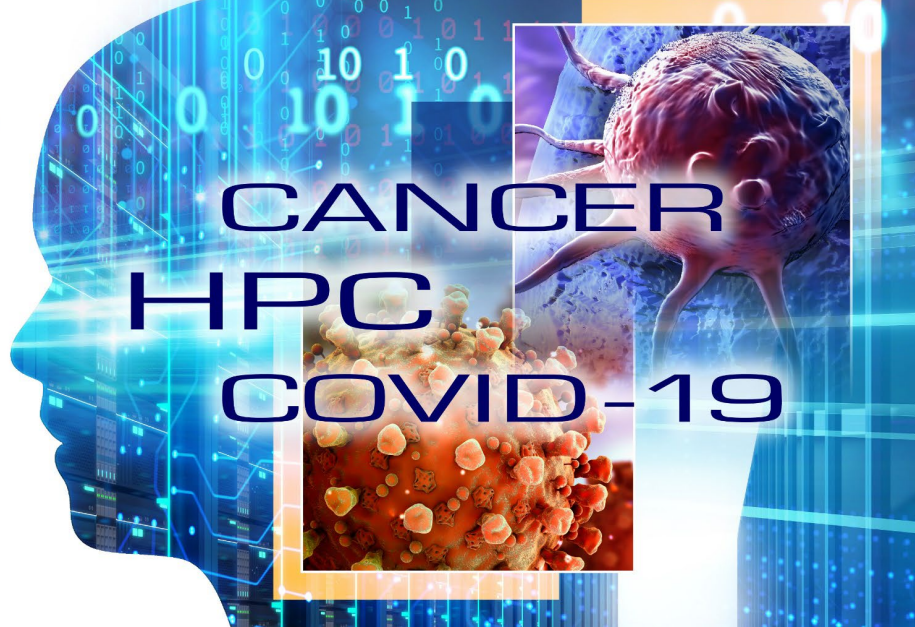


COMPUTATIONAL APPROACHES FOR CANCER WORKSHOP

CAFCW20

November 13, 2020



Welcome to the Sixth Computational Approaches for Cancer Workshop

Panels and Presentations will be moderated.

To ask a question, type it in the Questions window. Attendees can vote on questions to be answered, and the moderators will take these votes into account when selecting questions to be answered by the speakers.

Links for Lunch (2:00 p.m.—3:20 p.m. ET)

Virtual Poster Session

URL: cafcw20.virtualpostersession.org

password: cafcw20

Presenters will be available during this time to discuss their posters, click on the “Chat with Presenter” button to meet the presenter.

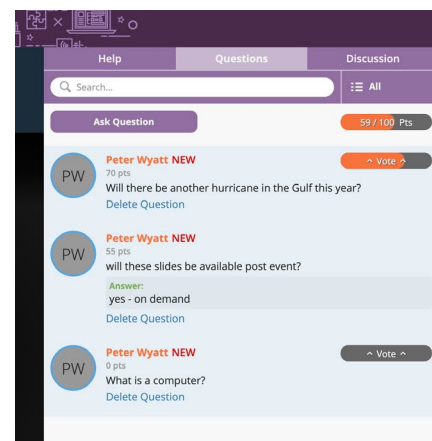
Informal Networking

URL: http://bit.ly/CAFCW20_Lunch_Networking

password: FsnmRZh5*75

This WebEx session will feature breakout sessions for anyone wishing to chat. All participants will be given co-host privileges to be able to move in and out of the breakout sessions. When leaving, **please do not end the meeting for all.**

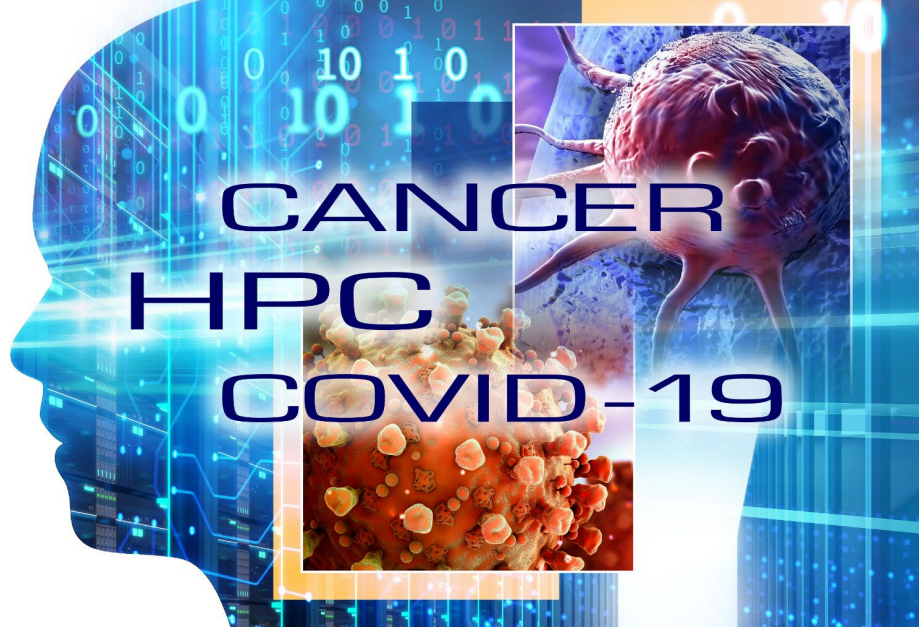
Following the workshop, we’d appreciate if you would take a few minutes to fill out a brief survey to help us improve our future workshops: https://bit.ly/CAFCW20_Eval



COMPUTATIONAL APPROACHES FOR CANCER WORKSHOP

CAFCW20

November 13, 2020



Program

Friday, November 13, 2020, 10:00 a.m.–6:30 p.m.
(All times listed are Eastern Time)

