UNIVERSITY OF DELAWARE

DATA SCIENCE INSTITUTE

DELAWARE Data Science DARWIN COMPUTING SYMPOSIUM

FEB 12, 2020

#DARWINcomputing2020

AGENDA

Registration	8:30-9:00am
Introduction	9:00-9:25am
Session I: DARWIN for Chemcal, Physical and Materials Science & Engineering	9:30-10:25am
Coffee Break	10:25-10:40am
Stefan Robila, NSF Program Manager: "Building a Diverse and Sustainable Scientific Cyberinfrastructure Ecosystem"	10:40-10:55am
Session II: DARWIN for Biomedical and Biological Sciences	10:55-11:50am
Lunch	11:50am-12:40pm
Poster Session	12:40-1:40pm
About DARWIN	1:40-2:10pm
Session III: DARWIN for Environmental Sciences	2:15-3:10pm
Coffee Break	3:10-3:25pm
Session IV: DARWIN for Business, Social Sciences and Educational Efforts	3:25-4:20pm
Panel and Discussion	4:20-4:50pm
Closing Remarks	4:50-5:00pm

INTRODUCTION

Rudolf Eigenmann

DARWIN Principal Investigator (PI)
Director, Data Intensive & Computational Science (DICoS) Core
Professor, Electrical and Computer Engineering
University of Delaware

Welcome to the DARWIN Computing Symposium!

With over 170 attendees, this Symposium shall be the beginning of a series of events that will boost computational and data-intensive (CDI) research and technology at the University of Delaware (UD) and in the greater Delaware region. CDI methods have become an important catalyst in almost all areas of science, and they create competitive advantages in the private sector.

I'm excited that we can introduce to you today - on Darwin Day, February 12 2020 - the new DARWIN compute cluster, sponsored by the National Science Foundation (NSF). While the cluster will be a key instrument for accelerating CDI projects at the University of Delaware and partners, it is not just a machine. DARWIN is the Delaware Advanced Research Workforce and Innovation Network - a project that will create synergy in the greater Delaware region to evolve this important technology broadly.

The DARWIN cluster has just been ordered, will be fully operational by mid-2020, and I look forward to seeing the progress of the more than 50 research teams that have contributed to winning the NSF award. Today, you will hear from several leaders of such efforts, and you will see presentations of many results enabled by the new resource at future DARWIN Computing Symposia.

In addition to NSF, the DARWIN Computing Symposium has been made possible by UD's Data Science Institute, host of this event, as well as by our industry sponsors Atipa, Dell, and Compass Red. We thank these sponsors for their support, and encourage everyone here today to stop by the Atipa and Dell industry tables stationed in the atrium outside this event during the breaks.

We would also like to thank the many external partners and supporting organizations that have made the trip to be here with us today, including our collaborators at the NSF, the NSA, ChristianaCare, Nemours Children's Hospital, CompassRed, the Chemours Company, Solvay, Wesley College, Delaware State University, Wilmington University, and Howard University.

Please take advantage of this event not just to learn about DARWIN, but also to meet new potential collaborators and, in this way, contribute to our mission.

REMARKS FROM NSF

Stefan Robila

Director, Office of Advanced Cyberinfrastructure (OAC)
Directorate, Computer and Information Science and Engineering (CISE)
National Science Foundation

"Building a Diverse and Sustainable Scientific Cyberinfrastructure Ecosystem"

Dr. Stefan Robila is Program Director in the Office of Advanced Cyberinfrastructure (OAC), Directorate for Computer and Information Science and Engineering (CISE), at the National Science Foundation. Dr. Robila is part of the OAC's Data group and manages a portfolio of proposals and awards that are part of the Cyberinfrastructure of Sustained Scientific Innovation (CSSI), OAC Research Core and Major Research Instrumentation (MRI) programs. Dr. Robila is also Professor of Computer Science in the Department of Computer Science at Montclair State University (MSU). His principal research interests are in computational sensing. This work has now expanded into more general research and applications for large data sets. More recently, he has also started working on greening the computing infrastructure by designing and deploying a network of sensors that monitors environmental and system utilization of the MSU's data center, and uses the generated data for visualization as well as input for the creation of an expert system aimed to assist IT staff in reducing energy and equipment costs. Stefan completed his M.S. and Ph.D. in Computer Science at Syracuse University and his B.S. at the University of Iasi (Romania). Apart from MSU, he also worked at University of New Orleans and held visiting appointments at University of Oxford and Cambridge University (UK), University of Geneva (Switzerland) and Uppsala University (Sweden). Dr. Robila has published over 70 refereed publications, held grants from National Science Foundation, PSEG Foundation, Hewlett-Packard, SPIE, and Sun Microsystems.

SESSION I:

DARWIN FOR CHEMICAL, PHYSICAL AND MATERIALS SCIENCE & ENGINEERING

Arthi Jayaraman *(Moderator)*

DARWIN co-PI
Professor, Chemical and Biomolecular Engineering
Professor, Material Sciences and Engineering
University of Delaware
https://www.che.udel.edu/people/faculty/arthij/

Professor Arthi Jayaraman received her B.E (Honors) degree in Chemical Engineering from Birla Institute of Technology and Science, Pilani, India in 2000. She received her Ph.D. in Chemical and Biomolecular Engineering from North Carolina State University in 2006, and from 2006-2008 conducted her postdoctoral research in the department of Materials Science and Engineering at University of Illinois-Urbana Champaign. In August 2008, she joined the faculty of the Department of Chemical and Biological Engineering at University of Colorado at Boulder, and held the position of Patten Assistant Professor. In August 2014, she joined the faculty at the University of Delaware as Associate professor of Chemical and Biomolecular Engineering and Materials Science and Engineering. Her research expertise lies in development of theory and simulation techniques and application of these techniques to study polymer functionalized nanoparticles and polymer nanocomposites, and to design macromolecular materials for biomedical applications. Her research has been recognized with the Dudley Saville Lectureship at Princeton University (2016), the AIChE COMSEF division Young Investigator Award (2013), ACS PMSE division Young Investigator (2014), University of Colorado Provost Faculty Achievement Award (2013), and Department of Energy (DOE) Early Career Research Award (2010). Her teaching has been recognized with the University of Colorado outstanding undergraduate teaching award (2011) and University of Colorado outstanding graduate teaching award (2014) in Chemical and Biological Engineering.

