National Big Data Health Science Conference 2022

February 11-12 Virtual

PROGRAM







National Big Data Health Science Conference 2022

The National Big Data Health Science Conference is a signature annual event of the Big Data Health Science Center (BDHSC). This 3rd annual conference will include innovative plenary sessions, panels, and workshops that emphasize the role of interdisciplinary collaboration in Big Data applications and advancements in the health sciences.

The UofSC BDHSC was funded by the UofSC Excellence Initiative in 2019. Its goals are to leverage the existing expertise and resources in Big Data science and healthcare research at UofSC, promote the utilization of Big Data analytics in healthcare research, academic training, and services improvement, and accelerate the transformation of UofSC into a global leader in Big Data health science. The BDHSC has five strategic objectives: Infrastructural and capacity development; Professional development; Community/industry engagement; Academic training; and Methodological advancement.

BDHSC consists of 5 content cores (Electronic Health Records, Genomics, Artificial Intelligence for Sensing and Diagnosis, Geospatial, and Social Media) and 2 functional hubs (Business/Entrepreneurship and Technology). It has assembled a multi-college, multi-disciplinary group of faculty that conduct cutting cutting-edge research and discovery, offer professional development and academic training, and provide service to the community and industry.

We are pleased to announce the theme of this year's conference is "Unlocking the Power of Big Data in Health: Developing an Interdisciplinary Response for Health Equity"

Access our conference program online at:



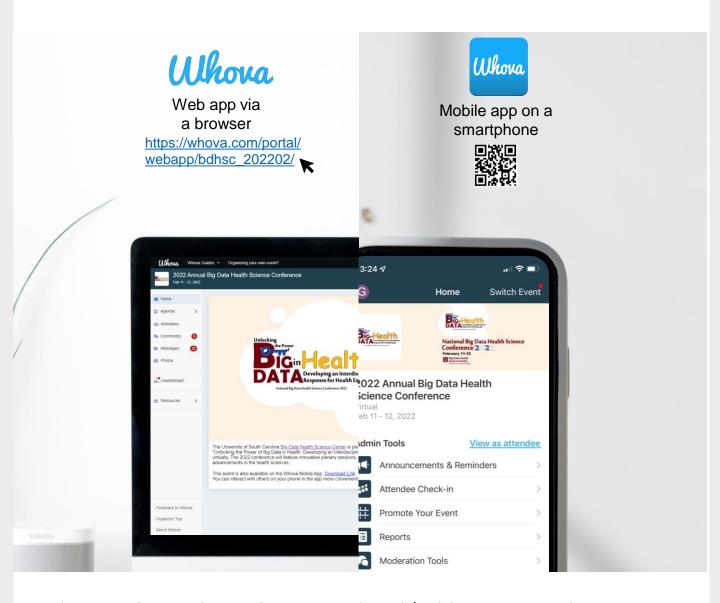
2022 Big Data Health Science Conference Sponsor:



We are using an online platform called

Whova

Attendees have two ways to access the conference:



Please note that attendees need to sign up on the Web/Mobile app to create Whova accounts prior to the conference. Attendees should use the same emails that were used to register for the conference.

Please visit our conference website for instructions:



https://www.sc-bdhs-conference.org/2826-2/

Not Registered Yet?

CLICK HERE TO REGISTER



Or Visit the Conference Website:

https://www.sc-bdhs-conference.org/



Friday, February 11, 2022 Plenary Presentations | 8:45am — 12:00pm (EST)

8:45am - 9:00am

Welcome and Opening Remarks



Stephen Cutler, Interim Executive Vice President for Academic Affairs and Provost, UofSC

9:00am - 9:10am

2022 Student Case Competition Announcement

9:20am - 10:05am

Fusing Psychological Biomarkers to Produce a Hemorrhagic Trauma Severity Score



Anthony Guiseppi-Elie (Anderson University)

