

The Ken Kennedy Institute for Information Technology

Fiscal Year 2017 Annual Report

Moshe Y. Vardi, Director

Keith D. Cooper, Co-Director

Jan E. Odegard, Executive Director

Victoria Langlais, Associate Director (editor)

Deborah Heath, Marketing and Events Specialist (editor)

E-mail: k2i@rice.edu

Phone: +1 713.348.5823

Anne and Charles Duncan Hall

Offices: 1084, 1085 & 1088

8/22/2017



RICE | **KEN
KENNEDY**
INSTITUTE FOR INFORMATION TECHNOLOGY

Executive Summary

During Fiscal Year 2017 (FY17), the Ken Kennedy Institute continued to develop new collaborative efforts and achievable returns on investment. As a “service institute,” we have been recognized for our consistent efforts to build and support a vibrant research community, and to continue developing and supporting, in partnership with the Office of Information Technology, a robust shared research cyberinfrastructure ecosystem for Rice’s research community. The Institute continues to provide multiple platforms for intellectual engagement amongst faculty across the university, with colleagues at other academic institutions, and with industry. In FY17, we facilitated many endeavors aimed to strengthen our organization and support the community:

- Partnered with faculty to prepare and submit three proposals to NSF.
- Partnered with IBB, SCI and RSI to develop LIFT proposal “Destination Rice.”
- Continued growing our graduate fellowship program:
 - Computational Science & Engineering Supplemental (recruiting) fellowships to help recruit high-achieving students in the application pool (funded by OG-HPC Conference revenues).
 - Continue running our well-established and highly successful industry-funded Graduate (enhancement) Fellowship program for students already at Rice (named fellowships funded by industry).
- Despite the downturn in the oil & gas industry we hosted another successful Rice Oil & Gas High Performance Computing Conference (10th anniversary) with over 440 attendees
- Hosted a Faculty Data-Science Meetup. This was the 3rd networking opportunity designed to help faculty discover and explore “data-science” research challenges, opportunities, tools, and infrastructure across Rice.
- Hosted day long Machine-Learning Workshop, which featured presentations from Rice machine learning experts and industry leaders with close to 400 attendees.
- Delivered two week-long training classes (Summer Boot Camps) open to external and internal participants:
 - HPC Boot Camp (the 8th in the series), May 22-26, 2017
 - Data Science Boot Camp (the 4th in the series), May 15-19, 2017
- Hosted/co-Hosted eight Distinguished Lectures (up from four the year before) and eight monthly member luncheons in support of community development.
- Co-Hosted the NorTex Data Science Digitalization Workshop, May 3, 2017
- Enhanced the Institute’s image to align with campus-wide online presence and visibility with the new Drupal based website platform, and added a YouTube channel presence to host online recordings of lectures and events.

Institute Space and Leadership

The Ken Kennedy Institute is located in Duncan Hall where it occupies three offices (rooms 1084, 1085 and 1088). No other space is currently dedicated to the Institute. The Directors of the institute are Moshe Y. Vardi, Karen Ostrum George Distinguished Service Professor in Computational Engineering and Keith Cooper, L. John & Ann H. Doerr Chair, Computational Engineering, Professor, Computer Science. Full time staff: Jan E. Odegard, Executive Director (80%) and since March of 2015 Associate Vice President (20%), Office of Information Technology in charge of research cyberinfrastructure; Victoria Langlais, Associate Director and Deborah Heath, Marketing and Events Specialist.

Institute Membership

The Ken Kennedy Institute's membership is comprised of about 160 Rice faculty members, 70 associate members who are part of the greater campus population, and 20 non-Rice "adjunct" members who are affiliated with the Texas Medical Center, industry, or other organizations outside of Rice University. In addition, we also maintain mailing lists to "friends" of the institute with over 1,900 contacts that are informed about public events organized by the Institute.

Internal Collaboration

Shared Computing Infrastructure

Since 2002, the Ken Kennedy Institute has successfully worked with faculty to fund and build Rice's shared cyberinfrastructure. Today, the Institute, in partnership with Rice's Office of Information Technology, Center for Research Computing, supports the computational research needs of more than 170 faculty members and over 600 shared-computing users.

