FLORIDA ATLANTIC UNIVERSITY

CHARLES E. SCHMIDT COLLEGE OF SCIENCE

DATA-DRIVEN SCIENCE AND AI CONFERENCE

Saturday, May 21, 2022

fau.edu/data

PROGRAM BOOKLET

Welcome

On behalf of the Charles E. Schmidt College of Science at Florida Atlantic University, we would like to welcome you to the Data-Driven Science and AI Conference, May 21, 2022. Before attending the conference, we would like for you to review the following information:

The plenary sessions will be presented in the Live Oak Pavilion located outside the Student Union Building just northeast of the newly renovated Student Union. Light breakfast, lunch and the Student Poster Session will also take place in the Live Oak Pavilion. Minisymposium Session talks will take place in the Grand Palm rooms (1, 2, 3, and 4) located inside the Student Union building. Your registration package will include a name badge and this conference program. Please wear your name badge all day.

Connection to Wifi

1. Look for and join the wireless network "FAU Guest" on your device.
2. In the pop-up window and under the "username and password" section, click on "Request a Login" at the bottom of the screen.
3. In the box, enter your 10-digit telephone number that can receive messages (standard rates apply). Click "Send SMS."
4. Once you receive the text message, login in with your 10-digit telephone number as your username and the 6-digit code (text message) as your password.
5. You are now connected to the FAUGuest wireless network. You will be able to connect for up to 72 hours after receiving your login information.
6. For more information or if you are experiencing issues, please call (561) 297 3999 or contact the helpdesk at http://helpdesk.fau.edu.

Parking

Parking is available in Lot 16, directly in front of the Student Union at Florida Atlantic University (FAU).   Parking in Lot 16 is available for you to park at no charge. Regular parking fees and charges do apply in all other lots on the FAU campus

- Please DO NOT PARK IN SPACES WITH METERS
- DO NOT BACK INTO PARKING SPOT

Index

Schedule

Saturday, May 21, 2022

Florida Atlantic University, Student Union, Boca Raton Campus

---

8:00 a.m.
Student Union Lobby

Sign-In and Registration
Complimentary light breakfast in Live Oak Pavilion.

8:45 a.m.
Live Oak Pavilion

Welcome

9:00 a.m.
Live Oak Pavilion

Plenary Talk
"The Impact of AI in Financial Market Investment Strategies"
Moustapha Awada, Ph.D., *Physicist and Chief Risk and Investment Sciences Officer at Alpha Innovations Global Investment, LLC*

9:50 a.m.
Live Oak Pavilion

Coffee Break

10:15 a.m.
Grand Palm Room

Minisymposium Sessions
Talks in this session will take place in parallel at three separate time slots of 30 minutes each: 10:15 a.m., 10:45 a.m., and 11:15 a.m., unless otherwise noted. Session topics and talks listed below.

South Florida Tech Hub Sponsored: Health Tech Panel Discussion

10:15-11:45 a.m.
Grand Palm Room 4

"Digital Transformation and Data-Driven Insight in Healthcare: Where We Are and Where We're Going"
Moderator: Meggie Soliman, *Director of Strategic Innovations at DSS, Inc.*
Panelists: Pete Martinez, *CEO of Sivotech Bioinformatics*; Aliya Aaron, *Founder of Nursing Innovation Hub;* Yenvy Truong*, Founder of LSM Group*; Christopher Kunney, *Chief of Strategy & Business Development at DSS, Inc.*

Urban Applications of Data Science

10:15 a.m.
Grand Palm Room 1

"Person-Activity Tracking with Smartphones as a Way to Measure the Impact of Government Lockdowns During the COVID-19 Pandemic"
Louis Merlin, Ph.D., *Assistant Professor, FAU Department of Urban and Regional Planning*

10:45 a.m.
Grand Palm Room 1

"Structured Additive Regression and Tree Boosting"
Michael Mayer, Ph.D., *Pricing Actuary, La Mobiliere, Bern Switzerland*

11:15 a.m.
Grand Palm Room 1

"How Does Street Space Influence Crash Frequency? An Analysis Using Segmented Street View Imagery"
Jonathan Stiles, Ph.D., *Postdoctoral Fellow, FAU Urban and Regional Planning*