Summary: The Hemorrhage Intensive Severity and Survivability (HISS) score, based on the fusion of multi-biomarker data; glucose, lactate, pH, potassium, and oxygen tension, is intended to serve as a patient-specific attribute in hemorrhagic trauma. One hundred instances of Sensible Fictitious Rationalized Patient (SFRP) data were synthetically generated, and the HISS score assigned by five clinically active physician experts (100[5]). The HISS score stratifies the criticality of the trauma patient as; low(0), guarded(1), elevated(2), high(3) and severe(4). Standard classifier algorithms: linear support vector machine (SVM-L), multi-class ensemble bagged decision tree (EBDT), artificial neural network with Bayesian regularization (ANN:BR) and possibility rule-based using function approximation (PRBF) were evaluated for their potential to similarly classify and predict a HISS score. SVM-L, EBDT, ANN:BR and PRBF generated score predictions with testing accuracies (majority vote) corresponding to 0.91 ± 0.06 , 0.93 ± 0.04 , 0.92 ± 0.07 , and 0.92 ± 0.03 , respectively, with no statistically significant difference (p>0.05). Targeted accuracies of 0.99 and 0.999 could be achieved with SFRP data size and clinical expert scores of 147[7](0.99) and 154[9](0.999), respectively. The predictions of the data-driven model in conjunction with an adjunct multi-analyte biosensor intended for point-of-care continual monitoring of trauma patients, can aid in patient stratification and triage decision-making.

10:05am - 10:15am

Coffee Break

10:15am - 11:00am

How to Use Findings from Big Data to Effectively Influence Public Health Policy



Linda Bell (SC DHEC Bureau of Disease Control)

Summary: We are realizing many new opportunities to link data sets and use new ways to analyze, visualize and disseminate information. With the explosion in what we are learning about populations at risk, and the relationships of adverse health outcomes to everything from our genetic make up, to our socioeconomic status and the environments we live in, consideration should be given to how we will apply what we learn to translate into the development of public policies that will actually change adverse health outcomes.

11:15am - 12:00pm

Leveraging Partnerships and Collaboration



Windsor Westbrook Sherrill (Clemson University, Prisma Health-Upstate)

Summary: It is imperative that the primary response to opportunities represented by big data be scholarship and collaboration. Now more than ever, we need to work together to leverage big data to improve healthcare practices locally, nationally, and even globally. Although combating barriers to big data is multifaceted, team-based approaches representing academics and clinical delivery will actualize results to improve population health. With a strong and vigorous integration of biomedical and healthcare data, healthcare organizations can achieve financial and clinical advantages, and researchers can provide key inputs. Not only can big data be discussed in the context of improving medical care through proper management, analysis, and interpretation, but its applications can catalyze the scholarship of health innovation including better health outcomes, facilitating the study of disease-related risk factors, equipping health systems with vital diagnostic tools, and more. The implications of leveraging all modalities of health data with an interdisciplinary approach can be revolutionary. National exemplars of innovation with big data will be explored. South Carolina universities and health care systems have an opportunity to be at the forefront of leveraging data to improve health.

Speed Networking | 12:15pm — 12:45pm (EST)

Participants will be randomly assigned to virtual speed networking rooms where they can introduce themselves and network with fellow attendees. After 10 minutes they will be automatically reassigned to new rooms with a new group of attendees to meet!

Friday, February 11, 2022 Breakout Sessions (Set 1) | 1:15pm — 3:15pm (EST)

Electronic Health Records Breakout Session I

Application of Novel Methodology in Big Data Public Health Research

Session Summary: With the increasing use of Big Data in health care comes growing challenges in Big Data analysis. In this session, a diverse panel of investigators will present on methods of machine learning, causal inference and network analysis for use in Big Data public health research. Each presenter will speak for 25 minutes, followed by a group Q&A.

Note: To participant in this interactive Breakout session, you will need access to Zoom.



