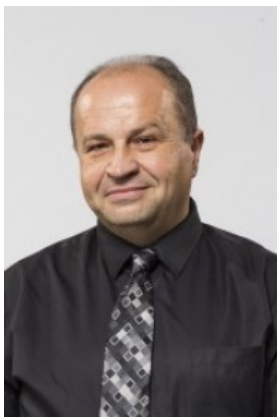# Workshop Speakers

Tutorial Instructors

Zoran Obradovic
Laura H. Carnell Professor of Data Analytics
Temple University

Zoran Obradovic is an Academician at the Academia Europaea (the Academy of Europe) and a Foreign Academician at the Serbian Academy of Sciences and Arts. He is a L.H. Carnell Professor of Data Analytics at Temple University, Professor in the Department of Computer and Information Sciences with a secondary appointment in Department of Statistical Science, and is the Director of the Center for Data Analytics and Biomedical Informatics. His research interests include data science and complex networks in decision support systems. Zoran is the editor-in-chief at the Big Data journal and the steering committee co-chair for the SIAM Data Mining conference. He is also the editorial board member at 13 journals and was the general chair, program chair or a track chair for 11 international conferences. He has published more than 370 articles and is cited more than 23,000 times (H-index 55). For more details see http://www.dabi.temple.edu/~zoran

Polo Chau
Associate Professor of Computing
Georgia Institute of Technology

Duen Horng (Polo) Chau is the Machine Learning Area Leader of the College of Computing at Georgia Tech. He co-directs Georgia Tech's MS Analytics program. His research bridges machine learning and visualization to synthesize scalable interactive tools for making sense of massive datasets, interpreting complex AI models and decisions, and solving real world problems. Cybersecurity is his primary application area.
Novel research from his group research has led to real-world impact: ShapeShifter (Intel) is the first physical adversarial attack against object detectors, ActiVis (Facebook) interprets complex deep learning models, GAN Lab (Google), Polonium and AESOP (Symmantec) protect 120M+ people from malware. He has received multiple academic, scholarship, and industry awards. His Ph.D. in Machine Learning is from Carnegie Mellon University.

Mladen Kezunovic
Regents Professor and Eugene E. Webb endowed Professor
Texas A&M University

Professor Kezunovic has been with Texas A&M University since 1986, and serves in several leading roles: Director, Smart Grid Center; Site Director, Power Systems Engineering Research Center (PSerc), and Director, Power Systems Control and Protection Lab. He is also the Principal Consultant, as well as President and CEO of XpertPower™ Associates, which has been providing consulting services for utility industry for over 25 years. He worked for Westinghouse Electric in the U.S.A. as a Systems Engineer on developing the first all-digital substation design during 1979-1980 and for Energoinvest Company in Europe as the Technical Lead for substation automation development during 1980-86. He served three terms (2009-2013) on the Board of Directors of the Smart Grid Interoperability Panel (SGIP). He was appointed by the US Secretary of Energy to the Electricity Advisory Committee for DOE in 2018. Dr. Kezunovic was a Principal Investigator on over 120 R&D projects, published more than 550 papers and gave over 100 invited lectures, short courses and seminars around the world. He is an IEEE Life Fellow and Distinguished Speaker, CIGRE Fellow, Honorary Member and Distinguished Member, and Registered Professional Engineer in Texas.

B. Aditya Prakash
Associate Professor of Computer Science
Virginia Tech.

Dr. Prakash obtained his Ph.D. in Computer Science at Carnegie Mellon University in 2012, and got his B.Tech (in CS) from IIT -- Bombay in 2007. He has published one book, more than 75 refereed papers in major venues (with multiple best paper/best-of-conference awards), holds two U.S. patents and has given five tutorials at leading conferences. His research interests include data science and applied machine learning, with emphasis on big-data problems in large real-world networks and time-series. His work has been funded through grants/gifts from the NSF, NSA, DoE and NEH and from companies like Symantec. Tools developed by his group have been in use in many places including ORNL, Walmart, and Facebook. He received a Facebook Faculty Award 2015, the NSF CAREER award 2018 and was named as one of '2017 AI Ten to Watch' by IEEE. Aditya's homepage is at: http://www.cs.vt.edu/~badityap.

