences, Janssen

2:00 Framework for Management, Analysis, and Visualization of Matched Genetic and Clinical & Real World Datasets

Alexandra Dumitriu, Ph.D., Manager, Computational Biomedicine, Genome Sciences and Technologies, Pfizer

Coupled Electronic Health Records (EHR) and genetic datasets have recently become more accessible for exploratory clinical research. This advancement brings opportunities for genotype-to-phenotype (G2P) queries, but also poses challenges for researchers who need to address similar issues around data management, analysis, and visualization. The presentation will describe our approach and learning related to building an analytical framework focused on EHR-based G2P resources, which allows for streamlined protocols, including semi-automated computational cohort definitions and PheWAS analyses.

2:30 Clinical Trials Innovations in the Age of Big Data and Advanced Analytics

Christel Chehoud, Ph.D., Scientist, Data Sciences, Janssen

At Janssen, we have pioneered the application of technologies such as machine learning, natural language processing, and artificial intelligence to create novel solutions which have resulted in data driven efficiencies realized from predictive and prescriptive analytics on clinical trials data. This presentation will delve on aspects of this work and present vignettes to highlight the challenges and successes.

3:00 Big Data Analysis of Human Gliomas Using Oncoscape

Eric Holland, M.D., Ph.D., Director, Seattle Tumor Translational Research, Fred Hutchinson Cancer Research Center and University of Washington Medicine

We have developed an open access on-line tool for visualizing and interacting with large clinical/molecular datasets of cancer patients. This tool (oncoscape) collapses large data using MDS and connects tumors with the molecular alterations found in them and with clinical outcome. The analytic tabs are websockets that can be written independently. Subsets of patients or genetic alterations can be identified in one tab and further refined in other tabs.

3:30 Data Visualization for Decision Making in Pharma

Geoff Kremer, Director, CMR Informatics - Strategy & Operational Effectiveness, Novo Nordisk

Data Visualization tools are a key component to analyzing the myriad of data pharmaceutical companies receive and create each day. The visual display of data can help executives, managers, and analysts quickly identify trends and find hidden patterns within the data. Powerful visual analytics will enable more informed decision making, generate more thought provoking questions, and potentially provide actionable information. This talk will summarize various aspects related to data visualization, including: best practices, performance, and demo the creation of an ad hoc analysis using a data visualization tool.

4:00 Conference Adjourns

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

25 YEARS Organized by Cambridge Healthtech Institute

9 Data Visualization & Exploration Tools

GENOMICS, DRUG DISCOVERY, AND CLINICAL DEVELOPMENT

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W4) Data Visualization to Accelerate Biological Discovery

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W9) Data Science Driving Better Informed Decisions
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing

Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing

Sponsored by InterSystems

DEVELOPING AND IMPLEMENTING DATA VIZ AND EXPLORATION TOOLS

10:50 Chairperson's Remarks
Nils Gehlenborg, Ph.D., Assistant Professor, Biomedical Informatics, Harvard Medical School

11:00 Deploying OMERO at Harvard Medical School
Jay Copeland, Manager, Harvard Medical School Image Management Core, HMS IT, Harvard Medical School
OMERO is an open source server platform for managing, viewing, and sharing microscopy images and metadata. The HMS IT Department is deploying OMERO as a major new service targeted to supporting researchers and science. The goal is to make it easier for researchers to manage and share their microscope data and simplify collaboration using large, multi-terabyte, microscope image datasets. We also aim for HMS to contribute to the development of OMERO for the benefit of broader international community of scientists using OMERO.

11:30 A Lean Spotfire Implementation for Streamlining, Visualizing, and Analyzing Discovery Research Data
Yi Lin, Ph.D., Principal Scientist, Data Science V, pRED Informatics, Roche Pharma Research and Early Development, Roche Innovation Center Shanghai
We developed a Spotfire-based solution integrating multiple internal and external data sources to support decision making for drug discovery projects. An automated process extracts, streamlines, and transforms complex scientific data into a chemistry-friendly interface with customized views enabling interactive analysis of both molecular structures and biological data.

12:00 pm Sifting through the Noise of mHealth Data in Clinical Research

Sponsored by ERT

Nick Neri, Platform Manager, ERT
As drug development organizations incorporate wearables and other mHealth sensors into clinical trials, they face key challenges in analyzing the large volume of data these devices produce, and integrating these data with other patient data to create a coherent picture of patient outcomes. Nick Neri will present use cases on analyzing and visualizing integrated mHealth data in clinical research.

12:15 Enjoy Lunch on Your Own

1:40 Session Break

VISUALIZATION FOR BIOMEDICAL DATA

1:50 Chairperson's Remarks
G. Elisabeta Marai, Ph.D., Associate Professor, Electronic Visualization Lab, University of Illinois

1:55 Galaxy as a Platform for Visual Analytics
Aysam Guerler, PhD, Software Engineer, Taylor Lab., Johns Hopkins University
Galaxy (https://galaxyproject.org) is an open-source, Web-based scientific gateway for analyzing large biomedical datasets that is used by thousands of scientists worldwide. In this talk, I will describe how Galaxy makes it simple to add new visualizations and combine them with Galaxy datasets, tools, and workflows. Visualizations implemented as static scripts, Web-based dynamic displays, and even client-server applications can all be integrated into Galaxy.

9 Data Visualization & Exploration Tools

GENOMICS, DRUG DISCOVERY, AND CLINICAL DEVELOPMENT

2:25 Multi-Scale Visualization Tools for Exploration of Chromosome Interaction Data

Nils Gehlenborg, Ph.D., Assistant Professor, Biomedical Informatics, Harvard Medical School

How do you visualize a 3 million x 3 million matrix and allow users to explore features across a wide range of different scales? We built HiGlass, a web-based visualization tool for analysis of Hi-C and other genome-wide chromosome interaction data that enables comparison of multiple contact matrices and integration of other data types. In my talk, I will discuss several use cases and describe how we architected HiGlass.

2:55 Delivering Computational Chemistry to Cheminformatics: Collaborative Drug Discovery with LiveDesign

Sponsored by SCHRÖDINGER

Erin Davis, Ph.D., Senior Product Manager, Enterprise Informatics, Schrödinger

Reducing attrition in drug discovery means better predictions earlier in the pipeline, drawing on the ideas of the team's various expertise. LiveDesign facilities this by enabling collaborative computational chemistry in near real time delivered through a user-friendly web-based cheminformatics platform. Here we will demonstrate the streamlining of drug discovery through several successful use cases.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

VISUALIZATION FOR BIOMEDICAL DATA (CONT.)

4:00 Co-Presentation: Multiple and Extensively Drug-Resistant Tuberculosis Data Exploration Portal (MXDR-TB DEPOT)

Michael Harris, Senior Informatics Scientist, Bioinformatics and Computational Biosciences Branch (BCBB), NIAID, NIH

Darren Schneider, Senior Manager, Analytics + Information Management, Deloitte Consulting LLP

The National Institute of Allergy and Infectious Diseases (NIAID) and multiple vendors worked collaboratively to design and develop an analytics portal to support hypothesis generation and testing aimed at improving TB patient diagnostics and outcomes. The publicly available solution enables clinicians and researchers to create and compare cohorts of patients based on clinical, socioeconomic, genomic, and diagnostic image data.

4:30 Co-Presentation: Multiple and Extensively Drug-Resistant Tuberculosis Data Exploration Portal (MXDR-TB DEPOT)

Michael Harris, Senior Informatics Scientist, Bioinformatics and Computational Biosciences Branch (BCBB), NIAID, NIH

Darren Schneider, Senior Manager, Analytics + Information Management, Deloitte Consulting LLP

The National Institute of Allergy and Infectious Diseases (NIAID) and multiple vendors worked collaboratively to design and develop an analytics portal to support hypothesis generation and testing aimed at improving TB patient diagnostics and outcomes. The publicly available solution enables clinicians and researchers to create and compare cohorts of patients based on clinical, socioeconomic, genomic, and diagnostic image data.

5:00 Journey to the Center of the Nucleus: Exploring 3D Genomic Datasets with Juicebox

Muhammad Saad Shamim, MD/PhD Candidate, Medical Scientist Training Program, Baylor College of Medicine/Rice University

Juicebox is a tool for exploring contact maps generated using Hi-C and other 3D genome-sequencing technologies; it allows users to zoom in and out interactively and supports a variety of annotation tools, enabling researchers to more accurately examine genomes and how they fold in 3D. This talk will explore ways to use Juicebox as well as the types of data sets that can be examined.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM

Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

CONSIDERATIONS FOR VISUALIZATION OF BIOINFORMATICS DATA

10:30 Chairperson's Remarks

Baohong Zhang, Ph.D., Director of Clinical Bioinformatics, Early Clinical Development, Pfizer, Inc.

10:40 Data Visualization: Make Every Data Point Alive

Baohong Zhang, Ph.D., Director of Clinical Bioinformatics, Early Clinical Development, Pfizer, Inc.