Presentation 1 The ABCs of Predictive Model Governance

Eric Poon (Duke Health)

Presentation 2 Using AI and EHR for Disease Risk Prediction

Cui Tao (University of Texas Health Science Center at Houston)

Presentation 3 Reliable and Trustworthy Clinical Diagnosis and Decision

Making with Causal AI

Pooyan Jamshidi (UofSC)

Presentation 4 Statistical Modeling of Brain Disorders with Topological Data Analysis



Friday, February 11, 2022

Breakout Sessions (Set 1) | 1:15pm — 3:15pm (EST)

Genomic Breakout Session

Application of Advanced Data Analytics in Biological and Health-Related Data

Session Summary: Advanced data analytics, including AI and ML techniques, have become the cornerstone of a broad spectrum of investigations. In this session, a number of investigations are presented to highlight the application of data analytics to health-related data. The analysis span the gamut of investigating molecular biomarker data to the application of ML techniques.

Note: To participant in this interactive Breakout session, you will need access to Zoom.



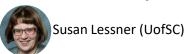
Presentation 1 Explainable Artificial Intelligence and the Rubik's Cube



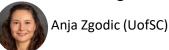
Presentation 2 Identification of Differentially Expressed Genes Involved in Thoracic Aortopathy



Presentation 3 Using Big Data Multiple Scales to Improve Diagnosis and Treatment of Peripheral Arterial Disease in South Carolina



Presentation 4 A Novel Approach to High-Dimensional Linear Regression in a Pharmacogenomic Setting



Presentation 5 Protein Data Mining Optimizations in PDBMine



Friday, February 11, 2022 Breakout Sessions (Set 1) | 1:15pm — 3:15pm (EST) Geospatial Breakout Session

Application of Geospatial Methods in a Variety of Scientific Domains and Techniques

Session Summary: This section will feature mini-talks from seven faculty members at UofSC. The goal is to provide audience members with ideas of how they might utilize distinct GIS methods to enhance their own research.

Note: To participant in this interactive Breakout session, you will need access to Zoom.



Presentation 1 Remote Sensing Application to Detect and Prioritize Areas for

Mosquito Control

Melissa Nolan (UofSC)

Presentation 2 Leveraging Social Media Data to Model Risk Perceptions and

the Politics of Mask Usage in the US

Diego Leal (UofSC)

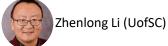
Presentation 3 Micro-Places, Neighborhoods, and Crime in Urban Areas

Cory Schnell (UofSC)

Presentation 4 Applying GIS Methods to Medication Use Research

Bryan Love (UofSC)

Presentation 5 Measuring Human Mobility Dynamics and Place Connectivity
Using Big Social Media Data



Presentation 6 Applications of GIS to Address Racial and Ethnic Equity in Health Care

Nathaniel Bell (UofSC)

Presentation 7 Deriving Fine-Scale Regional Product of Population Grid from Census Records and Open-Source Satelite



Friday, February 11, 2022

Breakout Sessions (Set 2) | 3:30pm - 5:30pm (EST)

Electronic Health Records Breakout Session II

Leveraging EHR Data to Understand Healthcare Utilization and Health Outcomes

Session Summary: Electronic health record (EHR) data offer a promising opportunity to understand resource allocation, assess patients' health outcomes, and improve health equity. Panels in this session will discuss various approaches and issues tackle with EHR data – including the Corona-Infectious-Virus Epidemiology Team (CIVETs), National COVID Cohort Collaborative (N3C), health system-wide EHR data, statewide Medicaid claims data linking to birth certificates, and national Medicare admission data – in the efforts to elucidate the variations in healthcare utilization and patient outcomes in a large health system, in a state, and nationally.

Note: To participant in this interactive Breakout session, you will need access to Zoom.

