UNIVERSITY of VIRGINIA

iTHRIV SCHOOL of DATA SCIENCE ARTS & SCIENCES
Department of Psychology

THE BIOMEDICAL
DATA SCIENCE
INNOVATION LAB
2016-2023

From the DIRECTOR

The Biomedical Data Science Innovation Lab (BDSIL) stands as a pioneering initiative, generously supported by the NIH, that converges the realms of biomedical sciences and data science. This unique platform brings together junior faculty members spanning diverse disciplines such as clinical medicine, bench research, biomedical engineering, computer science, statistics, and mathematics. Its primary objective is to foster interdisciplinary collaboration and stimulate innovation by addressing specific challenges in the biosciences through the application of cutting-edge data science methods.

At the heart of BDSIL's mission is the cultivation of new, multidisciplinary partnerships. These collaborations are meticulously facilitated to tackle challenges in biomedicine where data science plays a pivotal role. The annual selection of unique but topical themes ensures a dynamic and evolving focus, reflecting the rapidly advancing landscape of biomedical research and computational applications.

The BDSIL experience is characterized by a multifaceted approach. Participants engage in weekly webinars featuring thought leaders in the field, gaining insights into the latest developments and best practices. The 'micro-lab' activities provide interactive opportunities to discuss data science methods to real-world challenges, advancing interdisciplinarity, and innovative thinking.

A cornerstone of BDSIL is the five-day, mentored, and facilitated grant and manuscript development workshop. This intensive collaborative setting serves as a catalyst for translating ideas into tangible outcomes. The emphasis on mentorship ensures that participants receive guidance from experienced professionals, enhancing the quality and viability of their proposals and manuscripts. With an amazing team of facilitators – who act as ring masters, cruise directors, and cheerleaders – provide overall structure to these events. An amazing team of BDSIL staff members coordinates all event logistics, ensuring that these events run smoothly.

Themes from previous years, such as the microbiome, the brain, mobile devices, environmental exposures, bioethics, and biomedical data science education in artificial intelligence, underscore the breadth of topics addressed. This diversity reflects the wide-ranging impact that data science can have across various facets of biomedicine.
The BDSIL takes pride in its role as a catalyst for innovation and creativity in the application of data science to biomedical challenges. The ultimate goal is to not only generate new knowledge but also to facilitate the development of projects that have the potential to secure NIH or NSF grant funding. Moreover, the emphasis on resulting in new peer-reviewed research articles underscores the commitment to contributing meaningfully to the scientific literature and the career paths of our participants.

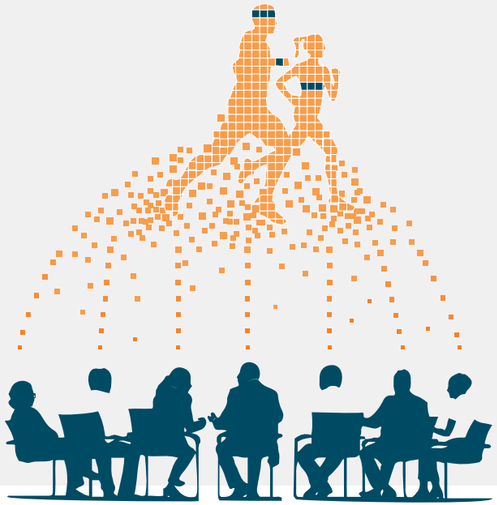
Throughout all BDSIL activities, the concept of Team Science is omnipresent. Recognizing that complex challenges in biomedicine require collaborative efforts, the program fosters a culture of teamwork. Participants are encouraged to leverage diverse expertise, perspectives, and skills, emphasizing that breakthroughs often emerge at the intersection of disciplines.

The Biomedical Data Science Innovation Lab represents a dynamic and forward-thinking initiative at the crossroads of biomedical science and modern computational applications. By nurturing collaboration, fostering innovation, and embracing the principles of Team Science, BDSIL stands as a beacon in the quest to unravel the complexities of biomedicine through the transformative power of data science.

John Darrell Van Horn, Ph.D., M.Eng.

2016

Biomedical Data Science Challenges of Wearable and/or Ambient Sensors



Achieving the potential of wearable and ambient sensors will necessitate the integration of different data formats, including structured and unstructured data, realtime and static data, from diverse populations interacting with a range of devices including but not limited to wearable electronics, mobile devices, and environmental sensors. This data requires the development of computational and analytical methods to enable high-confidence predictions that relay just in time information to individuals through a personalized user interface and experience (UI/UX) for the greatest impact. The teams developed at the lab created project ideas that addressed this data science challenge.

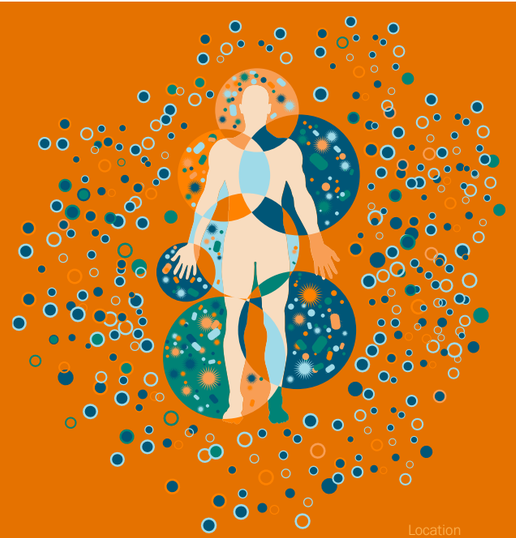
Location
Lake Arrowhead, CA

Date
June 15-19, 2016

2017

Quantitative Approaches to Biomedical Data Science Challenges in our Understanding of the Microbiome

The study of the microbiome is a rapidly developing area with the aim of identifying, treating, and preventing disease as well as promoting health by understanding the interaction of microbes with humans in a variety of ecological niches both within the human body and the external environment. The projects generated at the lab approached the relevant data associated with the microbiota with a health or biomedical research objective.



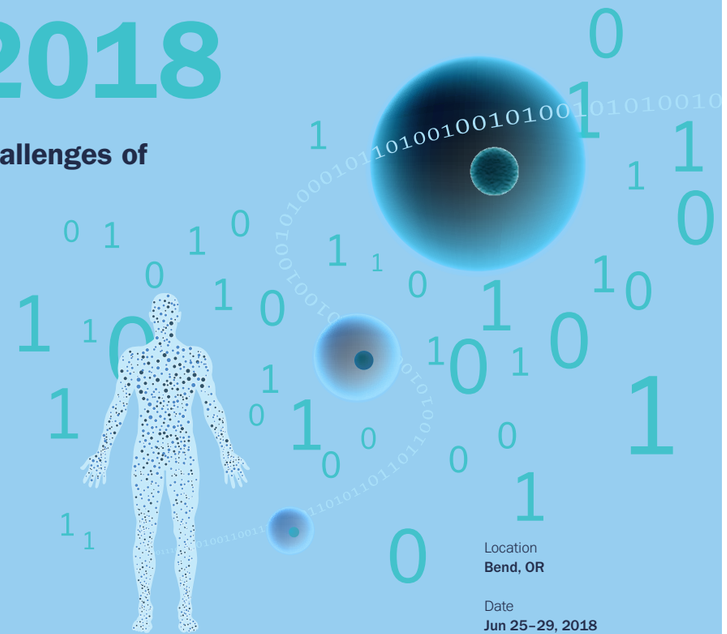
Location
Beverly, MA

Date
June 19-23, 2017

2018

Mathematical Challenges of Single Cell Dynamics

Understanding the implications of single cell heterogeneity is critical for developing personalized treatments. The Lab promoted collaboration between mathematicians, statisticians, and biomedical researchers towards the development of novel or significantly adapted models, methods, and approaches for overcoming difficult data science challenges arising from the collection and analysis of single cell big data.