Santiago Grijalva
Southern Company Distinguished Professor
Georgia Institute of Technology

Dr. Grijalva is the Southern Company Distinguished Professor of Electrical and Computer Engineering and Director of the Advanced Computational Electricity Systems (ACES) Laboratory at The Georgia Institute of Technology. His research interest is on decentralized power system control, power system analytics and economics, and future sustainable energy systems. He has been principal investigator for research under DOE, ARPA-E, EPRI, PSERC, NSF and numerous other sponsors. From 2002 to 2009 he was with PowerWorld Corporation. From 2013 to 2014 he was with the National Renewable Energy Laboratory (NREL) as founding Director of the Power System Engineering Center (PSEC). Dr. Grijalva is a Member of the NIST Federal Smart Grid Advisory Committee. Dr. Grijalva's holds a Ph.D. in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign.

Main Session Speakers

Mary Ann Weitnauer
Senior Associate Chair, Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Mary Ann Weitnauer has been a member of the Electrical and Computer Engineering faculty since completing her doctorate at Georgia Tech in 1989. Her research since the mid 1990s has been focused on the lower three layers of MIMO wireless networks that have virtual or distributed antenna arrays, with emphasis on wireless LAN, ad hoc, mesh, and sensor networks. She established the Smart Antenna Research Laboratory, which performs both experimental and theoretical studies. Recent SARL activities include synchronization for distributed or virtual arrays, nonlinear precoding and interference alignment for wireless LANs with distributed MIMO access points, modeling the residual from interference cancelation, distributed array-based network time synchronization, and millimeter wave communications. Most protocols are demonstrated on a 20-node network of software defined radios (SDRs) in practical environments and topologies.

David Pope
Sr Manager, US Energy
SAS

David leads a team of data scientists and technical architects who enable customers to gain value from implementing analytics in the US Energy and Manufacturing industries. He earned a BS in Industry Engineering and a Computer Programming Certificate from North Carolina State University and he has over 27 years of business experience, holds 12 patents for SAS, and wrote a book entitled: Big Data Analytics with SAS.

Keynote: Smart Grid Edge Analytics
This presentation will discuss why smart grid edge analytics matter and present a prioritized list of areas within the utility industry where analytics are being applied and which of these areas will be directly impacted by smart grid edge analytics. It will show how the traditional analytics lifecycle has evolved to include streaming or edge analytics and provide key drivers for smart grid edge analytics, examples of what is being solved, and why they are being solved. Finally, artificial intelligence (AI) and machine learning (ML) will be discussed highlighting the 3 main area in utilities where AI should be applied and end with some AI/ML customer examples.

Valentine A. Emesih
Vice-President, Operational Technology & Markets
CenterPoint Energy

Valentine Emesih is currently Vice President of Operational Technology & Markets at Centerpoint Energy. He is the responsible for smart grid systems that enable reliable and efficient electricity delivery, and market services for 2.4 million metered customers in the Houston metroplex, and our customers in our Indiana Electric division. He is also responsible for corporate advanced analytics capabilities, Technology Operations strategic planning, innovation and delivery. He has worked in utility and utility automation business for over 30 years. Mr. Emesih, who is licensed Professional Engineer, has worked for Centerpoint Energy since 1997. Prior to joining Centerpoint Energy, he held engineering, system development and project management positions at electric utility automation systems vendors – Ferranti International Controls in Sugar Land, Texas (currently ABB Enterprise Software); and Johnson Yokogawa Controls/Syseca Inc., in Carrollton, Texas (currently ARINC, Inc.).
He earned a bachelor's and a Master's degree in electrical engineering from The University of Texas at Austin and Auburn University, respectively.

Keynote: Using Grid Edge Information to Thrill Customers and Drive Asset Performance

CenterPoint Energy deployed 2.4 million smart meters and thousands of intelligent grid devices on its electricity distribution feeders. These smart devices generate an unprecedented amount of data on customer consumption, outage status, power quality, and asset condition. The combination of smart grid data with data from various sources, and advanced analytics capabilities have enabled CenterPoint energy to create new value for our customers and our business. I will briefly share our story and sample use cases on how we are continuously searching for ways to deliver better value to all our stakeholders.