Dionisios Vlachos

Director, Delaware Energy Institute
Director, Catalysis Center for Energy Innovation
University of Delaware
https://www.che.udel.edu/people/faculty/vlachos/

"Data Science and Multiscale Modeling in Sciences and Engineering"

Dionisios (Dion) G. Vlachos is the Allan and Myra Ferguson Professor of Chemical & Biomolecular Engineering, a Professor of Physics and Astronomy at the University of Delaware, the Director of the University of Delaware Energy Institute (UDEI) and of the Catalysis Center for Energy Innovation (CCEI), an Energy Frontier Research Center (EFRC). He obtained a five-year diploma in Chemical Engineering from the National Technical University of Athens, Greece in 1987, his M.S. and Ph.D. from the University of Minnesota in 1990 and 1992 respectively, and spent a postdoctoral year at the Army High Performance Computing Research Center in Minnesota. Professor Vlachos is the recipient of the 2016 Catalysis Club of Philadelphia Award, the R. H. Wilhelm Award in Chemical Reaction Engineering from AIChE (2011) and is an AAAS Fellow (since 2009). He also received a NSF Career Award and an Office of Naval Research Young Investigator Award. He is a member of AIChE, ACS, the Combustion Institute, MRS, the North American Catalysis Society (NACS) and the Society for Industrial and Applied Mathematics (SIAM). He has given named lectures: the Doumas Lecture, Department of Chemical Engineering, Virginia Tech, 2016, the ICI Distinguished Lecturer, DB Robinson Lectureship Series at the University of Alberta, 2014-2015, and the J. D. Lindsay Lecture Series, Chemical Engineering Department, Texas A&M University, Oct. 8, 2014. Dr. Vlachos' main research thrust is multiscale modeling and simulation along with their application to catalysis, crystal growth, portable microchemical devices for power generation, production of renewable fuels and chemicals, catalyst informatics, detailed and reduced kinetic model development and process intensification.

Jodi Hadden-Perilla

Assistant Professor, Chemistry & Biochemistry
University of Delaware
https://www.chem.udel.edu/people/full-list-searchable/jhadden?uid=jhadden&Name=Jodi%20Hadden-Perilla

"Revealing the mechanisms of biological machines through the computational microscope"

Dr. Jodi Hadden-Perilla is an Assistant Professor in the Department of Chemistry and Biochemistry at the University of Delaware. She completed her B.S. in Chemistry from Armstrong Atlantic State University in 2007 and her Ph.D. in Computational Chemistry from the University of Georgia in 2014. From 2014 to 2017, she was a Postdoctoral Research Associate at the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign, where she also served as the Technology Training Organizer for the National Institutes of Health Center for Macromolecular Modeling and Bioinformatics. From 2017-2019, she was the Chair's Postdoctoral Fellow in Chemistry and Biochemistry, also at the University of Delaware. Dr. Hadden-Perilla's research employs all-atom molecular dynamics simulations -- often referred to as the computational microscope -- to study biological machines, including viruses and molecular motors.

SESSION II:

DARWIN FOR BIOMEDICAL AND BIOLOGICAL SCIENCES

Cathy Wu *(Moderator)*

DARWIN co-PI
Director, Data Science Institute
University of Delaware
https://dsi.udel.edu/

Dr. Cathy Wu has conducted bioinformatics and data analytics research for 25 years in areas encompassing genomic annotation and digital curation, biomedical text mining and natural language processing, biomedical ontology and semantic computing, genome-phenome network modeling, big data analytics and machine learning. She has led/co-led numerous grants from the NIH, NSF, DOE, and other agencies, totaling over $60 million funding support for her lab. The effort includes the international UniProt Consortium with over 6 million page views per month from 750,000 unique sites worldwide. Recognized as a “Highly Cited Researcher” (top 1%) since 2014, she has published more than 270 peer-reviewed papers, with over 35,000 citations and an h-index of 65. She established the Center for Bioinformatics and Computational Biology (CBCB) in 2009, serving as the founding director of the Bioinformatics master’s, doctoral and graduate certificate programs and launched the Data Science Institute (DSI) as the founding director in 2018. The DSI serves as a nucleating effort to accelerate research in data science, catalyzing interdisciplinary research collaborations across fields impacting our society—from engineering, health sciences, physical sciences and environmental sciences, to behavioral and social sciences and public policy—across UD campus and partner organizations.

Jeffrey Caplan

Associate Professor, Plant & Soil Sciences
Director Bioimaging Center
University of Delaware
https://www.udel.edu/academics/colleges/canr/departments/plant-and-soil-sciences/faculty-staff/jeffrey-caplan/

“Bioimaging and image analysis within biomedical research at UD and in the greater Delaware region"”

Dr. Caplan directs the Bio-Imaging Center at the University of Delaware that provides microscopy access to over 100 research groups each year. His main goal is to ensure that the Bio-Imaging Center’s growing user group continues to obtain high quality assistance and microscopy data that complements, augments, or elevates the research conducted within each project. His role as the Director is to match research projects with teams of highly specialized research scientists and instrumentation within the Bio-Imaging Center. A rapidly growing need is to include assistance with image analysis to handle the larger, multidimensional image data sets from new imaging technologies.

Mia Papas, PhD

Corporate Director
Christiana Care Value Institute
https://research.christianacare.org/valueinstitute/people/mia-papas-ph-d-ms/

"Recognition of spatiotemporal patterns in data: A modern day application of a Darwinian methodology"

Dr. Mia Papas is the Corporate Director of the ChristianaCare Value Institute where she is responsible for oversight of research and operations. She is also the site Principal Investigator for the Delaware Clinical and Translational Research (DE-CTR ACCEL) Program focused on building a research structure that allows for the rapid translation of research into practice targeting improved health outcomes for all Delawareans.

Dr. Papas' expertise is the design and conduct of research that utilizes real-world evidence to understand disease, evaluate treatments, and demonstrate the impact of healthcare innovations and interventions. She is actively engaged in research that leverages data from electronic health records, administrative claims, registries, and government databases to answer critical questions about disease epidemiology, burden, and costs. Through a focus on real world evidence, she works toward bridging the gap between clinical and translational research, advancements in the quality of care, and improved population health outcomes.

Dr. Papas received her PhD in Epidemiology from the Bloomberg School of Public Health at Johns Hopkins University, her Masters of Science in Biostatistics and Epidemiology from the Arnold School of Public Health at the University of Massachusetts, Amherst and her Bachelor of Science in Mathematics from Fairfield University. She has authored over 60 peer-reviewed articles, obtained funding from the National Institutes of Health and the Centers for Disease Control and Prevention, and presented her research at over 120 national and international conferences. Dr. Papas has analytic expertise in experimental and observational study designs, the assessment of validity and reliability of screening tools, generalized and hierarchical linear modeling, and the use of geographical information systems in understanding the effect of place on health and health behaviors. She has been an active member of the American Public Health Association and the American College of Epidemiology for over 15 years and is a Founding Board Member for the Delaware Public Health Association. She has taught the principles of epidemiology throughout the world and is committed to developing good scientific practice in medical research. She is a native Delawarean and lives in Newark with her husband and two daughters.