Location
Bend, OR

Date
Jun 25–29, 2018

2019

Data Science Challenges in Rural Health and Environmental Exposures

Location
Winston-Salem, NC

Date
June 24-28, 2019

Rural health and environmental exposures specifically concerns communities in less populated areas of the country where exposure to pollutants, toxins, chemicals, etc have a disproportionate impact on citizen health and wellbeing. The immediate and long-term consequences of these exposures have a direct impact on life quality due to reduced access to health care, policies of local government, health disparities, changes to diet and exercise, and corporate interests. Collaborations between biomedical and quantitative scientists lead to project ideas with better approaches to implementation and interpretation of such data.

The brain is recognized as a major source of microscopic, systems-level, spatial, and temporal datatypes in health as well as in disease. This BDSIL aimed to highlight the challenges of working with such datatypes and how data might be integrated to gain insights into brain form, function, and connectivity and the understanding of major clinical disorders.

2021

Brain Analytics and Data Integration

Location
Virtual

Date
June 21-25, 2021

Artificial Intelligence (AI) has great potential to assist in biomedical decision making. However, such systems are not immune from making erroneous recommendations, struggling to maintain patient privacy, and which give rise to vexing questions about their suitability across genders, ethnic, or cultural communities. The lab formed new interdisciplinary collaborations that generated creative strategies for addressing ethics of AI in biomedicine.

2022

Ethical Challenges of Artificial Intelligence in Biomedicine

Location
Charlottesville, VA

Date
June 13-17, 2022

2023

Data Science and the Public Health Consequences of the COVID-19 Pandemic

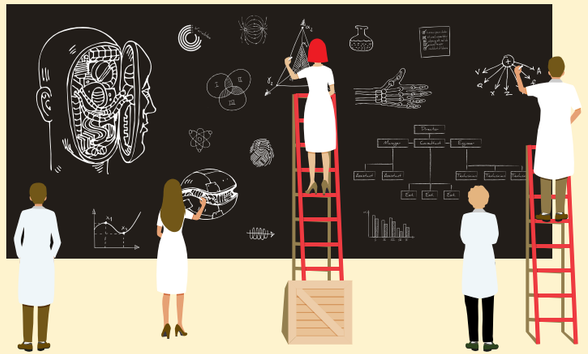
The COVID-19 pandemic has upset our lives in ways unimaginable. We are only now, just contending with the secondary effects of COVID exposures – so-called 'long-haul COVID' – but also the effects of deferred regular health check-ups, poor mental health, etc. This lab fostered the formation of new interdisciplinary collaborations to generate creative strategies on the use of data science approaches for predicting secondary health effects of the COVID-19 pandemic.

Location
Kenmore, WA

Date
June 26-30, 2023

Building Partnerships for Generative AI Training in Biomedical and Clinical Research

2024



How can those with expertise in the development of AI, large-language, and generative content models join forces with biomedical science educators to provide insight and guidance on training about, using, and leveraging these tools?

The goal of the 2024 Biomedical Data Science Innovation Lab (BDSIL) is to foster the formation of new interdisciplinary collaborations which will generate creative strategies on the use of data science approaches for creating new strategic partnerships between AI developers, biomedical researchers, and educators.

Location
San Diego, CA

Date
June 16-20, 2024

Process

Microlabs

Making initial connections, developing the foundation, and exploring the problems space prior to the in-person Innovation Lab

Knowledge in the Room

Learning about attendees, their expertise and bringing together thoughts, ideas, insights and questions

Day 1

Stewarding Rounds

Progect idea rounds of exploration, working through obstacles, and receiving guidance

Day 3

Seeds of Research Ideas

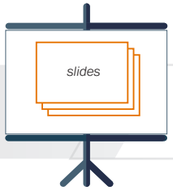
Aspirational challenges and identifying potential research projects

Day 2

Presentations & Feedback

Refining projects, and final presentations of the 6-10 groups

Days 4 & 5



Post-Lab

Mentor feedback, 4 month post-lab team check-in, and continued scientific outcomes

Roles

Participants are early career level advanced post-doctoral fellows, coming from biomedical or data science research, seeking to take on interesting, data-driven challenges engaging with new, multidisciplinary teams in biomedicine.

Mentors are scientifically established senior investigators able to provide experience, encouragement, and support. They do not control – they merely meet with the forming teams to provide feedback, guidance, and direction to orient them, reinforcing the positive aspects of potential team projects while asking clarifying questions about areas in need of improvement.

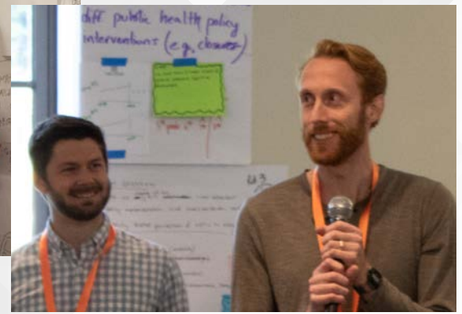
Provocateurs are domain experts who give short talks to challenge, push, encourage, and disrupt, expanding the bounds of thinking. Provocateurs have been inventors, entrepreneurs, astronauts, microbiologists, and neurogeneticists. Each has the goal to off balance in a good way to help see problems in a new light.

Facilitators are a combination or tour guides, cruise directors, and ring masters. They run the day-to-day agenda, coordinate our activities, and bring their own positive attitudes to bear on all things we hope to achieve.

The UVA-based **Leadership Team** provides direction, coordination, and assistance throughout. Theirs is a year-long level of participation, planning, promotion, support, and number one points of contact before, during, and after each event.

know innovation

Knowledge in the Room



The photographs in this book are a compilation of all the pictures taken during the BDSIL events from 2016 through 2023.

