

# Memphis DATA – A Data Science Conference

[www.memphis-data.org](http://www.memphis-data.org)



**March 29, 2019**  
**Fedex Institute of Technology**  
**The University of Memphis**  
**Memphis, TN, USA**

*Organized by the Data Science Cluster of the Fedex Institute of Technology*  
*Sponsored by The University of Memphis*

## **Organizing Committee**

### *General Chairs*

David M. Kemme (Economics) & Ebenezer Olusegun George (Mathematical Sciences, Statistics)

### *Program Chairs*

Hongmei Zhang (Public Health) & Vasile Rus (Computer Science)

### *Local Arrangements Chair*

Deepak Venugopal (Computer Science) & Mark Gillenson (BIT)

### *Speakers Chairs*

Dale Bowman (Mathematical Sciences, Statistics) & Bernie J. Daigle, Jr. (Biological Sciences and Computer Science)

### *Tutorials Chairs*

Vasile Rus (Computer Science) & Deepak Venugopal (Computer Science)

## Conference Keynote Speakers

Friday, March 29, 9:00am – 10:00 am

### **Dr. Henry Kautz**

National Science Foundation

#### **Toward a Roadmap for AI Research & Development (Everything Old is New Again)**

**Abstract:** Prominent AI researcher and founder of Coursera, Andrew Ng, has said, “AI is the new electricity”. Indeed, AI is powering profound changes in commerce, healthcare, defense, entertainment, and education. Yet only a few years ago the field of AI was met with derision in many quarters, and there is still much confusion about what the phrase “artificial intelligence” denotes. In this talk, I will provide a brief history of AI and speculate on its trajectory into the next decade - which, I will argue, will revisit unsolved problems from decades past. I will describe current efforts to expand and coordinate support by government and industry in AI R&D. These efforts include the development of a roadmap for AI research being led by the Community Computing Consortium with support from the National Science Foundation, and the development of the US National AI R&D Strategic Plan and Progress Report.

**Biographical Sketch:** Henry Kautz is Director for the Division Information & Intelligent Systems in the Directorate for Computer & Information Science & Engineering at the National Science Foundation. He is professor at the University of Rochester, where he was founding director of the Goergen Institute for Data Science. He is a past president of the Association for the Advancement of Artificial Intelligence and past chair of the Section on Information & Computing at the American Association for the Advancement of Science. His research is in stochastic algorithms for logical inference, social media analytics, and healthcare applications.

## Conference Keynote Speakers

Friday, March 29, 1:15pm – 2:15 pm

### **Dr. Bala Vaidyanathan**

Director, Advanced Analytics & Industry Research at FedEx  
Greater Memphis Area

#### **Bridging the Gap between Theory and Practice**

**Abstract:** Analytics is expected to drive trillions of dollars in value and positively impact lives all around the world. While academia traditionally values novelty of research contributions and complexity, simplicity and business impact are primary measures of success in the industry. Therefore, bridging the gap between academic theory and industry practice is crucial to delivering the promise of analytics. The five pillars that support the bridge between theory and practice are simplify, partner, iterate, integrate, and deploy. This talk will focus on these pillars and provide examples from the business world.

**Biographical Sketch:** My passion is to deliver breakthroughs in business performance through the application of Advanced Analytics, Artificial Intelligence, and Technology. My philosophy to delivering value is to integrate business thinking, data, and the scientific method to discover new insights that drive impact. A native of India, I earned a Bachelor's degree in Engineering from the Indian Institute of Technology and a PhD in Operations Research from the University of Florida. My contributions have been published in premier journals including Transportation Science, Operations Research, Networks, IBM Journal of Research & Development, and Encyclopedia of Optimization. Professional societies such as Transportation Science and Logistics Society, Transportation Research Forum, and Railways Applications Society have also recognized the value of my work through awards. From 2010 to 2017, I served as an Associate Editor of the international journal, Networks, managing the peer review process for submissions in the network optimization area.

## Panel on Opportunities and Challenges for Data Science in Biomedical Research

Friday, March 29, 3:45 pm – 4:45 pm

**Dr. Natasha Sahr** received her BS in Mathematics from Marquette University (Milwaukee, WI) and PhD in Biostatistics from the Medical College of Wisconsin (Milwaukee, WI). She is currently an Assistant Faculty Member at St. Jude Children's Research Hospital and serves as the study statistician for the clinical trials in the Solid Tumor Division at St. Jude. Her research interests include high dimensional data analysis, variable classification, variable screening, variable selection, and statistical learning/machine learning.