The Institute could not operate and continue to build upon its shared computing infrastructure without the technical support of Rice's Office of Information Technology, the financial support of Rice University, funding from the National Science Foundation (NSF) and some donated equipment by industry. Donated equipment, while new and can be nice, is not generally large enough to matter and often comes with hidden cost since it usually is different from existing infrastructure requiring specialized support, and on average has from experience a higher/hidden staff cost of ownership. While we continue partnering with faculty to write infrastructure grant proposals, the competition is getting tough and we were recently turned down for a \$2 million NSF MRI proposal for a much needed upgrade and expansion of our infrastructure in support on increasing data science research needs. What is really concerning is that more than 50% of our shared computational infrastructure is approaching its sixth anniversary, which is in reality one year past it expected useful life for effectively supporting researchers needs. Since 2002 the Institute has helped fund, on average, \$1 million per year of research computing infrastructure. This has helped Rice sustain a healthy computational infrastructure, serving as a catalyst for research and maintaining Rice's ability to be competitive for research funding.

Building on the success enabled since 2002 by the Institute, the Office of Information Technology has made a significant step towards longer-term infrastructure sustainability with the approval of \$1.6 million over 4 years (funds made available starting in FY17). This

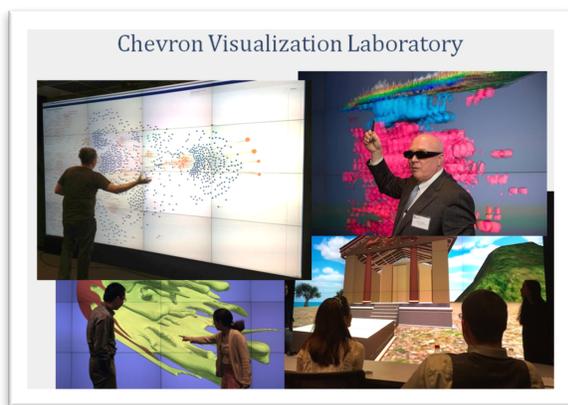
funding will support building a scalable research data-storage infrastructure (slated to go into production early in FY18), a more scalable research computing (HPC and VM) infrastructure (investments in VM infrastructure and ability to more aggressively supporting faculty funded condominium HPC investments). This funding complements annual service-center funding as well as any shared-infrastructure awards and faculty investments in condos. The emphasis of this initial funding will be focused on exploring a more scalable infrastructure supporting a broader set of users, enabling Rice to be even more competitive for grant funding.

Research CyberInfrastructure

Direct government funding and research grants continue to be the Institute's largest funding source for HPC resources, and we estimate that more than 60% of Rice's annual research expenditures is significantly enabled by the shared research cyberinfrastructure we help bring to Rice. Infrastructure awards as a result of the Ken Kennedy Institute leadership and coordination since 2002 have been a critical enabler for Rice's growth in research expenditures and have also helped Rice maintain the recognition as a *Tier 1 HPC institution*. The Ken Kennedy Institute, in collaboration with the Vice President and CIO at Rice University, is working with faculty to expand Rice's shared research cyberinfrastructure reach and services. Current projects include exploratory cloud offerings, research VMs and scalable research data-storage services.

Chevron Visualization Laboratory

Rice's Chevron Visualization Laboratory, funded in part by a grant from the National Science Foundation (NSF) and a gift from Chevron, enables researchers to interact with



data in more effective ways and to probe details and develop understanding. The 200-inch wall (measured diagonally) lets users display and analyze images of all types, from atoms to galaxies, to archeological structures and sites, to urban developments and planning. This "studio" is not only able to help researchers broadly across all fields of study, only limited by user's imagination. On average the space is booked 10-15 hours per week with recurring events and frequented by others as needed to support visualization needs. The lab services extend past the

physical facility and we help support and develop in-situ data visualization workflows. Due to additional space needed for students and faculty, The Visualization Lab will be relocated from Dell Butcher Hall to the 3rd floor of the Keith Weiss Earth Science Building.

Research Computing Service Center Operation

K2I continues to assume the responsibility of administering, managing and charging for access to Rice's shared-computing infrastructure. During FY17 this included operating of the Shared Computing Service Center. During FY15 we successfully deployed the Billing

Application Service Center Software (BASCS) web application that helped automate large portions of the billing process and includes web pages for department staff or principal investigator (PIs) to manage assignment for funds that will be charged for different PIs and users. In an effort to ensure a seamless transition and collect feedback on the software solutions, the Institute continues to administer BASCS training to all department administrators who have faculty using the shared computing infrastructure. The BASCS software is in need of a significant enhancements, this is a high priority for FY18.