Ajay Madwesh
Senior Manager, Analytics and Cognitive
Deloitte

Ajay has over 28 years of experience and has operated in various capacities from hands-on roles to being a trusted advisor to clients within Energy, Utilities, Resources, and Industry. He has helped teams conceptualize and realize solutions using practical ways of applying new technologies using his deep knowledge of emerging, operational and IT technologies. For utilities, Ajay has applied Big Data analytics, ML/AI to develop advanced asset analytics solutions, demand response, vegetation ignition propensity, revenue assurance, outage analytics and analytics for DER penetrated feeders to analyze capacity utilization and forecasting. He has been a panelist and speaker at the CPUC Wildfire Big Data Analytics summit, Utility Analytics IoT session, and Soft Grid Renewable energy panels. He has led Thought Leadership webinars with Siemens, SAP, and Accenture. He holds a Bachelor's degree in Electronics and Telecommunications and a Master's degree in Business Administration. He was selected for ABB's Corporate Executive program at Duke University and GE's Crotonville Leadership program.

Abstract: Application of Advanced Analytics

Energy, Utilities, Resources and Industry are in the process of a Digital Transformation process through the application of IoT devices, LiDar data,

images, weather, social data, and telematics. These are within the domain of operational analytics. There is a plurality of underlying technologies that have now matured and industrial grade that have become available. However, the technology selection depends on a best-fit analysis which in turn is based on the application space. There is a need for a practical process to identify the right technologies and applying them to architect a robust and production-grade solution. Often, these technology projects become stranded assets or remain as proof's-of-concept without seeing the light of day. Due diligence on Security, Scalability, Performance and Budgets are tend to critical to building a robust solution. This session will focus on providing insights on technology, architectural and process considerations to enable optimal organizational decisions.

Ratanak Heng
Senior Manager, Advanced Analytics
Exelon Utilities

Ratanak Heng currently serves as the Senior Manager of Advanced Analytics for Exelon Utilities (EU). Ratanak leads the Grid Analytics organization at Exelon Utilities to drive sustained value by improving reliability and customer satisfaction, decreasing operational cost, and capturing new revenue opportunities. In his role, Ratanak collaborates with business leaders and experts across the Exelon family of utility companies to further develop EU's analytics strategy, lead innovative projects and grow EU's grid analytics portfolio. Ratanak has over eleven years of utility experience in operations, planning, and engineering & standards with Baltimore Gas & Electric, a subsidiary of Exelon. Ratanak earned a bachelor's degree in Electrical Engineering and a master's degree in Business Administration from University of Maryland, College Park.

Abstract: Exelon Utilities Grid Analytics
Analytics is key to Exelon Utilities' (EU) long-term strategy for grid modernization and delivering sustained value to its customers. This session will provide insights on how EU is developing grid analytics capabilities to provide more reliable energy, quicker outage response and more efficient solutions for its customers. To support this effort, EU has partnered with GE to develop a new digital analytics application that will use information such as historical data and weather, asset and area conditions to help predict when and where an outage might occur. The technology will also produce real-time reports on equipment, integrating data from intelligent assets Exelon has deployed on its grid to provide greater precision in determining when maintenance is required. The utilities can then implement preventative solutions and also recover faster from storms and other incidents affecting the grid, improving overall customer reliability.

Brian Smith
Data Engineer and Analyst
Electric Power Board of Chattanooga

Brian is a Data Engineer and Analyst at EPB, Tennessee, the electric and fiber optic distributor behind the "Gig City" nickname. His work covers the full spectrum of data analytics, from statistical modeling to data visualization. He is primarily interested in finding methods for pulling needles from haystacks, in the form of geographic information system errors, capacitor bank failures, energy inefficiency hotspots, and more. In his free time, he enjoys playing chess, making music, reading and writing, and epic adventures. He holds a degree in Economics from Colorado State University.

Abstract: **Clustering Applications: Customer Usage Archetypes**
There are many benefits of learning more about the usage patterns of our customers, from marketing campaign design to peak-shaving. Since the time of usage is often as important as the actual amount, we can create standardized load profiles for each customer by averaging out meter reads over time. We then utilize our data infrastructure to apply a K-means clustering algorithm, grouping customers with similar load profiles together. We are then able to analyze how customer behavior changes through the seasons, and how that behavior relates directly to margin. This allows us to design cost-effective peak-shaving programs and marketing campaigns, predict future behavior, and gain insight towards load profile seasonality.