Computational Perspectives on Brain Function and Behavior

10:15 a.m.
Grand Palm Room 2

"What Can and Cannot Be Concluded About Neural Representations From Neuroimaging Data"
Fabian A. Soto, Ph.D., *Associate Professor, Florida International University*

10:45 a.m.
Grand Palm Room 2

"Normalization in Neuroscience and Deep Neural Networks"
Odelia Schwartz, Ph.D., *Associate Professor, University of Miami*

11:15 a.m.
Grand Palm Room 2

"Thalamocortical Function in Sleep and Wakefulness"
Carmen Varela Castro, Ph.D., *Assistant Professor, Florida Atlantic University*

AI in Chemistry and Molecular Medicine

10:15 a.m.
Grand Palm Room 3

"Incorporating Biological Context into Data-Driven Models"
Mohammed Ali Eslami, Ph.D., *Chief Data Scientist, Co-founder, Netrias, LLC*

10:45 a.m.
Grand Palm Room 3

"Data-Guided Protein Stabilization"
Jedediah M. Singer, Ph.D., *Principal Research Scientist, Two Six Technologies*

11:15 a.m.
Grand Palm Room 3

"Topological Data Analysis Algorithms in Robot Motion Planning"
Chinwe Pamela Ekenna, Ph.D., *Assistant Professor, Department of Computer Science, University at Albany*

11:45 a.m.

Coffee Break

12:00 p.m.
Live Oak Pavilion

Plenary Talk
"Using Data and AI to Discover New Solar Energy Materials in the Laboratory"
Joshua Schrier, Ph.D., *Kim B. and Stephen E. Bepler Chair of Chemistry at Fordham University*

12:50 p.m.

Break

1:00 p.m.
Live Oak Pavilion

Complimentary Lunch and Student Poster Session

Minisymposium Sessions

South Florida Tech Hub Sponsored: Career and Internship Panel

2:15-3:45 p.m.
Grand Palm Room 4

"Data/AI and Tech Career and Internship Experiences Panel"
Rich Viens, *Chief Financial Officer, PeakActivity*; Valeria Tineo, *Incoming Account Executive, Organizational Business Development at Cleveland Cavaliers, and recent FAU Alumna*; Jose Delgado, *Software Developer, Office Depot and recent FAU Alumnus*; Lakshamana Sankarakuttalam, *Senior Manager, IT, Enterprise Intelligence, Office Depot*

Medical Physics and Image Processing

2:15 p.m.
Grand Palm Room 1

"A Socio-technical Approach to Reduce the Digital Divide: A Case of Biomedical Content Authoring and Publishing Retrieval"
Syed Ahmad Bukhari, Ph.D., *Assistant Professor and Director of Healthcare Informatics at St. John's University, New York*

2:45 p.m.
Grand Palm Room 1

"Designing an IT-Based Decision Support System for Optimizing Cancer Treatment"
Adnan Muhammad Shah, Ph.D., *Data Scientist, Postdoc Research Associate, Department of Physics, Florida Atlantic University*

3:15 p.m.
Grand Palm Room 1

"Finding the Trees in the Forest: Bulk Analysis of Scientific Literature"
Silke Dodel, Ph.D., *Founder, DeepScience, LLC;* Sophie Schuelke and Radleigh Santos, *Department of Mathematics, Nova Southeastern University*

Geoscience Data Mining using Machine Learning Techniques

2:15 p.m.
Grand Palm Room 2

"Machine Learning Applications in Remote Sensing and Environmental Management"
Sara Durgan, Ph.D., *Remote Sensing & UAS Data Scientist at AECOM*

2:45 p.m.
Grand Palm Room 2

"Augmenting Geological and Image Analysis Techniques with Machine Learning for Sand Characterization"
Molly Smith, Ph.D., *Scientist and Teacher, Omni Middle School*

3:15 p.m.
Grand Palm Room 2

"Permafrost Characterization in Alaska and Greenhouse Gases Estimation in the Everglades using AI and Multiple Earth Data Sources"
Caiyun Zhang, Ph.D., *Professor, Department of Geosciences, Florida Atlantic University*

Mathematical Tools for Biological Networks and Neuroscience

2:15 p.m.
Grand Palm Room 3

"Comparing Topological Measures of Neural Structure and Function"
Chad Giusti, Ph.D., *Assistant Professor of Mathematical Sciences, University of Delaware*

2:45 p.m.
Grand Palm Room 3

"Computational Characterization of Large-scale Brain Dynamic Landscapes and Attractor Transition Networks"
Mengsen Zhang, Ph.D., *Postdoctoral Researcher, University of North Carolina*

3:15 p.m.
Grand Palm Room 3

"Efficient Selection and Falsification of Network Models Via Combinatorial Dynamics"
Shane Kepley, Ph.D., *Postdoctoral Researcher, Vrije Universiteit Amsterdam*

3:45 p.m.