Moderator



Peiyin Hung (UofSC)

Presentation 1

Impacts of COVID-19 Pandemic on Racial/Ethnic Disparities in Severe Maternal Morbidity: Data from the South Carolina COVID Cohort



Jihong Liu (UofSC)

Presentation 2

Temporal Trends of Telehealth Use for Prenatal Care among Medicaid Beneficiaries and during COVID by Maternal Race/Ethnicity and Urban/Rural Residence



Chelsea Norregaard (UofSC)

Presentation 3

On the Use of Importance Sampling to Improve Generalization Under Covariate Shift for Predictive Models Built on Electronic Health Records



Richard Jordan (Health Data Analytics Institute)

Presentation 4

Impact of Discontinuity in EHR Data Pipelines on Inferences that are Contingent upon SARS-CoV-2 Testing or Vaccination Data: Lessons Learned from the CIVETs Collaboration of 5 Healthcare Systems



Keri Althoff (Johns Hopkins)

Presentation 5

Implementing a Social Determinants of Health Screening and Electronic Health Records-Based Platform to Assess Healthcare Utilization and Community Health Resource Use: an Observational Study Based in a Large Health System in South Carolina



Friday, February 11, 2022 Breakout Sessions (Set 2) | 3:30pm – 5:30pm (EST)

Al for Sensing and Diagnosis Breakout Session

Opportunities for AI in Biomedicine

Session Summary: We will highlight research efforts to build statistical models of health-related datasets, develop open-source algorithms, utilize publicly available datasets on platforms such as Kaggle, and transparency in the model predictions when utilizing data for clinical and health-related decisions. After hearing several presentations from active researchers on these topics, we will then have a general discussion on SC state-wide infrastructure needs and opportunities for AI in biomedicine.

Note: To participant in this interactive Breakout session, you will need access to Zoom.





Christopher Sutton (UofSC)



Qian Wang (UofSC)

Presentation 1

Explainable Artificial Intelligence for Decoding and Modulating Neural Circuit Activity Linked to Behavior



Nina Hubig (Clemson University)

Presentation 2

What are you Doing? An Open Source Machine Learning Accelerometry and Gyroscopy Analysis Toolbox



Fluri Wieland (University of Bern)

Presentation 3

Don't let the Foxes Guard the Henhouse: "Explainable" Al as a Mask for Cosmetic Compliance



Marilyn Gartley (UofSC)

Presentation 4

Medsensor: Medication Adherence Monitoring Using Neural Networks On Smartwatch Accelerometer Sensor Data



Chrisogonas Odhiambo (UofSC)

Presentation 5

False Discovery Rate Control for Lesion Symptom Mapping with Heterogeneous data via Weighted P-values



Siyu Zheng (UofSC)

Friday, February 11, 2022 Breakout Sessions (Set 2) | 3:30pm – 5:30pm (EST)

Social Media Breakout Session

Applications of Social Media Analysis for COVID-19 Pandemic

Session Summary: During the COVID-19 pandemic, people and government officials have utilized social media to regularly share policies and news related to COVID-19. For example, millions of tweets were posted on Twitter during the pandemic. A growing research trend argues that social media platforms play an important role in public health. Social media can help to provide communication in real-time and at a relatively low cost and monitor public response. In this session, a diverse panel of investigators will discuss social media studies focused on mental health, vaccination, substance use disorder, and discriminatory speech during the COVID-19 pandemic.

Note: To participant in this interactive Breakout session, you will need access to Zoom.



Presentation 1

Global Comparison of 'COVID-19 Vaccination' Sentiments and Emotions on Twitter: Findings from 192 Countries



Jungmi Jun (UofSC)

Presentation 2

Assessment of Mental Health using Twitter data for pre-, during-, receding-COVID



Yesoda Bhargava (Indian Institute of Information Technology)

Presentation 3

The Platformization of Organizational (Counter-) Discriminatory Speech on Facebook during the Pandemic



Yingying Chen (UofSC)

Presentation 4

Exploring Sentiment and Communication Exchange Patterns of Substance Use Disorder (SUD) on Twitter Associated with Pregnant Women During COVID-19 Pandemic



Saturday, February 12, 2022

Poster Presentations | 8:30am — 9:30am (EST)

Please visit the Poster Hall to talk with poster presenters and learn about their research.