**Hui Zhang Ph.D.** is an Associate Member at the Department of Biostatistics of St. Jude Children's Research Hospital. Hui received his Ph.D. in statistics from University of Rochester in 2010. Prior to it, he received M.S. in Pharmacology and B.S. in Biology from University of Rochester and Nankai University, China, respectively. Hui's statistical methodology expertise includes categorical data, longitudinal data, missing data and computational neuroscience. In St. Jude, Hui has supported clinical and biomedical investigators from various backgrounds including psychology, surgery, bone marrow transplantation, immunology and structural biology. Hui has published ~100 peer reviewed papers including ~40 statistical methodology manuscripts, with total ~4,000 citations. He is currently serving as the associate editor for the Journal of Statistical Computation and Simulation, and the president for the American Statistical Associate West Tennessee Chapter.

**Dr. Meredith Ray** is Assistant Professor of Biostatistics in the Division of Epidemiology, Biostatistics, and Environmental Health within the School of Public Health at the University of Memphis. She also has an affiliation with the Bioinformatics Program at UofM. Dr. Ray received her PhD in Biostatistics from the University of South Carolina in 2014, her MPH in Biostatistics from the University of Georgia in 2009, and her BS in Mathematics from Piedmont College in 2007. Her research has two focuses, methodology and applied methodology. Her methodological research focuses on spatial clustering using Bayesian and semi-parametric approaches and clustering across high-dimensional mixed data using nonparametric approaches. Her applied methodology focuses on lung cancer research, fMRI, and epigenetics.

**Dr. Mehmet Kocak** earned his MSc degree in applied statistics from Michigan State University and a PhD in statistics from the University of Memphis. He has been a study biostatistician for numerous Phase-I and Phase-II clinical trials conducted by St. Jude Children's Research Hospital from 2002-2011 and by Pediatric Brain Tumor Consortium (PBTC) from 2002-present, and for clinical and observational studies conducted by University of Tennessee Health Science Center (UTHSC) since 2011. His areas of research have been time-course gene expression data analysis, meta-analysis of p-values, Phase-I/II clinical trial design, Survival analysis, Categorical data analysis, and Statistical Simulations. He is an expert in the SAS programming language as well as SAS/Graph.

**Conference Program**  
**Friday, March 29, 2019**

**REGISTRATION DESK OPENS – 8:00am**

**BREAKFAST: 8:00am – 8:45am**

**Friday, March 29, 8:45am – 10:00am**

**Session 1**

**Welcome, 8:45am – 9:00am**

**Dr. David Kemme & Dr. Thomas Nenon, Provost**

**Session 1A: Plenary Session (Room: The Zone)**

**Chair: Vasile Rus**

9:00am

**Henry Kautz**

National Science Foundation

**Toward a Roadmap for AI Research & Development  
(Everything Old is New Again)**

**BREAK, 10:00am – 10:15am**

**Friday, March 29, 10:15am – 11:00am**

**Session 2**

**Session 2A: Statistical Learning (Room: Methodist Theater)**

**Chair: Deepak Venugopal**

10:15 am     *EM Bayesian variable selection for clustered discrete and continuous outcomes*

Yunusa Olufadi

10:30 am     *Dissecting the Gene-environment Interactions via Bayesian Hierarchical models*

Yu Jiang

10:45 am     *Identifying patterns of STROM images via standardized statistics in spatial point process*

Xueyan Liu, Luhang Han, Clifford S Guy, and Hui Zhang

**Session 2B: Biomedical Data Science (Room: Fishbowl)**

**Chair: Dale Bowman**

10:15 am     *A 2nd Opinion based on Automated Clinical Diagnosis*

Vivek Datla<sup>1</sup>, Sadid Hasan, Kathy Lee, Ashequl Qadir, Joey Liu, Shamsuzzaman Mohd., Hafiz Khan and Shahed Sorower

10:30 am     *Patterns of multi-genetic/epigenetic factors identified by a novel non-parametric clustering approach*

Meredith Ray, Lauren Sobral, S. Hasan Arshad, John Holloway, Wilfried JJ Karmaus, and Hongmei Zhang

10:45 am     *Identification of clinically relevant glaucoma biomarkers on fundus images using deep learning*