- 10:00 a.m.–10:10 a.m. **Welcome — Sixth Computational Approaches for Cancer Workshop (CAFCW20)**
Eric Stahlberg, PhD, Frederick National Laboratory for Cancer Research
- 10:10 a.m.–11:00 a.m. **Keynote: *Data Science Initiatives at the National Cancer Institute***
Norman “Ned” Sharpless, MD, National Cancer Institute Director
Introduced by Sean Hanlon, PhD, National Cancer Institute
- 11:00 a.m.–11:45 a.m. **Panel: *HPC, Cancer and COVID-19***
Organizer and Moderator: Patricia Kovatch, Icahn School of Medicine, Mount Sinai
Panel:
Jim Brase, Lawrence Livermore National Laboratory
Diane Del Valle, Icahn School of Medicine at Mount Sinai
Sacha Gnjatic, PhD, Icahn School of Medicine at Mount Sinai
Sharon Nirenberg, MD, Icahn School of Medicine at Mount Sinai
Carlos Simmerling, PhD, Stony Brook University
- 11:45 a.m.–12:00 p.m. ***Scalable Human Pharmacokinetics Property Prediction for Cancer Drug Discovery at ATOM***
Presenter: Benjamin Madej, PhD, Frederick National Laboratory for Cancer Research

- 12:00 p.m.–12:15 p.m. **CAFCW20 Morning Break**
- 12:15 p.m.–12:30 p.m. ***Scaffold-Induced Molecular Subgraphs (SIMSG): Effective Graph Sampling Methods for High-Throughput Computational Drug Discovery***
Rick Stevens; Ashka Shah; Arvind Ramanathan, PhD; Max Zvyagin; Austin Clyde
Presenters: Ashka Shah, University of Chicago, and Max Zvyagin, Argonne National Laboratory
- 12:30 p.m.–12:45 p.m. ***Causal Deconvolution of a Mechanistic Model of EGFR and ERK Signaling Explains Adaptive and Genetic Resistance in Melanoma***
Presenter: Fabian Fröhlich, PhD, Harvard Medical School
- 12:45 p.m.–1:00 p.m. ***An Efficient, Data-Driven Approach to Model Specific Cancer Cell Lines***
Peter Balogh, PhD; John Gounley, PhD; Amanda Randles, PhD
Presenter: Peter Balogh, PhD, Duke University
- 1:00 p.m.–1:15 p.m. ***Deep Learning Based Prediction of the Temporal Behavior of RAS Protein Conformations on Simulated Cell Membrane Surfaces***
Brian C. Van Essen, PhD; James Glosli, PhD; Sam Adé Jacobs, PhD; Timothy S. Carpenter, PhD; Peer-Timo Bremer, PhD; Harsh Bhatia, PhD; Helgi I. Ingólfsson, PhD; Gautham Dharuman, PhD; Adam Moody
Presenter: Adam Moody, Lawrence Livermore National Laboratory
- 1:15 p.m.–2:00 p.m. **Panel #2: *Digital Twins for Cancer Care***
Organizers: Emily Greenspan, PhD, National Cancer Institute and Eric Stahlberg, Frederick National Laboratory for Cancer Research
Moderator: Eric Stahlberg, PhD, Frederick National Laboratory for Cancer Research
Panel:
Richard Arthur, GE Research
Emily Greenspan, PhD, National Cancer Institute
Tina Hernandez-Boussard, PhD, Stanford University
- 2:00 p.m.–3:20 p.m. **Lunch Break (Networking and Posters)**
Poster and Presenters
A Multistaged Hyperparallel Optimization of the Fuzzy-Logic Mechanistic Model of Molecular Regulation, Paul Aiyetan, MD, Frederick National Laboratory for Cancer Research

Analyzing Tracking Graphs to Better Understand RAS Diffusion, Torin McDonald, Scientific Computing and Imaging Institute, University of Utah

Analyzing Tumor Heterogeneity thru Advanced Proteogenomic Bioinformatic Approaches, Don Johann, MD, University of Arkansas for Medical Science

Bayesian Deep Learning for Robust Information Extraction from Cancer Pathology Reports, Devanshu Agrawal, University of Tennessee at Knoxville

Developing Predictive Random Forest Models of MRP3 and MRP4 Transporter Inhibition, Amanda Paulson, PhD, Accelerating Therapeutics for Opportunities in Medicine (ATOM) Consortium

Exploring Strengths and Weaknesses of PQSAR 2, Stewart He, PhD, Lawrence Livermore National Laboratory

Large-Scale Photoacoustic Tomography: A Trade-off Between Image Quality and Computational Cost, Gabriel Bordovský, Brno University of Technology

Leveraging a Hybrid and Multiscale Model to Assist Androgen Deprivation Therapy in Recurrent Prostate Cancer Patients, Mengdi Tao, University of Pennsylvania

Long Time Evolution of Lipid Membranes, Tomas Ooppelstrup, Lawrence Livermore National Laboratory

Machine Learning for Detection of Circulating Tumor Cells by Optical Imaging without Label, Yaling Liu, PhD, Lehigh University

Making the Most of Genomic Data for Small and Medium-Sized Data Commons, Jong Cheol Jeong, PhD, University of Kentucky

Molecular Classification of Pediatric High-Risk Leukemias Using Expression Profiles of Multimodally Expressed Genes, Sneha Jariwala, University of California, Santa Cruz

Neural Collaborative Filtering Method for IC50 Prediction of drug compounds, Sookyung Kim, PhD, Lawrence Livermore National Laboratory

Performance Evaluation of Graph Convolutional Networks for Information Extraction from Cancer Pathology Reports, Hong-Jun Yoon, PhD, Oak Ridge National Laboratory

Probabilistic-based Interactive Machine Learning for Tomogram Segmentation and Annotation, Talita Perciano, PhD, Lawrence Berkeley National Laboratory