Day 1


HASH TAGS
#DSIL2019
#InnovationLab2019
#RuralHealth
#EnvironmentalExposures
TAG US @
BD2KTcc

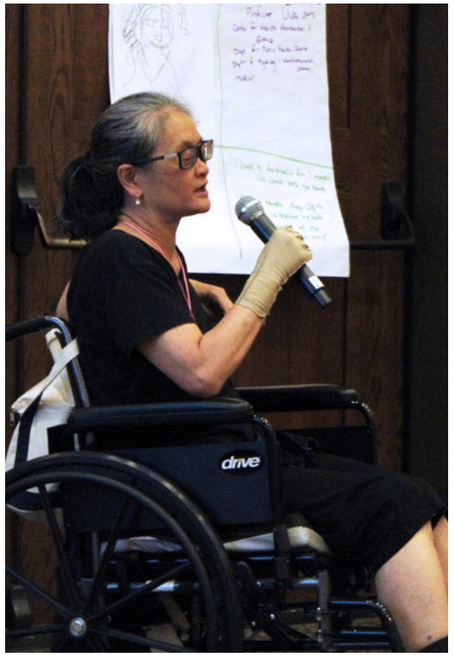
"We help to encourage interdisciplinary teams, comprised of biomedical and computational scientists, to think differently, spark new ideas, develop novel research projects, challenge themselves, and to be innovative."
-Jack Van Horn, Director

"It's really exciting. I never had this kind of engaging, interactive activities."
-2019 Participant

"This was an experience unlike anything I've had in my career and it was great!"
-2018 Participant

"The whole workshop is so dynamic."
-2022 Participant

"It is an outstanding opportunity to get to know some brilliant minds and connect."
-2022 Participant



"In setting up new collaborations, it's a great opportunity... I think one of the big reasons for coming to one of these things is definitely to talk about science and learn new things, but really from a community building standpoint, this is an amazing experience." -Laura Heiser, Mentor 2018


AHA GARDEN

Seeds of Research Ideas

Day 2

"It's really exciting. I never had this kind of engaging interactive activities that made my brain constantly thinking. This is really accelerated learning...Whenever you have questions, you directly get answers."
- 2018 Participant

“It was an opportunity to build collaborations that would take a year or more, in a single week.” -Anonymous

“One of the best academic experiences I have ever participated in!” -2018 Participant

“I believe a team comprising of people with a diverse background, minds and ideas has the potential to solve any complex problem. I was impressed with the ability of the lab organizers to bring computational and experimental researchers to brainstorm about science for a few days in a single room.” -Anonymous

“The BDSIL was such a great opportunity for me to be a part of. Meeting other early investigators from around the country interested in common research themes has been magical for me.” -2021 Participant

Stewarding Rounds

Day 3

"One of the things that we just love about this event is that everybody here was not only really interested in collaborating and working with each other, brainstorming and coming up with novel ideas; everybody is also very nice." -2018 Participant

"It was the most productive workshop I have attended." -2022 Participant

"It's super exciting and very stimulating. It's a very unique opportunity to make a lot of progress very quickly." -2019 Participant


Learn/collaborate
validation
Sonia

“What a great experience this was! Excellent speakers and mentors! But the best part was meeting all the incredibly smart and motivated participants, each bringing ideas and different expertise to the table!”
-2021 Participant

Day 4 & 5

Presentations and Feedback

Day 4
Presentations and Feedback

Day 5
Final Presentations

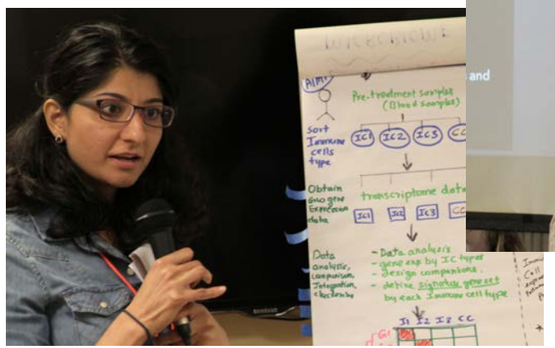
"It was an extremely positive experience that both increased my research network and prompted novel and exciting collaborations."
-2018 Participant


OMICS TO
DYNAMIC MODEL
Maximize efficacy
of chemotherapy

"I have had an invigorating week at BDSIL 2022 working on solving big problems in AI and AI Ethics in healthcare. It was such a productive week. Our team has already submitted an LOI!"
-2022 Participant

“The whole process was magical. At the end you have an amazing set of projects, and people that you legitimately want to work and hangout with.”
-2022 Participant

"Thanks to the BDSIL for the wonderful experience. Great networking and collaborative learning opportunity."
-2023 Participant



“I made so many great connections and left bursting with energy.” -2023 Participant

“It was an eye-opening experience for me. I would definitely recommend it to my colleagues.” -2019 Participant

knowinnovation

"What I really get excited about is when you talk to people five years after they've been to one of these events and they say, "Yeah, I've been working with so-and-so on two different projects now for the last three years and we met five years ago at an Innovation Lab!" -Tim Dunne, KI Facilitator

THANK YOU

THE NATIONAL INSTITUTES OF HEALTH

The National Institute of General Medical Sciences (NIGMS) plays a pivotal role in advancing the mission of the Biomedical Data Science Innovation Lab (BDSIL) and its primary objective of cultivating collaborative relationships between junior faculty in biomedical sciences and those specializing in data science, statistics, mathematics, computer science, and related fields. NIGMS funding has been instrumental in realizing the fundamental idea behind BDSIL, as it facilitates the formation of multidisciplinary partnerships through Team Science. Importantly, NIGMS funding directly contributes to the promotion of the career paths of junior faculty, providing them with invaluable opportunities to experience the latest in cutting-edge biomedical data science research, foster cross-disciplinary collaborations, and make significant contributions to the evolving landscape of biomedical data science. In essence, NIGMS support ensures the sustainability and success of the BDSIL's mission to advance knowledge at the intersection of biomedical and data sciences. Additional contributions to BDSIL success have also come from the National Institute for Environmental and Health Sciences (NIEHS) and from the National Cancer Institute (NCI). We sincerely thank these NIH institutes for their invaluable support.

Appendix

Scientific Outcomes and Measures of Success

PUBLICATIONS AND CONFERENCE PRESENTATIONS

Alambo, A., et al. (2020) Measuring Pain in Sickle Cell Disease using Clinical Text. Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE Engineering in Medicine and Biology Society, Annual International Conference; 2020:5838-584.

Alekseyenko, A., et al. (2021) Each patient is a research biorepository: informatics-enabled research on surplus clinical specimens via the living BioBank. *Journal of the American Medical Informatics Association*; 28(1): 138–143.

Choi, K., et al. (2020) Bayesian model selection reveals biological origins of zero inflation in single-cell transcriptomics. *Genome Biology*; 21, 183.

Chung, A., et al. Precision VISSTA: Bring-Your-Own-Device (BYOD) mHealth Data for Precision Health. Podium presentation. American Medical Informatics Association Annual Meeting; November 2019; Washington, DC.

Clifton, S., et al. (2017) Hybrid Statistical and Mechanistic Mathematical Model Guides Mobile Health Intervention for Chronic Pain. *Journal of Computational Biology* : a journal of computational molecular cell biology; 24(7):675-688.

