Dear friends of RDI²,

Welcome to the inaugural edition of the RDI² Newsletter. As we celebrate our 5th anniversary, we are excited to share with you reflections of the last five years, significant RDI² milestones and the latest RDI² news.

The era of big data and extreme computing is upon us with the potential to transform science and society and impact every aspect of our lives and our environment. Computing and data are critical to understanding and managing natural, engineering and human systems, from climate change and smart infrastructure to personalized medicine and healthcare. Consequently, it has become essential for academic institutions to have access to leading computing and data capabilities.

In 2012, the Rutgers Discovery Informatics Institute (RDI²) was founded in response to this reality. RDI², a unit of the Office of Research and Economic Development at Rutgers University, strives to stimulate new thinking and new practices across all disciplines to address the world’s greatest challenges. To do this, it integrates research, education and advanced technologies nationally and globally.

Our overarching mission is to accelerate discovery and drive innovation through advanced computing and data. We’ve come a long way these past five years, and I invite you to join us on our journey as we continue to push the boundaries and break barriers of multidisciplinary research.

Sincerely,

Manish Parashar
Founding Director of RDI²
Distinguished Professor of Computer Science
Rutgers Discovery Informatics Institute (RDI²) established as a university-wide multidisciplinary institute for computation and data

Awarded $10 million by the State of New Jersey for Caliburn, the State’s largest Supercomputer

Formed the New Jersey Big Data Alliance (NJBDA), a unique consortium of 12 universities across NJ collaborating on data challenges

Awarded $11.5 million to design, deploy and operate CI for NSF Ocean Observatories Initiative (OOI), the largest oceaning observing program in the world

Led cyberinfrastructure strategic planning across Rutgers and established the Rutgers Office of Advanced Research Computing

Awarded $4M from NSF to establish a Regional Cyberinfrastructure for Collaborative Data Intensive Science

Launched Caliburn, a statewide resource for computing and data analytics

Created the RDI² Fellowship for Excellence in Computational and Data Science
INTERDISCIPLINARY COLLABORATION TO FURTHER INNOVATIVE RESEARCH

Combining leadership in research, expertise, advanced tools and algorithms and advanced cyber infrastructure ecosystems, RDI² engages leading academic and industry researchers, both nationally and internationally, in innovative, interdisciplinary collaborations. Since its inception, RDI² has led several strategic initiatives including the NSF Ocean Observatories Initiative, the NSF Virtual Data Collaboratory, the New Jersey Big Data Alliance, Green HPC and more. Additionally, RDI² developed several innovative software systems including CometCloud, Data Spaces/DART and Fenix.

ADVANCED CYBER INFRASTRUCTURE LEADERSHIP

Research at RDI² is making a difference and the world knows it. During its first five years, RDI² received a total of $49.9 million from 61 awards in grants and sponsorships, including a $10 million state grant to build a supercomputer.

As a result, Rutgers is now home to Caliburn, the most powerful supercomputer system in the state, and its capabilities benefit a broad spectrum of industry sectors and academic disciplines.

2017 AWARDS & HONORS

Dr. Manish Parashar
- Elected to IEEE Computer Society’s Golden Core
- Received 2017 TCPP Outstanding Service Award

Qian Sun & Melissa Romanus
Best Paper Award, 2017 IEEE SoSE

Mengsong Zou & Ali Zamani
Best Paper Award, ACM/IEEE International Workshop on Extreme Scale Programming Models and Middleware

90 CONFERENCE/WORKSHOP PRESENTATIONS
73 JOURNAL PUBLICATIONS
5 BOOK CHAPTERS
2 PATENTS

RESEARCH IN ACTION
Dr. Pradeep Subedi is currently part of the DataSpaces group in RDI² where he is involved in researching extreme-scale data management and deep-memory hierarchy data staging for HPC workflows. Before joining RDI² in March 2017, Subedi received his Ph.D. in Computer Engineering from Virginia Commonwealth in 2016. His thesis work focused on the exploration of erasure-coded storage systems for high performance, reliability and interoperability. He is also involved in other research activities with Oak Ridge National Laboratory, Sandia National Laboratories and Intel.

Dr. Daniel Balouek-Thomert is a postdoctoral associate whose research focuses on how distributed systems can benefit from modern infrastructures and facilities as well as transversal concerns such as energy efficiency and data management. Balouek-Thomert moved to Rutgers after getting his Ph.D. from École normale supérieure de Lyon in France in 2016 where he actively collaborated with industries in national projects.

Philip Davis is an application developer on the DataSpaces project at RDI². His research work is focused on extreme-scale data transfer in HPC applications. Additionally, he is involved in numerous research efforts at RDI² involving large-scale data movement and scientific simulation. This includes work on the Exascale Computing Project in collaboration with research partners at various national labs (e.g. ORNL, ANL, LLNL, BNL). He also provides research support for fusion simulation science in collaboration with PPPL and is engaged in research collaborations with various groups at Rutgers University.

