Dear Colleagues,

Greetings from the Rutgers Discovery Informatics Institute (RDI²). I am delighted to share our many exciting activities, events and achievements with you in this newsletter.

RDI² continues to drive computational and data-enabled research across a broad range of disciplines, providing leadership in Big Data and Advanced Cyberinfrastructure at the university, state, and national levels. We provide expertise, research partnerships, services, and access to resources to the broader New Jersey academic community, with the overarching goal of driving discovery and innovation.

RDI²’s activities span foundational and translational research, advanced cyberinfrastructure services, and education and outreach. Education and community outreach are central components of RDI²’s mission. Our education and training programs aim to foster the next generation of data science researchers and professionals through classroom teaching and laboratory training. This Spring 2019 newsletter highlights our education and outreach activities during the year. This includes welcoming the 2019-2020 student awardees of the RDI² Fellowship for Excellence in Computation and Data Science, presenting the achievement of our students and staff, and summarizing our many events and activities.

The rest of the year promises to be exciting and eventful, as we at RDI² continue to advance research, discovery and innovation, and transform society through computation and data. I hope you will join us in our journey.

Sincerely,

Manish Parashar

Founding Director of RDI²
Distinguished Professor of Computer Science
RESEARCH WITH IMPACT

RDi² dynamically integrates research, education, and advanced technologies to broaden academic industry access to state-of-the-art computing and data analytics capabilities. RDi² researchers are contributing to large-scale and streaming data analytics, edge computing and IoT, and computational and data-driven science and engineering, and advanced cyberinfrastructure more broadly.

RDi² has deployed a comprehensive ecosystem at Rutgers, which includes research and production facilities, high performance computing, cloud and data infrastructure, research instruments, and experimental platforms. RDi²’s research cyberinfrastructure is open to all faculty and students across Rutgers, as well as academia and industry throughout the state, and is part of the XSEDE national cyberinfrastructure.

New Jersey’s most powerful supercomputer, Caliburn, is delivering hundreds of millions of hours of massively parallel computing to researchers across New Jersey and enabling cutting-edge research and innovations in a wide range of disciplines that would not be possible otherwise. Recent innovations have added hardware and software support for machine learning and data analytics as well as seamless Cloud integration.

We are ensuring our continued growth and excellence by providing the best facilities we can to educate our students and conduct world-class research.

NJBDA RESEARCH COLLABORATION SYMPOSIUM

The New Jersey Big Data Alliance (NJBDA) Research Collaboration Forum, hosted by RDi², convened close to 60 New Jersey researchers representing 18 institutions including industry, academia, and government. The event highlighted the NJBDA Research Committee mission, NJ Research collaboration tools, and some specific data sets followed by individual group discussion focused on NJ research collaboration opportunities.

One of the outcomes of the event was a multi-institutional research collaboration on the NSF Data Science Corp project proposal led by Rowan University.

ABOUT NJBDA—Launched by RDi², the New Jersey Big Data Alliance (NJBDA) was established in 2013 by Rutgers University along with 8 higher education partners to catalyze collaboration among New Jersey academia, industry and government in building advanced computation and data analytics capabilities and expertise. NJBDA currently has 16 higher education members, six government partners and several industry members.

CC* INTEGRATION DYNAMO: DELIVERING A DYNAMIC NETWORK-CENTRIC PLATFORM FOR DATA-DRIVEN SCIENCE

In collaboration with the University of North Carolina at Chapel Hill, University of Southern California, and University of Massachusetts at Amherst, RDi² research will support programmable, on-demand access to high-bandwidth, configurable network paths from community data repositories to national CI facilities. This will help satisfy data, computational and storage requirements of CASA and OOI workflows.

RDi² researchers Ivan Rodero and J. J. Villalobos will work to enable researchers to test new algorithms and models in real time with live streaming data. These efforts, referred to as DyNamo, will allow atmospheric scientists and hydrologist to improve short and long-term weather forecasts and aid the oceanographic community. DyNamo’s workflows will allow for adaptability by other CI facilities, as well as management systems and middleware.

To learn more, visit bit.ly/2u2KOrt
RUTGERS DAY: DIVE INTO BIG DATA AND EXPLORE WITH RDI²

For this year’s Rutgers Day, attendees at the RDI² display learned about Big Data capabilities and research. Among the highlights were an autonomous Open Source robot platform, as well as a raspberry pi raffle and a bean bag toss game.