Break

4:00 p.m.
Live Oak Pavilion

Plenary Talk
"BioNFTs: Enabling Decentralized Consented Genomics in the Metaverse"
Daniel Uribe, MBA, *Co-Founder & CEO, GenoBank.io*

4:00 p.m.
Live Oak Pavilion

Wrap Up

Abstracts

Plenary Talks

Live Oak Pavilion

The Impact of AI in Financial Market Investment Strategies

9:00 a.m.

Moustapha Awada, Ph.D.
Physicist and Chief Risk and Investment Sciences Officer at Alpha Innovations Global Investment LLC

Quantitative Finance is one of the most important pillars in financial economics. Major financial institutions consider it as a critical domain of its operation particularly when it comes to investment in financial markets. We will discuss how the advent of data science and machine learning in the last two decades gave rise to predictive analytics and AI trade execution that has generated billions of dollars in profitability and cost effectiveness.

****

Using Data and AI to Discover New Solar Energy Materials in the Laboratory

12:00 p.m.

Joshua Schrier, Ph.D.
Kim B. and Stephen E. Bepler Chair Professor of Chemistry, Fordham University

New materials are essential for many clean-energy technologies, but the materials discovery process is often a tedious process of trial and error. In this talk, I'll discuss ways that we have been able to combine laboratory data collection and automation with data-science and AI tools to accelerate the synthesis and characterization of novel halide perovskites, a class materials with applications for photovoltaics and sensors.

****

BioNFTs: Enabling Decentralized Consented Genomics in the Metaverse

4:00 p.m.

Daniel Uribe, MBA
Co-Founder & CEO, Genobank.io

Biosamples and their multi-omics data (genomics, proteomics, epigenomics, etc.,) pass through a sequence of events, a supply chain of custody, from preanalytical biosample collection to post analytical results interpretation, dissemination, and application to vendor commercialization. This presentation examines risks in this chain of custody and the asymmetries of power and knowledge between parties from donors to controllers and the multitude of ways our biosamples are accessed inevitably create opportunities for expanded risks and benefits across stakeholders. The purpose of this presentation is to serve as an

introduction to the biodata provenance journey with an ethical risk analysis of who benefits, who is at risk, and who decides on the biological assets (biosamples + biodata). Lastly, we present a use case for promoting transparent consent in the supply chain with the introduction of Biological Non-Fungible Tokens or BioNFTs that resolve in a public blockchain. We offer a description of how BioNFT, a decentralized biosample consent platform, could serve as a sentinel mechanism that mitigates risks and expands benefits to stakeholders. BioNFTs are proposed as a dynamic blockchain solution which provides liability coverage for the stakeholders at the top and a negotiation mechanism for data originators and donors at the bottom of the supply chain.

Minisymposium Sessions

Grand Palm Rooms 1, 2, 3, 4

South Florida Tech Hub Sponsored: Health Tech Panel Discussion

Grand Palm Room 4

Digital Transformation and Data-Driven Insight in Healthcare: Where We Are and Where We're Going

Moderator: Meggie Soliman, *Director of Strategic Innovations at DSS, Inc.*
Meggie Soliman serves as the Director of Strategic Innovations, Applications for DSS, Inc. As an experienced Healthcare Information Technology Leader and Strategist, Meggie has worked with hospitals, health care leaders, project management teams, and executive leadership on emerging technology offerings, strategic solutions, and technology adoption. Prior to moving into this role, Meggie's experience includes leading teams in Project Management Operations across the healthcare technology spectrum within DSS, Inc. working with the largest hospital system in the world. Mrs. Soliman has a track record for developing and implementing innovation solutions across the Healthcare Network, identifying the technical opportunities, and serving as a liaison between clinical and technical teams. Meggie is also actively involved in the community, she serves as the Chair for the Healthcare & Life Science Committee with both the Palm Beach North Chamber of Commerce and Tech Hub South Florida.