SESSION III:
DARWIN FOR ENVIRONMENTAL SCIENCES

Jeffrey Buler *(Moderator)*

Associate Professor, Wildlife Ecology
University of Delaware
https://www.udel.edu/academics/colleges/canr/departments/entomology-and-wildlife-ecology/faculty-staff/jeffrey-buler/

Jeff Buler is an Associate Professor in the Department of Entomology and Wildlife Ecology. He earned his PhD in Biology from the University of Southern Mississippi and MS in Wildlife from Louisiana State University. He established the Aeroecology Program at the University of Delaware in 2011. He is an international leader in the development of novel methods and software for using weather surveillance radars to study the broad-scale distribution, movement, and habitat use patterns of birds, insects, and bats.

Kent Messer

Professor, Economics
University of Delaware
https://www.udel.edu/canr/departments/applied-economics-and-statistics/faculty-staff/kent-messer/

"Harnessing Big Data and High Performance Computing to Protect Water in the Changing Coastal Environment"

Kent Messer is the S. Hallock du Pont Professor of Applied Economics at the University of Delaware. As a behavioral economist, Kent's research interests are in the nexus of agriculture and the environment. He received his BA from Grinnell College, his MS from the University of Michigan, and his PhD from Cornell University. He has been an investigator on interdisciplinary proposals worth over $70 million, and published over 80 publications and two textbooks. Kent also co-directs the Center for Behavioral and Experimental Agri-Environmental Research (CBEAR), a USDA Center of Excellence; and he also is the Project Director of the $23 million NSF-EPSCoR project entitled, "Water in the Changing Coastal Environment of Delaware" (aka. Project WiCCED).

Pinki Mondal

Assistant Professor, Geography
University of Delaware
https://www.udel.edu/academics/colleges/ceoe/departments/gss/faculty/pinki-mondal/

"Earth observation for sustainable ecosystem and livelihood: A geospatial data science approach"

Dr. Pinki Mondal leads the "EArth observation for Sustainable Ecosystem and Livelihood" (EASEL) research group at the University of Delaware. She is an Assistant Professor in the Department of Geography and Spatial Sciences with a joint appointment in the Department of Plant and Soil Sciences. Dr. Mondal is a Resident Faculty at the UD Data Science Institute and is an affiliate to the Delaware Environmental Institute (DENIN), the Water Science and Policy Program, and the Center for Food Systems and Sustainability (CENFOODS). Dr. Mondal is a geospatial data scientist interested in the dynamics of coupled natural and human systems. She received her PhD in Land Change Science from the University of Florida. Prior to joining UD, she was a Senior Research Associate at Columbia University.

SESSION IV:

DARWIN FOR BUSINESS, SOCIAL SCIENCES, AND EDUCATIONAL EFFORTS

Benjamin Bagozzi *(Moderator)*

DARWIN co-PI
Associate Professor, Political Science & International Relations
University of Delaware
https://www.poscir.udel.edu/people/faculty/Bagozzib

Benjamin Bagozzi (PhD, Penn State University, 2013) is an Associate Professor of Political Science & International Relations at the University of Delaware and an affiliated faculty with the University of Delaware's Data Science Institute, MS in Data Science Program, and Center for Political Communication. His primary areas of specialization are political methodology and international relations. Within international relations, his research and teaching interests include environmental politics, international political economy, and the study of intra and international political violence. Methodologically, he teaches and conducts research in computational social science, text analysis, event data, and the analysis of rare events.

Malcolm D'Souza

Professor, Chemistry
Dean, Interdisciplinary/Collaborative Sponsored Research
Wesley College
https://wesley.edu/faculty-staff/faculty-staff-directory/malcolm-dsouza

"Guiding Principles & Interdisciplinary Nature of Wesley's Data Science Initiatives"

Dr. Malcolm J. D'Souza is Professor of Chemistry & Dean of Interdisciplinary Collaborative/ Sponsored Research at Wesley College (Wesley). Dr. D'Souza laid the groundwork to establish the Wesley Sponsored Research Office and by developing an inclusive and successful mentoring program, he launched undergraduate research (UR) activities to inspire undergraduates from marginalized groups to be resilient, collaborative, and successful. His academic training in chemistry, physics, and mathematics has allowed him to develop UR projects, presentations, and publications in the areas of chemical informatics, chemometrics, chemical synthesis and mechanistic pathway studies, STEM-education, (societal) data-driven analytics, and for the design of commercial databases that assist in the development of new pharmaceutical and agricultural products.

Dr. D'Souza has successfully administered NIH, NSF, NOAA, NIFA and NASA projects (e.g. staffing, research, budget), collaborated with other researchers and produced 98 peer-reviewed journal publications. Dr. D'Souza has served as a faculty mentor to 44 faculty and has been a research mentor to 287 undergraduates (68% underrepresented). 130 Wesley College undergraduates trained in his laboratory have earned national and regional recognitions. The Delaware Bioscience Association, a Delaware consortium of pharmaceutical and biotechnology firms, medical device manufacturers, agricultural biotech and chemical companies, has recognized Dr. D'Souza in 2016, 2018, and 2019 for supporting innovation work and encouraging creativity with STEM undergraduate involvement. In addition, Dr. D'Souza has received the 2016 NIH-NIGMS Sidney A. McNairy Jr. Mentoring Award, the 2012 American Chemical Society E. Emmett Reid Award for Excellence in Teaching at a Small College in the Mid-Atlantic Region, and the 2009 NIU Golden Anniversary Alumni Award, for which he was nominated and identified by the selection committee as one of fifty distinguished alumni (among its nearly 70,000 alumni) spanning the 50-year history of the college.

Xiao Fang

Professor, Management Information Systems
University of Delaware
https://lerner.udel.edu/faculty-staff-directory/xiao-fang/

"Machine Learning for Business and Social Good: Social Network Analytics, FinTech, and Health Analytics"

Xiao Fang is Professor of Management Information Systems and JPMorgan Chase Senior Fellow at Lerner College of Business & Economics and Institute for Financial Services Analytics, University of Delaware. He also holds courtesy appointments at Department of Computer and Information Science as well as Department of Electrical and Computer Engineering. He studies business and social network analytics with research methods and tools drawn from reference disciplines including Management Science (e.g., Optimization) and Computer Science (e.g., Machine Learning). He has published in elite business journals including Management Science, Operations Research, MIS Quarterly, and Information Systems Research as well as computer science outlets such as ACM Transactions on Information Systems and IEEE Transactions on Knowledge and Data Engineering. His research has been featured in MIT Technology Review, MIT Sloan Management Review, and London School of Economics Business Review. Recent research awards received by Professor Fang include the 2017 INFORMS Design Science Award and the 2019 INFORMS Workshop on Data Science Best Paper Award.

ABOUT DARWIN

Rudolf Eigenmann

DARWIN Principal Investigator (PI)
Director, Data Intensive & Computational Science (DICoS) Core
Professor, Electrical and Computer Engineering
University of Delaware

William Totten

DARWIN co-PI
Enterprise Architect
Information Technologies
University of Delaware

This presentation will describe the architecture of the DARWIN cluster and the timeline for installation and operation. The mechanisms for requesting computational allocations by UD members, partners in academia and the private sector will be explained. Services for obtaining assistance with developing, optimizing, and running computational and data applications will also be discussed.