This presentation will describe interactive visualization of next-generation sequencing data using the latest web 2.0 techniques, such as jQuery, D3 and many other javascript libraries. I will be using single cell RNA-seq data as showcase to demonstrate the easy-to-share, publication-ready, server-less, internet-less, reproducible data exploration tool.

9 Data Visualization & Exploration Tools

GENOMICS, DRUG DISCOVERY, AND CLINICAL DEVELOPMENT

11:10 Big Display Visualization of Bioinformatics Data

G. Elisabeta Marai, Ph.D., Associate Professor, Electronic Visualization Lab, University of Illinois

Visualization is an increasingly important component in the effective analysis of large biological datasets. However, visualization of large datasets also suffers from scalability issues. While multi-tiled and high-resolution displays have the potential to address scalability issues, new approaches are needed to take advantage of such environments, in order to enable the effective visual analysis of large biological datasets. In this talk, I describe our group's work on designing novel and scalable systems for the visual analysis on I

11:40 Sponsored Presentation *(Opportunity Available)*

12:10 pm Session Break

12:20 Luncheon Presentation *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

DATA VISUALIZATION FOR CLINICAL AND TRANSLATIONAL RESEARCH

1:55 Chairperson's Remarks

Christel Chehoud, Ph.D., Scientist, Data Sciences, Janssen

2:00 Framework for Management, Analysis, and Visualization of Matched Genetic and Clinical & Real-World Datasets

Alexandra Dumitriu, Ph.D., Manager, Computational Biomedicine, Genome Sciences and Technologies, Pfizer

Coupled Electronic Health Records (EHR) and genetic datasets have recently become more accessible for exploratory clinical research. This advancement brings opportunities for genotype-to-phenotype (G2P) queries, but also poses challenges for researchers who need to address similar issues around data management, analysis, and visualization. The presentation will describe our approach and learning related to building an analytical framework focused on EHR-based G2P resources, which allows for streamlined protocols, including semi-automated computational cohort definitions and PheWAS analyses.

2:30 Clinical Trials Innovations in the Age of Big Data and Advanced Analytics

Christel Chehoud, Ph.D., Scientist, Data Sciences, Janssen

At Janssen, we have pioneered the application of technologies such as machine learning, natural language processing, and artificial intelligence to create novel solutions which have resulted in data driven efficiencies realized from predictive and prescriptive analytics on clinical trials data. This presentation will delve on aspects of this work and present vignettes to highlight the challenges and successes.

3:00 Big Data Analysis of Human Gliomas Using Oncoscape

Eric C. Holland, M.D., Ph.D., Senior Vice President, Director, Human Biology, Seattle Translational Tumor Research, Fred Hutchinson Cancer Research Center

We have developed an open access on-line tool for visualizing and interacting with large clinical/molecular datasets of cancer patients. This tool (oncoscape) collapses large data using MDS and connects tumors with the molecular alterations found in them and with clinical outcome. The analytic tabs are websockets that can be written independently. Subsets of patients or genetic alterations can be identified in one tab and further refined in other tabs.

3:30 Data Visualization for Decision Making in Pharma

Geoff Kremer, Director, CMR Informatics - Strategy & Operational Effectiveness, Novo Nordisk

Data Visualization tools are a key component to analyzing the myriad of data pharmaceutical companies receive and create each day. The visual display of data can help executives, managers, and analysts quickly identify trends and find hidden patterns within the data. Powerful visual analytics will enable more informed decision making, generate more thought provoking questions, and potentially provide actionable information. This talk will summarize various aspects related to data visualization, including: best practices, performance, and demo the creation of an ad hoc analysis using a data visualization tool.

4:00 Conference Adjourns

Final Weeks to Register

CLICK HERE TO
Register Online!

Organized by Cambridge Healthtech Institute

10 Pharmaceutical R&D Informatics

COLLABORATION, DATA SCIENCE AND BIOLOGICS

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Morning Pre-Conference Workshops*

12:30 – 4:00 pm Afternoon Pre-Conference Workshops*

** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION

Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing

Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION

Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing

Sponsored by InterSystems

DATA INTEGRATION, ANALYTICS & COLLABORATION

10:50 Chairperson's Remarks

Yuriy Gankin, Vice President, CSO, Life Sciences, EPAM

11:00 Pharma R&D: Challenges & Opportunities

Ingrid Akerblom, Ph.D., Executive Director, Analytics and Knowledge Management, Amgen

The challenges and opportunities for pharmaceutical R&D have never been greater. The drive to bring value to stakeholders whether patients, payors or healthcare providers from our therapies is transforming our approach to data and analytics for insight generation. This presentation will explore the nature of that change and how it is leading to greater global collaboration internally and with partners.

11:30 From Big Data to Smart Data: Advancing a Data First Strategy

M. Hall Gregg, Ph.D., Vice President, Research & Development, Business Technology, Pfizer, Inc.

The pressure on pharmaceutical R&D organizations to accelerate innovation, to deliver quality medicines at the lowest possible cost, and to demonstrate the value of their medicine has never been greater. Enabling a culture where data is an enterprise asset is critical to being responsive to these challenges. A data first strategy focused on delivering data liquidity, agility, reuse and self-service allows data to be an enterprise asset and big data becomes smart data.

12:00 pm A Paradigm Shift in Informatics to Platform-Based Solutions

Sponsored by CORE INFORMATICS part of Thermo Fisher Scientific

John Stalker, Director, Core Informatics (part of Thermo Fisher Scientific)

The pace of scientific change is accelerating far beyond the capabilities of traditional data management systems. A laboratory informatics platform should make life easier by removing barriers to progress and supporting a company's unique workflows. This talk will discuss the benefits of a platform-based approach to informatics and how it adds value to organizations developing the next generation of biotherapeutics.

12:15 Smart Data Lakes - Transforming Pharma R&D, Customer Success Stories

Sponsored by CAMBRIDGE SEMANTICS

Jim LaPointe, Managing Director, Pharma & Life Sciences, Cambridge Semantics

We will present use cases in production at some top tier Pharma companies and how they are leveraging Smart Data Lakes in cutting drug development times by years and realizing mammoth savings.

12:30 Session Break

12:40 Luncheon Presentation I: The Five Key Technologies Required to Deliver Global Mega-Scale Biomedical Data Analysis

Sponsored by databiology

Georges Heiter, Founder, Databiology

Dealing with the variety of biomedical data and real-world evidence by managing scale and complexity. Ensuring agility and simplicity to conduct any analysis on any data at any time. Delivering end-to-end provenance and reproducibility. Managing integrity & security of data and access. Enabling the flexibility to collaborate at any level and at any location.

1:10 Luncheon Presentation II: Next Gen Research Platform - NGRP

Sponsored by HCL

Anil Verma, Vice President, Life Sciences, HCL

1:40 Session Break

DATA ANALYTICS & DATA INTEGRATION

1:50 Chairperson's Remarks

Peter Covitz, Ph.D., Senior Director, Research & Translational IT, Biogen

1:55 If Data Is the Lifeblood of Scientific Research, Can We Breathe New Life into Pharmaceutical R&D by Improving the Circulation and Use of Data?

Bryan Takasaki, Ph.D., Director, R&D Information US Lead, AstraZeneca

Progress in the drive to correlate biological mechanism and disease, preclinical and clinical results as well as clinical and real world evidence requires effective data sharing and collaboration between academia and multiple industries. This talk will explore the opportunities and challenges in addressing the shortcomings of the current scientific data sharing and collaboration paradigm.

2:25 Integration and Use of Thousands of Preclinical and Clinical Studies at Novartis

Philippe Marc, Ph.D., Director, Global Head Integrated Data Sciences, Novartis Institute for Biomedical Research

10 Pharmaceutical R&D Informatics

COLLABORATION, DATA SCIENCE AND BIOLOGICS

The Novartis translational medicine study data-warehouse became, in the past years, an essential tool for hundreds of researchers. It is fed every day with data from Novartis-sponsored preclinical and clinical studies, generated in house or at CROs. Basic architecture and use will be described.

2:55 From HELM to Applications: ChemAxon's Way into the World of Large Molecule Informatics

Sponsored by ChemAxon

Aurora Costache, Ph.D., Application Scientist, ChemAxon

HELM is a success story in pre-competitive collaboration. Uptake is increasing in the industry to improve description of biological entities in informatics systems. ChemAxon contributes since the beginning, defining extensions, providing support for HELM user community. This talk will give an overview about ChemAxon's involvements in the HELM project, and introduce the tools built on these grounds through use cases.

3:10 Drive Innovation, Accelerate Research and Shorten Drug Time-to-Market with Cognitive Search & Analytics

Sponsored by SINEQUA

Laurent Fanichet, Vice President, Marketing, Sinequa

Martin Leach, Ph.D., Vice President, R&D IT, Enterprise Data Management & Analytics, Alexion Pharmaceuticals

In the highly competitive world of biopharma, your organization is likely to cope with hundreds of millions of documents, including lab and clinical trial reports, publications, patent filings, and emails, as well as billions of database records from internal and external trade sources. Learn how Cognitive Search & Analytics can help drive innovation, speed up submission of new drugs.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

DATA ANALYTICS & DATA INTEGRATION

4:00 Informatics Enabling Chemoproteomics and Knowledge Driven MOA Analytics

Xudong Qiao, Senior Specialist, Chemistry Pharmacology & HTS Informatics IT, Merck Research Labs IT.