Panelist: Pete Martinez, *CEO of Sivotech Bioinformatics*
Pete Martinez is Chairman and CEO of the SIVOTEC family of companies; high tech and scientific innovators at the massive intersection of Health, Sports, Technology, Big Data Analytics, and Artificial Intelligence. An international entrepreneur, he has led numerous high impact initiatives in the areas of high technology, business consulting, healthcare, and education. Martinez is a 32 year veteran of IBM, where he was Vice President of Global Business Services and Senior Executive for IBM South Florida. He is also a research affiliate at MIT, and visiting professor at Florida Atlantic University and Florida International University. With over 30 patents and publications he is a frequent international speaker while consulting to leading edge companies. He is a graduate of the University of Miami College of Engineering and past Chairman of the Florida Research Consortium. June 22, 2007 was proclaimed 'Pete Martinez Day' by the City of Boca Raton.

Panelist: Aliya Aaron, *Founder of Nursing Innovation Hub;*
Aliya brings 20+ years of healthcare, information technology, consulting, healthcare security, and innovation

experience forward in healthcare leadership. Director of Health Informatics for Trinity Health, which spans a geographic reach over 25 states and 17 clinically integrated networks. Founded Nursing Innovation Hub, Inc.(NIHUB) in 2019 to drive support for entrepreneurs and innovators developing new health tech ventures. NIHUB is a nationally recognized health tech business accelerator focused on three themes (technology, innovation, and entrepreneurship) and three pillars (Launch, Grow, Scale). NIHUB "the hub" provides support to a diverse community of healthcare startups and innovators globally. Cofounded, Center for Health Tech, Research, Innovation & Policy (CHTRIP), a non-profit service, research, and education center based in Miami, focused on promoting improved access to public health services for underserved communities. Founded AMR Healthcare Consulting, LLC (AMR), an award-winning boutique healthcare technology consulting firm providing IT management, software development, telehealth, remote patient monitoring, and EHR implementation services to health systems since 2010. Enjoys supporting the entrepreneurial community as an advisor, mentor, judge, speaker, and board member for various organizations including Texas Global Health Security Innovation Consortium (TEXGHS) and Together.Health. Earned a Master of Science in Health Systems Management from George Mason University, and a Bachelor of Science in Nursing from Winston-Salem State University

Panelist: Yenvy Truong, *Founder of LSM Group*;
Yenvy Truong is a 15+ year veteran of the healthcare and wellness industries—a serial entrepreneur, investor and thought leader. Currently, Yenvy is CEO to Upswing Health. A virtual care technology company servicing individuals with acute and chronic musculoskeletal conditions. Previous, she was founder and CEO of HealthSnap, a lifestyle remote patient monitoring healthcare company. She has founded and co-founded start-ups in the arenas of digital health and remote lifestyle monitoring, as well as lead companies in the metabolic, functional, anti-aging, and regenerative medicine space.

Panelist: Christopher Kunney, *Chief of Strategy & Business Development at DSS, Inc.*
Christopher Kunney serves as Chief Strategist & Business Development for DSS, Inc.'s Juno EHR product line. As an experienced healthcare information technology executive and strategist Kunney works with hospitals, health systems, accountable care organizations, clinically integrated networks, ancillary service providers, and physician-owned entities to provide guidance and insights specific to emerging healthcare technology offerings, compliance-related issues, and technology adoption. Before joining DSS, Inc., Kunney served as Senior Vice President of Healthcare Information Technology for Coker Group and Managing Partner and IT Strategist of IOTECH CONSULTING. He also served as Entrepreneur in Residence at The Combine, Director of Emerging Healthcare Technology for AT&T, Vice President and COO for Health Innovations, and Vice President and CIO for Piedmont Healthcare. Most recently Christopher has launch at podcast titled, "Straight Outta Health IT."