PANEL DISCUSSION

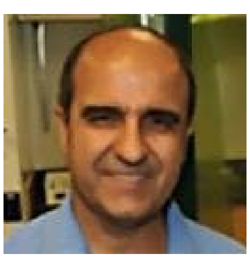
Benjamin Bagozzi *(Moderator)*
DARWIN co-PI
Associate Professor, Political Science & International Relations
University of Delaware

Cathy Wu
DARWIN co-PI
Director, Data Science Institute
University of Delaware

William Totten
DARWIN co-PI
Enterprise Architect, Information Technologies
University of Delaware

Juan Perilla
Assistant Professor, Chemistry & Biochemistry
University of Delaware

Jean-Yves Delannoy
Manager, Statistical Modeling and Artificial Intelligence Team
Solvay Leader in Modeling and Simulation
Solvay S.A.



Hacene Boukari
Associate Professor, Physics and Engineering
Director, Graduate Program
Delaware State University

POSTER SESSION

1. **Utkarsh Kapoor, Postdoc/Researcher, Chemical & Biomolecular Engineering**
   Title: Computational Studies of Soft Materials Using High Performance Computing
   Abstract: This poster presents the ongoing computational projects in the Jayaraman research group based in the Departments of Chemical and Biomolecular Engineering and Materials Science and Engineering at the University of Delaware. The various research projects are focused on understanding molecular-level phenomena governing complex biological processes and material science problems using high performance computing, theory and simulation techniques .
2. **Kapildeb Dolui, Postdoc/Researcher, Physics & Astronomy**
   Title: Spin-orbit-torque-driven magnetic phase transition in van der Waals heterostructure of CrI3/TaSe2
   Abstract: The recently discovered 2D magnetic insulator CrI3 is an intriguing case for basic research and spintronic applications since it is a ferromagnet in the bulk, but an antiferromagnet in bilayer form, with its magnetic ordering amenable to external manipulations. Using first-principles quantum transport approach, we predict that injecting unpolarized charge current parallel to the interface of bilayer-CrI3/monolayer-TaSe2 van der Waals heterostructure will induce spin-orbit torque (SOT) and thereby driven dynamics of magnetization on the first monolayer of CrI3 in direct contact with TaSe2. By combining calculated complex angular dependence of SOT with the Landau-Lifshitz-Gilbert equation for classical dynamics of magnetization, we demonstrate that current pulses can switch the direction of magnetization on the first monolayer to become parallel to that of the second monolayer, thereby converting CrI3 from antiferromagnet to ferromagnet while not requiring any external magnetic field.
3. **Dixit Bhatta, PhD Student, Computer & Information Sciences**
   Title: Cost-Aware Cloudlet Placement in Edge Computing Systems
   Abstract: One of the well-known challenges in Edge Computing is strategic placement of cloudlets. The fundamental goals of this challenge are to minimize the deployment cost and to guarantee minimum latency for the users of edge services. We address this challenge by designing a cost-aware cloudlet placement approach that fully maps user applications to appropriate cloudlets while ensuring their latency requirements. We investigate the effectiveness of our proposed approach by performing extensive experiments based on New York City OpenData. The results show that our approach obtains close to optimal cost solutions with significantly reduced execution time.
4. **Yuqi Wang, Staff, Computer & Information Sciences**
   Title: Clustering of Universal Protein Sequences
   Abstract: Protein sequence clustering and family classification are very important steps for characterizing the protein functions, which is essential to most research areas in biology and precision medicine aimed to understand the mechanism of human disease. As a member in the UniProt consortium that provides a single, centralized, authoritative resource for protein sequences and functional information, this team is tasked to perform computational clustering of the protein sequences into the UniRef clusters in three levels and the proteomes into the reference groups, and classify the proteins into PIRSF families. These tasks require infrastructure support for high-performance computation and ample memory and storage DARWIN system could offer. The DARWIN system will enable the current clustering procedures to scale up and reduce overall computation time. Key features of DARWIN system that enable this project include scaling, large memory/storage, co-location, system access, and training access.
5. **Charles Cheung, PhD Student, Physics & Astronomy**
   Title: Development of the Online Portal for High-Precision Atomic Physics Data and Computation
   Abstract: There has been a demonstrated need for high-quality atomic data and software in several scientific communities, including atomic, plasma, and astrophysics. Having an accurate atomic theory is indispensable for experiments involving studies of fundamental interactions, astrophysics, atomic clocks, plasma science, quantum degenerate gases, quantum information, precision measurements, and others. The goal of this project is to provide the scientific community with an online portal with access to a database of high-precision atomic properties and a package of application codes that can be used to compute these properties. The code package will feature MPI parallelization and can be run on modern large-scale computational facilities to enable precision modeling of complex atoms not currently possible with any existing codes.

6. **Parinaz Barakhshan, PhD Student, Electrical & Computer Engineering**
Title: Xpert Network: Exchanging Best Practices and Tools for Computational and Data-Intensive Research
Abstract: The Xpert Network is part of an NSF project that aims to advance science by increasing the productivity of researchers who use computational and data-intensive (CDI) methods for pushing the frontiers. CDI-based methods have been recognized as high-impact science enablers, and investments in computational experts who help domain researchers to increase the productivity of CDI applications have been recognized as having a high return. The project builds on these opportunities by creating synergy among teams of expert assistance. Furthermore, the effort aims to enhance the toolbox of such experts by bringing them together with those who develop the tool environments of CDI applications. Teams of expert assistants have emerged as essential components of many projects conducting CDI research. The proposed work will enable the exchange of best practices and open problems, and the discussion of supporting tool environments among all these teams.

7. **Yifan Wang, PhD Student, Chemical & Biomolecular Engineering**
Title: Statistical-learning-assisted Structure Optimization for Catalysts in Subnanometer Regime
Abstract: Supported metal nanoparticles are widely applied as environmental catalysts. The catalyst activity is strongly affected by the nanoparticle size. Recently, single atom catalysts are being explored as effective and selective catalysts for several chemistries. However, predicting catalyst stability and activity has remained elusive due to the inability to observe the dynamics of the catalysts operando. Density functional theory (DFT) is a powerful tool to predict the energetics of metal-support and adsorbate-metal interactions. However, the computational time needed to describe numerous clusters and sites and the long-time scales for sintering make direct first-principles calculations impractical. To address this multiscale problem, here we present a statistical-learning framework to enable atomistic level structure determination for single atom and subnanometer cluster catalysts. The methodology is applicable to any metal/support system.