Corlin, L., et al. Sex as an effect modifier of the association between co-exposure to multiple toxic metals and the risk of type 2 diabetes. NIH Office of Research on Women's Health BIRCWH meeting; December 2020; virtual.

Cullen, C., et al. (2020) Emerging Priorities for Microbiome Research. *Frontiers in Microbiology*; 19;11:136.

Irimia, A., et al. (2017) Mobile monitoring of traumatic brain injury in older adults: challenges and opportunities. *Neuroinformatics*; 15(3): 227–230.

Johnson, A., et al. (2019) Use of Mobile Health Apps and Wearable Technology to Assess Changes and Predict Pain During Treatment of Acute Pain in Sickle Cell Disease: Feasibility Study. *Journal of Medical Internet Research Publications*, 7(12), Article e1367.

Jonassaint, C., et al. (2017) Understanding patterns and correlates of daily pain using the Sickle cell disease Mobile Application to Record Symptoms via Technology (SMART). *British Journal of Haematology*, 183 (2).

Marai, G. E., et al. (2019) Precision Risk Analysis of Cancer Therapy with Interactive Nomograms and Survival Plots. IEEE transactions on visualization and computer graphics, 25(4):1732-1745.

Morrorw, B., et al. (2019) Periphery Plots for Contextualizing Heterogeneous Time-Based Charts. IEEE Transactions on Visualization and Computer Graphics; 2019:1-5.

Riseberg, E., et al. (2021) Development and application of an evidence-based directed acyclic graph to evaluate the associations between metal mixtures and cardiometabolic outcomes. *medRxiv* 2021.03.05.21252993.

Riseberg, E., et al. A Multipollutant, longitudinal analysis of the associations between urinary tungsten and incident diabetes in a rural population. International Society of Environmental Epidemiology; August 2020; Virtual 2020.

Stingone, J., et al. (2021) Interdisciplinary data science to advance environmental health research and improve birth outcomes. *Environmental Research*, 197:111019.

Tosado, J., et al. (2020) Clustering of Largely Right-Censored Oropharyngeal Head and Neck Cancer Patients for Discriminative Groupings to Improve Outcome Prediction. *Scientific Reports*; 10(1):3811.

Wentzel, A., et al. (2020) Cohort-based T-SSIM Visual Computing for Radiation Therapy Prediction and Exploration. IEEE Transactions on Visualization and Computer Graphics; 26(1):949-959.

Wentzel, A., et al. (2020) Precision toxicity correlates of tumor spatial proximity to organs at risk in cancer patients receiving intensity-modulated radiotherapy. Radiotherapy and oncology : *Journal of the European Society for Therapeutic Radiology and Oncology*, 148:245-251.

Yang, F., et al. (2018) Improving Pain Management in Patients with Sickle Cell Disease from Physiological Measures Using Machine Learning Techniques. *Smart Health* (Amsterdam, Netherlands); 7-8:48-59.

Yang, F., et al. (2019) Continuous Pain Assessment Using Ensemble Feature Selection from Wearable Sensor Data. Proceedings. IEEE International Conference on Bioinformatics and Biomedicine; 2019: 569-576.

Zdilar, L., et al. (2018) Evaluating the Effect of Right-Censored End Point Transformation for Radiomic Feature Selection of Data From Patients With Oropharyngeal Cancer. *JCO Clinical Cancer Informatics*; 2:1-19.

AWARDS

GRANT NUMBER	PROJECT TITLE
1557576; 1557593; 1557642; 1557668; 1557589	QuBBD: Collaborative Proposal: Interactive Ensemble Clustering for Mixed Data with Application to Mood Disorders
1557765; 1557722; 1557727; 1557730; 1557712	QuBBD: Collaborative Research: Advancing mHealth using Big Data Analytics: Statistical and Dynamical Systems Modeling of Real-Time Adaptive m-Intervention for Pain
1557733; 1557739; 1557742	QuBBD: Collaborative Research: Precision medicine and the management of infectious diseases
1557679; 1557565; 1557559; 1557578	QuBBD: Collaborative Research: SMART -- Spatial-Nonspatial Multidimensional Adaptive Radiotherapy Treatment
1557466	QuBBD: Estimating drug-drug and drug-disease interactions for nursing homes residents
5R01CA214825	SMART-ACT: Spatial Methodologic Approaches For Risk Assessment And Therapeutic Adaptation In Cancer Treatment
5R01CA225190	QUBBD: Precision E –Radiomics For Dynamic Big Head & Neck Cancer Data
R01EB025024	QUBBD: Statistical & Visualization Methods For Pghd To Enable Precision Medicine
3R01AT010413	SCH: INT: Collaborative Research: Development And Analysis Of New Mathematical And Statistical Models For Chronic Pain
1R21TR002513	Increasing Access To Clinical Microbiome Specimens Via A Living μbiome Bank
1R01CA258827	Longitudinal Spatial-nonspatial Decision Support For Competing Outcomes In Head And Neck Cancer Therapy
1R01HL155945	Multiscale Modeling Of Right Ventricular Fibrotic Remodeling In Pulmonary Arterial Hypertension
1R01ES032612	The Impact of Drought on Arsenic Exposure and Cardiometabolic Outcomes in a Rural Aging Population

Attendees

2023 ATTENDEE LIST

PARTICIPANTS

Maria Alva *Georgetown University*
Prajakta Bedekar *Johns Hopkins University*
Felix Bradbury *MITRE Corporation*
Zachary Butzin-Dozier *University of California, Berkeley*
Glenda Canderan *University of Virginia*
Edward Chiyaka *Wingate University*
Pei-Ying Chuang *University of Maryland*
Nicholas DeFelice *Icahn School of Medicine at Mount Sinai*
Kevin Desai *The University of Texas at San Antonio*
Arjun Dirghangi *University of Virginia School of Medicine*
Scott Greenhalgh *Siena College*
Aya Haghamad *Northwell Health*
Tomás León *California Department of Public Health*
Chang Li *University of South Florida*
Haidong Lu *Yale University*
Rayanne Luke *Johns Hopkins University*
Pulong Ma *Clemson University*
Chelsea Marie Braun *University of Virginia*
Abolfazl Mollalo *Baldwin Wallace*
Olga Morozova *University of Chicago*
Lyndsey Muehling *University of Virginia*
Mayuri Panditrao *California Department of Public Health*
Albert Park *University of North Carolinea at Charlotte*
Jimmy Phuong *University of Washington*
Geetha Saarunya Clarke *Stanford School of Medicine*
Heman Shakeri *University of Virginia*
George Shaw, Jr. *University of North Carolina at Charlotte*
Benjamin Sines *University of North Carolina at Chapel Hill*
Faith Summersett Williams *Northwestern University*
Maria Sundaram *Marshfield Clinic Research Institute*
Hannah Thompson *University of California, Berkeley*
Mengxi Zhang *Virginia Tech Carilion School of Medicine*

MENTORS

Vivien Bonazzi *Deloitte*
Kathy Kim *MITRE*
Madhav Marathe *University of Virginia*
Vinay Pai *U.S. Food and Drug Administration*
Julia Roper *Panorama Strategy*