Urban Applications of Data Science

Grand Palm Room 1

Organized by: Steven Bourassa, Ph.D., *FAU Department of Urban and Regional Planning*

Person-Activity Tracking with Smartphones as a Way to Measure the Impact of Government Lockdowns During the COVID-19 Pandemic

Louis Merlin, Ph.D., Assistant Professor, *Florida Atlantic University, Department of Urban and Regional Planning*

This study uses data from the Mobintel Platform, which tracks smartphone location along Clematis Street in Downtown West Palm Beach from January 1, 2020 to September 30, 2021. The data provides a proxy for person-activity and is examined longitudinally during the COVID-19 pandemic to examine mandated public health lockdowns on street-level activity. This study helps illustrate how activity patterns change in response to different phases of a pandemic and is one of the first to examine how big data from mobile phone can be used to test person-level activity during a public health crisis.

Structured Additive Regression and Tree Boosting
Michael Mayer, Ph.D., *Pricing Actuary, La Mobiliere, Bern Switzerland*

Structured additive regression (STAR) models are a rich class of regression models that include the generalized linear model (GLM) and the generalized additive model (GAM). Using feature interaction constraints, we show that such models can be implemented by the gradient boosting powerhouses XGBoost and LightGBM, thereby benefiting from their excellent predictive capabilities. We illustrate the methodology with case studies pertaining to house price modeling, with very encouraging results regarding both interpretability and predictive performance.

How Does Street Space Influence Crash Frequency? An Analysis Using Segmented Street View Imagery
Jonathan Stiles, Ph. D., *Postdoctoral Fellow, FAU Urban and Regional Planning*

Road crashes in metropolitan areas are challenging to prevent because they stem from the interactions of multiple modes of travel within multifaceted built environments. In this study we segment street view imagery into discrete objects to consider the relationship of individual street elements and clusters of elements with crash frequency on Columbus Ohio roads. Through modeling we identify that street views of trashcans and trees are associated with fewer crashes, and that arterial class roads with more road surface, open sky and signage are the least safe.

Computational Perspectives on Brain Function and Behavior

Grand Palm Room 2

Organized by: William Alexander, Ph.D., *FAU Department of Psychology and Center for Complex Systems and Brain Sciences*

What Can and Cannot Be Concluded About Neural Representations From Neuroimaging Data
Fabian A. Soto, Ph.D., *Associate Professor, Florida International University*

It is a common practice in cognitive neuroscience to test hypotheses about neural representation using human neuroimaging data (fMRI, EEG). Such data are not direct measures of neural activity, but rather measures that have been transformed through a measurement model (the process that transforms the activity of a neural population into observed measurements). I will show that studying the impact of the measurement process on

the outcome of neuroimaging studies, through theoretical and computational research, is key to reach valid conclusions about neural representation.

Normalization in Neuroscience and Deep Neural Networks
Odelia Schwartz, Ph.D., *Associate Professor, University of Miami*

Neural responses and perception of visual inputs strongly depend on the spatial context, i.e., what surrounds a given object or feature. I will discuss our work on developing a statistical model of contextual influences in vision, resulting in a generalization of a ubiquitous nonlinear computation in neuroscience known as normalization. I will further discuss how we are incorporating such nonlinearities into deep neural networks, and implications for biological and artificial systems.

Thalamocortical Function in Sleep and Wakefulness
Carmen Varela Castro, Ph.D., *Assistant Professor, Florida Atlantic University*

I will summarize experimental work in rodents and computational models that we use to gain insight into the functions of thalamocortical networks. The results suggest mechanisms that contribute to sparse neuronal firing during sleep, and provide a theoretical foundation for context-dependent input discrimination during wakefulness.

AI in Chemistry and Molecular Medicine

Grand Palm Room 3

Organized by: Andrew Terentis, Ph.D., *FAU Department of Chemistry and Biochemistry* and Francis Motta, Ph.D., *FAU Department of Mathematical Sciences*

Incorporating Biological Context into Data-Driven Models
Mohammed Ali Eslami, Ph.D., *Chief Data Scientist, Co-founder, Netrias, LLC*

Biology is all about context and the data generated by experiments also follows that same principle. Thus, representation and integration of that context into AI models is required to enhance their predictions and reason across the system. In this talk, we will define what we mean by "biological context" and how models can benefit from their inclusion.

