Dear Colleagues,

As we wrap up the 5th anniversary of RDI², I am pleased to share some highlights of our accomplishments from the past year, as we continue on our mission to accelerate discovery and drive innovation through advanced computing and data at Rutgers and beyond.

Education and community outreach are central to our mission at RDI² as we address the grand challenges in science, engineering and society on a global scale. Last September, RDI² led the organization of the NSF Large Facilities Cyberinfrastructure (CI) workshop bringing together experts to share insights on CI models, challenges, and best practices to increasing the efficiency and impact of shared-use facilities, which represent some of NSF’s largest investments in science and engineering. RDI² hosted Paul Messina, as the Fall Distinguished Seminar Speaker, Project Director for the U.S. Department of Energy Exascale Computing Project who talked about the “Challenges of Exascale Computing”. RDI² hosted the University of Derby (UK) to explore joint projects and collaboration opportunities.

RDI² had another active year of education and training activities. During National Engineers week, we launched its Diving into Big Data: Large Scale Computing workshop for high school students. We also welcomed undergraduate students into the newly launched Research Internship program that provides students with opportunities to conduct research in computational and data-enabled science and engineering. We launched a new initiative on 2018, the Data Science Practitioner Roundtable series, with the goal of introducing data science career paths and insights to graduate and undergraduate students by exposing them to industry professionals. We also hosted the Introduction to Data Management workshop for early career researchers.

RDI² has continued its research activities and leadership in across computational and data-enabled science engineering. This includes new and existing research projects, collaborations and grants, as well as research publications, keynotes, presentations and awards. RDI² also continues its leadership in research CI through its computational and data services to researchers in Rutgers and across NJ, through its roles as CI lead for the NSF Oceans Observatories Initiative (OOI) and through the NSF-funded Virtual Data Collaboratory (VDC) project.

On a personal note, I am excited to announce my recent appointment at the National Science Foundation as the Office Director for the Office of Advanced Cyberinfrastructure (OAC). I am excited to be part of the CI leadership at NSF at a time when computation and data are transforming science and society, and I look forward to bringing my experiences and insights back to RDI² and Rutgers.

Sincerely,

Manish Parashar

Founding Director of RDI²
Distinguished Professor of Computer Science
RDI² DIRECTOR MANISH PARASHAR APPOINTED TO NATIONAL SCIENCE FOUNDATION

RDI² Director Manish Parashar was appointed Office Director for the Office of Advanced Cyberinfrastructure at the National Science Foundation. In addition to his new appointment at NSF, Manish Parashar continues his position at Rutgers University as Distinguished Professor of Computer Science and the founding Director of the Rutgers Discovery Informatics Institute and the Applied Software Systems Laboratory.


ADVANCED COMPUTING & DATA CYBERINFRASTRUCTURE

This has been an exciting year for RDI², particularly in the area of Advanced Computing and Data Cyberinfrastructure (ACI). RDI² has continued delivering access to cutting-edge supercomputing resources and services to the Rutgers and New Jersey academic community. Active projects include research in a broad range of disciplines, including biochemical engineering, finance, and economics as well as physics and cosmology.

RDI² invites applications for the allocation of computing resources on Caliburn, New Jersey's largest academic supercomputing facility that provides high-performance computing capabilities to academic researchers across the state to accelerate research programs that use or develop highly scalable computing applications.

For more information about the call for proposals and information sessions, please visit http://rdi2.rutgers.edu/access-use.

RESEARCH INTERN:

SRUJANA SURE
B.S. Electrical and Computer Engineering

Srujana Sure is a senior at Rutgers University. During her time at RDI², Srujana researched microservices architectures to find patterns and to make it easier to analyze large data sets. This research work aimed at characterizing and describing advanced management features to face upcoming challenges of edge computing and complex software architectures to deal with online data.
RDI² welcomes the 2018-2019 student awardees of the RDI² Fellowship for Excellence in Computation and Data Science. 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. Each fellowship appointment is for one year and comes with $30,000 towards Graduate Assistant support, with potential for renewal.