Presenters



Andrew Smith (UofSC)

Application of Machine Learning to Sleep Stage Classification



Yang Ran (UofSC)

Using D-Dimer as biomarker to

Predict Disease Severity and Clinical

Outcomes with a Large COVID-19

Inpatient Cohort: A Machine Learning

Approach



Yuqi Wu (UofSC)

Using Machine Learning Methods to Identify Risk Factors of Low-Weight-Birth Outcomes in a US Hospital System



Julia Certa (UnitedHealth Group)

Differences in COVID-19 Testing,
Positivity, and Outcomes by Race, Sex,
and Health System Setting in a Large
Diverse US Cohort



Christopher Lee (UofSC)

Analysis of Protein Sequence
Similarities within Non-Structural
Protein 1 (NSP-1) of SARS-CoV-2



Tianchu Lyu (UofSC)

A Rule-based Algorithm to Infer

Gestational Age Using Electronic

Health Records: A Pilot Study



Xiuming Feng (Guangxi Medical University)

Association between Healthy

Lifestyles, Genetic Modifiers and

Thyroid Cancer Risk: a Prospective

Cohort Study in the UK Biobank



Ruilie Cai (UofSC)

Using Twitter Data to Estimate

Prevalence of Mental Disorders in the
United States during COVID-19

Pandemic: a Comparison with
National Prevalence Data



Jacqueline Purtell (UofSC)

Analysis of Vaccine Misinformation on
Social Media



Serena Harn (UofSC)

Exploring Patterns of Cancer

Misinformation Diffusion on Social

Media



Ran Xu (University of Connecticut)

Deriving Nutrition Information of
Restaurant Food from Crowdsourced
Food Images: Case of Hartford



Huan Ning (UofSC)

Analysis of COVID-19 Spreading on Public Places in South Carolina Metropolitan Areas



Yi-Wen Shih (UofSC)

Variations in County-Level Teen Birth
Rates and Low Birth Weight by

Multidimensional Racial Residential
Segregation



Casey Cole (UofSC)

Calculating Gene Rhythmicity using

Sparse Time Points over Biological

Replicates



Zichun Meng (UofSC)

Using Machine Learning to Capture

Complexities of Dietary Patterns for

Prediction of Health Status



Olivia Moran (UofSC)

Differences in Gene Expression,
Rhythmicity and Pathways Due to
Prenatal KYNA Elevation in Wistar
Rats

Saturday, February 12, 2022 Plenary Presentations | 9:35am — 12:50pm (EST)

9:35am - 10:20am

Understanding Ticks and Tick-borne Diseases through Surveillance and Modeling



Holly Gaff (Old Dominion University)

Summary: Ticks and tick-borne diseases are on the move worldwide. Tracking these movements is critical to understanding the risks of the diseases. Mathematical models can be used to explore tick-borne pathogen dynamics, quantify risk of tick-borne disease, and identify optimal strategies to reduce that risk. Long-term surveillance data are needed to understand the range expansion of a population and to build ecologically valid models. In order to build such a dataset, ticks have been collected since 2009 using a variety of techniques from a set of locations in Virginia, USA. This surveillance project has collected, identified, and cataloged more than 200,000 ticks. These data have been used to parameterize and validate models for range expansion and control of Lyme disease, and while the models are helpful they also highlight additional data needs.

10:30am - 11:15am

Using Artificial Intelligence and Knowledge Graphs for Precision Medicine at Scale



Sergio Baranzini (University of California at San Francisco)

Summary: Spoke is a massive knowledge network containing more than 40 specialized databases and spanning multiple disciplines within biomedicine. This talk will focus on the creation of this resource and some of its cutting-edge applications, that include predicting outcomes in multiple sclerosis and Parkinson's disease, unraveling the pathogenesis of COVID-19 and efforts in drug development and repurposing.