Data-Guided Protein Stabilization
Jedediah M. Singer, Ph.D., *Principal Research Scientist, Two Six Technologies*

Proteins, the building blocks of life, are long molecular chains of amino acids that fold up into complex structures. When these structures are more stable, they last longer and perform their functions more reliably.

This talk describes data-driven approaches to predicting protein stability, and the use of the resulting models to guide stabilizing refinements to protein sequences.

Topological Data Analysis Algorithms in Robot Motion Planning
Chinwe Pamela Ekenna, Ph.D., *Assistant Professor, Department of Computer Science, University at Albany*

This talk will discuss research on theoretical foundations for improving motion planning algorithms. These algorithms have been used in computational biology and geometry problems to plan robot motions as they traverse different start and goal positions. I will discuss the algorithms we developed and the innovations that have resulted from using various topology methods such as simplicial complexes, Vietoris rips complex, and discrete Morse theory. Our algorithms allow us to capture critical points and define singularities in configuration space, and we demonstrate that it works in higher dimension spaces. The results of our algorithm's simulation and real-world application to ground robots will be discussed.

South Florida Tech Hub Sponsored: Career and Internship Panel

Grand Palm Room 4

Data/AI and Tech Career and Internship Experiences Panel

Rich Viens, Chief Financial Officer, PeakActivity
South Florida native Richard Viens (BSEE '99) is an Entrepreneur at heart. Rich's passions include building technology companies, practicing Stoicism and adventure traveling. (He once spent a week living in an igloo!) He currently is the Chief Financial Officer at PeakActivity and proud to be Treasurer of South Florida Tech Hub.

Valeria Tineo, Incoming Account Executive, Organizational Business Development at Cleveland Cavaliers, and recent FAU Alumna
Valeria Tineo attended Saint Thomas Aquinas High School in Fort Lauderdale, FL where she participated in basketball, speech & debate, and the Spanish team. Due to her academic success, she earned a full academic scholarship to Florida Atlantic University where she double-majored in Marketing and International Business. During her time at FAU, she's been able to gain experience within the marketing agency, technology, and sports industries. These experiences have enhanced her professional experience and have given her a leg up in her career. She recently graduated from FAU and has decided to take a full-time role with the Cleveland Cavaliers as an Account Executive for their Organizational Business Development sector.

Jose Delgado, Software Developer, Office Depot and recent FAU Alumnus
Jose Delgado graduated from Lake Worth Community Highschool at the top of his class with a 4.0 GPA. He was involved in multiple extracurricular and student-led activities including forming BroCode to expose students to computer programming. Throughout highschool, Jose was dual enrolled at Palm Beach State College where he graduated as an early admit student, was a Medical Assistant at the Center for Bone and Joint Surgery of the Palm Beaches and was also a summer intern at the Max Planck Neuroscience Institute. He had a love for both medicine and technology. Jose took part in the Path To College program during highschool which helped him obtain a scholarship to Florida Atlantic University's Wilkes Honors College. During his time at FAU, Jose interned with South Florida Tech Hub, a nonprofit supporting tech throughout the South Florida region, where

he helped drive tech talent initiatives, prepare virtual community events, and built dynamic stats in website through API's. Tech Hub introduced Jose to one of their member companies, Office Depot, where he accepted an internship as a Business Intelligence Intern. He was also accepted to Florida Atlantic University's Master Degree program in IT & Management with a focus on Data Analytics. Jose recently obtained his Bachelor of Science in Data Analytics this May and continues to tackle his Masters.

Lakshamana Sankarakuttalam, Senior Manager, IT, Enterprise Intelligence, Office Depot

Medical Physics and Image Processing

Grand Palm Room 1

Organized by: Luc Wille, Ph.D., *FAU Department of Physics* and Wazir Muhammad, Ph.D., *FAU Department of Physics*

A Socio-technical Approach to Reduce the Digital Divide: A Case of Biomedical Content Authoring and Publishing Retrieval

Syed Ahmad Bukhari, Ph.D., *Assistant Professor and Director of Healthcare Informatics at St. John's University, New York*