System for High Intensity Evaluation During Radiation Therapy (SHIELD-RT): A Prospective Randomized Study of Machine Learning-Directed

- Clinical Evaluations During Outpatient Cancer Radiation and Chemoradiation*, Julian Hong, MD, University of California, San Francisco
- Training Data Error Impacts on Deep Neural Networks for Classifying RNA-seq Gene Expressions*, Rajeev Jain, Argonne National Laboratory
- 3:20 p.m.–3:25 p.m. **Afternoon Welcome**
Patricia Kovatch, Icahn School of Medicine, Mount Sinai
- 3:25 p.m.–3:40 p.m. ***Machine Learning Driven Importance Sampling Approach for Multiscale Simulations***
Presenter: Harsh Bhatia, PhD, Lawrence Livermore National Laboratory
- 3:40 p.m.–3:55 p.m. ***A Metapath Approach to Predicting Drug Response in Cancer Cell Lines***
Judith D. Cohn, PhD, Los Alamos National Lab; Benjamin H. McMahon, PhD, Los Alamos National Lab; Nicholas W. Hengartner, PhD, Los Alamos National Lab; Marian Anghel, PhD, Los Alamos National Lab; Tudor Oprea, MD, PhD, University of New Mexico School of Medicine
Presenter: Judith D. Cohn, PhD, Los Alamos National Laboratory
- 3:55 p.m. – 4:10 p.m. ***Towards a Data-Driven System for Personalized Cervical Cancer Screening***
Presenter: Geir Severin Elvatun Rakh Langberg, Cancer Registry of Norway
- 4:10 p.m.–4:40 p.m. ***Integration of Domain Knowledge Using Medical Knowledge Graph Deep Learning for Cancer Phenotyping***
Presenter: Mohammed Alawad, PhD, Oak Ridge National Laboratory
- 4:40 p.m.–4:55 p.m. **Afternoon Break**
- 4:55 p.m.–5:10 p.m. ***Deciphering Hallmarks of Resistance in Breast Cancer***
Presenter: Amrita Basu, PhD, University of California, San Francisco
- 5:10 p.m.–5:40 p.m. ***Why I'm Not Answering: Understanding Determinants of Classification of an Abstaining Classifier for Cancer Pathology Reports***
Sayera Dhaubhadel; Benjamin McMahon, PhD
Presenter: Sayera Dhaubhadel, Los Alamos National Laboratory

5:40 p.m.–6:25 p.m.

Panel #3: *Translating Cancer Research Advances in Artificial Intelligence into Clinical Practice*

Organizer and Moderator: Sally Ellingson, PhD, University of Kentucky

Panel:

Douglas W. Blayney, MD, Stanford University Medical Center

Cody Bumgardner, PhD, University of Kentucky

Stephanie Harmon, PhD, National Cancer Institute

Yaling Liu, PhD, Lehigh University

6:25 p.m.–6:30 p.m.

CAFCW20 Wrap-Up

Keynote



Norman “Ned” Sharpless, MD, National Cancer Institute

Dr. Ned Sharpless has served as the director of the National Cancer Institute since October 2017. Previously, Dr. Sharpless served as acting commissioner of the U.S. Food and Drug Administration, director of the University of North Carolina Lineberger Comprehensive Cancer Center and on the faculty of Harvard Medical School.

Dr. Sharpless earned his medical degree from UNC and completed his internal medicine residency at Massachusetts General Hospital and a hematology/oncology fellowship at Dana-Farber and Partners Cancer Care.

Dr. Sharpless is a member of the Association of American Physicians as well as the American Society for Clinical Investigation (ASCI), the nation’s oldest honor society for physician–scientists. He served on the ASCI council from 2011 to 2014.

He has written more than 150 original scientific papers, reviews and book chapters and holds 10 patents. He co-founded two clinical-stage biotechnology companies: G1 Therapeutics and HealthSpan Diagnostics.

In addition to serving as Director of NCI, Dr. Sharpless is Chief of the Aging Biology and Cancer Section in the National Institute on Aging’s Laboratory of Genetics and Genomics, where he continues his research on the biology of the aging process that promotes the conversion of normal self-renewing cells into dysfunctional cancer cells. Dr. Sharpless has made pivotal contributions to the understanding of the relationship between aging and cancer and in the preclinical development of novel therapeutics for melanoma, lung cancer and breast cancer.

Thank you to our CAFCW20 Program Committee:

Frank Alexander, PhD, Brookhaven National Laboratory

Orly Alter, PhD, University of Utah

John Baldoni, PhD, Integral Health, Inc.

Jeff Buchsbaum, MD, PhD, National Cancer Institute

Fernanda Foertter, BioTeam

Emily Greenspan, PhD, National Cancer Institute

Ryuji Hamamoto, PhD, National Cancer Center Japan/RIKEN, Tokyo

Abdul Hamid Halabi, NVIDIA

Florence Hudson, Northeast Big Data Innovation Hub, Columbia University

Steven Litster, PhD, Amazon Web Services

Michael McManus, PhD, Intel

William Richards, PhD, Brigham and Women's Hospital, Harvard Medical School

Ilya Shmulevich, PhD, Institute for Systems Biology

Thomas Steinke, PhD, Zuse Institute Berlin

Kristin Swanson, PhD, Mayo Clinic Arizona

George Zaki, PhD, Frederick National Laboratory

CAFCW20 Organizing Committee:

Eric Stahlberg, PhD, Frederick National Laboratory for Cancer Research

Sean Hanlon, PhD, National Cancer Institute

Sally Ellingson, PhD, University of Kentucky

Patricia Kovatch, Icahn School of Medicine, Mount Sinai

Petrina Hollingsworth, Frederick National Laboratory for Cancer Research

Lynn Borkon, Frederick National Laboratory for Cancer Research

Presenters



Devanshu Agrawal, University of Tennessee, Knoxville

Devanshu Agrawal is a PhD student in the Data Science and Engineering program in the Bredesen Center at the University of Tennessee, Knoxville and works as a research assistant in the Computational Sciences and Engineering division at Oak Ridge National Laboratory. He received bachelor's degrees in mathematics and physics and a master's degree in mathematics from East Tennessee State University, where he taught calculus after that. Agrawal's research focuses on the theory and applications of Bayesian deep learning. (dagrawa2@vols.utk.edu)

Paul Aiyetan, MD, Frederick National Laboratory for Cancer Research

Dr. Paul Aiyetan obtained his MD from the College of Medicine, University of Ibadan. He completed his intern year as a house physician at the University College Hospital and the next two years as a resident medical officer. He earned a Master of Science degree in bioinformatics from Johns Hopkins University and a PhD from the George Mason University School of Systems Biology. After a graduate fellowship in oncology, bioinformatics and biostatistics, Dr. Aiyetan completed a postdoctoral fellowship in pathology at the Johns Hopkins University School of Medicine where he was a member of the NIH/NCI-funded Clinical Proteomics Tumor Analysis Consortium. He was also a member of the NIH/NHLBI-funded Programs of Excellence in Glycosciences (PEG). He received the Johns Hopkins Pathology Young Investigators Excellence Award and the inaugural NIH-funded Big Data to Knowledge fellowship at the University of Minnesota and the Mayo Clinic. Dr. Aiyetan's research interest is at the intersection of medicine, computational biology and quantitative science, investigating physical and biochemical science principles and computational and quantitative methods to elucidate mechanism of disease processes, particularly neoplastic processes (cancer), with an aim to identify and develop more precise diagnostic and therapeutic approaches. (paul.aiyetan@nih.gov)



Mohammed Alawad, PhD, Oak Ridge National Laboratory

Dr. Mohammed Alawad is a research scientist in the Advanced Computing for Health Sciences Section at Oak Ridge National Laboratory. He holds a PhD degree in computer engineering from the University of Central Florida. His research interests include AI/DL for natural language processing, energy-efficient computing, and deep learning acceleration and security. (alawadmm@ornl.gov)



Richard Arthur, GE Research

Rick Arthur leads pathfinding efforts in applying computational methods at GE Research, supporting products and services that spanned diverse industrial sectors such as healthcare, air and rail transportation, media, finance, defense and energy. He represents GE in government policy and project discussions, including as co-chair of the U.S. Council on Competitiveness Advanced Computing Roundtable and as a member of the ECP Industry Council technical advisors. He holds a BS and MEng in computer engineering and an

MBA and is a senior member of the Association for Computing Machinery (ACM).

(arthurr@ge.com)

Peter Balogh, PhD, Peter Balogh, Duke University

Dr. Peter Balogh received his PhD in mechanical engineering from Rutgers University in 2018. He developed a method for modeling blood cells flowing through complex capillary networks in 3D, for which he was awarded the Acrivos Dissertation Award in Fluid Dynamics from the American Physical Society. He is currently a postdoctoral associate in the Department of Biomedical Engineering at Duke University. His research interests include computational fluid dynamics

modeling of biological flows, numerical methods for complex fluid-structure interfaces and code development for high performance computing. His current research is focused on modeling cancer cells and their transport through the circulatory system and on investigating the remodeling of microvasculatures through comparisons with experiments. (peter.balogh@duke.edu)



Amrita Basu, PhD, University of California, San Francisco

Dr. Amrita Basu has over 15 years of experience in bioinformatics and computational biology, leading biomedical informatics initiatives that improve clinical outcomes and support biomarker discovery. She is currently an assistant professor at the University of California, San Francisco in the Department of Surgery. Her group develops computational tools for biomarker discovery in treatment-resistant breast cancer patients. She works closely with the Athena Breast Health Network and I-SPY clinical trial (adaptive trial for early-stage high-risk breast cancer patients). Her group is involved

in identifying biomarker and imaging predictors in ductal carcinoma in situ (DCIS) through the Consortium for Molecular and Cellular Characterization of Screen-Detected Lesions. Dr. Basu has a PhD from Rockefeller University and an undergraduate degree in electrical engineering from Cornell University. (amrita.basu@ucsf.edu)



Harsh Bhatia, PhD, Lawrence Livermore National Laboratory

Dr. Harsh Bhatia is a computer scientist at the Center for Applied Scientific Computing at Lawrence Livermore National Laboratory (LLNL). His research spans the broad area of visualization and computational topology, with special focus on scientific data. Dr. Bhatia is also interested in machine learning-based approaches for scientific applications and use of visual analytics to understand performance data. Prior to joining LLNL, he earned his PhD from Scientific Computing and Imaging Institute at the University of Utah in 2015, where he worked on the feature extraction for vector fields. (bhatia4@llnl.gov)

Douglas Blayney, MD, Stanford University School of Medicine

Dr. Douglas Blayney is a professor of medicine at the Stanford University School of Medicine. His clinical interest is breast cancer, and his research focuses on quality improvement in cancer care systems, new drug development and improvement of the patient experience.