Soorya Kuloor
President & CEO
Symbiant Analytics

Soorya Kuloor is the President & CEO of Symbiant Analytics. Soorya has more than 23 years of experience in the utility industry. Most of his career, Soorya has focused on founding and building technical teams that are focused on development and delivery of software, analytics and simulation solutions for the utility industry. He has led multiple analytics solutions deployment projects at utilities and partner vendors. He has also led and performed multiple analysis studies for transmission and distribution utilities, and regulatory authorities for DER integration and optimization. He has published numerous articles in industry publications and magazines. An avid speaker, he has presented at multiple industry forums such as FERC, IEEE, Distributech and other conferences on distribution system analytics and optimization. Before founding Symbiant, Soorya was Practice Director for Distribution Operations and CTO for Advanced Grid Analytics at Landis+Gyr. Soorya was the co-founder and CTO of GRIDiant Corporation, where he led the development of a suite of Grid Analytics products, which was acquired by Landis+Gyr in 2014. His experience includes CTO position at Optimal Technologies, and lead software development positions at Hewlett-Packard, Agilent Technologies and Verano. Soorya holds a Ph.D. in Power System Optimization from the University of Calgary, Canada.

Abstract: **Analytics Use Cases using Smart Meter Data**
 More than 50% of US utilities have already deployed smart meters and early deployers coming up for second refresh cycle of smart meter deployments. Smart meters generate the largest amount of data of any system in a distribution utility. Modern meters typically measure multiple parameters including interval kWh

consumption, kVAh consumption, various voltage parameters, outage and restoration events and peak demand parameters. This large rich cache of data can be used in multiple analytics use cases starting from billing and customer analytics to grid analytics. This short presentation will give a quick overview of how meter data is used for these varied analytics use cases.

Tao Hong
Associate Professor
University of North Carolina at Charlotte

Dr. Tao Hong is Associate Professor and Research Director of Systems Engineering and Engineering Management Department, Director of BigDEAL (Big Data Energy Analytics Laboratory), NCEMC Faculty Fellow of Energy Analytics, and associate of Energy Production and Infrastructure Center at University of North Carolina at Charlotte. He is the Founding Chair of IEEE Working Group on Energy Forecasting, Director at Large of International Institute of Forecasters, General Chair of Global Energy Forecasting Competition, and author of the blog Energy Forecasting. Dr. Hong received his B.Eng. in Automation from Tsinghua University in Beijing and his PhD with co-majors in Operations Research and Electrical Engineering from North Carolina State University.

Abstract: Data Cleansing in Smart Grid Applications
Garbage in, garbage out. As the utility industry is embarking data-driven decision-making processes, data quality becomes increasingly important. In this talk, I will start with three case studies on data cleansing. The first one is on cleansing outliers in the load data. The second one is on detecting load transfers at distribution level. The third one is to detect anomalies in the commercial-grade weather data. At the end, I will introduce several robust forecasting models that can stay accurate even if the input data are contaminated.

Claude Godin
Director Data Analytics
DNV GL

Mr. Godin has over 40 years of experience in the electric utility industry specializing in load research and advanced energy data analytics; advanced metering infrastructure product & services development and implementation specializing in meter data acquisition; meter data management; non-intrusive load monitoring (NILM); and application development using advanced metering and sensor data.
Mr. Godin recently completed two of the world's largest end-use data development efforts in a large middle eastern country—one for electricity and the other for water. He is one of the principal architects of these projects, which focused on combining billing information, survey data, whole facility load metering and non-intrusive load monitoring to develop information for end-use forecasting. Also, responsible for conceiving and leading the development of advanced deep learning neural network algorithms for the disaggregation of high frequency premise level water flow into applicable end uses.
Mr. Godin was one of the principal authors of "Predictive Analytics – Beyond the Hype," September 2013, a contributing author on the Northeast Energy Efficiency Partnership (NEEP), "Changing the EM&V Paradigm," December 2015, and a

contributing author on the current update of the Association of Edison Illuminating Companies (AEIC) Load Research & Data Analytics Manual.

Abstract: Global Best Practices in Interval Load Data Analytics
DNV GL recently completed a study on Global Best Practices in utility driven Interval load data analytics by surveying twenty (20) utilities with mature Load Research & Interval Data Analytics Departments where most of the surveyed utilities were from jurisdictions that have Smart Metering systems fully deployed. The survey found that, besides big data analytics and management, the most pressing future analysis undertakings will be driven by use cases related to distribution network operations and planning, specifically in the areas of understanding the impacts of distributed energy resources and electric vehicles on distribution networks. The presentation will illustrate key findings of the study.