PROVOCATEURS AND OTHER GUESTS

Alex Bui *University of California, Los Angeles*
Jennifer Couch *National Institutes of Health*
Samantha Finstad *National Institutes of Health*
Shawn O'Neil *University of Colorado*
Micaela Parker *Academic Data Science Alliance*
William Petri *University of Virginia*
Carol Shreffler *National Institutes of Health*

KI FACILITATORS

Tim Dunne
Jocelyn Tejeda

2022 ATTENDEE LIST

PARTICIPANTS

Md Ashad Alam *Tulane University*
Andrew Barros *University of Virginia*
Ganapati Bhat *Washington State University*
Anne Brown *Virginia Tech*
Benjamin Collins *Oregon Health and Science University*
Laura Corlin *Tufts University*
Arjun Dirghangi *University of Virginia*
Afsaneh Doryab *University of Virginia*
Beth Epstein *University of Virginia*
Diana Ferro *Children's Mercy Institute*
Mingchen Gao *University at Buffalo, SUNY*
Eleanor Gilmore-Szott *Baylor College of Medicine*
Eric Hixson *Cleveland Clinic*
Athmeya Jayaram *Johns Hopkins University*
Allison Kelli Coffelt *Clinical Humanities Commons*
Stephanie Kraft *University of Washington*
Jundong Li *University of Virginia*
Jenn Maccormack *University of Virginia*
Kellie Owens *University of Pennsylvania*
Jonathan Pinney *Gryphon*
Susannah Rose *Cleveland Clinic*
Nawar Shara *MedStar Health Research Institute*
Alexis Walker *Columbia University*
Alexander Weston *Mayo Clinic Florida*
Nicole Wilson *The University of Rochester*
Chenglong Ye *University of Kentucky*
Miaomiao Zhang *University of Virginia*

MENTORS

Elie Alhajjar *United States Military Academy*
Donna Chen *University of Virginia*
Renée Cummings *University of Virginia*
Caitlin Donahue Wylie *University of Virginia*
Brian Wright *University of Virginia*

PROVOCATEURS AND OTHER GUESTS

Jennifer Couch *National Institutes of Health*
Judy Illes *University of British Columbia*
Carol Shreffler *National Institutes of Health*
Walter Sinnott-Armstrong *Duke University*
Daniel Steger *Center for Open Science*

KI FACILITATORS

Tim Dunne
Effie Kistner

2021 ATTENDEE LIST

PARTICIPANTS

Alyssa Ailion *Boston Children's Hospital*
Elie Alhajjar *United States Military Academy*
Ayham Alkhachroum *University of Miami*
Benedict Anchang *National Institutes of Health*
Moritz Armbruster *Tufts University*
Christine Cooper *Medical University of South Carolina*
Judith Dexheimer *University of Cincinnati*
Tanya Evans *University of Virginia*
Shannon Farris *Virginia Tech*
Ignacio Fernandez Mata *Cleveland Clinic*
Mark Fiecas *University of Minnesota, Twin Cities*
Brandon Foreman *University of Cincinnati*
William Howe *Virginia Tech*
Jennifer Kim *Yale University*
Sarah Shizuko Morimoto *University of Utah School of Medicine*
Chia-Ling Phuah *Washington University School of Medicine in St. Louis*
Renato Polimanti *Yale University*
Meghan Puglia *University of Virginia*
Nathan Rowland *Medical University of South Carolina*
Karen Schliep *University of Utah Health*
Jeffrey Thompson *University of Kansas Medical Center*
Daniel Tward *University of California, Los Angeles*
Brian Wright *University of Virginia*
Kirsten Wright *Oregon Health & Science University*
Zhengwu Zhang *University of Rochester*
Yi Zhao *Indiana University*

MENTORS

Aidong Zhang *University of Virginia*
Chongzhi Zang *University of Virginia*
Don Brown *University of Virginia*

PROVOCATEURS AND OTHER GUESTS

Phil Bourne *University of Virginia*
Jennifer Couch *National Institutes of Health*
Randy Gollub *Harvard University*
Karen Johnston *University of Virginia*
Jaideep Kapur *University of Virginia*
Carol Shreffler *National Institutes of Health*

KI FACILITATORS

Annemarie Boss
Tim Dunne

2019 ATTENDEE LIST

PARTICIPANTS

Tanya Alderete *University of Colorado Boulder*
Elie Alhajjar *Unites States Military Academy*
Marissa Baker *University of Washington*
Andres Cardenas *University of California, Berkeley*
Laura Corlin *Boston University*
Hadiza Galadima *Old Dominion University*
Pierce Greenberg *Creighton University*
Matthew Gribble *Emory University*
Lynda Hardy *The Ohio State University*
Susan Hutfless *Johns Hopkins University*
Katherine James *Colorado School of Public Health*
Jeremy Jay *University of North Carolina at Charlotte*
Jay Kitt *University of Utah*
David Kline *The Ohio State University*
Alexandra Larsen *Duke University School of Medicine*
Gang Liu *Arbor Research Collaborative for Health*
Valerie Mac *Emory University*
Judit Marsillach Lopez *University of Washington*
Rachel Melamed *University of Chicago*
Jennifer Moss *Penn State*
Amanda Nguyen *University of Virginia*
Ai (Andy) Ni *Ohio State University*
Giorgio Quer *Scripps Research*
David Russell *Appalachian State University*
Lincoln Sheets *University of Missouri*
Jeanette Stingone *Columbia University*
Ellen Swanson *Centre College*
Sofia Triantafillou *University of Pittsburgh*
Kai Zhang *The University of Texas Health Science Center at Houston*
Jiwei Zhao *State University of New York at Buffalo*

MENTORS

Arlene Chung *University of North Carolina at Chapel Hill*
Gerald Denis *Boston University*
Tyrone Hayes *University of California, Berkeley*
Kathy Kim *University of California, Davis*
Herbert Sauro *University of Washington*

PROVOCATEURS AND OTHER GUESTS

Jennifer Couch *National Institutes of Health*
Chris Duncan *National Institutes of Health*
Eric Johnson Chavarria *National Institutes of Health*
Carol Shreffler *National Institutes of Health*
Lance Waller *Emory University*