11:15am - 12:05pm

Lunch Break

12:05pm - 12:50pm

Flexible and Interpretable Learning for High-Dimensional Health Data



Hao Zhang (University of Arizona)

Summary: Modern health data are often massive, high-dimensional, and intrinsically complex. When learning hidden patterns from such data, it is desired to build a decision rule which is both flexible and interpretable, in the sense that the final model should be able to capture nonlinear patterns hidden in data and is also easy to explain and understand. Due to the curse of dimensionality, traditional statistical tools often experience theoretical and computational difficulties in achieving these goals for high-dimensional data. In this talk, I will present modern nonparametric and sparse learning methods which can extract meaning low dimensions and build nonlinear models for high-dimensional data simultaneously. As a showcase, a class of regularization frameworks will be discussed in the context of high-dimensional nonparametric regression and classification, including their theoretical guarantee, scalable computation, and applications to learning various types of data in health sciences.

Saturday, February 12, 2022 Featured Workshops | 1:15pm — 3:15pm (EST) NIH NIAID R25 Big Data Health Science Fellow Program

Workshop Summary: This workshop features 2021-2022 Big Data Health Science Fellows and their mentors. Supported by NIAID (R25Al164581-01), the UofSC Big Data Health Science Center implemented its Big Data Fellow program in 2021. The program aims to provide fellows with skills development in Big Data Science research, hands-on research experience and proposal development, and rich mentoring experiences in Big Data Science research and professional development. The 2021-2022 cohort includes six fellows representing a diverse range of UofSC departments including biostatistics, nursing, medicine, sociology, education, and health promotion, education, and behavior. Each fellow will deliver a short presentation about their experience with the Big Data Fellow Program and their plans moving forward. Special guest Dr. Susan Gregurick, Director of NIH Office of Data Science Strategy, as well as other invited NIH Program Officers will be present for an open floor discussion with fellows and attendees. Prospective applicants to the 2022-2023 Big Data Fellow cohort are strongly encouraged to attend. Applications for the next cohort of trainees are due April 15, 2022.

Moderators



Xiaoming Li (UofSC)



Jiajia Zhang (UofSC)

Special Guest



Susan Gregurick (NIH)

Panelists



Christopher Goodman (UofSC)



Sara Donevant (UofSC)



Diego Leal (UofSC)



Gregory Trevors (UofSC)



Xueying Yang (UofSC)



Stella Self (UofSC)

Computational Modeling of Protein Structure and Dynamics Workshop

Workshop Summary: This workshop will demonstrate use of Visual Molecular Dynamics (VMD) and Nanoscale Molecular Dynamics (NAMD) applications. VMD is a molecular visualization program for displaying, animating, and analyzing biomolecular systems using 3D graphics and scripting. NAMD is a parallel molecular dynamics program designed for high-performance simulation of large biomolecular systems in conjunction with VMD. Both applications are freeware for non-commercial use by individuals, academic faculty and students. Workshop participants will first be introduced to basic functions of VMD including loading protein structures obtained from Protein Data Bank (PDB), protein structure manipulation, and navigating through visualization tools. Next, we will work with NAMD to perform basic molecular dynamic simulations and analysis. We will go over how to generate a protein structure file (PSF), solvate the protein structure, perform energy minimization, and finally run the protein structure through molecular dynamic simulations. Upon completion of the simulations, we will analyze the results by looking at structural similarities and the energies of the structures. Finally, we will explore the application of the presented topics in molecular interaction such as the spike protein of SARS-CoV-2 and the human ACE protein. More specifically, we will examine the molecular binding properties between the spike and ACE protein from the South African and UK variants of the SARS-CoV-2.