Despite recent advancements in information retrieval techniques, significant barriers exist towards seeking quick access to online biomedical information. It is mainly because most biomedical contents are unstructured with minor or no explicit machine-interpretable semantic (context-aware) markups or annotations. Adding semantic annotations to biomedical content is critical for improving their search and retrieval precision rate by enhancing the context-aware indexing of search engines, e.g., Google. Over the years, a few good Semantic Content Authoring (SCA) frameworks have been developed for elite users in non-biomedical domains. However, the biomedical information domain lacks any such framework and process, even for novice users. Thus, there is a critical need to bridge the digital divide by developing a framework enabling novice users to author and publish biomedical contents that are easily accessible and exhibits improved accuracy. This talk will explain an out-of-the-box socio-technical approach that our team has developed to enhance medical and healthcare content FAIRness (Findability, Accessibility, Interoperability, and Reusability). Our approach democratizes the process of semantic content authoring currently available to sophisticated users by extending cutting-edge semantic web infrastructure to research and education communities at large. Furthermore, our system enables peer-to-peer knowledge creation that will open up career progression opportunities for a socio-economically diverse pool of learners, especially the under-represented minorities in research and academia.

Designing an IT-Based Decision Support System for Optimizing Cancer Treatment
Adnan Muhammad Shah, Ph.D., *Data Scientist, Postdoc Research Associate, Department of Physics, Florida Atlantic University*
Our IT-based clinical decision support system will be able to provide cancer patients with the best and specific treatment plan and recommendation system for a specific cancer type.

Finding the Trees in the Forest: Bulk Analysis of Scientific Literature
Silke Dodel, Ph.D., *Founder, DeepScience, LLC; Sophie Schuelke and Radleigh Santos, Department of Mathematics, Nova Southeastern University*
We are living in the information age. Whereas in the past access to information was limited (think libraries with physical books and journals) nowadays the challenge is to pick relevant information from the relentless information flow that surrounds us. This is especially critical in the biomedical sciences where the abundance of new literature often outpaces the capacity of individual researchers to digest and take into account new results, thereby hampering scientific progress. Here we discuss a system that is designed to represent information efficiently and present an approach to extract and group information from a large number of biomedical publications in a hierarchical manner. Specifically, we combine clustering at the publication level with clustering at the sentence level. This allows researchers to gain a high-level overview of the publication contents before zooming in on more specific results of their interest. We demonstrate our approach using hundreds of publications that appear as search results from a literature search performed on PubMed. We show that our approach reveals distinct patterns of scientific developments as they are reflected in the biomedical literature. We discuss the implications of our results for efficient information processing and knowledge graph generation.

Geoscience Data Mining using Machine Learning Techniques

Grand Palm Room 2

Organized by: Caiyun Zhang, Ph.D., *FAU Department of Geosciences*

Machine Learning Applications in Remote Sensing and Environmental Management
Sara Durgan, Ph.D., *Remote Sensing & UAS Data Scientist at AECOM*
This presentation will cover three case studies where machine learning algorithms improved data products for environmental management. These applications include wetland classification with unmanned aircraft system (UAS) photogrammetry, mapping impervious surfaces for stormwater assessment, and predicting mercury contamination in sediment, surface water, and fish tissue for different remediation activities.

Augmenting Geological and Image Analysis Techniques with Machine Learning for Sand Characterization
Molly Smith, Ph.D., *Scientist and Teacher, Omni Middle School*
Traditional sand analysis is labor and cost-intensive, entailing specialized equipment and operators trained in geological analysis. Digital image analysis techniques and computer vision have been well developed and applied in various fields but rarely explored for sand analysis. This research explores capabilities of remote sensing digital image analysis techniques, machine learning, and photogrammetry techniques to automate or semi-automate the traditional sand analysis procedure. A framework was developed by combining object-based image analysis and machine learning classification of microscope imagery to model mineral composition of ten terrigenous beach sand samples.

Permafrost Characterization in Alaska and Greenhouse Gases Estimation in the Everglades using AI and Multiple Earth Data Sources
Caiyun Zhang, Ph.D., *Professor, Department of Geosciences, Florida Atlantic University*
Field data collection of active layer depth, and snow depth of permafrost in cold regions, and greenhouse gases (CO_2 and CH_4) in wetlands are challenging due to the harsh environments. We combined remote sensing datasets with limited field measurements to model and estimate active layer depth, snow depth in Fairbanks, Alaska, and upscale greenhouse gases in the Florida Everglades using artificial intelligence (AI) techniques. Spatial and temporal patterns of these valuables were delineated to document their response to climate and environmental change.