8. **Akshay Bhosale, PhD Student, Electrical & Computer Engineering**
Title: Compile Time Parallelization of Subscripted Subscript Patterns
Abstract: An increasing number of scientific applications are making use of irregular data access patterns. An important class of such patterns involve subscripted subscripts, where in an array value appears in the index expression of another array. Even though the information required to parallelize loops with such patterns is available in the class of programs that we analyze, present compiler techniques fall short of analyzing that information. In this paper we present a study of subscripted subscripts, the properties that define the subscript arrays, and an algorithm based on symbolic range aggregation, that will help prove the presence of some of the properties of the subscript array in the program. We show that, in an important class of programs, the algorithm can boost the performance from essentially sequential execution to close to fully parallel.

9. **Priyanka Mondal, PhD Student, Physics & Astronomy**
Title: Quantum spin transfer torque induced nonclassical magnetization dynamics and entanglement
Abstract: The classical standard spin transfer torque (STT) explains the spin dynamics driven by quantum electrons interacting with classical spin magnetization when injected spins are noncollinear to the ferromagnetic layer. However, recent experiments observing magnetization dynamics in spin valves at cryogenic temperatures, even when electron spin is collinear to magnetization, point at overlooked quantum effects in STT that can lead to highly nonclassical magnetization states. Using quantum many-body treatment, where an electron injected as a spin-polarized wave packet interacts with local spins comprising the anisotropic quantum Heisenberg ferromagnetic chain, we define quantum STT as any time evolution of local spins due to an initial many-body quantum state not being an eigenstate of an electron+local-spins composite system and that the spin dynamics driven by quantum STT results in magnetization entanglement between electron and local spin subsystems.

10. **Wei Li, PhD Student, Materials Science**
Title: Predicting photovoltaic materials band gap using high throughput and machine learning
Abstract: A machine learning model is developed that can accurately predict the HSE06 band gap correction to the PBEsol level of pervoskite oxide family compounds, by quantitatively analyzing the valence band maximum (VBM) and conduction band minimum (CBM) offset. Not only does this resulting tool provide the ability to accurately predict the HSE06 band gap based on the PBEsol results but also the speed of the prediction based only on the cubic structure will make this a great resource to screen functional pervoskite materials. And this design represents a powerful database and tool for mapping the vast materials landscape and accelerating discovery.

11. **Carolina Perez Segura, PhD Student, Chemistry & Biochemistry**
Title: Revealing the mechanisms of biological machines through the computational microscope
Abstract: At the technological interface of chemistry, physics, biology, and computing, there exists the only scientific instrument capable of characterizing the conformational dynamics of macromolecules, and the functional properties they give rise to, at atomistic resolution. The computational microscope, realized through all-atom molecular dynamics (MD) simulations, has emerged as a powerful tool to investigate the complex cellular machinery that supports life, as well as the pathogenic systems that threaten it. The Hadden Lab leverages the computational microscope to study biological machines, including viruses and molecular motors. We dissect machines to reveal the inner-workings of key components and how the components cooperate to drive overall action. By elucidating the mechanisms by which machines function, we aim to identify strategies to inhibit undesired functions (e.g., in viral infection) and prevent dysfunction (e.g., in essential cellular processes), ultimately to treat disease.

12. **Ali Salimi Tarazouj, PhD Student, Civil & Environmental Engineering**
Title: Two Phase Modeling of Sand Ripple Bed Under Oscillatory Flow using sedFoam
Abstract: Oscillatory motions of surface wave near the seabed can generate various bed forms, mostly notably known as orbital ripples and an-orbital ripples. The importance of ripples arises from their effect on the seabed roughness and net onshore/offshore sediment transport. The bed roughness directly affects the near-bed boundary layer hydrodynamics, which in turn controls wave attenuation and sediment transport in coastal areas. Understanding the geometry and migration of such ripples driven by a range of wave forcing is the main objective of the present study. As a first step, a Eulerian two-phase model, SedFOAM, is utilized to study ripple dynamics under oscillatory flow. SedFOAM is able to resolve the full profile of sediment transport without the priori assumption of bedload and suspended load. Model verification shows SedFOAM's ability to model characteristics and dynamics of orbital ripple bed both for flow field and sediment concentration.

13. **Yashar Rafati, PhD Student, Civil & Environmental Engineering**
Title: Euler-Lagrange modeling of graded sand transport by nearshore waves
Abstract: Wave nonlinearity plays a key role in driving coastal sediment transport. However, it remains unclear how sediment size gradation can affect wave-driven onshore/offshore transport in the coastal zones. Field and laboratory observations suggest wave-driven sediment transport generates inverse grading, where the coarser grains become exposed in the surface layer and finer grains become armored just below the surface. Using a two-phase Eulerian-Lagrangian model, where the fluid phase is based on SedFoam developed in the OpenFOAM framework and the particle phase is simulated with the open-source discrete element method solver LIGGGHTS, we quantified the inverse grading effects of wave-driven onshore transport during sheet flow conditions for medium and coarse grains. Numerical simulations presented in this study were carried out using the Caviness cluster at University of Delaware.

14. **Subash Adhikari, PhD Student, Physics & Astronomy**
Title: Magnetic Reconnection from a Turbulence Perspective
Abstract: The inter-relationship between turbulence and magnetic reconnection has been the focus of increasing scrutiny, with significant research on the statistics of reconnection as an element of turbulence, the generation of turbulence during reconnection, and the effect of reconnection on the cascade of energy. However, a basic question is largely unexplored: What are the energy spectral features of laminar magnetic reconnection, independent of self-generated turbulence? In this study, we perform a 2.5D kinetic particle in cell (PIC) simulation of magnetic reconnection with periodic boundary conditions to explore the spectral properties of the system. Most notably, we find that quasi-steady and laminar magnetic reconnection exhibits a power law consistent with the Kolmogorov law of turbulence. In the process, reconnection isotropizes the distribution of magnetic energy in k-space.

15. **Cesar Claros, Master Student, Electrical & Computer Engineering**
Title: Active Error Analysis and Mitigation for Synergistic Machine Learning Infrastructure
Abstract: Harnessing machine learning models has been shown to accelerate data-intensive science. However, these models are often brittle and can benefit from human interventions to correct for spurious cases. To maximize the benefit, intelligent infrastructure is needed to combine human expertise and machine learning models. In particular, we investigate how already trained machine learning models can be boosted in deployment by analyzing their performance and using the discrepancy between model predictions in the training and validation sets. Even if the performance meets specifications, there may be cases of systematic errors caused by model underfitting or poor model design. We propose to perform error analysis of the training and validation set during deployment to alert human experts when instances similar to previous systematic errors arise. Triggering human vigilance during deployment will improve the synergistic operation of the machine and the expert.