Siamak Yosefi

**BREAK, 11:00am – 11:15am**

**Friday, March 29, 11:15am – 12:00pm**

**Session 3**

**Session 3A: Data Science Platforms (Room: Methodist Theater)**

**Chair: Hongmei Zhang**

- 11:00 am     *Analysis and Visualization of Genomic Data*  
Vinhthuy Phan, Diem-Trang Pham, Caroline Melton, Adam Ramsey, Bernie J Daigle Jr., and Jennifer R. Mandel
- 11:15 am     *mCerebrum and Cerebral Cortex: Data Science Platforms for High-frequency Mobile Sensor Data*  
Timothy Hnat, Syed Monowar Hossain, Nasir Ali, Anandathirtha Nandugudi, Nusrat Nasrin, and Santosh Kumar
- 11:30 am     *Toward a Platform for Educational Data Mining in Computer Programming Courses*  
Vinhthuy Phan, Eric Hicks, Kriangsiri Malasri

**Session 3B: Applied Data Science (Room: Fishbowl)**

**Chair: Segun George**

- 11:00 am     *Anthropology by Data Science: The EPIC Project with Indicia Consulting as an Exploratory Case Study*  
Paff, Stephen
- 11:15 am     *Coh-MatrixML: Multilingual Semantic and Syntactic Analysis*  
Leah Windsor, Grayson Cupit, Zhiqiang Cai, Andrew Tackett, and Alistair Windsor
- 11:30 am     *Dynamic Real Time Data Driven Speed Bump in the World of Finance*  
Bidisha Chakrabarty, Jianning Huang, Pankaj K. Jainb

**LUNCH, 12:00pm – 1:15pm**

**Friday, March 29, 1:15pm – 2:15pm**

**Session 4**

**1:00 PM Data Science Cluster - Proposal Winners Announcement**

**Cody Behles, David Kemme**

**Session 1A: Plenary Session (Room: The Zone)**

**Chair: David Kemme**

1:15pm

**Dr. Bala Vaidyanathan**

FedEx

**Bridging the Gap between Theory and Practice**

**BREAK, 2:15pm – 2:30pm**

**Friday, March 29, 2:30pm – 3:30pm**

**Session 5**

**Session 5: POSTERS Session (Room: FIT lobby)**

*BI-GRU Capsnet for Students Answers Assessment*  
Nisrine Ait Khayi and Vasile Rus

*Dimension Reduction Methods for Metagenomic Data*  
Quang Tran, Diem-Trang Pham and Vinhthuy Phan

*Variational AutoEncoders for Unsupervised*  
Ott, Andrew

*Lung Cancer Mortality among Nonsmokers in the United States: Estimating Smoking-Attributable Mortality with Nationally-Representative Data*  
Joseph T. Lariscy, Robert A. Hummer, and Richard G. Rogers

*Automatic Evaluation of Cellular Lifespan from Time-lapsed Microscopic Images*  
Haobo Guo, Weiwei Dang, and Hong Qin

*A Narcissist's Jihad: An Automatic Text Analysis of Jihadi Literature*  
Joshua Tschantret

*Monitoring glaucoma worsening using manifold learning and unsupervised clustering*  
Yosefi, Siamak

*Cointegrated Cryptocurrencies? An Exploration of Price Movements*  
Stephen Lee

*An NLP Approach to Analyzing the Voynich Manuscript*  
Leah Windsor, Zhiqiang Cai, and Ivan Zelinka

*Large Scale Indoor LiDAR and Image Data Processing using Deep Neural Networks*  
Mazharul Hossain, Tianxing Ma, Junaid Ahmed Khan, Eddie Jacobs, and Lan Wang

*Improving Protein Modeling Through Residue Interaction Network Analyses*  
Thomas J. Summers, Qianyi Cheng, and Nathan J. DeYonker

*CDS Premiums, Bond Yields and Equity Returns: Causality, Co-movement and Price Discovery*  
Yi Lu

*Visualizing Data*  
Orrin Cooper and Mike Racer

*Auditor Style and Financial Reporting Similarity*  
Joseph A. Johnston and Joseph H. Zhang

*Singular Value Decomposition: Data-Driven Feature Selection for Transcriptomics Data Sets*



Ted Ling Hu and Bernie J. Daigle Jr.