The application of chemical biology techniques, tools, and analyses to the investigation of disease biology can aid in the search for novel therapeutic targets and a much increased understanding of target and pathway biology. Merck has developed information technology solutions to automate chemoproteomics experiment workflows, and enable data processing and analysis, in order to establish a knowledgebase needed for data mining and decision-making.

4:30 Using Salesforce for Outsourcing Scientific Data Curation

Angelika Fuchs, Ph.D., Head, Large Molecule Workflows, Roche pRED Informatics

We have developed a novel data curation application based on the popular Salesforce platform to enable the outsourcing of data curation tasks to CROs with the required expert knowledge. The application provides easy access to both internal and external users and is fully integrated with in-house terminology services and down-stream systems to facilitate data availability for review and further processing by in-house scientists.

5:00 Transforming Disparate Data into Collective Insights

Sponsored by ELSEVIER

Frederik van den Broek, Ph.D., Consultant, Research & Development Solutions, Elsevier Gregory Landrum, Ph.D., Vice President, Life Sciences, KNIME.com AG

Big data in R&D is simply today's data. Efficiently and reproducibly learning from this data requires us to combine different types of information from various sources and to apply advanced analytics and visualisation to the resulting data set(s) using integrated workflow and analytics platforms. This talk will present use cases showing how one can obtain these insights which are vital for R&D decision-making in today's competitive drug discovery environment.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM

Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

FROM TRANSLATIONAL TO REAL WORLD - THE IMPACT ON R&D

10:30 Chairperson's Remarks

Tom Plasterer, Ph.D., Director, US Cross-Science Lead & Open PHACTS Lead, AstraZeneca

10:40 Transforming Early Pharmaceutical R&D Strategies with Real World Evidence

Cliona Molony, Ph.D., Senior Director & Head of Computational Biomedicine, Computational Biology & Biomedicine, Pfizer

Pfizer is capitalizing on maturing healthcare informatics, expanding sources of real world evidence, and deeply characterized populations to adopt value-driven and precision medicine R&D strategies. Together these resources (a) fuel a deeper understanding of disease subtypes to dissect novel mechanisms, (b) help to assess the burden of patient co-morbidities and adjacencies for drug development opportunities, and (c) enable precision medicine approaches to test therapeutic hypotheses matched to the pathology of patients.

11:10 Co-Presentation: rHEALTH (Real World Evidence Health Analytics Hub): Transition Janssen from Opportunistic to Systematic Strategic Use of Real World Evidence

Xiaoying Wu, M.D., MS, Director, RWE IT CoE & Medical Informatics, Data Sciences, Janssen IT, Johnson & Johnson

Asha Mahesh, Senior Manager, Emerging Technologies, Janssen R&D IT, Johnson & Johnson

rHEALTH is a Janssen developed enterprise Real World Data (RWD) and analytics

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

Organized by Cambridge Healthtech Institute

platform to address challenges that the Janssen RWE community is facing. The ecosystem consists of four components: a global RWE knowledge portal, a Smart Catalog, a data platform that ingests data from disparate data sources and transforms the data to a Common Data Model, and a tailored analytical environment. With the new enterprise RWE-generating platform rHEALTH, we reduced the time for data access and analysis from months to days, enabling business partners to deliver valuable research insights faster than previously possible.

11:40 Accelerating Research through Full Text Semantic Enrichment and Data Integration

Sponsored by Copyright Clearance Center

Mike Iarrobino, Product Management, Copyright Clearance Center
Anna Lyubetskaya, Data Scientist, Copyright Clearance Center

R&D-focused organizations are processing scientific content at scale to extract relationships between biological features. You will gain insights from diverse cases for applying full-text semantic enrichment integrated with other data sets at your organization. In this session, you'll learn how to: use a network environment where text, raw data, and structured data are easily connected to help you make informed decisions, save time and increase efficiency by uncovering and integrating expert knowledge computationally to address potential challenges that result from an increased volume of information and novel facts present in unstructured text and experimental output.creased volume of information and novel facts present in unstructured text.

12:10 pm Session Break

12:20 Luncheon Presentation I: Addressing the Pain Points of Bioinformatics R&D in the Post-Genomic Era: Lessons Learnt from the Pharmaceutical Industry

Sponsored by Genestack

Misha Kapushesky, Ph.D., CEO, Genestack

A key challenge in pharmaceutical R&D is to leverage high-throughput multi-omics data, combined with clinical data to inform the drug discovery process. This requires data federation, data governance, and scalable storage and compute, enabled by a modular environment which would fast-track genomics data analysis. Here, we present case studies illustrating how Genestack has helped its customers achieve these goals.

12:50 Luncheon Presentation II *(Sponsorship Opportunity Available)*

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

FROM TRANSLATIONAL TO REAL WORLD - THE IMPACT ON R&D

1:55 Chairperson's Remarks

Ralph Haffner, Head Discovery Workflows Basel - pRED Informatics, Roche Pharma Research and Early Development, F. Hoffmann-La Roche Ltd.

2:00 Digital Biomarker Development at Roche pRED - Using mSensors to Gain New Insights into Neurological Disorders

Christian Gossens, Ph.D., Global Head Early Development Workflows, Roche pRED Informatics, F. Hoffmann-La Roche Ltd.

Roche is pioneering a smartphone-based remote monitoring system for patients with e.g. Parkinson's disease. It complements the conventional physician-led assessments in the clinical trial, which are resource-intensive and represent only a snapshot in time. The data collected provide information on patients' symptom fluctuations and disease impact on daily living. This presentation will discuss what it took to collect such data and convert raw data into clinical insights.

IDENTIFICATION AND PRIORITIZATION OF TARGETS IN DRUG DISCOVERY & DEVELOPMENT

2:20 Pfizer's Causal Reasoning Engine: Generating Specific and Testable Hypotheses from Large-Scale 'Omics Data

Enoch S. Huang, Ph.D., Executive Director, Head of Computational Sciences, Pfizer Worldwide Research and Development

Ease of large-scale generation of different 'omics data types for model systems and patients has driven the need for computational methods that offer ease-of-interpretation as well as robust results. In this talk, I will present Pfizer-developed approaches based on causal networks assembled from prior knowledge for the prediction of upstream regulators, boosting genetic signals, and robust patient stratification.

2:40 Navigating through an Ocean of Targets—How Informatics Can Help to Advance an R&D Pipeline

Jonathan Dewey, M.D., Associate Director of R&D Information Technology, Biogen

Drug discovery organizations advancing a portfolio of targets through an R&D pipeline require evidence beyond genetic associations and genomic correlations to disease states. Biogen scientists accelerate the process via target prioritization using network analysis, 'omics data integration, and computational pathway engines to discover novel hypotheses and mechanisms of action critical to advancing each target.

ENSURING HIGH QUALITY DATA FOR GOOD DECISION-MAKING

3:00 BioPharma Adoption of FAIR Data, a Collaborative Advantage

Tom Plasterer, Ph.D., Director, US Cross-Science Lead & Open PHACTS Lead, AstraZeneca

The concept of FAIR (Findable, Accessible, Interoperable and Reusable) data is becoming a reality as stakeholders from industry, academia, funding agencies and publishers are embracing this approach. For BioPharma being able to effectively share and reuse data is a tremendous competitive advantage, within a company, with peer organizations, key opinion leaders and regulatory agencies. A few key drivers, success stories and prognostications for FAIR data will be presented.

3:30 Enabling Collaborative Analysis and Decision-Making between Business and R&D Groups in the Pharmaceutical Industry

Robert Cain, Ph.D., Informatics Investigator, Allergan plc

In this talk I will discuss tools to bring together pharmacology, genomic and chemistry research data along with competitive intelligence, patent, financial and other data and make it computable. I'll also describe our open science model and our distributed decision making and tools we use to share insights between groups. Steps to define drug landscapes and collaboratively mine target, indication and drug information will be discussed along with our repurposing tools.

4:00 Conference Adjourns

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO
Register Online!

Organized by Cambridge Healthtech Institute

Clinical Genomics

DATA ANALYTICS IN A DYNAMIC LANDSCAPE

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W2) An Intro to Blockchain in Life Sciences

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W12) Leveraging
Technologies to Enable Large-Scale Integration of Human Genome and Clinical Outcomes Data
(W13) Proteogenomics: Integration of Genomics and Proteomics Data
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing
Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing
Sponsored by InterSystems

CLASSIFYING VARIANTS: PERSONAL > POPULATION GENOMICS

10:50 Chairperson's Remarks
Narges Bani Asadi, Ph.D., Vice President and Life Cycle Leader, Roche Sequencing Solutions

11:00 KEYNOTE PRESENTATION: Analysis of Personal Genomes
Mark Gerstein, Ph.D., Albert L. Williams Professor, Biomedical Informatics, Molecular Biophysics & Biochemistry, and Computer Science; Co-Director, Computational Biology & Bioinformatics, Yale University
I discuss how we interpret and prioritize the many variants in a personal genome. I concentrate on rare and somatic variants and discuss ways of assessing their impact in coding regions using protein structure. For non-coding regions, I present ways to analyze the overall burden of mutations, finding allelic elements differentially affected by variants in paternal and maternal chromosomes and using network connectivity to prioritize variants.