2019-2020 UPCOMING EVENTS

BIG DATA WORKSHOP FOR UNDERGRADUATE STUDENTS
“Knowledge Discovery and Data-Driven Decisions”
When: June 4, 2019
Where: Rutgers University Inn

HIGH SCHOOL FUTURE SCHOLARS PROGRAM
When: July 29 to August 2, 2019
Where: Busch Campus, CoRE Building

RDI² ANNUAL OPEN HOUSE
Join us for a showcase of research enabled by RDI².
When: October 25, 2019
Where: Busch Campus

RDI² DISTINGUISHED SEMINAR

The RDI² distinguished seminar series regularly hosts leading researchers from academia, governments and industry. This spring, Margaret Martonosi, PhD. Director of the Keller Center for Innovation in Engineering Education, the Hugh Trumbull Adams ’35 Professor of Computer Science at Princeton University, presented “What is the Role of Architecture and Software Researchers in Making Quantum Computing Practical?”

RDI² ANNUAL OPEN HOUSE

RDI² opened its doors to all during its annual Open House event during the Fall semester in order to share RDI² accomplishments and research over the past year. Sanja Padhi, PhD., Principal, Amazon Web Services’ Global Scientific Computing, delivered the keynote, “Predictive Analytics using Amazon Web Services.”
Throughout the academic year, educational programs through RDI² range from the Big Data workshop for High School students, to undergraduate Career Panels in Data Science, to Data Management seminars for graduate students and researchers.

In addition to hosting educational programs for students at all levels, including K-12 outreach programs, RDI² offers research internships for high-school and undergraduate students, summer workshops, as well as a fellowship program for graduate students.

**UPCOMING EDUCATIONAL MODULES FOR 2019-2020**
- Big Data Tools and Techniques
- Scientific Computing and HPC for Graduate Students
- Scientific Computing and HPC for Undergraduate Students
- Streaming Data Analytics
- Data Stewardship and Life Cycle Management

**UNDERGRADUATE WORKSHOP:**
**KNOWLEDGE DISCOVERY AND DATA-DRIVEN DECISIONS**

RDI², in collaboration with Rutgers Center for Critical Intelligence, presents workshop on Knowledge Discovery and Data Driven Decisions as part of the Rutgers IC CAE Certificate in Intelligence and National Security-Critical Technology Studies, June 2019. This workshop provides students with an introduction to the skills and knowledge to apply data science and technology skills in a national security context. The program is open to students from the Rutgers IC CAE Consortium (Rutgers University, Borough of Manhattan Community College, City College of New York, an New Jersey City University).

**SUMMER 2019 HIGH SCHOOL FUTURE SCHOLARS PROGRAM**

In partnership with Library Services and Department of English Writing Program, RDI² will participate in the Rutgers Future Scholars program by delivering a Big Data Workshop during the summer of 2019.

Through the Future Scholars Program, 11th grade students will come to Rutgers for an intensive writing course followed by a one-week internship experience designed to inspire students, boost confidence, and help them explore careers and majors. The internship will take place from July 29 to August 2.

RDI² is pleased to contribute to this important program, which provides academic and social support for students living in New Jersey’s most impoverished areas with the goal of increasing high school and college retention and graduation rates.

**RDI² HOSTS CAMDEN COUNTY TECHNICAL SCHOOL STUDENTS:**
**DIVE INTO THE WORLD OF BIG DATA WORKSHOP**

Earlier this year, Camden County Technical School students were provided the opportunity to tour of state-of-the-art facilities at RDI² and learn about fundamentals of data science and research projects at the institute, including the Virtual Data Collaboratory and the Ocean Observatory Initiative. Students also had the opportunity to experiment with live oceanographic data, visualize the results of simple data transformations, and develop a web-based dashboard that continuously displayed data streaming real-time from under water sensors on the Pacific Ocean seafloor.

To schedule a tour or program for your class or group, contact info@rdi2.rutgers.edu
In collaboration with Rutgers Institute for Women’s Leadership, Johnson & Johnson, as well as several other Rutgers University organizations, RDI\textsuperscript{2} helped to develop and deliver the RU Sparks WiSTEM2D workshop for high school girls. In addition to exposing students to women leaders and innovators in STEM fields, STEM demos including VR/AR, Big Data research, and MedTech innovations, students worked with a mentor to create a program in their communities for raising awareness to STEM fields. Forough Ghahramani, RDI\textsuperscript{2} Associate Director highlighted career opportunities in computer science and data science and served as a mentor to students from Edison High School throughout the day to develop a STEM career awareness program for middle-school student.

### RDI\textsuperscript{2} 2019-2020 Fellowship for Excellence

The Rutgers Discovery Informatics Institute is pleased to announce the 2019-2020 student awardees of the RDI\textsuperscript{2} Fellowship for Excellence in Computation and Data Science.