Proposal Submissions

The Ken Kennedy Institute assisted in writing, coordinating and submitting the following NSF proposals.

NSF Campus Cyber Infrastructure SciDMZ Proposal (Submitted 08.23.16) – \$499,646

The project proposes to deploy the Rice Science DMZ (Rice SciDMZ), a cyber infrastructure that will be distinct from the general-purpose campus network at Rice University. The Rice SciDMZ will be specifically designed and optimized for friction-free data movement to and from Rice. It will support direct peering to R&E networks using IPv4 and IPv6 and deliver 100 gb/s friction-free connectivity to Globus-enabled data transfer facilities. The proposed Rice SciDMZ is vital to the progress of data-intensive research at Rice and will foster collaboration and facilitate discovery in ongoing, high-priority projects.

NSF MRI Proposal (Submitted 1.11.17) – \$1,975,431

Computation is a cornerstone of scientific inquiry, as important to science and engineering as the traditional approaches of theory and experimentation. To facilitate progress in various areas of research and scholarship across Rice University, the Institute proposes to procure and deploy the Heterogeneous Accelerated Data Computing Platform (HADoP) infrastructure. HADoP will support the increasing demand for computing across science, social science and engineering, that will benefit from nodes with GPU accelerators, nodes with large memory, as well as compute nodes that can support traditional high-performance computing (HPC) workloads.

NSF TRIPODS (Submitted 03.14.17) – \$600,000

With modern large data sets, finding answers to seemingly simple questions often turns into an intractable computational problem. To address this, the Texas TRIPOD for Research on Algorithmic Foundations of Data Science will develop data-analysis algorithms where users can explore the trade-offs between accuracy, computational complexity, and statistical efficiency. The center will bring together three top research institutions in Texas (Rice University, Texas A&M University, and the University of Texas at Austin), and researchers from three disciplinary areas (Computer Science, Mathematics, and Statistics) to conduct research on the algorithmic foundations of data science, motivated by problems arising in the biomedical arena, such as high-throughput genomics, in the energy arena, such as reflection seismology, in the transportation arena, such as traffic analytics, and in the finance arena, such as predictive market models.

Data Science Faculty Meetup (October 13, 2016)

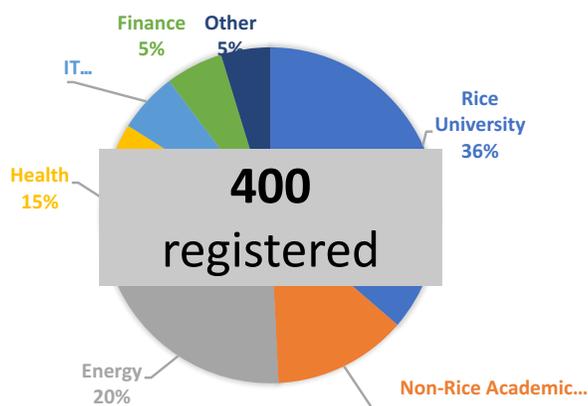
Approximately 60 people attended the third Data Science MeetUp on October 13, 2016. There were over 25 posters on display representing research from the School of Engineering, Jones Business School, Natural Sciences, Social Sciences, Humanities, the Kinder Institute, and Fondren Library. The main focus of the meetup was on researchers sharing their research ideas and challenges with colleagues with the idea of finding potential areas of collaboration. <http://dsmeetup.rice.edu/fall-2016-poster-listings>

External Collaboration

The Ken Kennedy Institute continues to engage with and support the local community at Rice and in Houston with a variety of educational workshops, public lectures, and monthly networking opportunities. The Institute's power lies in its ability to build and strengthen relationships at the forefront of intellectual ambitions. Critical to its mission is the Institute's efforts to continue building and maintaining strong relationships across the community at Rice, the Texas Medical Center, and the greater Houston area.

Machine Learning Workshop (January 24, 2017)

Presentations at the Machine Learning Workshop ranged from Rice machine learning researchers speaking about the latest theory and algorithms to industrial leaders speaking about applications in the energy, medical, financial, and legal industries. This was the first event of its kind hosted by the Ken Kennedy Institute and given the overwhelming attendance of approximately 400 people, the Institute plans to build off this successful event by hosting the inaugural Data Science Conference October 9 & 10, 2017.