Renata Rawlings-Goss
Executive Director
South Big Data Regional Innovation Hub

Dr. Renata Rawlings-Goss is a Data Strategic Coach and Author of the upcoming book "Data Careers, Data Programs, and Data Talent" by Springer Press. She is the current Executive Director of the South Big Data Regional Innovation Hub, one of only four Big Data Innovation Hubs for the nation, serving 16 states – Delaware through Texas – in data science partnerships between industry, academia, and government. She is also the Director of Industry Partnerships for the Institute of Data Engineering and Science at the Georgia Institute of Technology. Previously, Dr. Rawlings-Goss worked with the White House Office of Science and Technology Policy, under President Obama, to create the National Data Science Organizers. She co-led the writing team for the Federal Big Data Strategic Plan, including 19 federal agencies, and was awarded a AAAS Big Data Science and Technology Policy fellowship with the National Science Foundation in the directorate of Computer and Information Science and Engineering (CISE-OAD). Dr. Rawlings-Goss lives in Atlanta, GA and received her training in physics, biophysics, and computational genomics from Florida A&M University, The University of Michigan, and The University of Pennsylvania. Her research interests include data-driven analysis of African and African-American genetic expression, as it relates to drug metabolism, as well as professional interests in data science education and workforce development in all communities.

Abstract: The South Big Data Innovation Hub
Promoting, Bridging, and Building Cross-sector Data-enabled Communities
The South Big Data Innovation Hub is part of a network of four regional Big Data Hubs, launched by the National Science Foundation to serve 16 states and the District of Columbia—from Texas to Delaware—with more than 1000 members from universities, corporations, foundations, and cities committing their support. The goal of the South Hub is to increase the capacity for data science workforce, innovation and partnerships across sectors to advance science, address societal challenges and spur economic development in the region.

Kyle Bradbury
Managing Director, Energy Data Analytics Lab
Duke University Energy Initiative

Kyle Bradbury is the Managing Director of the Energy Data Analytics Lab at the Duke University Energy Initiative where he leads applied research projects at the intersection of machine learning techniques and energy problems. His research includes developing techniques for automatically mapping global energy infrastructure and access from satellite imagery; transforming smart electric utility meter data into energy efficiency insights; and exploring the reliability and cost trade-offs of energy storage systems for integrating wind and solar power into the grid. He received both a Ph.D. in energy systems modeling and an M.S. in electrical and computer engineering from Duke University, as well as a B.S. in electrical engineering from Tufts University. He has worked for ISO New England, MIT Lincoln Laboratories, and Dominion.

Abstract: Distributed energy resource and infrastructure assessment using overhead imagery
Distributed energy resources are rapidly being deployed globally. Distributed solar photovoltaic (PV) systems add uncertainty to generation forecasts since the precise location and capacity of all systems are not always known by system operators tasked with maintaining balance between supply and demand. Similar challenges are faced in regions of the world increasing electricity access as small-scale distributed resources are used to meet nascent demand – in this context, accurate information on the location of the nearest grid connection is needed to determine whether a distributed solution or a grid connection is optimal. We work to fill these knowledge gaps by automatically mapping both solar PV and transmission and distribution infrastructure using overhead imagery and machine learning techniques. We apply convolutional neural networks for semantic segmentation and object detection to explore how accurately data on energy infrastructure can be assessed while also examining potential hurdles to scaling these techniques.

Matthew Lave
Senior Member of Technical Staff
Sandia National Laboratories

Dr. Matthew Lave is a Senior Member of the Technical Staff in the Renewable and Distributed Systems Integration Department at Sandia National Laboratories. Matthew is an expert at monitoring, analyzing, and modeling PV power production, both for performance assessment and for grid integration studies. His recent work has focused on island energy resiliency, especially in the US Virgin Islands and Puerto Rico, and the grid integration of electric vehicles. Previous work has included developing new methods for analysis of bifacial PV module performance, parameter and topology estimation of distribution grid secondary circuits (i.e., transformer to customer resistance and reactance), and evaluation of PV forecasts' performance and their value to grid operations.

Abstract: Towards High-Fidelity DER Integration
The integration of PV and other DER into the electric grid has challenges including variability, lack of utility monitoring of behind-the-meter systems, and varied technologies and performances. Analysis related to PV integration, including instantaneous hosting capacity, year-long voltage regulator tap change impact studies, and advanced forecasting and modeling has paved the way for

evaluation of integration of other DER into distribution grid operations. For example, electric vehicle (EVs) adoption may result in a sharp rise in variable feeder loads. If uncontrolled, could lead to significant impacts on voltage and current flow. Analyses originally used for PV integration can be easily adapted understand in-depth EV and other DER impacts on distribution grid operations.