11:30 The Clinical Cloud: An Industry View
Pamela Duffy, MS, Director, Core Clinical Solutions, IT, Pfizer
Aggregation of clinical data in the cloud to create a single source of the truth is here and maturing. Is it hype or is the business case for savings and insights proving out?

12:00 pm Clinical Genomics for Research Questions
Sponsored by CONGENICA
Alan Martin, Head, Innovation & Data Science, Congenica
Bringing high throughput DNA sequencing into clinical practice presents an opportunity to analyse aggregate patient data to find new biological understanding. Potential analyses will be presented as well as the pitfalls of working with heterogeneous data sources and some data integration methods used by Congenica.

12:15 Sponsored Presentation *(Opportunity Available)*

12:30 Session Break

12:40 Luncheon Presentation I: Low Abundance Calls: Using Group Variant Calling to See Through the Noise and Find Value that Drives Translational Discovery
Sponsored by paradigm4
Zachary Pitluk, Ph.D., Vice President, Life Sciences and Healthcare, Paradigm4
Large-scale group-based somatic mutation calling using tunable ROC curves produces high quality identification of low abundance calls. These low abundance calls support accurate determination of genomic and transcriptional heterogeneity, providing a window into cell population changes from cell line cloning, disease progression, treatment response and adaptation, and biomanufacturing environmental factors. New scalable computational platforms like SciDB enable computationally efficient and validatable support for executing this method over groups of thousands of aligned sequences, as well as for other common variant analyses on a million whole exomes.

1:10 Luncheon Presentation II *(Sponsorship Opportunity Available)*

1:40 Session Break

CLASSIFYING VARIANTS: ADVANCING INTO THE CLINIC

1:50 Chairperson's Remarks
Leonard Lipovich, Ph.D., Associate Professor, Center for Molecular Medicine and Genetics, Wayne State University

1:55 Taking the Next Step: Computational Challenges Overcome in Rapidly Scaling Output from a Carrier Screening Program within 6 Months
Andrew McLellan, Ph.D., Assistant Professor, Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai
The Genetic Testing Laboratory at MSSM has increased its volume of carrier-screening five-fold this past year. Ramping up to processing data from over 200 HiSeq2500 rapid run flow cells per month has presented numerous challenges regarding data analysis, storage, automation and development of the tools required to provide a timely service without compromising quality. This talk documents our journey of discovery as we embraced high-throughput NGS-based genetic testing.

2:25 A Holistic Approach to Pediatric Rare Disease: Integration of Genomic Data for Research, Diagnostics, and Therapy
Timothy Yu, M.D., Ph.D., Instructor, Pediatrics, Harvard Medical School; Medicine Research, Division of Genetics and Genomics, Boston Children's Hospital
The talk will explore how genomic systems can be applied to manage unsolved rare disease cases. Whole exome sequencing has a diagnostic yield of 20-40% in complex genetic cases which leaves up to 80% of the cases unsolved. To address this challenge, it requires new genomic medicine systems that create an integration between laboratory diagnostics, clinical care and ongoing collaborative clinical research.

2:55 Big-Data Analytics and Genomics to Optimize Infertility Treatment
R. Mark Adams, Ph.D., CIO, Celmatix Inc.
Dr. Adams will discuss how Celmatix Inc., a next-generation company focused on women's health, is leveraging genomics and big data to explore the underlying biological causes of infertility. Celmatix has leveraged these capabilities to create and market an industry-leading genetic test that reveals how DNA may influence a woman's ability to have a baby today and in the future.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

CLASSIFYING VARIANTS: UNDERSTANDING AND TREATING COMPLEX DISEASE

4:00 Genetic Data Sharing to Advance Complex Disease Research for Type 2 Diabetes
Jason Flannick, Ph.D., Senior Group Leader, Medical and Population Genetics, Broad Institute of Harvard and MIT
I focus on our efforts to understand the genetic and biological basis of type 2 diabetes. I describe progress from large-scale sequencing studies to characterize the genetic basis of T2D, as well as to translate association signals to biological or clinical insights. I discuss a forward paradigm to maximize the utility of these data: through data sharing via a public and broadly accessible knowledge portal of T2D genetics.

4:30 From GWAS to Post-Genomic Personalized Therapeutics: Integrated Variant Annotation and Validation Reveals a Functional and Targetable Long Non-Coding RNA in Type 2 Diabetes
Leonard Lipovich, Ph.D., Associate Professor, Center for Molecular Medicine and Genetics, Wayne State University
We report computational discovery and experimental validation of an exonic polymorphism in a novel, primate-specific, evolutionarily non-conserved, long non-coding RNA (lncRNA), a leading type 2 diabetes direct causal candidate from GWAS. This lncRNA is expressed in the liver, where it regulates glycogen storage and hence fasting glucose levels. Since *in vivo* liver delivery of therapeutic oligonucleotides targeting this lncRNA is feasible, our results herald a new treatment for type 2 diabetes.

5:00 KEYNOTE PRESENTATION:
Using Networks to Link Genotype to Phenotype
John Quackenbush, Ph.D., Professor, Biostatistics, Harvard T.H. Chan School of Public Health; Professor, Biostatistics and Computational Biology, Dana-Farber Cancer Institute
Network and systems methods, combined with unprecedented large-scale data sets, are providing us with a unique opportunity to explore the associations between genotype and phenotype. Using data from the Genotype-Tissue Expression project (GTEx), we will explore the factors that influence the manifestation of complex traits.

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM
Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

Clinical Genomics

DATA ANALYTICS IN A DYNAMIC LANDSCAPE

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

CLASSIFYING VARIANTS: CANCER BIOMARKER INTERPRETATION AND CLASSIFICATION

10:30 Chairperson's Remarks

Honey Reddi, Ph.D., FACMG, Clinical Laboratory Director, The Jackson Laboratory

10:40 Fusion RNAs and Their Implications in Cancer Diagnosis

Hui Li, Ph.D., Associate Professor, Pathology, University of Virginia

Gene fusions and fusion products have been thought to be cancer-unique features. However, our recent work on intergenic splicing challenged this prevailing view. These findings complicate the usage of fusion RNAs as a whole for cancer diagnosis and treatment. On the other hand, the novel mechanism represents a new repertoire for the discovery of new biomarkers and drug targets.

11:10 Somatic Variants – Approaches to Predicting Actionability

Honey Reddi, Ph.D., FACMG, Clinical Laboratory Director, The Jackson Laboratory

There are currently no defined algorithms in place for the interpretation and classification of somatic variants, unlike the ACMG guidelines for germline variants. We are currently evaluating a series of steps that can be implemented for the interpretation and classification of somatic cancer variants keeping in mind the factors that define actionability.

11:40 Unlock Asian Giant's Precision Medicine Computing Potential – Intel's Bio-IT Know-How and Footprints in China

Jian Li, Ph.D., Life Science Business Lead, Intel Health & Life Science, Greater Asia

Chang Yu, Life Science Solution Architect, Intel Health & Life Science, Greater Asia

China has become a new battlefield for global precision medicine, which could be largely attributed to it having the largest patient population in the world, the government's endorsement and huge funding support, the booming market needs on NIPT and cancers, and also the emerging genomics giant players, e.g., Beijing Genomics Institute and Novogene

12:10 pm Session Break

12:20 Luncheon Presentation *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

CLASSIFYING VARIANTS: ANALYZING > INTERPRETING > STORING > SHARING DYNAMIC DATA

1:55 Chairperson's Remarks

Liz Worthey, Ph.D., Faculty Investigator & Director, Software Development and Informatics, HudsonAlpha Institute for Biotechnology

2:00 KEYNOTE PRESENTATION: Challenges and Opportunities for Global Genomic Sharing: Lessons from Current Successes

John Mattison, M.D., Assistant Medical Director, Chief Medical Information Officer, Kaiser Permanente

2:30 FEATURED PRESENTATION:

A Comprehensive Approach to Rare Disease Interpretation: Moving beyond Identification of a Causal Variant

Liz Worthey, Ph.D., Faculty Investigator & Director, Software Development and Informatics, HudsonAlpha Institute for Biotechnology

We discuss application of methods for analyzing, storing, and interpreting variant data that allows us to move beyond MDx related to a single variant to support a refined understanding of human health and disease. The discussion includes methods to: 1) support increased definitive diagnosis rates, 2) explore genotype to phenotype associations that can be used to predict disease course in a particular individual, and 3) identify modifiers that alter clinical presentation.

3:00 A Curated Database of 150,000 Variants from Clinical Whole-Genome Sequencing: What Does It Take to Maintain and Donate to ClinVar?