16. **Marko Petrovic, Postdoc/Researcher, Mathematical Sciences and Physics & Astronomy**
    Title: Annihilation of magnetic domain walls induced spin wave burst intertwined with spin pumping
    Abstract: Motivated by recent experiments [S. Woo et al., Nat. Phys. 13, 448 (2017)] on magnetic-field-driven annihilation of two magnetic domain walls (DWs) and the ensuing burst of spin waves (SWs) of very short wavelengths, we predict that the same process is accompanied also by pumping of electronic spin current whose time-dependence encodes information about the moment of the burst. The pumped currents carry spin-polarized electrons which, in turn, exert spin-transfer torque on localized magnetic moments, so that the whole process cannot be captured by standard classical micromagnetic simulations. Instead, we employ recently developed [MD Petrovic et al., Phys. Rev. Applied 10, 054038 (2018)] quantum-transport/classical-micromagnetics approach, combining time-dependent electronic nonequilibrium Green functions with the Landau-Lifshitz-Gilbert equation.
17. **Carlos H. Mendoza Cardenas, PhD Student, Electrical & Computer Engineering**
    Title: X Learning waveforms of neural oscillations in pre-ictal EEG recordings
    Abstract: The waveform shape of the neural oscillations in electroencephalographic (EEG) signals provides valuable information about the physiology and the disease and behavior states of the brain. In this work, we present a data-driven methodology to learn the waveforms of neural oscillations from pre-ictal segments in the Temple University Hospital EEG Seizure Corpus (TUSZ). The waveforms are learned using a shift-invariant k-means algorithm and inferred by dividing each pre-ictal segment into a set of overlapping windows and assigning each window to a cluster whose centroid is one of the learned waveforms. In this way, we can localize waveform occurrence, identify neural oscillation sources and their characteristics, and use waveform occurrence as a feature to distinguish different conditions. The TUSZ data set, with multi-channel EEG signals, long-duration recordings, and hundreds of patients, makes our problem of learning and inference a challenging large-scale machine learning problem.
18. **Mauricio Ferrato, PhD Student, Computer & Information Sciences**
    Title: Applying Machine Learning Techniques on Acute Lymphoblastic Leukemia
    Abstract: Our project focuses on using both clinical and genetic data of Acute Lymphoblastic Leukemia (ALL) patients to determine molecular changes that drive childhood cancers. The goal is to find what mix of feature selection techniques and classification models provide the most accurate and generalizable results. Our project started with Decision Trees and progressed to more complex mathematical predictors like Support Vector Machines (SVMs). We obtained an prediction accuracy of 65% with Decision Trees, a 75% with both Random Forests and XGBoost, and 88% with SVMs. Some further questions we have include: Will the compare and contrast of different feature selection algorithms applied to our genetic data report a similar group of genes? With 200 cases, what is the optimal number of features to increase accuracy but not overfit. More research is also needed to better understand the sensitivity and specificity of the models rather than just their prediction accuracy.
19. **Brad Altmiller, Undergraduate Student, Computer & Information Sciences**
    Title: Applying Machine Learning Techniques on Acute Lymphoblastic Leukemia
    Abstract: Our project focuses on using both clinical and genetic data of Acute Lymphoblastic Leukemia (ALL) patients to determine molecular changes that drive childhood cancers. The goal is to find what mix of feature selection techniques and classification models provide the most accurate and generalizable results. Our project started with Decision Trees and progressed to more complex mathematical predictors like Support Vector Machines (SVMs). We obtained an prediction accuracy of 65% with Decision Trees, a 75% with both Random Forests and XGBoost, and 88% with SVMs. Some further questions we have include: Will the compare and contrast of different feature selection algorithms applied to our genetic data report a similar group of genes? With 200 cases, what is the optimal number of features to increase accuracy but not overfit. More research is also needed to better understand the sensitivity and specificity of the models rather than just their prediction accuracy.
20. **Farid Qamar, Master Student, Graduate Interdisciplinary Program and Biden School of Public Policy & Administration**
    Title: Segmentation of Ground-based Hyperspectral Images to Identify Vegetation in Urban Environments
    Abstract: There is a growing need to develop a better understanding of plant health in urban environments including the impact of air quality on the health of public green spaces. This need has led to the use of Hyperspectral Imaging (HSI) as a potential rapid and non-destructive sensor to assess plant health remotely and in an automated fashion. However, due to the geometric complexity and strong diversity of materials in urban areas, separating the vegetation from other materials using remote, ground-based imagery remains a challenge. Here we examine and compare various machine learning methods performing image segmentation on ground-based hyperspectral imagery in order to distinguish the materials and make use of their spectra for further analysis.