Dr. Anthony Simonet received his Master’s degree from the University of Bordeaux and his Ph.D. in Computer Science from the École normale supérieure de Lyon in France in 2015. After a first postdoctoral position at Inria in Nantes, he joined RDI² as a postdoctoral associate in September 2017. His research topics include designing and implementing Distributed Data Management systems, Fog/Edge Computing and Energy. At RDI², Simonet will contribute to data stream processing for data intensive science and edge computing.
INTERNATIONAL COLLABORATION
VISITING SCHOLAR

The international reputation of RDI² attracts distinguished visitors from around the world. The RDI² program for visiting scholars invites scholars of distinction to engage in postdoctoral research and scholarship as well as scholarly and cultural interaction with RDI²’s faculty, staff and students.

In May 2017, RDI² welcomed visiting scholar Dr. Eugenio Cesario, a researcher at the Institute of High Performance Computing and Networks of the National Research Council of Italy (ICAR-CNR). Cesario’s research interests include data mining, urban data analysis, energy-aware cloud computing and cloud/grid services architectures.

During his time at RDI², Cesario collaborated on the design and development of a “Hierarchical Exascale Framework for Iterative Parallel Data Analysis Algorithms.” This research activity was focused on the design of a hierarchical framework for parallel algorithms, like tree or graph structures, to provide better scalability for exascale systems.

INTERDISCIPLINARY RESEARCH
SCIENCE AT EXASCALE

The overarching goal of the DataSpaces project at RDI² is to address data management challenges and enable scientists to run simulations at extreme-scales. Funded by the U.S. Exascale Computing Program (ECP) and other DOE and NSF programs, the open-source DataSpaces software allows computational scientists to easily manage their data and accelerate the end-to-end simulation performance without requiring an advanced knowledge of the underlying cyberinfrastructure.

As a result, DataSpaces is used in scientific applications doing significant fundamental research in areas such as nuclear fusion with the edge fusion simulation code XGC at the Princeton Plasma Physics Lab; combustion with the direct numerical simulation code S3D from Sandia National Lab; and seismic analysis with the IPARS code from the Center for Subsurface Modeling at UT Austin.

Additionally, the DataSpaces team collaborates closely on systems research with groups from many DOE National Labs, such as Oak Ridge National Lab, Argonne National Lab, and Lawrence Livermore and Lawrence-Berkeley National Labs as part of its contributions to the ECP project.

COLLABORATION
ON CAMPUS

Since its founding, RDI² has engaged in interdisciplinary research with researchers across Rutgers University, as well as with researchers at other universities, national laboratories and industry in New Jersey, the U.S. and internationally.

Additionally, RDI²’s research cyber-infrastructure is open to all faculty and students across Rutgers. As a result, the institute collaborates with over seven different schools at the university including the School of Engineering, the School of Environmental and Biological Sciences, the School of Communication and Information, Rutgers Business School, Rutgers New Jersey Medical School and more.
FELLOWSHIP FOR EXCELLENCE IN COMPUTATIONAL & DATA SCIENCE

The RDI$^2$ Fellowship for Excellence in Computational & Data Science was launched in 2017 to support Ph.D. Students and Postdoctoral Associates in all areas of Computational and Data-enabled Science and Engineering (CDSE), with a specific focus on research in Big Data and Extreme Scale computing.

Humna Awan, Ph.D. candidate in Physics & Astronomy, and Guangyan Hu, Ph.D. candidate in Computer Science, were selected as the first recipients of this exclusive fellowship.

Awan’s project is titled, “Big Data in Astrophysics: Clustering Analysis of Partial Galaxies,” under supervision of Dr. Eric Gawiser, Associate Professor in the Department of Physics and Astronomy.

Hu, supervised by Dr. Thu Nguyen, Professor and Chair of Computer Science, is currently focusing on approximate computing targeting large-scale data analytics applications in Spark.

GRADUATE ADVANCEMENT IN CLOUD COMPUTING & BIG DATA

Ivan Rodero, Associate Director for Technical Operations at RDI$^2$, designed and delivered a new graduate-level course, “Introduction to Cloud Computing and Big Data.” This course introduced fundamental concepts, technologies and innovative applications of Cloud and Big Data systems. It also included engineering aspects such as bridging the gap between analytics and data-driven platforms, performance evaluation and benchmarking.

CONNECTING STUDENTS, STAFF & FACULTY CAMPUS-WIDE

RDI$^2$ launched a new Research Colloquium and Working Group Discussion series in April 2017, open to all members of the university. The series, designed to connect students, faculty and staff, aimed to facilitate discussion about current research projects as well as identify areas of interest for future collaborations.