KI FACILITATORS

Annemarie Boss
Tim Dunne
Effie Michailidis
Donnalyn Roxey

2018 ATTENDEE LIST

PARTICIPANTS

Benedict Anchang *Stanford University*
Rhonda Bacher *University of Florida*
Michael Blinov *University of Connecticut*
Bryan Bryson *Harvard University*
Young Chang *Oregon Health and Science University*
Bin Chen *University of California, San Francisco*
Yang Chen *University of Michigan*
Haiyan Cheng *Willamette University*
Kwangbom Choi *The Jackson Laboratory*
Roy Dar *University of Illinois at Urbana-Champaign*
Audrey Hendricks *University of Colorado-Denver*
Stephanie Hicks *Johns Hopkins*
Sahand Hormoz *Harvard University/ Dana-Farber Cancer Institute*
Yuchao Jiang *University of North Carolina, Chapel Hill*
Michael Konopka *The University of Akron*
Minjoung Kyoung *University of Maryland, Baltimore County*
Javier Lopez *University of California, Davis*
Jing Ma *Fred Hutch Cancer Research Center*
Michael Mak *Yale University*
Aaron Meyer *University of California, Los Angeles*
Eran Mukamel *University of California, San Diego*
Mónica Nadal-Quirós *Inter American University of Puerto Rico*
Idoia Ochoa *University of Illinois at Urbana-Champaign*
Yu Qian *J. Craig Venter Institute*
Tara Sigdel *University of California, San Francisco*
Tamar Sofer *Harvard University*
Daniela Valdez-Jasso *University of California, San Diego*
Mohit Verma *Purdue University*
Roy Wollman *University of California, Los Angeles*
Chongzhi Zang *University of Virginia*

MENTORS

James Eberwine *University of Pennsylvania*
Laura Heiser *Oregon Health and Science University*
John Lowengrub *University of California, Irvine*
Vinay Pai *JanusView*
Nicholas Tatonetti *Columbia University*
Navin Varadarajan *University of Houston*

PROVOCATEURS AND OTHER GUESTS

Graham Atkinson *JKTG Foundation for Health and Policy*
Jennifer Couch *National Institutes of Health*
Markus Covert *Stanford University*
Rick Horwitz *Allen Institute*
Shannon Hughes *National Institutes of Health*
Eric Johnson Chavarria *National Institutes of Health*
Carol Shreffler *National Institutes of Health*

KI FACILITATORS

Tim Dunne
Stavros Michailidis
Donnalyn Roxey

2017 ATTENDEE LIST

PARTICIPANTS

Xi Kathy Zhou *Weill Cornell Medical College*
Alexander Alekseyenko *Medical University of South Carolina*
David Alvarez-Ponce *University of Nevada, Reno*
Matthew Anderson *The Ohio State University*
Sinem Beyhan *J. Craig Venter Institute*
Lisa Bramer *Pacific Northwest National Laboratory*
Clara Cho *Utah State University*
Arlene Chung *University of North Carolina at Chapel Hill*
Matte Convertino *University of Minnesota*
Bhakti Dwivedi *Emory University*
David Gonzalez *University of California, San Diego*
Michele Guindani *University of California, Irvine*
Lisa Karstens *Oregon Health & Science University*
Miranda Kroehl *University of Colorado Anschutz Medical Campus*
Li Ma *Duke University*
Sophie McCoy *Florida State University*
Christine Peterson *The University of Texas*
Jennifer Powers *Duke University*
Gail Rosen *Drexel University*
Ananya Sen Gupta *University of Iowa*
Rohita Sinha *University of Nebraska, Lincoln*
Ekaterina Smirnova *University of Montana*
Anthony Sung *Duke University*
ZhengZheng Tang *Vanderbilt University*
Katherine Thompson *University of Kentucky*
Erik Wright *University of Pittsburgh*
Michael Wu *Fred Hutchinson Cancer Research Center*
Masanao Yajima *Boston University*
Yanyan Zhang *New Mexico State University*
Angela Zivkovic *University of California, Davis*

MENTORS MENTORS

Kjersti Aargaard *Baylor College of Medicine*
Pieter Dorrestein *University of California, San Diego*
Charles Epstein *University of Pennsylvania*
Curtis Huttenhower *Harvard University*
Katherine Kim *University of California, Davis*
Hongzhe Li *University of Pennsylvania*

PROVOCATEURS AND OTHER GUESTS

Jennifer Couch *National Institutes of Health*
Michael Fischbach *Stanford University*
Nandini Kannan *National Science Foundation*
Vinay Pai *National Institutes of Health*
Gabe Rosenfled *National Institutes of Health*
Carol Shreffler *National Institutes of Health*
Larry Smarr *University of California, San Diego*

KI FACILITATORS

Tim Dunne
Julia Figliotti
Stavros Michailidis

2016 ATTENDEE LIST

PARTICIPANTS

Laura Anderson *University at Buffalo*
Tanvi Banerjee *Wright State University*
Sharmodeep Bhattacharyya *Oregon State University*
Vrushali Bokil *Oregon State University*
Robert Canales *University of Arizona*
Arlene Chung *University of North Carolina at Chapel Hill*
Matteo Convertino *University of Minnesota*
Ellen Eischen *University of Oregon*
Nils Gehlenborg *Harvard Medical School*
Vivian Genaro Motti *George Mason University*
Kimberly Glass *Brigham and Women's Hospital*
Erica Graham *Bryn Mawr College*
Subhanshu Gupta *Washington State University*
Andrei Irimia *University of Southern California*
Spyros Kitsiou *University of Illinois at Chicago*
Dongmei Li *University of Rochester*
Ping-I Lin *Cincinnati Children's Hospital Medical Center*
Nanshu Lu *University of Texas at Austin*
Roummel Marcia *University of California, Merced*
Lucas Mentch *University of Pittsburgh*
Gail Rosen *Drexel University*
Nirmish Shah *Duke University*
Thomas Stewart *Vanderbilt University*
Adam Szpiro *University of Washington*
Yi Wang *Indiana University*
Susan Wei *University of Minnesota, Twin Cities*
Wenyao Xu *State University of New York at Buffalo*
Wenlong Zhang *Arizona State University*
Yuping Zhang *University of Connecticut*
Vadim Zipunnikov *Johns Hopkins University*

MENTORS

Todd Coleman *University of California, San Diego*
John Harer *Duke University*
Katherine Kim *University of California, Davis*
Xihong Lin *Harvard University*
Yazhen Wang *University of Wisconsin, Madison*

PROVOCATEURS AND OTHER GUESTS

Jennifer Couch *National Institutes of Health*
Scott Delp *Stanford University*
Michelle Dunn *National Institutes of Health*
Michael Gernhardt *NASA*
Jody Holtzman *AARP*
Nandini Kannan *National Science Foundation*
Joy Ku *Stanford University*
Vinay Pai *National Institutes of Health*
Andrew Stern *Smartfin*

KI FACILITATORS

Tim Dunne
Julia Figliotti
Stavros Michailidis

Virtual Seminar Series

The Foundations of Biomedical Data Science virtual seminar series consists of regularly scheduled webinar presentations covering the basics of data management, representation, computation, statistical inference, data modeling, & other topics relevant to "big data" bio-medicine. The seminar series provide essential topic introductions suitable for individuals at all levels of the biomedical and computational sciences community. All video presentations are streamed for live viewing, recorded, & posted online for future viewing & reference.