*Blockchain Forks*

Achma Akther and Dr. Konstantin Sokolov

*Machine learning approach to early identification of international graduate students at risk of academic probation*

Divya Choudhary and Abhishek Ganguly

*CDS Premiums, Bond Yields and Equity Returns: Causality, Co-movement and Price Discovery*

Yi Lu

**Friday, March 29, 3:30pm – 4:30pm**

**Session 6**

**Session 6: Panel – Opportunities and Challenges for Data Science**

**in Biomedical Research (Room: Methodist Theater)**

**Chair: Dale Bowman**

**4:30 pm – 4:45 pm CLOSING REMARKS AND FAREWELL (Methodist Theater)**

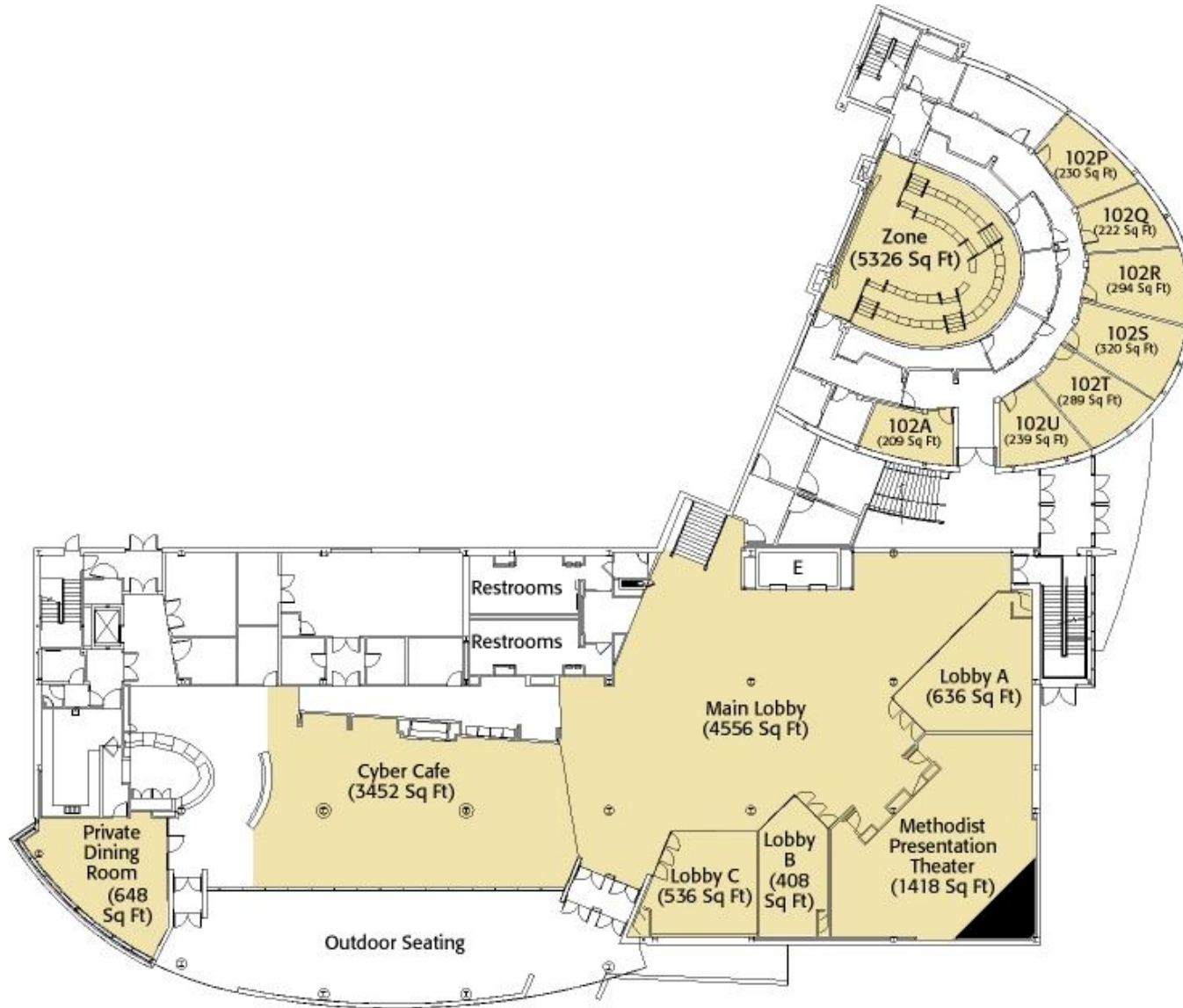