Presentations and brainstorming sessions at this bi-monthly meeting bring fresh perspectives to the table and encourage constructive dialogue about topics of importance to academics.
DISTINGUISHED SEMINAR SERIES

RDi2’s Distinguished Seminar series is designed to bring members of the Rutgers community together to explore Data Science topics and regularly hosts leading researchers from academia, governments and industry.

The next Distinguished Seminar event will take place on Wednesday, November 8th in the Busch Student Center. Dr. Paul Messina will present “Challenges of Exascale Computing.” To register and learn more about this free event, visit http://conta.cc/2yFhwzf.

THE 2017 NATIONAL SCIENCE FOUNDATION LARGE FACILITIES CYBERINFRASTRUCTURE WORKSHOP

Cyberinfrastructure is a critical component of NSF facilities as it continues growing in scale and complexity. To address current and future CI needs, RDi2 Director Manish Parashar organized the 2017 NSF Large Facilities Cyberinfrastructure Workshop on September 6th and 7th enabling industry experts to connect and communicate with each other and share insights on CI models, challenges and best practices.

ANNUAL OPEN HOUSE SYMPOSIUM

RDi2 held its annual Open House Symposium on March 28. The full-day event featured an introduction to the institute and its initiatives as well as several exciting sessions and workshops. It also included a student poster showcase that featured research collaborations of the institute. The event was well attended by Rutgers University students, faculty and staff and attracted representatives from New Jersey industry and academia.

CALIBURN LAUNCH EVENT

In December 2016, RDi2 held a launch event to celebrate Caliburn, Rutgers’ new supercomputer. Caliburn is the most powerful computer system in the region and currently ranks #2 among Big Ten universities and #8 among U.S. academic institutions. The launch party was open to the public and featured a talk from Rochelle R. Hendricks, Secretary of the New Jersey Office of Higher Education.
Dr. Xiaolin Andy Li received his doctorate degree in Computer Engineering from Rutgers University in 2005. His research interests include Cloud Computing, Big Data, Deep Learning, Internet-of-Things, and Security and Privacy. During his time at Rutgers, Li worked under the guidance of RDI² Director, Manish Parashar.

Since then, Li has accomplished a lot. Aside from his role as Associate Professor at the University of Florida, he is also the founding director of the Large-scale Intelligent Systems Lab and the National Science Foundation (NSF) Center for Big Learning. Through research sponsored by NSF, National Institutes of Health and the Department of Homeland Security, Li has published over 100 peer-reviewed papers, five books and four patents.

"Professor Parashar distilled the unique research gene into my research and mentoring style as a professor," Li said as he reflected on how RDI² helped to shape his future. "The large group of lab mates and teammates created the best environment for learning, creating, coding and entertaining."

Due to his outstanding work, Li has received several prestigious awards including the NSF CAREER Award in 2010, the Internet² Innovative Application Award in 2013, NSF I-Corps Top Team Award in 2015, the CAGI Challenge on Detecting Bipolar Disorder Top Team Award in 2016 and best paper awards (IEEE UbiSafe 2007, ACM CAC 2013, IEEE SECON 2016 and IEEE ICMLA 2016).

His current efforts focus on creating large-scale open intelligent platforms and deep learning models for a broad spectrum of applications in business, health, science and engineering.

"My experience at RDI² helped my research tremendously," Li said. "The hands-on experiences with large-scale system software, runtime management strategies, and computational engine for real-world sophisticated scientific applications was invaluable."
FURTHERING INNOVATIVE RESEARCH

RDI² engages in multidisciplinary co-discovery with partners in application areas ranging from healthcare and medical informatics to finance, insurance, food, oil & gas, logistics and more. The RDI² Industry Partnership Program provides industry members with access to a comprehensive set of services and resources including world-class technical expertise, RDI²’s advanced infrastructure, training in Computational and Data Sciences and Advanced Cyberinfrastructure, resources and expertise at affiliated centers, and participation in annual RDI² events.

SONY ERICSSON

Since 2014, RDI² has collaborated with Ericsson Research (San José, CA) on developing innovative strategies for cloud orchestration. While cloud hosting promises several advantages such as flexible deployment and elastic scaling, not all applications may see a straightforward migration path to clouds due to the complexity of application requirements and the cloud platforms.

The focus of this partnership is to enable an application-platform interface that allows applications to expose sufficient information to the platform to facilitate, and possibly automate, migration tasks. For example, the platform may build and use models of application behavior, including predictive behavior, to enable automation and optimization.

Applications under consideration include end-to-end media flows (consumer media, video conferencing, etc.), IoT-based control (autonomous vehicles, etc.), consumer applications and edge services. Above all, this partnership encourages cross-fertilization of ideas and equips our students with industrial perspective for their research work.
RUTGERS DISCOVERY INFORMATICS INSTITUTE

http://rdi2.rutgers.edu/