Dr. Blayney received the first Ellen Stovall Award for Innovation in Patient Centered Care, National Coalition for Cancer Survivorship, and the Joseph V. Simone Award and Lecture for Excellence in Quality and Safety in the Care of Patients with Cancer, American Society of Clinical Oncology (ASCO). He is past president of ASCO, was a founding member of the National Comprehensive Cancer Network Growth Factor Guideline panel and is a past member of the U.S. Food and Drug Administration's Oncology Drugs Advisory Committee. Dr. Blayney has worked in academic and community practice settings and has three daughters, one of whom is a physician. He received training at Stanford University; University of California, San Diego; and the National Cancer Institute. (DBlayney@stanfordhealthcare.org)



Gabriel Bordovský, Brno University of Technology

Gabriel Bordovský is a PhD candidate on the faculty of Information Technology, Brno University of Technology. He is a member of the Supercomputing Technologies Research Group where he focuses on improving the efficiency of spectral simulation codes running on clusters. He closely collaborates with PAMMOTH consortium on the development of a clinical prototype of a dedicated photo-acoustic breast imager. (ibordovsky@fit.vutbr.cz)



Jim Brase, Lawrence Livermore National Laboratory

Jim Brase is the deputy associate director for Computing at Lawrence Livermore National Laboratory. He leads LLNL research in the application of high-performance computing, large-scale data science and simulation to a broad range of national security and science missions. Jim is co-lead of the ATOM Consortium for computational acceleration of drug discovery and on the leadership team of the COVID-19 HPC Consortium.

Jim's research interests focus on the intersection of machine learning, simulation and high-performance computing. He is currently leading efforts in large-scale computing for life science, biosecurity and nuclear security applications. In his previous position as LLNL's Deputy Program Director for Intelligence, Jim led efforts in intelligence and cybersecurity R&D. He was formerly LLNL's Division Leader for Optical Sciences and led programs in laser and imaging research at LLNL. (brase1@llnl.gov)

V. K. Cody Bumgardner, PhD, University of Kentucky

V. K. Cody Bumgardner, PhD, is division chief of Pathology Informatics in the Department of Pathology and Laboratory Medicine. Dr. Bumgardner serves as assistant professor in the College of Medicine and by joint appointment College of Engineering at the University of Kentucky. Over the past 20 years he has held leadership roles in the areas of cyber infrastructure architecture, software development, telecommunication, research computing and healthcare. Dr. Bumgardner holds a bachelor's degree in computer engineering and a PhD in computer science from the University of Kentucky. He completed his academic training under Professor Victor Marek, with the dissertation titled "Contributions to Edge Computing." His work focuses on the applied use of edge computing, distributed inference and clinical-grade privacy-preserving AI in healthcare. (cody@uky.edu)



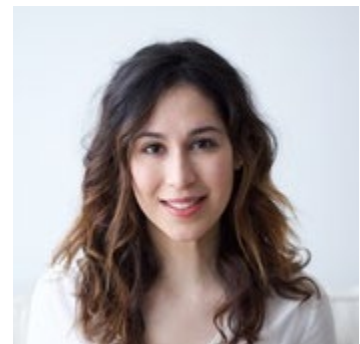


Judith Cohn, PhD, Los Alamos National Laboratory

Judith Cohn is a staff scientist at Los Alamos National Laboratory. She received a PhD in molecular biology from Cambridge University in 1981 and completed post-doctoral studies at the Salk Institute for Biological Studies and the Johnson Space Center. Following her post-doctoral studies, she spent seven years as a technical consultant for BBN Software Products, developing and supporting statistical software for scientific and engineering applications. During this period, she was also a part-time instructor in computer science, mathematics and statistics. In 1996, she joined the bioinformatics team for the Human Genome Project in Los Alamos. Since 2006, she has been a member of the Information Sciences group in the Computer, Computational and Statistical Sciences Division, where she designs, develops and runs software and databases for high-throughput data analysis in a variety of domain areas, primarily focused on biological problems. She was part of a small team that won an R&D 100 award in 2012 for Sequedex, a software application to classify phylogeny and function of metagenomic sequences. Currently, she is working on two large machine learning projects, which are collaborations between DOE and other federal agencies. The work presented here in particular is part of a pilot project with the National Cancer Institute under the auspices of the Joint Design of Advanced Computing Solutions for Cancer (JDACS4C) program. (jcohn@lanl.gov)

Diane Marie Del Valle, Icahn School of Medicine at Mount Sinai

Ms. Del Valle is a project director at Gnjatich Laboratory and Mount Sinai Human Immune Monitoring Core (HIMC). She is involved in the design and coordination of medical research with a focus on cancer immunology. (diane.delvalle@mssm.edu)



Sayera Dhaubhadel, Los Alamos National Laboratory



Sayera Dhaubhadel is a post-master's research associate at Los Alamos National Laboratory. Her background is in computer science, specializing in machine learning. She has been working on applying machine learning techniques on medical data to develop predictive models for a variety of tasks such as automatically predicting the type of cancer using cancer pathology reports, outcomes prediction and analysis for CVD, suicide attempts, etc. and understanding the importance of particular genes in the development of a tissue. (sayeradbl@lanl.gov)



Sally Ellingson, PhD, University of Kentucky

Dr. Sally Ellingson is a computational scientist working at the intersection of computational biology and high-performance computing. She has undergraduate degrees in computer science and mathematics from Florida Institute of Technology. She obtained her doctoral degree at the University of Tennessee and Oak Ridge National Laboratory under a fellowship funded by the National Science Foundation in computational biology. She is an assistant professor in the Division of Biomedical Informatics at the University of Kentucky College of Medicine. In her additional role as the

manager for High-Performance Computing Services for the Markey Cancer Center's Cancer Research Informatics Shared Resource Facility, she facilitates high throughput genomics and big data processing for precision medicine resulting in targeted cancer therapies. Recently she has been combining simulations, big data, and machine learning to increase the accuracy of drug binding predictions. With her passion for high-performance computing, her research goals lie in harnessing computational power for discoveries otherwise not possible in biomedical areas of high societal importance.