Aditi Chawla, Ph.D., Senior Scientist, Clinical Services Laboratory, Illumina

We present the workflow for variant classification, including an autoscore system that calculates a variant's potential to have caused the associated disease, search for literature on variants, and manual clinical curation based on ACMG guidelines. We discuss strategies for maintaining and updating variant classifications over time, laboratory policy for updating clinical reports as variant classification changes, process for donating to ClinVar, and summary of our ClinVar donation.

3:30 Community-Driven Approaches to Gene Curation

Marina DiStefano, Ph.D., Variant Scientist, Laboratory for Molecular Medicine, Partners HealthCare

As precision medicine expands and whole exome sequencing becomes standard, it is critical to define a common framework to assess the evidence for gene-disease associations. Testing this framework involves reaching out to the experts in various specialties for validation. With input from the genetic community, genes can be transparently and systematically evaluated and prioritized for analysis in various clinical contexts.

4:00 Conference Adjourns

12 Cancer Informatics

APPLYING COMPUTATIONAL BIOLOGY TO CANCER RESEARCH & CARE

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Morning Pre-Conference Workshops*

12:30 – 4:00 pm Afternoon Pre-Conference Workshops*

** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION

Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing

Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION

Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing

Sponsored by InterSystems

BIG DATA ANALYTICS ON CANCER GENOMIC DATA

10:50 Chairperson's Remarks

Bhanu Bahl, Ph.D., Director, Clinical and Translational Science Centre, Harvard Medical School

11:00 Deep Analytics of Cancer Genomics Data

Eric Neumann, Ph.D., Vice President, Knowledge Informatics, Technology, Foundation Medicine

The very large collection of cancer genomic data within Foundation Medicine allows us to see not only breadth of mutational variation, but re-occurring and significant patterns in many different cancers. These different mutational modes potentially have great relevance to more actionable diagnoses, as well as providing deeper insights for cancer research and drug development. We are exploring a new method to map each case into a statistically-defined multidimensional space along with functional and response associations.

11:30 Building a Harmonized Clinical Data Resource for Cancer Research and Translational Medicine: Challenges, Goals and Progress

John Methot, Director, Health Informatics Architecture, Dana-Farber Cancer Institute

Oncology's pioneering role in precision medicine has led to routine tumor sequencing. However, the incompleteness and inconsistency of data from EMRs and other sources presents significant challenges to building the patient-specific diagnosis, treatment and outcome "stories" that are required as correlates to genomic data to inform outcomes and clinical quality research. At DFCI, we are striving to improve the quality and coverage of clinical data for research and translational medicine, including decision support in the clinic.

12:00 pm Graph Genome Tools for Precision Medicine

Sponsored by SevenBridges

Jack DiGiovanna, Senior Scientist, Director of Program, Seven Bridges

Constructing a genome graph captures variations within a population as branches which diverge from the common reference but later rejoin it. This approach is fundamentally different than alignment and variant calling against a single linear reference, but does it matter for precision medicine? Here, we explore two key features of graph technology - improved variant calling and capturing distributions of variations in a population. We project these features onto the current state of precision medicine.

12:30 Session Break

12:40 Luncheon Presentation I: NGS in the Fast Lane: How a Collaborative Approach between the Mayo Clinic and Illumina Is Advancing Genetic Testing

Sponsored by illumina

Michael Ball, Vice President, EIBU Commercial, Illumina

Despite progress in sequencing technologies, obstacles still prevent a systematic method for deriving actionable information from an individual's genome. The BaseSpace® Suite from Illumina was designed to eliminate obstacles and enable users to easily store, manage, and analyze genomic data. In this session, learn how the Mayo Clinic is using BaseSpace Suite to expedite the delivery of its genomic expertise, and how Illumina is applying feedback from the Mayo Clinic to make the BaseSpace Suite even more efficient.

1:10 Making TCGA Tumor Datasets Accessible for Real-Time Multiomics Analysis and Visualization

Sponsored by WuXiNextCODE Leading Genomics

Jim Lund, Ph.D., Director, Tumor Product Development, WuXi NextCODE

The collection of cancer datasets in The Cancer Genome Atlas (TCGA) dataset has provided a wealth of insight into cancer biology; however, analyzing such a complex dataset is challenging for most investigators. This talk will demonstrate the Wuxi NextCODE approach using our proprietary genomically ordered relational (gor) database architecture to rapidly and simultaneously explore the diverse data types in TCGA.

1:40 Session Break

DESIGNING CLOUD FLEXIBILITY WHILE MAINTAINING SECURITY

1:50 Chairperson's Remarks

Lance Smith, MBA, Associate Director, IT, Celgene

1:55 How BMS Uses cBioPortal for Cancer Genomics Research

Isaac Neuhaus, Ph.D., Director, Computational Genomics, Bristol-Myers Squibb

BMS has started using cBioPortal for visualizing cancer genomics datasets early

2016, supported by The Hyve, an open source bioinformatics company based in The Netherlands. The cBioPortal server runs on Amazon AWS and is tied to the company Active Directory for authentication. Currently loaded datasets are TCGA, CCLE and M2Gen and we will be loading clinical trial data later this year. Concurrently we are running a pilot for cBioPortal with mouse data.

2:25 Enabling Celgene's Innovation While Protecting the Enterprise in the Cloud

Lance Smith, MBA, Associate Director, IT, Celgene

Celgene uses Amazon Web Services (AWS) to run R&D workloads, enterprise systems, and collaboration environments securely and all controlled from a single pane. Topics covered include network design, HPC and job scheduling, workload isolation, storage, automated security, and environment auditing. We also discuss business use cases, challenges overcome (technical, legal, security, and organizational), and Turbot guardrails for management.

2:55 Automated High-Scale Computational Infrastructure

Sponsored by OpenEye Scientific

Jharrod LaFon, Chief Cloud Engineer, OpenEye Scientific

Large-scale computation is an important aspect of drug design, and increasingly is cloud-based. The true value of the cloud comes not just from computation but by enabling company-wide collaboration, the ability to build, modify, share, manage versions, and evaluate methods. OpenEye's cloud-native platform, Orion, and its workflow engine, Floe, provide the tools to create innovative approaches for computation and collaboration.

3:10 Research Informatics: Get Ready for the Cloud!

Sponsored by Dassault Systèmes

Ton van Daelen, ScienceCloud, Product Director, Dassault Systemes, BIOVIA

The life sciences industry is looking to the cloud to support outsourcing and external collaboration initiatives. At the same time the cloud can give research IT the 'agility' to innovate more rapidly and significantly reduce their IT footprint. BIOVIA is helping research organizations implement 'hybrid cloud' solutions with its ScienceCloud platform to facilitate this critical transition.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

GLOBAL COLLABORATIONS: CONNECTING CANCER TREATMENT AND RESEARCH CENTERS

4:00 Computational Approaches to Cancer: Cooperation between India and the United States

Kenneth Buetow, Ph.D., Director, Computational Sciences and Informatics, Complex Adaptive Systems Initiative (CASI), Arizona State University

Timothy Lance, Distinguished Service Professor Emeritus, University at Albany; President Emeritus and Chief Research Officer, NYSERNet

Amit Saxena, Senior Technical Officer, Bioinformatics Group, C-DAC

Anil Srivastava, President, Open Health Systems Laboratory (OHSL)

IUCKA: Indo-US Cancer Knowledge Alliance is being designed as an integrated biomedical informatics cyberinfrastructure for cancer treatment and research in India. It will be a true translational research platform from bench to bedside connecting cancer treatment and research centers across the country with access and connection to global centers of research, especially in the United States. IUCKA is being implemented as a PPP (public private partnership) and is bringing together technology products and service providers and cancer treatment and research centers in an ecosystem to directly benefit cancer patients in India and contribute to global research collaboration, especially between cancer centers in India. ICTBioMed is a group of life sciences supercomputing centers brought together by OSHL. ICTBioMed members have been working together for almost four years to create a shared global cyberinfrastructure as a seamless and friction-free platform for the researchers worldwide for their collaborative research in consistent with the tenets of team science. The backbone research and education network in India and the United States are now connected by a direct 5+5 gig optical fiber links between Mumbai and New York making it possible for cooperation in biomedical research leveraging computational biology in a big way. OHSL and its IUCKA and ICTBioMed initiatives are playing a big role in this area. This panel session will discuss the biomedical applications and hardware engineering components, as well as the status, plans and prospects for US-India collaboration.

5:00 Sponsored Presentation *(Opportunity Available)*

5:30 – 6:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM

Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

GENOMIC VARIANTS

10:30 Chairperson's Remarks

Honey Reddi, Ph.D., FACMG, Clinical Laboratory Director, The Jackson Laboratory

10:40 Fusion RNAs and Their Implications in Cancer Diagnosis

Hui Li, Ph.D., Associate Professor, Pathology, University of Virginia

Gene fusions and fusion products have been thought to be cancer-unique features. However, our recent work on intergenic splicing challenged this prevailing view. These findings complicate the usage of fusion RNAs as a whole for cancer diagnosis and

treatment. On the other hand, the novel mechanism represents a new repertoire for the discovery of new biomarkers and drug targets.