21. **Prasad Dhurjati, Faculty, Chemical & Biomolecular Engineering**
Title: An agent-based model to investigate microbial initiation of Alzheimer's via the olfactory system
Abstract: Alzheimer's disease (AD) is a degenerative brain disease. A novel agent-based modelling framework was developed in NetLogo 3D to provide fundamental insights into the potential mechanisms by which a microbe (eg. Chlamydia pneumoniae) may play a role in late-onset AD. The objective of our initial model is to simulate one possible spatial and temporal pathway of bacterial propagation via the olfactory system, which may then lead to AD symptoms. The model maps the bacteria infecting cells from the nasal cavity and the olfactory epithelium, through the olfactory bulb and into the olfactory cortex and hippocampus regions of the brain. The goal of our initial model is to guide further hypothesis refinement and experimental testing to better understand the dynamic system interactions present in the etiology and pathogenesis of AD.
22. **Abdul Qadir, Master Student, Data Science**
Title: Forest Degradation in South Asia: A Vegetation Trend Analysis in the Context of Climate Variability
Abstract: Sustainable Development Goal (SDG) indicator 15.1.1 proposes to quantify "Forest area as a proportion of total land area" in order to achieve SDG target 15.1. While area under forest cover can provide useful information regarding discrete changes in forest cover, it does not provide any insight on subtle changes within the broad vegetation class, e.g. forest degradation. Continental or national-level studies, mostly utilizing coarse-scale satellite data, are likely to fail in capturing these changes due to the fine spatial and long temporal characteristics of forest degradation. Using a multi-scale, satellite-based monitoring approach, our goal is to provide an easy-to-implement reporting framework for South Asian forest ecosystems.
23. **Vishruta Yawatkar, Master Student, Geography and Mathematical Sciences**
Title: Machine learning and cloud computing for environmental research
Abstract: The EArth observation for Sustainable Ecosystem and Livelihood" (EASEL) research group at the University of Delaware uses multi-scalar satellite and aerial images and Geographic Information Systems (GIS) tools to investigate a wide range of environmental issues such as crop responses to weather variability, smallholder farming, vegetation trend, and coastal sustainability. We conduct our research in different parts of the world including Asia, Africa and the USA. Here we present some of our ongoing work on croplands and farmlands, and how machine-learning can be used to inform environmental policies.
24. **Rayanne Luke, PhD Student, Mathematical Sciences**
Title: Parameter Estimation for Evaporative Tear Film Breakup
Abstract: The tear film (TF) protects the ocular surface and promotes clear vision. Tear film breakup (TBU) occurs when a dry spot forms on the eye. Elevated osmolarity in TBU is likely a key factor in developing dry eye syndrome (DES). We infer TBU events from video recordings of the TF imaged using fluorescein (FL) dye, which glows green under blue light; physical and chemical quantities of interest cannot be measured directly at this time. We build and analyze nonlinear PDE models for TF height, osmolarity and FL concentration for evaporation-driven TBU. Theoretical FL intensity is fit via nonlinear least squares optimization to circular and linear TBU instances from experimental imaging data gathered from normal subjects' TFs to estimate parameters such as TF thinning rates. Our estimates for variables that cannot be measured in vivo during TBU fall within accepted experimental ranges, help clarify the mechanisms for TBU, and help medical professionals better understand TF function and DES.
25. **Yulin Zhang, PhD Student, Electrical & Computer Engineering**
Title: Fast Convolutional Neural Networks with Fine-Grained FFTs
Abstract: The Convolutional Neural Networks (CNNs) architecture is one of the most widely used deep learning tools. The execution time of CNNs is dominated by the convolution steps. Most CNNs implementations adopt a simple yet efficient im2col (image to column) +GEMM approach to implement convolution. This approach lowers the convolution into matrix multiplication that can be easily parallelized with highly efficient BLAS libraries. In this work, we observe significant but intricately patterned data redundancy in this matrix representation of convolution, and we reveal a new data pattern to more mathematically concisely describe the matrix representation for convolution. Based on this redundancy-minimized matrix representation, we implement a FFT-based convolution with finer FFT granularity. It achieves on average 23% and maximum 50% speedup on the ILSVRC2017 benchmark over the regular FFT convolution from NVIDIA's cuDNN library.
26. **Alison Leonard, Postdoc/Researcher, Physics & Astronomy**
Title: Preferential Lipid Solvation and Acyl Chain Entropy in the Vicinity of a G-protein Coupled Receptor
Abstract: The local lipid environment of the A2A adenosine receptor was recently reported to be enriched in unsaturated phosphatidyl choline (PC) during 30 μsec of all-atom simulation. However, even with tens of μsec of trajectory, it is challenging to converge the first shell lipid composition of an integral membrane protein. This motivated the construction of a Markov State Model (MSM) for local lipid composition. The resulting equilibrium distribution is used to verify the lipid composition of the receptor's first solvation shell. The thermodynamics of lipid-protein interactions are also assessed. The entropy and enthalpy associated with bringing a lipid from bulk to the first solvation shell of the receptor are obtained for only the acyl chains. The analysis yields the entropic contribution to preferential solvation by unsaturated lipids, at the same time providing a new way to conceptualize the link between lipid structure and thermodynamics.

27. Shalini Sundar, Undergraduate Student, Chemical & Biomolecular Engineering

Title: An agent-based model to investigate microbial initiation of Alzheimer's via the olfactory system

Abstract: A novel agent-based modelling framework was developed in NetLogo 3D to provide a potential mechanism by which Chlamydia pneumoniae may infect the olfactory system, possibly leading to AD symptoms. Simulated randomized infection led to the formation of beta-amyloid (Aβ) plaque and neurofibrillary (NF) tangles. Our initial simulations demonstrated that breathing in C. pneumoniae can result in infection propagation and significant buildup of Aβ plaque and NF tangles in the olfactory cortex and hippocampus. Our model also indicated how mucosal and neural immunity can play a significant role. The modelling framework provides an organized visual representation of how AD progression may occur via the olfactory system to better understand disease pathogenesis. The model confirms current research but can be easily adjusted to match future evidence. The goal of our initial model is to ultimately guide hypothesis refinement and experimental testing to better understand the pathogenesis of AD.

28. Emma Stell, Master Student, Geography and Plant & Soil Sciences

Title: Mapping global Rs at fine-scale resolution and optimizing future sample site locations

Abstract: Soil respiration (Rs) is commonly known as the efflux of carbon dioxide (CO2) from the soils to the atmosphere. It composes the second largest flux in the carbon cycle, after gross primary productivity (GPP) (Raich & Schlesinger 1992). However, global estimates of Rs and other components of the terrestrial carbon cycle are poorly constrained, leading to uncertainties in global climate projections (Friedlingstein et al. 2006, 2014). My research aims to not only map and quantify global Rs at fine-scale resolution, but also strives to identify optimal locations for future Rs sampling regimes.

29. Utkarsh Bajpai, PhD Student, Physics & Astronomy

Title: Spatio-temporal dynamics of shift current quantum pumping by femtosecond light pulse

Abstract: Using the formalism of time dependent non-equilibrium Green's functions (TD-NEGF) we investigate the transport properties of Rice-Mele model when irradiated with ultrashort light pulse in the absence of any bias voltage. We then illuminate the Rice-Mele chain with an ultrashort light pulse whose central frequency is equal to the energy gap of Rice-Mele model. While the time-dependent current oscillates following the pulse, there is an average current and a total charge is injected into the leads. Such a shift current exists even when Rice-Mele chain is short enough that evanescent wave functions from the leads fill its gap and make the whole device metallic because the whole process can be viewed as non-adiabatic charge pumping where the key requirement is broken left-right spatial symmetry. Finally, we quantify the velocity of propagation of shift current generated in the center of clear or disordered Rice-Mele chain and compare with recent experiments.

30. Hang Chen, Master Student, Computer & Information Sciences

Title: HPC Usage in Transfer Learning for Analyzing fMRI Images

Abstract: We will present our work done in evaluating transfer learning usage on a CNN-LSTM neural network while analyzing fMRI brain images. It serves as an extension to our previous work - using fMRI images to predict the BMI value of the patients. Since the fMRI dataset with BMI labels is so scarce on the internet, we leveraged transfer learning to borrow useful features from other fMRI datasets and had increased the accuracy of our proposed CNN-LSTM model while testing on our dataset with the BMI values. During the whole process, we used Google Cloud to generate augmented fMRI images and extensively used the HPC resources from PSC(Pittsburgh Supercomputing Center) to train our neural network and do transfer learning.

31. Adam Wallace, Faculty, Geological Sciences

Title: Forward Flux Sampling Simulations of Water Exchange Reactions on Pico- to Micro-Second Timescales

Abstract: Molecular dynamics simulations are invaluable tools in the study of the mineral-water interface. Although standard simulation methods routinely address timescales that are sufficient to characterize relatively small-amplitude fluctuations in the structure of the interface, more infrequent large-scale fluctuations that drive events such as ion attachment and detachment typically remain under-sampled. Yet, it is these "rare events" that beg our attention, because they are ultimately responsible for the chemical behavior we observe at much longer time- and length-scales. At present, our rare event simulation capabilities are limited by our ability to characterize complex multi-dimensional energy landscapes and to identify transition pathways between important stable and metastable states. We detail our efforts to develop a novel Forward Flux Sampling method for rare events. We demonstrate the validity of our approach over at least 6 orders of magnitude, well into the microsecond regime.