www.youtube.com/@biodatascilab

Seminar Topic	Speaker	Speaker Institution
Introduction to Big Data and the Data Lifecycle	Dr. Mark Musen	Stanford University
Data Indexing and Retrieval	Dr. William Hersh	Oregon Health & Science University
Finding & Accessing Datasets, Indexing & Identifiers	Dr. Lucila Ohno-Machado	University of California San Diego
Data Curation and Version Control	Dr. Pascale Gaudet	Swiss Institute of Bioinformatics
Ontologies	Dr. Michel Dumontier	Stanford University
Provenance	Dr. Zachary Ives	University of Pennsylvania
Metadata Standards	Dr. Susanna-Assunta Sansone	University of Oxford
Data Representation Overview	Dr. Anita Bandrowski	University of California San Diego
Databases & Data Warehouses, Data: Structures, Types, Integrations	Dr. Chaitan Baru and Dr. Elena Zheleva	National Science Foundation
Data Wrangling, Normalization & Preprocessing: Part I Signals	Dr. Joseph Picone	Temple University
Data Wrangling Normalization & Preprocessing: Part II Text	Dr. Sanda Harabagiu	University of Texas at Dallas
Exploratory Data Analysis	Dr. Brian Caffo	Johns Hopkins
NLP: Natural Language Processing	Dr. Noémie Elhadad	Columbia University
Computing Overview	Dr. Patricia Kovatch	Icahn School of Medicine at Mount Sinai
Data Workflows & Pipelines	Dr. Rommie Amaro	University of California San Diego
Running a Data Science Laboratory: Adventures of a Network Biologist	Dr. Trey Ideker	University of California San Diego
Modern Computing: Cloud, Distributed, & High Performance	Dr. Umit Catalyurek	Georgia Institute of Technology
Commons: Lessons Learned, Current State	Dr. Vivien Bonazzi	National Institutes of Health (NIH)
Data Modeling and Inference Overview	Dr. Rafael Irizarry	Harvard University
Supervised Machine Learning	Dr. Daniela Witten	University of Washington
Unsupervised Machine Learning	Dr. Ali Shojaie	University of Washington
Algorithms & Optimization	Dr. Pavel Pevzner	University of California San Diego
Bayesian Inference	Dr. Michael Newton	University of Wisconsin, Madison
Data Issues: Multiple Testing, Bias, Confounding, Missing...	Dr. Lance Waller	Emory University
Causal Inference	Dr. Joseph Hogan	Brown University
Data Visualization Tools & Communication	Dr. Nils Gehlenborg	Harvard University
Open Science	Dr. Brian Nosek	Center for Open Science (COS)
Why Data Sharing & Reuse Are Hard To Do?	Dr. Christine Borgman	University of California Los Angeles
Ethical Issues in Data Science	Dr. Bartha Maria Knoppers	McGill University
Reproducibility	Dr. John Ioannidis	Stanford University
Considerations & Limitations for Clinical Data	Dr. Isaac Kohane	Harvard University
DataScience@NIH: Current State, Future Directions	Dr. Patricia Flatley Brennan	National Institutes of Health (NIH)
Collaborative & Scalable Open Source Data Science: Deep Learning, Optimization & Education	Jonathan Entwistle	IBM
Avoiding the Tower of Babel: the Role of Data Description Standards in Biomedical Imaging	Dr. Chris Gorgolewski	Stanford Center for Reproducible Neuroscience

Big Data Technologies for Biomedical Knowledge Discovery	Dr. Ravi Madduri	University of Chicago
Principles of Scientific Knowledge Engineering	Dr. Gully Burns	USC Information Sciences Institute
Leveling the Playing Field Applying FAIR Principles to Your Daily Data Tasks	Dr. Carl Kesselman	University of Southern California
Big Brain Data Science & Predictive Health Analytics	Dr. Ivo D. Dinov	University of Michigan
"Why the Cloud Matters for Data Science"	Dr. Ian Foster	University of Chicago
Molecular Data & the Microbiome	Dr. Curtis Huttenhower	Harvard University
Theoretical Foundations & Software Infrastructure for Biological Network Databases	Dr. Mehmet Koyuturk	Case Western Reserve University
Imaging Informatics	Dr. Daniel Marcus	University of Washington
Genomic Fingerprinting & Representation	Dr. Gustavo Glusman	Institute for Systems Biology
Computationally & Statistically Efficient Distributed Inference with Theoretical Guarantees	Dr. Xiaoming Huo	Georgia Tech
Biomedicine & the Foundations of Data?	Dr. Michael Mahoney	University of California, Berkeley
How to do Data Science	Dr. Don Brown	University of Virginia
Neuroinformatics and Cognitive Ontologies	Dr. Lucina Uddin	University of Miami
Mapping Structure, Function and Connectivity in Humans and NonHuman Primate Cerebral Cortex	Dr. David Van Essen	Washington University
Assembling Global Neural Networks of the Mouse Brain	Dr. Hong-Wei Dong	University of California Los Angeles
Toward a Culture of Computational Reproducibility	Dr. Russell A. Poldrack	Stanford University
Data Structure, Disease Risk, GXE, and Causal Modeling	Dr. Robert W. Williams	University of Tennessee Health Science Center
ENIGMA and COINSTAC: Turning small datasets into big ones	Dr. Jessica Turner	Georgia State University

Neuroimaging of Pain and Affect: Generalizable models for prediction and explanation	Dr. Tor Wager	Dartmouth College
Brain Aging, Inflammation and Cardiovascular Health: Insights from a society in South America	Dr. Andrei Irimia	University of Southern California
Building Precision Medicine for Autism(s)	Dr. Kevin Pelphrey	Harrison-Wood Jefferson Scholars Foundation
Clinical Trial Readiness for Neurodevelopmental Disorders: On the Road to Precision Health	Dr. Shafali Jeste	University of California Los Angeles
The Use of Ontologies for FAIR Neuroscience	Dr. Maryann Martone	University of California, San Diego
Investigating Mechanisms of Exercise-Induced Neuroplasticity in Cog-Motor Systems in Parkinson's	Dr. Giselle Petzinger	University of Southern California
ReproNim: Towards Reproducible Neuroimaging Neuroscience	Dr. David Kennedy	University of Massachusetts
Functional Imaging - Individual Differences Social Cognition	Dr. James Morris	University of Virginia
Big Neuroscience, Data Sharing & Predictive Health Analytics	Dr. Ivo Dinov	University of Michigan
The Shifting Dunes of Data and Computation	Dr. Satrajit Ghosh	Massachusetts Institute of Technology
Brain-Based Biomarkers - A Focus on the Heterogeneous	Dr. Vince Calhoun	Center for Translational Research in Neuroimaging and Data Science
Integrative Data Science Approaches for Studying Transcriptional Regulation in the Genome	Dr. Chongzhi Zang	University of Virginia
Progress Towards Large-Scale, Consensus-Based Models of Human Brain Function: Contributions from fMR	Dr. Angela Laird	Florida International University
Advancing Neuroimaging Genomics with Continuous Team Science	Dr. Neda Jahanshad	University of Southern California
Meta-Learning for Biological Knowledge Transfer	Dr. Aidong Zhang	University of Virginia
Neuron Morphology, Connectivity, and Classification	Dr. Giorgio Ascoli	George Mason University