Dr. Ellingson engages in mentoring and outreach, especially for underrepresented groups in computational sciences. She has been actively engaged in the organization of the student programs at Supercomputing since 2014. She has helped organize the Broader Engagement program, spent several years helping with the Student Volunteer program, ran the Mentor-Protégé program and School Tours. This year she is Chair of Student Programming and managing the virtualization of this program. (sel228@uky.edu)

Fabian Fröhlich, PhD, Harvard Medical School

Dr. Fabian Fröhlich obtained his PhD in mathematics from Technische Universität München, Munich, where he worked in the groups of Jan Hasenauer and Fabian Theis. He developed mathematical methods for the efficient calibration of large kinetic models in the context of stochasticity and cell-to-cell variability and applied them to develop novel approaches for cancer precision medicine. He received a Master of Science degree from Technische Universität München in Mathematics in Biosciences. He is a HFSP postdoctoral fellow in Peter Sorger's group in the Laboratory of Systems Pharmacology at Harvard Medical School. He is developing methods to construct, calibrate and understand large kinetic models with hundreds to thousands of state variables and parameters, applying these methods to study adaptive resistance in melanoma where he integrates structural, molecular and omics data to build detailed mechanistic models. (fabian_froehlich@hms.harvard.edu)





Sacha Gnjatic, PhD, Icahn School of Medicine at Mount Sinai

Dr. Sacha Gnjatic is an associate professor and the co-director of the Human Immune Monitoring Center at Mount Sinai. After obtaining his PhD evaluating p53 as a tumor antigen in Paris, he trained with Dr. Lloyd Old at the Ludwig Institute in New York where he studied T cell and antibody responses specific to NY-ESO-1, in particular for antigen presentation and tolerance mechanisms. As a tumor immunologist, he specializes in studying immune correlates, both circulating and tissue-based, to define biomarkers of immunotherapies from human clinical trials.

sacha.gnjatic@mssm.edu

Emily Greenspan, PhD, National Cancer Institute

Dr. Emily Greenspan is a biomedical informatics program director in the Center for Biomedical Informatics and Information Technology (CBIIIT) at the National Cancer Institute (NCI). She serves as the NCI federal program lead for the NCI-Department of Energy (DOE) Collaborations focused on applying advanced computing and artificial intelligence (AI) technologies to specific areas of cancer research. She has supported and promoted predictive oncology and AI strategies across NCI. emily.greenspan@nih.gov



Sean E. Hanlon, PhD, National Cancer Institute



Dr. Sean E. Hanlon is acting deputy director of the NCI Center for Strategic Scientific Initiatives (CSSI) where he provides leadership in planning, developing and implementing initiatives with a focus on emerging areas of science with potential impact across the cancer research continuum. Dr. Hanlon is also the lead program director for the NCI's Provocative Questions Initiative that aims to foster research in understudied areas. Additionally, he provides scientific leadership to collaborative transdisciplinary programs, including the NIH Common Fund's 4D Nucleome program and the

NCI's Human Tumor Atlas Network. Dr. Hanlon also serves as a representative on NCI, NIH and inter-agency committees, including Cancer Moonshot Implementation teams, the trans-NCI Data Sharing working group and the trans-NCI Artificial Intelligence working group.

sean.hanlon@nih.gov



Stewart He, PhD, Lawrence Livermore National Laboratory

Dr. Stewart He joined Lawrence Livermore National Laboratory in 2017 after receiving his PhD in computer science from the University of California, Davis. Since then he has applied machine learning to several problems in the field of drug discovery, from predicting IC50 properties of drug compounds to using MD simulation data to predict cell membrane perturbations. Currently, Dr. He is part of the Accelerating Therapeutics for Opportunities in Medicine (ATOM) Consortium, with the mission of bringing together public and private groups to make advances in drug discovery. (he6@llnl.gov)

Tina Hernandez-Boussard, PhD, Stanford University

Dr. Hernandez-Boussard is an associate professor in medicine (biomedical informatics), biomedical data science and surgery at the Stanford University School of Medicine. Dr. Hernandez-Boussard's background and expertise is in the field of computational biology and epidemiology, with concentration on clinical informatics, population health and health policy. A key focus of her research is the application of novel methods and tools to large clinical datasets for hypothesis generation, comparative effectiveness research and the evaluation of quality healthcare delivery. (boussard@stanford.edu)



Julian Hong, MD, MS, University of California, San Francisco

Dr. Julian Hong is an assistant professor in the Department of Radiation Oncology and Bakar Computational Health Sciences Institute at the University of California, San Francisco. His research focuses on developing and implementing actionable analytics in oncology, including machine learning and artificial intelligence for prediction and decision making, data extraction and natural language processing and imaging analytics. Prior to joining UCSF, he completed his residency in radiation oncology at Duke University, where he collaborated with the Department of Biostatistics and Bioinformatics to complete a randomized controlled clinical study evaluating a machine learning-directed clinical intervention. He completed his Bachelor of Science in biomedical computation and Master of Science in bioengineering at Stanford University, his MD at the University of Wisconsin-Madison, and an internship in internal medicine at Yale-New Haven Medical Center. (julian.hong@ucsf.edu)



Rajeev Jain, Argonne National Laboratory

Rajeev Jain received his first master's degree in structural engineering from Arizona State University in 2009 and his second master's degree in computer science from the University of Chicago in 2020. He has been working at Argonne National Laboratory, since 2009 as a principal software developer in mathematics and computer science. He primarily develops simulation software and tools tailored for high performance systems. Over the past 11 years at ANL, he has been a part of development of tools for mesh generation, simulation of nuclear reactors, urban infrastructure and astrophysics. He also

develops deep learning algorithms, testing infrastructure and optimization routines for exascale deep-learning-enabled precision medicine for cancer treatment. (jain@mcs.anl.gov)

Sneha Jariwala, University of California, Santa Cruz

Sneha is a fourth-year undergraduate student at the University of California, Santa Cruz studying molecular, cellular and developmental biology. Sneha received a grant from Alex's Lemonade Stand Foundation to pursue a 10-week research project on gene expression analysis of pediatric high-risk leukemias. She completed this research with the Treehouse Childhood Cancer Initiative at the University of California, Santa Cruz Genomics Institute. This research means a lot to her because her sister passed away a few years ago from B-cell precursor acute lymphoblastic leukemia. Sneha is planning to graduate in March and apply for medical schools in summer to pursue a career in pediatric oncology. (sjariwal@ucsc.edu)