11:10 Somatic Variants – Approaches to Predicting Actionability

Honey Reddi, Ph.D., FACMG, Clinical Laboratory Director, The Jackson Laboratory

There are currently no defined algorithms in place for the interpretation and classification of somatic variants, unlike the ACMG guidelines for germline variants. We are currently evaluating a series of steps that can be implemented for the interpretation and classification of somatic cancer variants keeping in mind the factors that define actionability.

11:40 Unlock Asian Giant's Precision Medicine Computing Potential – Intel's Bio-IT Know-How and Footprints in China

Jian Li, Ph.D., Life Science Business Lead, Intel Health & Life Science, Greater Asia

Chang Yu, Life Science Solution Architect, Intel Health & Life Science, Greater Asia

China has become a new battlefield for global precision medicine, which could be largely attributed to it having the largest patient population in the world, the government's endorsement and huge funding support, the booming market needs on NIPT and cancers, and also the emerging genomics giant players, e.g., Beijing Genomics Institute and Novogene

12:10 pm Session Break

12:20 Luncheon Presentation *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

LOOKING BEYOND THE GENOME OF THE PATIENT: DATA, ANALYSIS AND TOOLS TO IMPROVE BETTER DISEASE UNDERSTANDING FOR CURRENT TREATMENTS AND DRUG DEVELOPMENT

1:55 Chairperson's Remarks

Michael N. Liebman, Ph.D., Managing Director, IPQ Analytics, LLC and Strategic Medicine, Inc.

2:00 Distinguishing between Precision Medicine and Accurate Medicine: Application to Heart Failure Patients and Clinical Practice

Michael N. Liebman, Ph.D., IPQ Analytics, LLC and Strategic Medicine, Inc.

Increasingly, patient stratification based on genomic analysis is being considered in disease management. Critically, the need to understand real world medical practice and real world patient complexities extends far beyond the genome of the patient. We have shown examples of this complexity in heart disease and how this impacts both development of clinical guidelines, trial design, and development of new patient management approaches.

2:30 CARPEDIEM - Comorbidity and Risk Profiles Evaluation in Diabetes and Heart Morbidities

Sabrina Molinaro, Psy.D., Ph.D., Head, Department of Epidemiology and Health Services, Institute of Clinical Physiology, National Research Council of Italy

Our project uniquely develops a patient record that includes clinical and individual factors (EHR-driven phenotyping) that will be validated through the comparison of existing standards for building new risk algorithms. An understanding of the current limitations and biases of risk profiling in heart disease and diabetes and how an extended, integrated database and automatic rule-based classification system can be used to improve patient management.

3:00 PANEL DISCUSSION: Precision Medicine vs. Accurate Medicine: The Need to Understand Real World Medicine and Real World Patients

Michael N. Liebman, Ph.D., IPQ Analytics, LLC and Strategic Medicine, Inc. (Moderator)

Charles Barr, M.D., MPH, Group Medical Director and Head, Evidence Science and Innovation, Genentech

Jonathan Morris, M.D., Vice President, Provider Solutions and Chief Medical Informatics Officer, Real World Insights

Krishnan Nandabalan, Ph.D., President, CSO and Co-Founder, BioXcel

Hal Wolf, Director, National Leader of Information and Digital Health Strategy, The Chartis Group

4:00 Conference Adjourns

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

25 YEARS Organized by Cambridge Healthtech Institute

Data Security

MEETING THE CHALLENGE IN THE BIG DATA-CENTRIC WORLD

TUESDAY, MAY 23

7:00 am Workshop Registration and Morning Coffee

8:00 – 11:30 Recommended Morning Pre-Conference Workshops*
(W2) An Intro to Blockchain in Life Sciences
(W3) Delivering Custom Mobile App Projects to the Cloud

12:30 – 4:00 pm Recommended Afternoon Pre-Conference Workshops*
(W15) Virtual Pharma: Evolving toward True Data Exchange Highways beyond Corporate Firewalls
** Separate registration required. Please see page 7 for details.*

2:00 – 6:00 Main Conference Registration Open

4:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

5:00 – 7:00 Welcome Reception in the Exhibit Hall with Poster Viewing

Sponsored by Genedata

WEDNESDAY, MAY 24

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION
Please see page 4 for details.

9:50 Coffee Break in the Exhibit Hall with Poster Viewing

Sponsored by InterSystems

NETWORKING & DATA SECURITY

10:50 Chairperson's Remarks
Tom Johnson, Senior Director, Healthcare and Life Sciences Solutions, Exostar

11:00 Science DMZ and Security around Moving Data between Sites
Eli Dart, Network Engineer on Science DMZ, ESNET

11:30 End-To-End Asymmetric Encryption of Biomedical Data In-Transit and At-Rest
Ryan Harrison, Ph.D., Head of Engineering, BioBright
We present a workflow for the end-to-end asymmetric encryption of biomedical data, that goes well beyond the typical (not encrypted at all) and conscientious (HTTPS/SSL in-transit, AES-256 at-rest) laboratory use-cases. The pipeline, which is commercially available to selected customers, allows the option of customizable metadata extraction, allowing efficient lab-workflow-related search to coexist with at-rest encryption.

12:00 pm Networking and Data Transfer in the Modern Life Sciences and Healthcare Era
Ari E. Berman, Ph.D., Vice President and General Manager of Consulting Services, BioTeam, Inc.
Data generation throughout the life sciences research and healthcare domains has risen at a rate far beyond that predicted by Moore's Law. As a result, organizations are accumulating 10's to 100's of petabytes (PB) of data, spending millions on storage systems, and doing it all in a manner consistent with out of date IT practices and policies. These practices include little to no data management, ineffective or non-existent data lifecycle policies, no metadata standards, and a dependence on network infrastructure and services that are unrealistic and unsustainable. In this presentation, we will discuss the general scope of the data generation landscape in Life Sciences as well as review generalized networking and security solutions that have successfully supported research missions throughout the industry and enable the of movement and sharing large amounts of data effectively and sustainably.

12:30 Session Break

12:40 Luncheon Presentation *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:40 Session Break

DESIGNING CLOUD FLEXIBILITY WHILE MAINTAINING SECURITY

1:50 Chairperson's Remarks
Lance Smith, MBA, Associate Director, IT, Celgene

1:55 How BMS Uses cBioPortal for Cancer Genomics Research
Isaac Neuhaus, Ph.D., Director, Computational Genomics, Bristol-Myers Squibb
BMS has started using cBioPortal for visualizing cancer genomics datasets early 2016, supported by The Hyve, an open source bioinformatics company based in The Netherlands. The cBioPortal server runs on Amazon AWS and is tied to the company Active Directory for authentication. Currently loaded datasets are TCGA, CCLE and M2Gen and we will be loading clinical trial data later this year. Concurrently we are running a pilot for cBioPortal with mouse data.

2:25 Enabling Celgene's Innovation while Protecting the Enterprise in the Cloud
Lance Smith, MBA, Associate Director, IT, Celgene
Celgene uses Amazon Web Services (AWS) to run R&D workloads, enterprise systems, and collaboration environments securely and all controlled from a single pane. Topics covered include network design, HPC and job scheduling, workload isolation, storage, automated security, and environment auditing. We also discuss business use cases, challenges overcome (technical, legal, security, and organizational), and Turbot guardrails for management.

Final Weeks to Register

- Cover
- Schedule-at-a-Glance
- Plenary Keynote Presenters
- Pre-Conference Workshops
- Awards Programs

- Data & Storage Management
- Data Computing
- Networking Hardware
- Software Applications & Services
- Cloud Computing
- Bioinformatics
- Next-Gen Sequencing Informatics
- Clinical Research & Translational Informatics
- Data Visualization & Exploration Tools
- Pharmaceutical R&D Informatics
- Clinical Genomics
- Cancer Informatics
- Data Security

- Sponsor & Exhibit Opportunities
- Hotel & Travel Information
- Current Sponsors
- Registration Information

CLICK HERE TO **Register Online!**

Organized by Cambridge Healthtech Institute

13

Data Security

MEETING THE CHALLENGE IN THE BIG DATA-CENTRIC WORLD

2:55 Automated High-Scale Computational Infrastructure

Sponsored by OpenEye Scientific

Jharrod LaFon, Chief Cloud Engineer, OpenEye Scientific

Large-scale computation is an important aspect of drug design, and increasingly is cloud-based. The true value of the cloud comes not just from computation but by enabling company-wide collaboration, the ability to build, modify, share, manage versions, and evaluate methods. OpenEye's cloud-native platform, Orion, and its workflow engine, Floe, provide the tools to create innovative approaches for computation and collaboration.

3:10 Research Informatics: Get Ready for the Cloud!