Data Activism and the Imagination of Biomedical Data Science	Renée Cummings	University of Virginia
Predictive Analytics Monitoring at the Bedside	Dr. Randall Moorman	University of Virginia
Introduction to use of Structured Medical Record Data	Johanna Loomba	University of Virginia
Ethics, Equity and Transparency in Biomedical Research and the Role of the Good Pharma Scorecard	Dr. Jennifer Miller	Yale University
Practical Tools for Developing Ethical AI (and why principles are not enough)	Dr. David Danks	University of California, San Diego
On the Necessity of Relational Ethics and Empathic Attunement for Data-Centric Technologies	Dr. Jarrett Zigon	University of Virginia
Utilizing Deep Learning to Create Tailored Summary Outcome Metrics for Clinical Populations	Dr. Mark Albert	University of North Texas
The Secret Lives of Predictive Models	Dr. Claudia Perlich	New York University
Current Practice, Challenges and Perspectives on Single Cell Data Science	Dr. Lana Garmire	University of Michigan
Ethical AI Requires Ethical Collaboration	Dr. Caitlin D. Wylie	University of Virginia
Promise or Peril?: Algorithmic Intelligence and Racial Discrimination in Health Care	Dr. LaTonya Trotter	University of Washington
Addressing Key Challenges in Genomic and Health Data Sharing	Dr. Yann Joly	McGill University
On Collective Wisdom When AI is Involved	Dr. Colin Allen	University of Pittsburgh
A Beginner's Mind	Dr. Donna Chen	University of Virginia
Team Science in Data Science: Bridging the Divide	Dr. Maritza Salazar Campo	University of California, Irvine
Challenges for Research Ethics Governance in the Era of AI and Data Science Research	Dr. Edward Dove	University of Edinburgh
Data Acumen in Action	Dr. Sallie Keller Dr. Stephanie Shipp	University of Virginia

Ethical Interactions in Scientific Teams as a Foundation for Ethical AI	Dr. Alison Antes	Washington University
Interdisciplinarity in the Health Sciences and the Pursuit of Health Equity	Dr. Sean Valles	Michigan State University
Radical Approaches to COVID-19 Testing and Predicting Susceptibility to Developing Long Covid/PASC	Dr. Orlando Lopez	National Institute of Dental and Craniofacial Research
Challenges in Estimating Vaccine Effectiveness Using Big Data	Dr. Maria Sundaram	Marshfield Clinic Research Institute
NHLBI Data Science Panel	Dr. Olga Brazhnik Dr. Asif Rizwan	National Heart, Lung, and Blood Institute
The Costs and Benefits of Different Covid-19 Interventions	Dr. Adam Atherly	Virginia Commonwealth University
Keeping Campus Open: Designing Mitigation Strategies with Biased Data	Dr. Shane Henderson	Cornell University
Data Enhancement and Portability for Healthcare	Dr. Melody Greer	University of Arkansas for Medical Sciences
Neurocognitive Decline in Young Adults with Mild COVID-19: Grounding Assessment in Pre-pandemic Data	Dr. Michael Lipton	Albert Einstein College of Medicine
Publishing Reproducible and Credible Models	Dr. Herbert Sauro	University of Washington
Enhancing Long-Term Forecasting: Learning from COVID-19 Models	Dr. Navid Ghaffarzadegan	Virginia Tech
Leveraging Big Health Data to Answer Questions about Program and Policy Effects	Dr. Jennifer Ahern	University of California, Berkeley
So You Want to Market Your Regulated Digital Health Device in the US?	Dr. Vinay Pai	Food and Drug Administration

After the Research: Using Policy, Advocacy, + Philanthropy to Address the Consequences of COVID-19	Julia Roper Autumn Lerner Darcy Sawatzki Jennifer Cho	Panorama Group
What can Data Science do in the Fight Against Infectious and Immune-mediated Diseases?	Dr. Wilbert van Panhuis	National Institutes of Health
Early-Pandemic Telehealth Implementation and Mid-Pandemic Post-COVID Neurobehavioral Symptoms	Dr. Paul Perrin	University of Virginia
Digital Health Equity	Dr. Katherine Kim	MITRE Corporation
Advancing the Practice of Transdisciplinary Team Science Research: Responding to COVID-19	Dr. Madhav Marathe	University of Virginia
Understanding the Impact of COVID-19 Pandemic on Mental Health Using Network Analysis	Dr. Folashade Agusto	University of Kansas
Choices Give Meaning to Uncertainty: Pandemic Emergence, Policy Advising and Personalized Medicine	Dr. Mike Famulare	Bill & Melinda Gates Foundation
Infectious Disease Surveillance and Modeling through Spatial Big Data	Dr. Shweta Bansal	Georgetown University
How Well Did the World do Modeling Covid?	Dr. Christopher Murray	University of Washington
Likelihood-Based Methods for Fitting Stochastic Epidemic Models to Noisy Data	Dr. Jason Xu	Duke University
Impact of COVID-19 Pandemic on Influenza	Dr. Sen Pei	Columbia University
Toward a Data Commons for Biomedicine	Dr. Vivien Bonazzi	Deloitte
Generative AI in Context	Dr. Brian Wright	University of Virginia
Can AI Uncover the Laws of Physics by Observing Apples Falling from Trees?	Dr. Stephen Baek	University of Virginia

Personalized Machine Learning for Clinical Natural Language Processing	Julian McAuley	University of California San Diego
Identifying Genetics Signals Adaptively and Reproducibly	Dr. Chiara Sabatti	Stanford University
Innovations and Challenges in AI and its Application to Health	Dr. Susan Gregurick	National Institutes of Health
Label-Efficient Learning for Biomedical and Educational Applications	Dr. Donald Brown	University of Virginia
Language, Knowledge, and LLMs in the Context of Biomedical Data Science Education	Dr. Rafael Alvarado	University of Virginia
Large Language Models for Healthcare: Where Are We Now, and Where Are We Headed?	Dr. Anna Baglione	University of Virginia
Acceding to the Use of AI Tools in Education	Dr. Prince Afriyie	University of Virginia
Biomedicine in the Age of Generative AI	Dr. James Zou	Stanford University
Clearing the Air for a Breathable Future: Unraveling the Biomedical Implications of Data Pollution	Dr. Amy Yeboah Quarkume	Howard University
Towards Sustainable Research Software	Dr. Daniel Katz	University of Illinois Urbana-Champaign
Psychometrics and Large Language Models: What Can Be Learned and What Can Go Wrong?	Dr. Steven Boker	University of Virginia
Interoperability at All Scales: from Data Bits and Bytes to National Public Health Surveillance	Dr. Melissa Haende	University of Colorado Anschutz Medical Campus

THE BIOMEDICAL DATA SCIENCE INNOVATION LAB

BDSIL

STORY PHOTOBOOK

Department of Psychology
College and Graduate School of Arts & Sciences
University of Virginia

NIH Grant
#R25GM139080

https://www.innovation.lab.virginia.edu/