Jong Cheol Jeong, PhD, University of Kentucky



Dr. Jong Cheol Jeong is an assistant professor at the University of Kentucky College of Medicine, Department of Internal Medicine, Division of Biomedical Informatics, and manager of Cancer Research Data Common (CRDC) in Markey Cancer Center. He has been extensively involved in technology development and has significant expertise in designing and implementing data analysis and management tools for large-scale genomics projects. His current research interest is focused on the development and application of methods in translational biomedical informatics to better understand

disease pathogenesis, leading to the development of precision medicine and improved diagnostics and therapeutics for cancer patients. (jongcheol.jeong@uky.edu)



Donald Johann, MSc, MD, FACP, University of Arkansas for Medical Sciences

Dr. Johann is a physician/scientist, an associate professor at the University of Arkansas for Medical Sciences and the Scientific Director of the UAMS Genomics Sequencing Facility. His scientific focus concerns the application of advanced molecular profiling and high-throughput technologies for the characterization of molecular alterations in cancer cells. Areas of emphasis include next-generation sequencing (NGS), high-resolution identity-based mass spectrometry (proteomics), laser capture microdissection (LCM),

bioinformatics and cancer biology. Previously, he was an assistant investigator at the National Cancer Institute, Center for Cancer Research, Medical Oncology Branch in Bethesda, MD.

Prior to attending medical school, he worked as an engineer for the Unisys Corporation for six years, where he directed a team of five engineers on projects involving advanced avionic software design and instrumentation. During this time, he also earned a graduate degree in computer science with distinction from Hofstra University. Dr. Johann received his MD from Case Western and graduated with distinction honors for computer applications in medicine. Following residency, he then became a postdoctoral research fellow at the NIH/NCI Lab of Pathology, under the mentorship of Dr. Lance Liotta, with a focus on clinical proteomics. He was twice selected for an AACR Scholar-in-Training Award for research work involving novel bioinformatics. He completed medical oncology/hematology fellowships at the National Institutes of Health in the National Cancer Institute and National Heart, Lung and Blood Institute. (don.johann@gmail.com)

Sookyung Kim, PhD, Lawrence Livermore National Laboratory

Dr. Sookyung Kim is machine learning researcher at Lawrence Livermore National Laboratory. She has an MS from Columbia University and PhD from Georgia Tech. Currently, she is pursuing various research projects on deep learning for physical science, including climate data and material data analysis and AI driven drug design. She mainly works on design and development of new deep learning model to analyze massive-scaled scientific model output. Specifically, her current projects include prediction of future climate events, extreme climate analysis using object detection techniques, tracking tropical cyclone in re-analysis data, crystal structure prediction using reinforcement learning and multi-task drug property prediction using collaborative filtering. (kim79@llnl.gov)





Patricia Kovatch, Icahn School of Medicine at Mount Sinai

Patricia Kovatch is the Senior Associate Dean for Scientific Computing and Data Science at the Icahn School of Medicine at Mount Sinai, founding the division in October 2011. She is also an associate professor for the Department of Genetics and Genomic Sciences, the Icahn Institute for Data Science and Genomic Technology and Pharmacological Sciences.

Yaling Liu, PhD, Lehigh University

Dr. Yaling Liu is a professor in the Mechanical Engineering and Mechanics department and Bioengineering department at Lehigh University. Dr. Liu receives his PhD from Northwestern University. After graduation, Dr. Liu worked as a senior engineer for a year at Seagate Technology. Dr. Liu is an elected ASME fellow and recipient of a National Science Foundation CAREER Award and Interdisciplinary Research Excellence Award. His research interests include liquid biopsy, nanomedicine, biosensing, and application of machine learning in cancer diagnosis and monitoring. (yal310@lehigh.edu)



Benjamin Madej, PhD, Frederick National Laboratory for Cancer Research

Dr. Ben Madej is a data scientist with the ATOM Consortium and the Frederick National Laboratory for Cancer Research, developing new computational methods for drug discovery. At ATOM, he is evaluating new machine learning techniques for QSAR models applied to pharmacokinetics (PK) property prediction. Previously, he completed a research fellow position with Novartis in the computer-aided drug discovery group advancing mechanistic and ML models for PK property prediction. Before that, he completed his PhD and a postdoc at the University of California, San Diego specializing in molecular dynamics simulations of membrane protein drug targets. Dr. Madej's research interests have centered around applying computational models to support and accelerate drug discovery. (benjamin.madej@nih.gov)



Torin McDonald, Scientific Computing and Imaging Institute, University of Utah

Torin McDonald is pursuing a Master of Science degree under Dr. Valerio Pascucci at the Scientific Computing and Imaging (SCI) Institute at the University of Utah. His current research focuses on using topological data structures for interacting with biological data and visualizing feature tracking for large time series data sets. (torin@sci.utah.edu)

Adam Moody, Lawrence Livermore National Laboratory

Adam Moody earned a Bachelor of Science degree in computer science and engineering and a Bachelor of Science degree in engineering physics from The Ohio State University in 2001. In 2003, he earned a master's degree in computer science and a master's degree in electrical engineering also from The Ohio State University. Since 2004, he has worked as a computer scientist at Lawrence Livermore National Laboratory, where he conducts research, development, and support of system software on large-scale supercomputers. His interests include distributed computing, machine learning and fault tolerance for high-performance computing systems. He co-leads the Scalable Checkpoint/Restart and mpiFileUtils open source projects. For the past two years, he has been a member of the DOE CANDL project working to apply machine learning models to output datasets from simulations of RAS proteins. (moody20@llnl.gov)