Sponsored by Dassault Systèmes

Ton van Daelen, ScienceCloud, Product Director, Dassault Systemes, BIOVIA

The life sciences industry is looking to the cloud to support outsourcing and external collaboration initiatives. At the same time the cloud can give research IT the 'agility' to innovate more rapidly and significantly reduce their IT footprint. BIOVIA is helping research organizations implement 'hybrid cloud' solutions with its ScienceCloud platform to facilitate this critical transition.

3:25 Refreshment Break in the Exhibit Hall with Poster Viewing

4:00 5 Steps to Vet and Manage Cloud Service Providers

Dianne Pacheco, CISM, Information Security Officer, Information Technology, The Jackson Laboratory

The use of cloud services is essential, but for transmitting, processing or storing data that way, buyer beware – not all service providers are created equal. Many consumers do not realize they bear full responsibility for the security of the data held by service providers. This talk walks the audience through vetting and managing their service providers to minimize risk of exposing legally protected or sensitive research data.

4:30 Cloud Enabling Networks for Pediatric Care

Marcia M. Nizzari, MS, Consulting CTO and Software Architect, Ksandr Software Consulting

Cloud infrastructure enables new networks of pediatric hospitals, labs, clinicians, researchers, and pharma that can improve a child's care. The economies of scale, ease of data movement, shared tools, and communication enrichment are significant game changers in the world of pediatric rare disease. We cover how cloud technology and related new software architecture constructs, such as microservices and containers, can drive advancement in precision medicine and facilitate a better future for our patients.

5:00 Presentation to be Announced

Sponsored by aspera an IBM company

5:30 15th Anniversary Celebration in the Exhibit Hall with Poster Viewing and Best of Show Awards

THURSDAY, MAY 25

7:00 am Registration Open and Morning Coffee

8:00 PLENARY KEYNOTE SESSION & AWARDS PROGRAM

Please see pages 3 & 4 for details.

8:05 Benjamin Franklin Awards and Laureate Presentation

8:35 Best Practices Awards Program

8:50 Plenary Keynote

9:45 Coffee Break in the Exhibit Hall and Poster Competition Winners Announced

STRATEGIES FOR MAINTAINING DATA PRIVACY

10:30 Chairperson's Remarks

Dianne Pacheco, CISM, Information Security Officer, Information Technology, The Jackson Laboratory

10:40 FEATURED PRESENTATION: Beyond the Firewall: Protecting Information Holistically with Advanced Data Analytics

Alexander D. Kent, Ph.D., Senior Executive Adviser, Cybersecurity, Los Alamos National Laboratory

Unfortunately, much of contemporary cybersecurity focuses on detection and mitigation of cyber threats at the perimeter of enterprise networks; the candy metaphor of "a hard shell with a soft, chewy center" applies well to modern cybersecurity. Today's cyber threats include traditional insider information theft as well as sophisticated external cyber actors where traditional protections have little value. Instead, modern defensive approaches must focus on understanding how information should and should not be used within the cyber environment through the use of advanced data analytics.

11:10 Creating & Leveraging an Industry Security Standard to Protect IP & Sensitive Data throughout the R&D Process

Matt King, Vice President, Global Policy, SAFE-BioPharma Association

Tom Johnson, Senior Director, Healthcare and Life Sciences Solutions, Exostar

We discuss the need for standards, the latest in SAFE-BioPharma standards development, and how solutions that leverage these emerging standards are being used today to issue, validate, and recognize industry-trusted certified credentials across the ecosystem. The standards, and credential providers that leverage them, provide secure access to the applications and data individuals need to collaborate seamlessly and productively.

11:40 From the Board Room Down: What Healthcare Executives and Managers Need to Know about Cyber-Security

James Slaughter, Principal, Wolf Den Associates

No healthcare organization can be completely immune from cyber-attacks and adversaries. The risk of cyber-attack is no longer limited to the IT department; it is a key business issue that must be addressed by the C-suite. In this we talk we explore the measures that can be taken to mitigate risks and integrate cyber-security into the business culture.

CLICK HERE TO **Register Online!**

Organized by Cambridge Healthtech Institute

13

Data Security

MEETING THE CHALLENGE IN THE BIG DATA-CENTRIC WORLD

12:10 pm Session Break

12:20 Luncheon Presentation *(Sponsorship Opportunity Available)* **or Enjoy Lunch on Your Own**

1:20 Dessert Refreshment Break in the Exhibit Hall with Poster Viewing

GENOMIC DATABASE MINING, MANAGEMENT, AND PRIVACY

1:55 Chairperson's Remarks

Liz Worthey, Ph.D., Faculty Investigator & Director, Software Development and Informatics, HudsonAlpha Institute for Biotechnology

---

2:00 KEYNOTE PRESENTATION: Challenges and Opportunities for Global Genomic Sharing: Lessons from Current Successes

John Mattison, M.D., Assistant Medical Director, Chief Medical Information Officer, Kaiser Permanente

---

2:30 FEATURED PRESENTATION: A Comprehensive Approach to Rare Disease Interpretation: Moving beyond Identification of a Causal Variant

Liz Worthey, Ph.D., Faculty Investigator & Director, Software Development and Informatics, HudsonAlpha Institute for Biotechnology

We discuss application of methods for analyzing, storing, and interpreting variant data that allows us to move beyond MDx-related to a single variant to support a refined understanding of human health and disease. The discussion includes methods to: 1) support increased definitive diagnosis rates, 2) explore genotype to phenotype associations that can be used to predict disease course in a particular individual, and 3) identify modifiers that alter clinical presentation.

3:00 A Curated Database of 150,000 Variants from Clinical Whole-Genome Sequencing: What Does It Take to Maintain and Donate to ClinVar?

Aditi Chawla, Ph.D., Senior Scientist, Clinical Services Laboratory, Illumina

We present the workflow for variant classification, including an autoscore system that calculates a variant's potential to have caused the associated disease, search for literature on variants, and manual clinical curation based on ACMG guidelines. We discuss strategies for maintaining and updating variant classifications over time, laboratory policy for updating clinical reports as variant classification changes, process for donating to ClinVar, and summary of our ClinVar donation.

3:30 Community-Driven Approaches to Gene Curation

Marina DiStefano, Ph.D., Variant Scientist, Laboratory for Molecular Medicine, Partners HealthCare

As precision medicine expands and whole exome sequencing becomes standard, it is critical to define a common framework to assess the evidence for gene-disease associations. Testing this framework involves reaching out to the experts in various specialties for validation. With input from the genetic community, genes can be transparently and systematically evaluated and prioritized for analysis in various clinical contexts.

4:00 Conference Adjourns

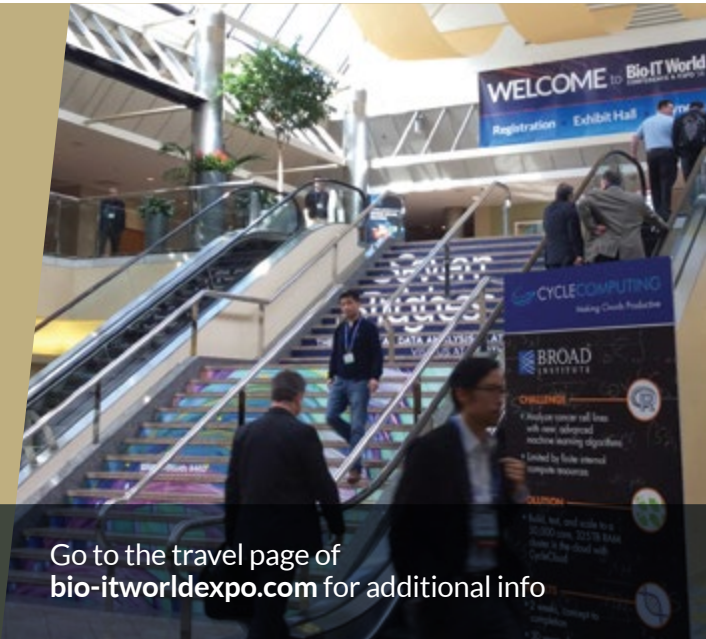
Hotel & Travel Information

CONFERENCE VENUE:
Seaport World Trade Center
200 Seaport Boulevard
Boston, MA 02210

HOST HOTEL:
Seaport Hotel *(Located directly across the street)*
One Seaport Lane,
Boston, MA 02210
Phone: 1-877-SEAPORT
(1-877-732-7678)

Reservations: Go to the travel page of **bio-itworldexpo.com**

Discounted Room Rate: $294 s/d | Discounted Cut-off Date: April 19, 2017



Go to the travel page of **bio-itworldexpo.com** for additional info

Organized by Cambridge Healthtech Institute

2017 Sponsor Opportunities

Comprehensive sponsorship packages allow you to achieve your objectives before, during, and long after the event. Signing on earlier will allow you to maximize exposure to hard-to-reach decision-makers.

PODIUM PRESENTATIONS
Available within the Main Agenda!

Showcase your solutions to a guaranteed, targeted audience through a 15- or 30-minute presentation during a specific conference program, breakfast, lunch, or separate from the main agenda within a pre-conference workshop. Package includes exhibit space, on-site branding, and access to cooperative marketing efforts by CHI. For the luncheon option, lunches are delivered to attendees who are already seated in the main session room. Presentations will sell out quickly, so sign on early to secure your talk!