Yingchen "YC" Zhang
Group Manager, Sensing and Predictive Analytics
National Renewable Energy Laboratory

Dr. YC Zhang is a principal researcher and manager of the Sensing and Predictive Analytics group at the National Renewable Energy Laboratory. He is also a visiting research assistant professor at the University of Denver and an adjunct faculty at Colorado State University. He authored and coauthored over 100 peer reviewed publications and holds 1 U.S. patent. His key areas of expertise lie in sensing and measurement for power systems, artificial intelligence, predictive analytics, advanced energy management system for future grids, the impact of large-scale integration of renewable energies on power system operations.
Dr. Zhang received his B.S. degree from Tianjin University, China in 2003 and his Ph.D. degree from Virginia Polytechnic Institute and State University in 2010. Dr. Zhang is a senior member of IEEE and serves as an editor for the IEEE Transactions on Sustainable Energy.

Abstract: Methods for behind-the-meter solar estimation
Distributed photovoltaic (PV) generation often occurs "behind the meter": a distribution grid operator can only observe the aggregate load, the sum of the true, masked load and the PV generation. This lack of observability poses a challenge to distribution grid management strategies, which typically require real-time or near-future disaggregated estimates of load and PV generation. In this presentation I will discuss use of local solar irradiance measurements to estimate and forecast solar power, to disaggregate the summed PV generation and true load signals.

Chris McClurg
Senior Product Manager
BuildingIQ

Chris McClurg has worked on all sides of the building delivery process, from facilities management to design, commissioning, nonprofit consultancy and data analytics software development. These experiences refined her focus on the application of cutting-edge technology in the deeply flawed reality of building design and operations. Chris is the Senior Product Manager for BIQ's services, where she drove the development of the Outcome Based Fault Detection service and the creation of new machine learning tools uncovering the deeper relationships in buildings. In this role, she has focused on applying cutting edge data analytics and IoT approaches to creating real, actionable results for facilities teams. Prior to BIQ, she worked at Rocky Mountain Institute, a non-profit think tank where she worked with large portfolios, campuses, and districts to integrate energy efficiency, energy systems, and innovative business models to create cost effective net-zero energy strategies.

Abstract: Harnessing Data Relationships with a Digital Twin
Every facilities person knows that buildings are complex systems which live and breathe around you. Subtle relationships between systems and the people with the environment drive the building's performance and these dynamics are never captured in traditional siloed construction documentation. These must be discovered through years of interactions with the building, and are constantly evolving as tenant fit outs or retrofits drastically alter the dynamics. When BuildingIQ approached building analytics, we knew we had to break free of the siloed, rules-based system specific approaches which dominate the industry. We had to build a digital twin capturing the true dynamics of the building. To achieve this, we created a suite of tools using machine learning to model and quantify the impacts of every system against every other system, including the whole building energy. Through these tools we can create a map of all the relationships driving a building's performance and identify what systems, decisions and processes truly drive energy and comfort.

Robert Rallo
Group Lead, Data Sciences
Pacific Northwest National Laboratory

Since 2017, Dr. Rallo leads the Data Sciences Group in the Advanced Computing, Mathematics, and Data Division at the Pacific Northwest National Laboratory. Prior to joining PNNL, he was an Associate Professor (2007-2016) in Computer Science and Artificial Intelligence and Director of the Advanced Technology Innovation Center (2012-2016) at the Universitat Rovira i Virgili, in Catalonia. He also served as chair of the Modeling WG for the European Commission Nanosafety Cluster (2012-2016) and as EU co-chair for the Community of Research of Predictive Modeling for the US-EU Dialogue on nanoEHS (2013-2015). His research interests focus on the applications of artificial intelligence at the intersection of science and engineering.

Abstract: Domain-aware Analytics
Analytics techniques leverage existing data to gain insight about the complex dynamic processes that occur in natural and engineered systems. However, there is a significant amount of scientific and expert knowledge, developed through years of research and development, that could be used to improve the accuracy and efficacy of data analytics, and in particular of machine learning techniques. In this talk we will discuss different approaches to embed domain knowledge into existing machine learning algorithms.