**Aidan Zabolo**  
*Physics & Astronomy*  
*Advisor: Jedediah Pixley, PhD*  
*Research: Stimulating Quantum Circuits Using Extreme-Scale Computing*

**Humna Awan**  
*Physics & Astronomy*  
*Advisor: Eric Gawiser*  
*Research: Big Data in Astrophysics: Clustering Analysis of Partial Galaxies*

**James Kelley**  
*Computational & Integrative Biology*  
*Advisor: Andrey Grigoriev, PhD*  
*Research: Accelerated Analysis of Variants in Multiple Cancer Genomes*

**Joseph Lubin**  
*Chemistry & Chemical Biology/IQB*  
*Advisor: Sagar Khare, PhD*  
*Research: Computational Data-Driven Methodology for Rapid, Rational Design of Proteolytic Enzymes*

Through this award, the Rutgers Discovery Informatics Institute supports students working on multi-disciplinary collaborative computational and data-enabled research projects in science and engineering, with a specific research focus on Big Data and Extreme Scale computing. The 2019-2020 Fellows will be recognized during the Rutgers Discovery Informatics Institute Open House event on October 25, 2019.
GLOBALLY ENGAGED COMMUNITY

In January 2019, RDI² welcomed visiting scholar, Gustavo Portella, who is currently working on finishing his PhD studies in Computer Science at the University of Brasilia in Brazil. Mr. Portella’s research interest is resource availability and cost constraints in cloud computing, and is currently working with statistics and utility-based models to provide high availability and low cost alternatives for cloud computing users.

During his time at RDI², Gustavo delivered a colloquium on “Statistical and Utility-Based Analysis of Cloud Transient Pricing and Availability” to the Rutgers research community. Gustavo was able to share his research results and identify common areas of interest for future collaboration, including IoT and cloud infrastructure applied for science workflow use cases.

“It was an incredible experience for me! I would like to keep in touch with Rutgers University’s research group and, perhaps in the future, return for a longer period so that we can develop some cloud computing research together.”

—GUSTAVO PORTELLA

In March 2019, RDI² welcomed Kevin Fauvel as a visiting scholar. Kevin Fauvel has a master’s degree in computer science and is currently working on finishing his PhD in computer science at INRIA in France. Kevin works in collaboration with INRA, which leads projects of targeted research for a sustainable agriculture, a safeguarded environment and a healthy and high quality food. Kevin’s research interest is machine learning and data mining. He is currently working on multivariate time series classification.

During his time at RDI², Kevin was able to share his research results and contribute to an interdisciplinary project aiming to propose a distributed streaming-based warning system for tsunamis.

“...
"Extreme Scales, Big Data, and the Transformation of Science."
Manish Parashar, Keynote Speaker, IEEE ICPADS 2018, Sentosa, Singapore.

Transformative Tools for Advancing Collaborative Research in New Jersey"
Forough Ghahramani, EdgeCon2019.


“The Future of Big Data: Artificial Intelligence and Machine Learning”
Manish Parashar, Keynote Speaker, NJBDA Annual Symposium 2019.

“Stacker: An Autonomic Data Movement Engine for Extreme-scale Data Staging-based In-situ Workflows”

“An Edge-Based Framework for Enabling Data-Driven Pipelines for IoT Systems”
E. Gibert Renart, D. Balouek-Thomert, M. Parashar. PAISE2019 Workshop at the 33rd IPDPS, Rio de Janeiro, Brazil.

“Distributed Operator Placement for IoT Data Analytics Across Edge and Cloud Resources”

For a full list of publications, please visit rdi2.rutgers.edu/publications

JULIA SCHNEIDMAN
HIGH SCHOOL INTERN
SENIOR AT BERGEN COUNTY ACADEMICS

At RDI², Julia performs research on algorithm optimization for data communication through a long range protocol, which provides much longer range than competing technologies, improves receiver sensitivity, and has a low bit error rate from inexpensive chips. Julia has recently been working on adapting a generic algorithm to this problem, for potential use in educational programs.

Going forward, Julia hopes to major in Computer Science and Mathematics in college. Congratulations to Julia on her upcoming graduation from high school!

“My internship at RDI² has provided me with many new skills in computer science, data processing, and working with researchers, and I am very grateful for this unique opportunity.”

ALI REZA ZAMANI ZADEH NAJARI
PHD CANDIDATE

In his research, Ali has addressed the challenge of designing a framework that uses new edge-processing approaches together in an integrated manner by developing a mathematical scheduling model to support edge and in-transit data analysis in cloud federation. In addition, Ali has designed and developed a subscription-based data streaming framework to deploy and execute stream oriented workflows, and a runtime management layer that leverages edge and in-transit resources.

Upon graduation, he will be working at LinkedIn as a Systems and Infrastructure Engineer. We wish Ali Reza the best.

“RDI² helped me to find my path and become independent researcher by providing opportunity to work with insightful and knowledgeable researchers around the world.”