Sharon Nirenberg, MD, Icahn School of Medicine at Mount Sinai

Dr. Sharon Nirenberg is a clinical informaticist with the Scientific Computing group at the Mount Sinai School of Medicine in New York, where she focuses on mining electronic health record data to support various research projects. Sharon has a medical degree from Rutgers University New Jersey Medical school and a Master of Science in medical informatics from Northwestern University. (sharon.nirenberg@mssm.edu)



Amanda Paulson, PhD, ATOM Consortium

Dr. Amanda Paulson received her PhD in cancer biology from the University of California, San Francisco in 2019. She previously studied the population dynamics of cells undergoing EMT state transitions and how this affected cell growth and response to chemotherapy. She is now a cancer drug discovery data science fellow at the ATOM Consortium. At ATOM, Amanda leads the drug-induced liver injury (DILI) team along with collaborators in the Arkin and Giacomini labs at UCSF. The team engages in a joint effort to develop and perform high throughput cell-based assays for hallmarks of DILI and to build predictive computational models from those assays. The team's ultimate goal is to create a comprehensive DILI prediction platform to identify signs of DILI early in the drug discovery process. (amanda.paulson@nih.gov)

Talita Perciano, PhD, Lawrence Berkeley National Laboratory

Dr. Talita Perciano acquired her PhD in computer science from the University of Sao Paulo, Brazil, in 2012. She is a research scientist in the Data Analytics and Visualization group and the Computational Biosciences group at Lawrence Berkeley National Laboratory. She conducts research in the areas of image analysis, machine/deep learning and high-performance computing motivated by the incredible challenges around scientific data generated by computational models, simulations and experiments. Her research focuses on mathematical foundations for new methods, implementation of scalable methods and platform-portability. Her goal is to develop powerful, mathematically grounded, scalable algorithms that meet the requirements needed to analyze current and future scientific datasets acquired in the U.S. Department of Energy user data facilities. She has built a diverse collaboration network throughout the years in fields such as materials science, biosciences, geosciences, among others. (tperciano@lbl.gov)



Geir Severin Elvatun Rakh Langberg, Cancer Registry of Norway

Severin Langberg is a PhD student at the Cancer Registry of Norway, working on personalized cervical cancer screening using matrix factorization and geometric deep learning techniques. His primary research interest entails computational modelling of biomedical phenomena and applied mathematics. Severin's previous experiences cover a broad range of topics spanning from image and video analysis using machine learning and deep learning methods, via mining of gene expression data to molecular diffusion modelling in chemical reactors. Severin has an MS degree in data science from the Norwegian University of Life Sciences and a BS degree in chemistry and biotechnology from Oslo Metropolitan University. (langberg91@gmail.com)



Carlos Simmerling, PhD, Stony Brook University

Dr. Carlos Simmerling obtained his bachelor's degree and PhD in chemistry at the University of Illinois at Chicago, performing early research on methods for computer modeling of biomolecules such as proteins. He went on to a post-doctoral fellowship in pharmaceutical chemistry at the University of California, San Francisco, where he became a lead developer of the Amber biomolecular simulation software that is used in thousands of research labs worldwide. In 1998, Dr. Simmerling joined the chemistry department at Stony Brook University, where he is currently a professor, and he became the associate director of SBU's Laufer Center for Physical & Quantitative Biology. His research is funded by the National Institutes of Health, National Science Foundation and Department of Energy. His work focuses on development of improved molecular simulation methods and models and using these tools to study biomolecular recognition mechanisms. His articles on improving the physics underlying biomolecular modeling have been cited nearly 10,000 times. Prof. Simmerling is currently the Marsha Laufer Chair of Physical & Quantitative Biology at Stony Brook University and a Fellow of the American Chemical Society. (carlos.simmerling@stonybrook.edu)

Ashka Shah, University of Chicago

Ashka Shah is a second-year PhD student at the University of Chicago and works on deep learning methods for virtual screening and explainable AI in medicine. She previously worked as a software engineer at Lawrence Livermore National Laboratory. Her background is in physics and computer science. (shahashka@uchicago.edu)



Eric Stahlberg, PhD, Frederick National Laboratory for Cancer Research

Dr. Eric Stahlberg is the director of Biomedical Informatics and Data Science (BIDS) at the Frederick National Laboratory for Cancer Research. Dr. Stahlberg is a founding co-organizer of the Computational Approaches for Cancer Research Workshop (CAFCW), held in conjunction with the SC conference since 2015. He has been instrumental in establishing the Frederick National Laboratory's high-performance computing initiative and in assembling scientific teams across multiple, complex organizations to advance predictive oncology. In 2017, he was recognized as one of FCW's Federal 100. (eric.stahlberg@nih.gov)





Mengdi Tao, University of Pennsylvania

Mengdi Tao is a PhD student of Dr. Ravi Radhakrishnan's lab, in the Department of Bioengineering at the University of Pennsylvania. She received her BS and MS degrees in biomedical engineering from Drexel University. A major focus of her ongoing research is to understand quantitative mechanistic characterization of structural biology and systems biology. The ultimate aim is to build hybrid and multiscale quantitative models of signaling networks while retaining sufficient molecular specificity, for predicting the interactions of therapeutic agents with biochemical signaling mechanisms.

(taome@seas.upenn.edu)

Hong-Jun Yoon, PhD, Oak Ridge National Laboratory

Dr. Hong-Jun Yoon is a research scientist at Oak Ridge National Laboratory. He is involved in the Joint Design of Advanced Computing Solutions for Cancer (JDACS4C), researching on natural language processing and information extraction from medical texts. His research interest is artificial intelligence, machine learning and population health. (yoohn@ornl.gov)



Max Zvyagin, Argonne National Laboratory

Max Zvyagin is a predoctoral appointee at Argonne National Lab working in the Data Science and Learning division. His background is in computer science and neuroscience. (vzvyagin@uchicago.edu)