Mathematical Tools for Biological Networks and Neuroscience

Grand Palm Room 3

Organized by: William Hahn, Ph.D., *FAU Department of Mathematical Sciences* and William Kalies, Ph.D., *FAU Department of Mathematical Sciences*

Comparing Topological Measures of Neural Structure and Function
Chad Giusti, Ph.D., *Assistant Professor of Mathematical Sciences, University of Delaware*
Topological methods provide us with measures of mesoscale, nonlinear organization in both structure and function of brain networks. Persistent homology, in particular, has been useful as a statistic for differentiation between populations and states. However, determining when two structures detected by topological means are similar -- homologous, in the non-mathematical sense -- remains a challenge. Here, I will discuss a new approach for comparing topological structure across modalities, subjects, or measurements; show initial results demonstrating their application to simulations of multi-system neural activity; and, discuss current and near-future challenges for applying topology to the study of brain networks. No prior knowledge of topological data analysis will be assumed.

Computational Characterization of Large-scale Brain Dynamic Landscapes and Attractor Transition Networks
Mengsen Zhang, Ph.D., *Postdoctoral Researcher, University of North Carolina*
The brain is a nonlinear dynamical system that affords multiple distinct patterns of activation, which may be considered as attractors mathematically. This talk addresses the scarcely studied problem of the global distribution of attractors in a high-dimensional, extensively multistable nonlinear dynamical model of the brain, and how such distribution relates to empirically observed correlation between brain regions in neuroimaging time series. A data-driven method is further developed to recover the transitions between attractors as a directed graph from both simulated and real neural time series.

Efficient Selection and Falsification of Network Models Via Combinatorial Dynamics
Shane Kepley, Ph.D., *Postdoctoral Researcher, Vrije Universiteit Amsterdam*
The goal of synthetic biology is to design biological networks which exhibit specified, possibly complex,

dynamical behavior both in vitro and in vivo. Many designs fail one or both requirements because the system parameters are high dimensional, difficult to measure, or due to the lack of first principles governing the dynamics. Since it often requires many expensive, time-consuming experiments to produce the data required to test a single design, failures are a very costly problem. In this talk we describe a modeling and computational framework that allows for rapid screening of thousands of potential network designs for particular dynamic behavior. This approach is based on analysis of all possible combinatorial dynamic phenotypes which a network design is capable of exhibiting and estimating the likelihood of witnessing each phenotype. The analysis of a single network is much coarser, but also much faster, than numerical simulation of differential equations models. We will demonstrate the approach using a recent example in which we performed a complete study of bistability and hysteresis in three node gene regulatory networks.

DATA-DRIVEN SCIENCE
AND AI CONFERENCE

Program Committee Members

William Kalies, Ph.D., *Conference Chair, FAU Schmidt College of Science*

Medical Physics and Image Processing Session
Luc Wille, Ph.D., *FAU Department of Physics*
Wazir Muhammad, Ph.D., *FAU Department of Physics*

AI in Chemistry and Molecular Medicine Session
Andrew Terentis, Ph.D., *FAU Department of Chemistry and Biochemistry*
Francis Motta, Ph.D., *FAU Department of Mathematical Sciences*

Computational Perspectives on Brain Function and Behavior Session
William Alexander, Ph.D., *FAU Department of Psychology and Center for Complex Systems and Brain Sciences*

Geoscience Data Mining and Satellite Image Analysis Session
Caiyun Zhang, Ph.D., *FAU Department of Geosciences*

Urban Applications of Data Science Session
Steven Bourassa, Ph.D., *FAU Department of Urban and Regional Planning*

Mathematical Foundations of Data Science and AI
William Hahn, Ph.D., *FAU Department of Mathematical Sciences*
William Kalies, Ph.D., *FAU Department of Mathematical Sciences*

Organizing Committee Members

William Kalies, Ph.D., *Conference Chair*
Zach Greathouse, MBA
Maria Provost, Ed.D.

Thank You

The organizers for the Data-Driven Science and AI Conference would like to thank everyone for contributing to the success of this conference. A special thanks goes out to our sponsors, the National Science Foundation, and South Florida Tech Hub.

Sponsors