ONE-ON-ONE MEETINGS

Select your top prospects from the pre-conference registration list. CHI will reach out to your prospects and arrange the meeting for you. A minimum number of meetings will be guaranteed, depending on your marketing objectives and needs. A very limited number of these packages will be sold.

INVITATION-ONLY VIP DINNER/ HOSPITALITY SUITE

Sponsors will select their top prospects from the conference pre-registration list for an evening of networking at the hotel or at a choice local venue. CHI will extend invitations and deliver prospects, helping you to make the most out of this invaluable opportunity. Evening will be customized according to sponsor's objectives. (i.e.: Purely social, Focus group, Reception style, Plated dinner with specific conversation focus)

ADDITIONAL BRANDING & PROMOTIONAL OPPORTUNITIES INCLUDE:

- Mobile App – SOLD!
- Hotel Room Keys – SOLD!
- Footprint Trails
- Staircase Ads – SOLD!
- Conference Tote Bags – SOLD!
- Literature Distribution (Tote Bag Insert or Chair Drop)
- Badge Lanyards – SOLD!
- Padfolios
- Program Guide Advertisement

EXHIBIT

RESERVE YOUR EXHIBIT SPACE BY FEBRUARY 23 AND SAVE $300!

Exhibitors will enjoy facilitated networking opportunities with 3,300+ qualified delegates, making it the perfect platform to launch a new product, collect feedback, and generate new leads. Exhibit space sells out quickly, so reserve yours today.

LOOKING FOR ADDITIONAL WAYS TO DRIVE LEADS TO YOUR SALES TEAM?

CHI's Lead Generation Programs will help you obtain more targeted, quality leads throughout the year. We will mine our database of 800,000+ life science professionals to your specific needs. We guarantee a minimum of 100 leads per program! Opportunities include:

- Live Webinars
- Whitepapers
- Research Surveys
- Podcasts and More!

FOR ADDITIONAL INFORMATION, PLEASE CONTACT:

Companies A-K: Katelin Fitzgerald
Sr Business Development Manager
781-972-5458 | kfitzgerald@healthtech.com

Companies L-Z: Uma Patel
Business Development Manager
781-972-1349 | upatel@healthtech.com

Final Weeks to Register

CLICK HERE TO
Register Online!

25 YEARS Organized by Cambridge Healthtech Institute

Current Sponsors & Exhibitors

(December 16, 2016)

Abbott
ACD Labs
Arxspan
Aspera, an IBM Company
Astrix Software Technology Inc.
Avere Systems
Bioinformatics Organization, Inc.
Benchling
Biobam
Biomatters
BioTeam
Bluebee
Brooks Life Science Systems
BSI
ByteGrid
Cambridge Computer
Cambridge Semantics
Certara
ChemAxon
Cloudera
Cloudian
Cloudy Cluster
Collaborative Drug Discovery
Congenica
Copyright Clearance Center
Core Informatics
Cray Inc.
Cycle Computing
Dassault Systemes, BIOVIA
Databiology
DDN Storage
Dell EMC
Deloitte
DeltaSoft Inc.
DNAnexus
DNASTAR, Inc.
Dotmatics
Eagle Genomics Ltd
Edico Genome
Elevada
Elsevier
EPAM Systems, Inc.
ERT
Exostar
FlowJo
Freezerworks, a Division of Dataworks Development, Inc.
Genedata
Genestack
Genospace
Globus
HGST a Western Digital Company
The Hyve
IBM
IBM Cloud Object Storage
Igneous Systems
Illumina, Inc.
Intel
International Society for Computational Biology (ISCB)
Internet2
Intomics
inviCRO
ITTM S.A.
Jisto
Komprise
LabAnswer
LabKey
Lab7 Systems, Inc.
LabVantage Solutions, Inc.
Lawrence Livermore
National Security, LLC
Liaison Technologies
Linguamatics
Microsoft
Nexan
NextMove Software Ltd.
Nodeum
NooBaa
OpenEye Scientific Software
Ovation
Panasas
Paradigm4
Partek Incorporated
PerkinElmer
Prysm
PetaGene
Pure Storage
QIAGEN
Qlucore
Quantum
Qumulo
RCH Solutions
Red River
Regeneron Pharamaceuticals
Roche Sequencing Solutions
SAP SE
Schrödinger
SciBite
Scigilian
Scilligence
Seven Bridges
SGI
ShareVault
Signet Accel
Silicon Mechanics
Simulations Plus, Inc.
Sinequa
SolveBio
Studylog Systems, Inc.
SwiftStack
TERALINK Solutions
Tessella
Thomson Reuters
Titian Software US Inc
Tonic Health
tranSMART Foundation
Twigkit
Univa Corporation
Validated Cloud Inc.
Wiley
WuXi NextCODE
Zifo Technologies

Organized by Cambridge Healthtech Institute

Pricing and Registration Information

WORKSHOP PRICING

	Commercial	Academic, Government, Hospital-affiliated	Student*
One Half-Day Workshop	$599	$299	$149
Two Half-Day Workshops	$899	$499	$249

Please refer to Workshop list on page 7.

MAIN CONFERENCE PRICING *(excludes workshops)*

Registrations after April 7, 2017, and on-site	$2,149	$999	$329

CONFERENCE TRACKS

TRACK 1: Data & Storage Management	TRACK 8: Clinical Research & Translational Informatics
TRACK 2: Data Computing	TRACK 9: Data Visualization & Exploration Tools
TRACK 3: Networking Hardware	TRACK 10: Pharmaceutical R&D Informatics
TRACK 4: Software Applications & Services	TRACK 11: Clinical Genomics
TRACK 5: Cloud Computing	TRACK 12: Cancer Informatics
TRACK 6: Bioinformatics	TRACK 13: Data Security
TRACK 7: Next-Gen Sequencing Informatics	

**Full time graduate students and Ph.D. candidates qualify for the student rate. Students are encouraged to present a research poster and receive an additional $50 off their registration fee. Student rate cannot be combined with any other discount offers, except poster discount. Students must present a valid/current student ID to qualify for the student rate. Limited to the first 100 students that apply.*

CONFERENCE DISCOUNTS

Exclusive Offer to Attend Medical Informatics World Conference*

Cambridge Healthtech Institute presents a series of informatics programs in Boston this spring with the goal of bridging the healthcare and life science worlds. Paid attendees of **Bio-IT World Conference & Expo** can attend the **Medical Informatics World Conference (May 22-23)** for a special discounted rate (20% discount off the registration fee for the main conference).

To receive this exclusive 20% discount, mention keycode **1720BITXP** when registering for Medical Informatics World.
Please note: Our records must indicate you are a paid attendee of **Bio-IT World Conference & Expo 2017** to qualify.

** Discount applies to paid attendees of Bio-IT World Conference & Expo 2017 only. Applies to new registrations only and cannot be combined with other discount offers, except poster discount. Discount does not apply to workshops. Discount taken off lowest priced item(s).*

Poster Submission-Discount ($50 Off)
Poster abstracts are due by April 7, 2017. Once your registration has been fully processed, we will send an email containing a unique link allowing you to submit your poster abstract. If you do not receive your link within 5 business days, please contact jring@healthtech.com.
** CHI reserves the right to publish your poster title and abstract in various marketing materials and products.*

International Society for Computational Biology (ISCB) Member-Discount (10% Off)

REGISTER 3 - 4th IS FREE: Individuals must register for the same conference or conference combination and submit completed registration form together for discount to apply. Additional discounts are available for multiple attendees from the same organization. For more information on group rates contact David Cunningham at 781-972-5472.

If you are unable to attend but would like to purchase the Bio-IT World Conference & Expo 2017 conference CD for $750 (plus shipping), please visit Bio-ITWorldExpo.com. Massachusetts delivery will include sales tax.

Please use keycode **1720 F** when registering!

ADDITIONAL REGISTRATION DETAILS

Each registration includes all conference sessions, posters and exhibits, food functions, and access to the conference proceedings link.

Handicapped Equal Access: In accordance with the ADA, Cambridge Healthtech Institute is pleased to arrange special accommodations for attendees with special needs. All requests for such assistance must be submitted in writing to CHI at least 30 days prior to the start of the meeting.

To view our Substitutions/Cancellations Policy, go to healthtech.com/regdetails

Video and or audio recording of any kind is prohibited onsite at all CHI events.

Final Weeks to Register

CLICK HERE TO
Register Online!

Organized by Cambridge Healthtech Institute

15th ANNIVERSARY 2002-2017

Bio·IT World
CONFERENCE & EXPO '17

Building a global network for precision medicine by uniting the Bio-IT community

Help Us Celebrate 15 Years of Progress and Discovery

as we look back and together envision the future of precision medicine.

THANK YOU FOR BEING PART OF OUR COMMUNITY.

MAY 23-25, 2017 | BOSTON, MA

SEAPORT WORLD TRADE CENTER