32. **Swapnil Baral, PhD Student, Physics & Astronomy**
    Title: Ultrafast formation of charge-transfer state reveals unique aspects of Prodan environment
    Abstract: Lipiophilic dyes such as Laurdan and Prodan are widely used in membrane biology due to a strong bathochromic shift in emission that reports structural parameters of the membrane such as area per molecule. Disentangling the factors which control the spectral shift is complicated by the stabilization of a charge-transfer-like excitation of the dye in polar environments. Predicting the emission therefore requires MM/QM approach to model both the relaxation of the environment and the corresponding evolution of the excited state.The QM steps are computed using many-body Green's function in the GW approximation and the Bethe-Salpeter Equation with the environment modeled as fixed point charges, sampled in the MD simulation steps. Comparison to ultrafast time resolved transient absorption measurements demonstrates that the iterative MM/QM approach agrees quantitatively with both the polarity dependent shift in emission and the timescale over which the charge transfer state is stabilized.
33. **Fateme Hosseini, PhD Student, Electrical & Computer Engineering**
    Title: X Adaptively Boosting Resilience Efficiency in the Face of Frequent, Clustered, and Diverse Faults
    Abstract: Technology scaling has equipped modern systems with higher performance and lower power consumption. However, delivering such features is challenged by the increased susceptibility of silicon devices to faults. In this poster, we propose an integrated fault-tolerant framework to address the reliability challenges caused by high fault rates and diverse fault behaviors and thus enable future computer systems to boost resilience efficiency. The key insights are to incorporate fine-grained adaptivity into the system, and to couple statically extracted application information with runtime optimizations to guide adaptation decisions. The proposed framework integrates three tightly-connected techniques: (1) adaptive checkpointing adaptive detection enabling rapidly adjusting fault detection and execution checkpointing granularity to match system reliability levels; (3) adaptive error recovery, capable of performing re-execution in a way that minimizes the recovery overhead.
34. **John Melkumov, PhD Student, Chemistry & Biochemistry**
    Title: Force Fields Automatically Constructed From Ab Initio Derived Interaction Potentials For Fragments
    Abstract: Molecular dynamics (MD) simulation is a powerful tool used to probe the physical properties of molecular systems. However, the validity of these simulations is ultimately dependent on the use of an accurate force field. The current state-of-the-art force fields (FFs) that are employed in MD simulations (e.g., AMBER, OPLS) are empirical ones with parameters fitted to reproduce a combination of available experimental data. In many cases, molecules are encountered for which FF parameters are missing. For such molecules, first-principles FFs may be derived by fitting an interatomic potential to quantum mechanical (QM) data. Since FF parameterization involves performing QM calculations, this task scales prohibitively with system size. Here we present results from an automatic, fragment-based, application of symmetry-adapted perturbation theory (SAPT) to derive highly accurate molecular FFs extensible to large systems such as proteins.
35. **Patrick Cronin, PhD Student, Electrical & Computer Engineering**
    Title: Low-Cost Sensor Enabled Explosive Detection to Protect High Density Environments
    Abstract: Large, open, public events, such as marathons and festivals, have always presented a unique safety challenge. These sprawling events can fill large open fields, take up entire city blocks, or stretch for many miles. Furthermore, these events are capable of drawing tens to hundreds of thousands of spectators and in some cases have completely open admission. At these events, it is not possible to subject every event-goer to a security screening, creating potential security threats. To this end, we propose a crowd-based explosive detection system which utilizes thousands of low-cost ChemFET sensors which are distributed to attendees. As the sensors only offer limited accuracy, we further propose a server based decision framework which utilizes a two-level feedback loop between the sensors and the server and explores spatial and temporal locality of the collected data to overcome the inherent low-accuracy of individual sensors.
36. **Chaoyi Xu, PhD Student, Chemistry & Biochemistry**
    Title: Molecular mechanisms of nucleotide translocation through HIV-1 CA hexamer
    Abstract: Reverse transcription (RT) in HIV-1 is a post entry event during which dNTPs fuel the cDNA synthesis. The latter requires the availability of dNTPs inside the virus core. It was postulated that the central cavity located within capsid protein (CA) hexamers could serve as a dNTP channel. However, it is unclear how nucleotides translocate through this channel. In addition, small molecules like IP6 and mellitic acid (MA) bind to CA and influence RT. Here, we designed and performed a series of molecular dynamics simulations to determine the molecular mechanism of the nucleotide import through the HIV-1 capsid. Results from our simulations support that a cooperative model where multiple nucleotides interact inside the cavity results in energetically favorable conditions for the dNTP import. In addition, the effect of IP6 and MA binding on dNTP translocation was determined. Altogether, our results reveal an atomistic description of the dNTP-import and regulatory activity of CA-hexamers.

37. Pooja Bhalode, PhD Student, Chemical & Biomolecular Engineering

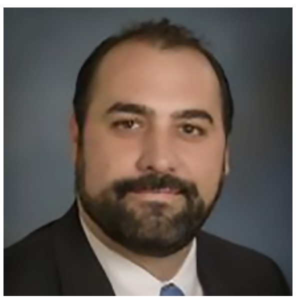
Title: Data-intensive Applications for Predictive Modeling in Continuous Pharmaceutical Manufacturing

Abstract: Continuous manufacturing (CM) has been a critical area in the pharmaceutical industry for the past decade aiming to develop more efficient and reliable solid-dosage based manufacturing processes. Recent emphasis on Industry 4.0 and improvement of process modeling in CM have led to the development of data intensive approaches using high-performance computing (HPC) resources. Large volumes of data generated from the plant are integrated into an Industry 4.0 framework, which is exploited using machine learning techniques to improve predictive modeling capabilities. Particle-scale approach can further be applied to improve CM models using discrete element modeling (DEM). Integrating these high-fidelity simulations within the process models allows for replication of pharmaceutical powder behavior within predictive models. Here, Industry 4.0 framework in a CM plant and DEM techniques for a feeder unit are demonstrated, with potential application of DARWIN HPC systems.

DARWIN COMPUTING SYMPOSIUM PLANNING COMMITTEE

Rudolf Eigenmann, PI

Arthi Jayaraman, co-PI



Benjamin Bagozzi, co-PI

William Totten, co-PI

Cathy Wu, co-PI

Andrea Trungold

THANK YOU TO OUR INDUSTRY SPONSORS

Gold Sponsor:

atípa technologies

Silver Sponsor:

DELL

Bronze Sponsor:

CompassRed

VISIT THE DATA SCIENCE INSTITUTE AT:

https://dsi.udel.edu/

JOIN US ON TWITTER

twitter @dsiudel

hashtag: #DARWINcomputing2020