Workshop Presenter



Saturday, February 12, 2022 Featured Workshops | 1:15pm — 4:15pm (EST)

End-to-End Modeling and Machine Learning with SAS Viya for Learners

Workshop Summary: This workshop will demonstrate use of SAS Visual Analytics and SAS Model Studio within the SAS Viya for Learners platform. SAS Viya for Learners is a suite of free, cloud-based software for teaching and learning data science skills available to academic faculty and students. Participants will be introduced to several methods to train supervised machine learning models to make better decisions on big data. Topics that will be demonstrated include predictive modeling techniques such as linear and logistic regression, decision tree and ensemble of trees such as forest and gradient boosting, neural networks, and support vector machines. A case study example will be used to guide participants through steps of an analysis including data preparation, model building, and model assessment. A general introduction and overview to the SAS Viya interface, including SAS Drive, will also be presented. Familiarity with predictive modeling techniques is helpful, but not necessary.

Note: This workshop requires pre-registration. Attendees are given the option to reserve a space upon registration (on a first come first serve basis) for the conference.

Presenters





Plenary Presentations | 3:30pm — 4:35pm (EST)

3:30pm - 4:15pm

What Can Data Science Do in the Fight Against Infectious and Immune-Mediated Diseases?



Wilbert van Panhuis (NIAID)

Summary: Infectious and immune-mediated diseases continue to pose a major threat to human health, as illustrated by the COVID-19 pandemic as well as the continued burden of long-term diseases such as HIV/AIDS, Tuberculosis, Malaria, and many others. Big data and data science have disrupted many segments of society and hold the promise to transform the biomedical landscape as well, e.g., through precision medicine, robotic surgery, and Al-supported decision making. Data Science has also accelerated research on infectious and immune-mediated diseases, leading to new knowledge, technologies, and tools to reduce the burden of these diseases on the global population. The impact of data science will depend on the data infrastructure, governance, and culture that our community will develop over the next decades. Dr. Van Panhuis will highlight the vision at the US National Institute of Allergy and Infectious Diseases for data science in the global fight against infectious and immune-mediated diseases.

4:20pm – 4:35pm Closing Remarks



Thomas Chandler, Dean of Arnold School of Public Health, UofSC

February 11 12:15pm - 12:45 pm





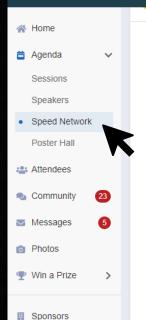
Please join us Speed Network

Whova Guides - Organi

Organizing your own event?



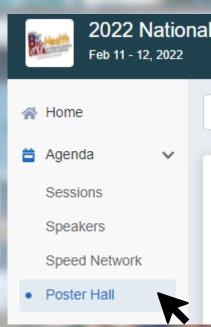
2022 National Big Data Health Science Conference Feb 11 - 12, 2022





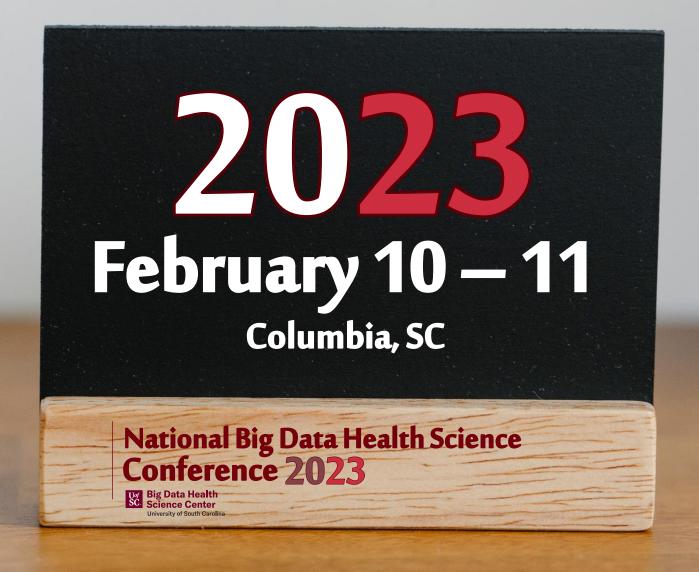






Don't Miss The Poster
Presentation Q&A from
8:30am — 9:30am on
February 12 at the Poster Hall!





Save The Date