**MALIHE ALIKHANI**  
*Computer Science—Natural Language Processing*  
**Advisor: Matthew Stone**  
**Research:** Deep data-driven modeling of multimodal communication

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

**JI HYE RYU**  
*Cognitive Psychology*  
**Advisor: Elizabeth Torres**  
**Research:** Tracking Brain-Body-Heart Pairwise Coupled Dynamics for Personalized Medicine

**COMMUNICATIONS INTERN:**  
**CHRISTA PRINCIPATO**  
*B.A. Journalism and Media Studies*  
Christa Principato is a senior in the School of Communication and Information at Rutgers University. At RDI², Christa was responsible for creating content to promote the institute through print and digital media. In addition to running the institute’s social media pages, Christa has helped RDI² transition to its new website. She was also responsible for designing and creating RDI²’s newsletters.
RDI² continues its quest for progressive technologies with new data-intensive initiatives. In 2016, we received a $4 million grant from the National Science Foundation to establish a regional data sharing network, the Virtual Data Collaboratory (VDC). The VDC is part of the Data Infrastructure Building Blocks (DIBBs) initiative, an NSF-funded project that brings together a deeply engaged interdisciplinary team of researchers and research & infrastructure organizations to build the next-generation data-centric cyberinfrastructure. The overarching goal is to promote collaboration and identify relationships among research products to facilitate deep and intuitive reuse of research data.

The participating organizations—led by Rutgers University and Pennsylvania State University—are working in partnership with KINBER, NJEdge, and the New Jersey Big Data Alliance (NJBDa) to transform shared data as a core modality for research and discovery.

Click here to learn more about VDC.

FURTHERING INNOVATIVE RESEARCH

PARTNERSHIP SPOTLIGHT: INTEL

Since 2017, RDI² has collaborated with Intel to develop cutting-edge data management systems for scientific workflows on HPC systems. As the processing power of leading supercomputers increases, data access costs become a major component of the overall time it takes to get results from scientific computing workflows.

The collaboration between Intel and RDI² focuses on synergistically integrating hardware and software systems produced by both organizations in order to provide more efficient data access technologies for HPC. We recently evaluated the performance of Intel Optane Drives for data-intensive HPC applications and focused on leveraging various machine learning techniques to enable autonomous data movement across such high performing hardware technologies.
**RDi² GLOBAL INVOLVEMENT**

RDi² engages in partnerships with international research groups and universities that have common expertise and interest. On April 30th RDi² hosted members from University of Derby, a university based in England, renowned for technology and innovation, to explore collaboration and partnership opportunities.

Also, Forough Ghahramani, Associate Director for RDi² Administration and Partnerships, was an invited to the First Future Engineering and Global Women’s Leadership Summit in Changsha, China. Forough was part of the delegation for Society of Women Engineers (SWE) commemorating a ground-breaking partnership between the Chinese Academy of Engineering (CAE) and SWE at Central South University (CSU). At the event she presented “Critical Success Factors for Building Inclusive Innovation Ecosystems.”

Additionally, RDi² is part of an international team with Inria Avalon called SUSTAM, Sustainable Ultra Scale Computing, Data and Energy Management. This long-term collaboration primarily focuses on Edge Computing and data-driven placement of operators for workflow applications, HPC and energy-efficient data management, scheduling and resource management using OOI data. RDi² Research Associate Daniel Balouek-Thomert has coordinated four visits since 2017.

**CONNECTING STAFF, STUDENTS & FACULTY CAMPUS-WIDE**

Complementing the existing Research Colloquia Series, aimed to facilitate discussion about current research projects, 2018 marked the start of the RDi² Distributed Systems Reading Group. In this group, students, postdocs, and staff meet regularly to read and discuss papers, share insight, and develop a critical understanding of important issues in the field. The reading group has focused on a mix of papers, studying both older papers that have demonstrated their value over time, and the latest papers that present new, cutting-edge research topics. Specific areas studied so far include associative memory storage systems and event-based execution systems, as well as distributed synchronization and GPU-based distributed memories.