**END OF CONFERENCE**

**We hope that you enjoyed the Memphis DATA conference!**

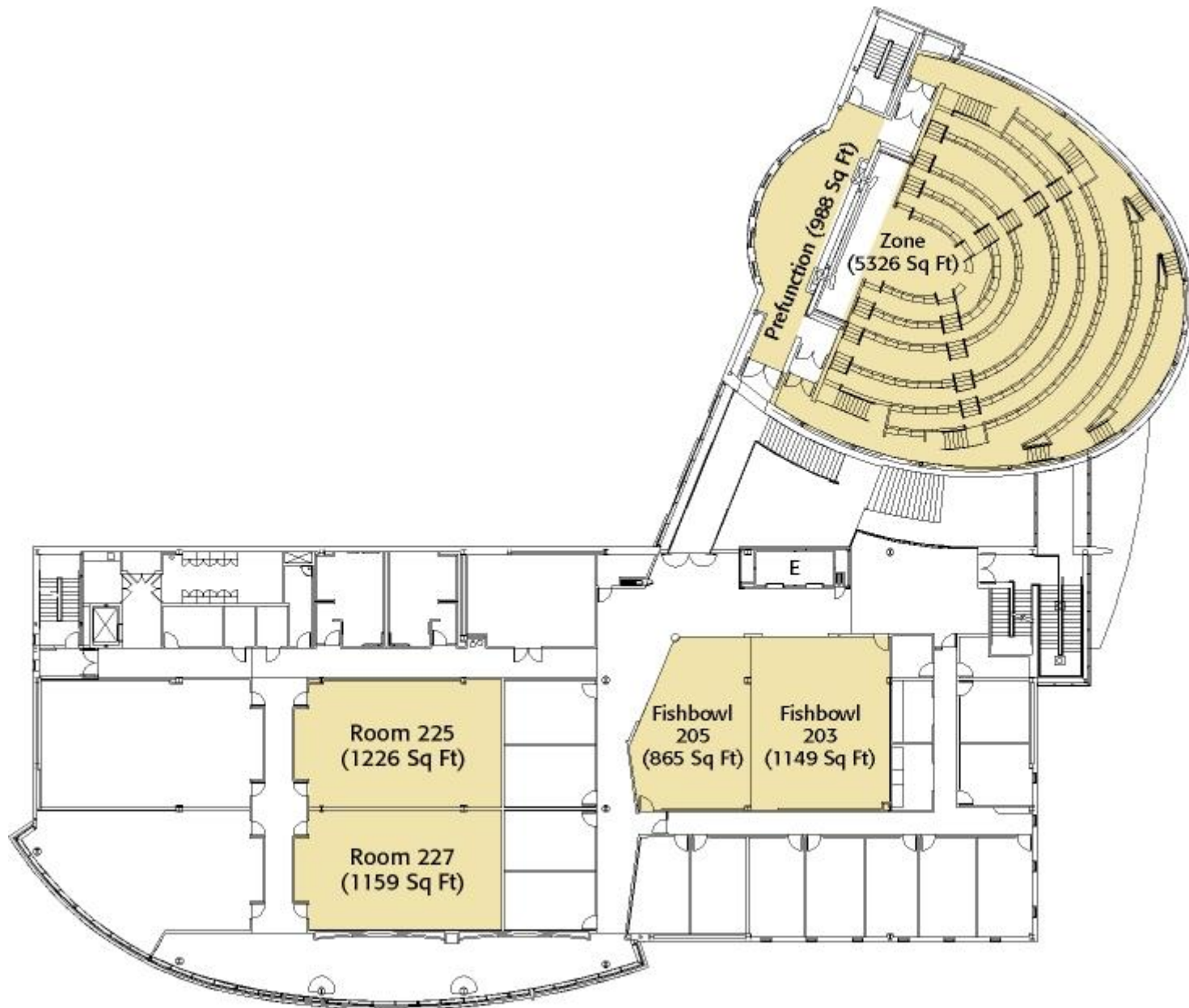
**Join us for Memphis DATA in 2020**

**<http://www.memphis-data.org>**

# Conference Center Map



FIRST FLOOR



**SECOND FLOOR**

## **Data Science Initiative's Mission**

*The goal of the Data Science initiative is to develop a vision for The University of Memphis and the local community, train future data scientists, create an environment to foster research and to build a Data Science community of practice that includes members from academia, government, and industry in West Tennessee, the Mid-South and beyond.*

[www.memphis-data.org](http://www.memphis-data.org)