**RD² TEAM MEMBER SPOTLIGHT**

**PAUL ARIAS**

Paul Arias is an Associate Research Scientist at RDi². He’s interested in developing novel information technology solutions that result in user satisfaction and adoption. He brings years of experience in combustion research and software development along with a personal interest in ensuring meaningful and useful data-driven story telling. His role at RDi² is to provide technical support for the user community and to provide policy advice for use of RDi² systems.

**WAYNE CHAN**

Wayne Chan is a Systems Administrator with RDi². Wayne supports the Ocean Observatories Initiative (OOI) & is looking forward to applying industry best practices at the university. Previously, Chan worked with JP Morgan Chase’s HPC group as well as the global 24/7 Unix break fix group. He is interested in all matters relating to Linux and open source software.
DISTINGUISHED SEMINAR: CHALLENGES OF EXASCALE COMPUTING

During the fall, Dr. Paul Messina, Former Director of the U.S. Exascale Computing Project, presented on “Challenges of Exascale Computing.” In this talk he discussed the motivations and challenges of achieving exascale computing from the perspective of the U.S. Department of Energy.

The RDI² Distinguished Seminar series regularly hosts leading researchers from academia, governments and industry and is designed to bring members of the Rutgers community together to explore Data Science topics.

CAREERS IN DATA SCIENCE SEMINAR

RDI², in collaboration with the Rutgers Department of Computer Science, launched the data science practitioner roundtable series in March to provide undergraduate and graduate students with the opportunity to learn about data science career paths and prepare for employment.

The event featured a roundtable discussion followed by networking with the data science industry experts.

OTHER RDI² EVENT HIGHLIGHTS

During the Fall, RDI² Director Manish Parashar organized the 2017 NSF Large Facilities Cyberinfrastructure (CI) Workshop, which convened national CI experts to address current and future CI needs including CI models, challenges, and best practices. Ivan Rodero, RDI² Associate Director for Technical Operations, shared the findings and reports of the workshop at the 2018 Large Facilities Workshop in May. The workshop reports and survey results are available at http://facilitiesci.org.

In December 2017, RDI² held an exclusive screening of the webinar, Gender in the Global Research Landscape, based on a recent report, headed by co-author Dr. Holly J. Falk-Krzesinski, Vice President of Global Strategic Networks at Elsevier, on research performance through a gender lens. This global study draws upon data and analytics, a unique gender disambiguation methodology, and involvement of global experts to provide powerful insight and guidance on gender equality policy for governments, funders, and institutions worldwide.

In May, RDI² hosted the first Introduction to Data Management seminar for early career researchers. This seminar introduces postdocs and early career researchers to best practices in data management through a tour of the research data lifecycle.
HIGH SCHOOL OUTREACH—
DIVING INTO BIG DATA WORKSHOP

RDI\(^2\) celebrated National Engineers Week by launching the *Diving into Big Data and Large Scale Computing* workshop. In February, RDI\(^2\) hosted 24 students from Hunterdon County Vocational Technical School District’s Software Engineering and Computer Science Academy, the Delaware Valley High School. In May, RDI\(^2\) welcomed students from the National Center for Girls Leadership Stuart at Stuart Country Day School. In addition to a tour of state-of-the-art facilities at RDI\(^2\) and an interactive workshop, students learned about fundamentals of data science and large scale computing and the impact and application of Big Data in research, commercial environments, and every aspect of our lives.

RDI\(^2\) researchers shared information about top research initiatives at the institute such as the Virtual Data Collaboratory and the Ocean Observatory Initiative. Students experimented with live oceanographic data and developed a web-based dashboard that continuously displayed real-time data from under water sensors on the Pacific Ocean seafloor. They also visualized the results of simple data transformations. By the end of the day, students had the opportunity to engage with members of RDI\(^2\) admin staff, technical operations, researchers, and graduate students to learn coding in python, career advice and more.