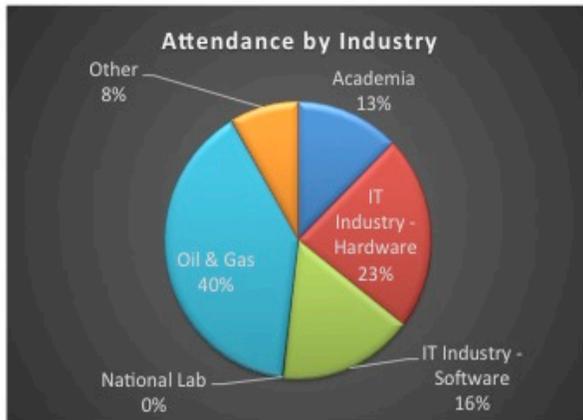


2017 Oil and Gas High Performance Computing Conference (March 14 & 15, 2017)

2017 marked the 10-year anniversary of the Rice Oil & Gas HPC Conference. The anniversary was a milestone in the Institute's continuing pursuit to provide participants with a platform to discover opportunities, broaden their knowledge, and advance their mission by engaging with some of the brightest and most talented leaders across oil and gas, technology, and academia.

Given the tumultuous times in the Oil & Gas industry, the program committee wasn't sure what to expect in terms of participation for this year's conference. After a slow start with

initial registration, surprisingly, concerns were unfounded. Registration was down only slightly with 444 people attending this year (down from 520 in 2016). Additionally, the sponsors for this event attended in force again this year, with a total of 28 sponsor companies represented. Even better, the demographic distribution of attendees improved with 41% attending from the Oil & Gas industry (up from 37% last year).



As a result of the sponsorships this year, the conference's gross was approximately \$175,000. After covering operational costs, all funds will be used to support the Computational Science & Engineering Enhancement (CS&E) Fellowship program. The Institute looks forward to the continued success of the conference with plans to host the 11th annual conference on March 12 & 13, 2018.

Data Science Boot Camp (May 15-19, 2017)

2017 marked the fourth time the Ken Kennedy Institute hosted a data-science training event. This year the Data Science Boot Camp was offered as a five-day training program, as opposed to four days in prior years. The program was designed to attract attendees that needed the skills required to start leveraging the latest data analytics methods and tools with their data. The 27 participants that completed the program were from companies in

the oil & gas industry: ExxonMobil, Schlumberger, and Statoil. Several Rice University and University of Houston students and staff members also completed the course. While participants from industry paid full cost to attend the course, the Ken Kennedy Institute heavily subsidized participation from academic institutions including Rice, TMC and the Gulf Coast Consortia.

HPC Boot Camp (May 22-26, 2017)

The 8th Annual 2017 HPC Boot Camp hosted by the Ken Kennedy Institute addressed continued training and education in the “art” of high-performance computing and scientific programming. While the main driver for the HPC Boot Camp has been participation from the oil and gas industry, the curriculum is broadly applicable to any field engaged in scientific computing where there is a need to harness more of the computing power offered by modern servers and clusters. This year 26 participants completed the HPC Boot Camp course. While participants from industry paid full cost to attend the course, the Ken Kennedy Institute heavily subsidized participation from Rice, TMC and the Gulf Coast Consortia.

NorTex Data Science Workshop (May 3, 2017)

This workshop was co-hosted by the Ken Kennedy Institute and the Global Center of Expertise Node and the International Institute of Stavanger as a continuation of discussions launched during OTC 2016 in Houston with focus on digitalization and effective interoperability. The goal was to share information about current projects across borders and entities as well as strengthening partnerships and collaboration. Approximately 50 people from Norway and the United States attended the workshop.

Student support

Graduate Fellowships (\$65,000 Awarded in FY17)

The Ken Kennedy Institute awarded nine supplemental graduate fellowships during FY17 to students across Computer Science (CS), Computational & Applied Mathematics (CAAM), Statistics (STAT), Electrical & Computer Engineering (ECE) and Earth Science (ESCI).

These fellowships are made possible by the sponsorship received from the energy industry (BP, ExxonMobil, Shell, and Schlumberger), from Cray as the Cray-Kennedy Graduate Fellowship, and the Andrew Ladd Memorial Excellence Fund.

Ken Kennedy Institute Computational Science & Engineering Fellowships

This year the Institute continued to expand its Supplemental Fellowship program by offering four-year “supplemental fellowships” each in the amount of \$7,500 to CS, CAAM, ECE, and ESCI, intended to recruit very well qualified graduate student applicants. The fellowships also require a department match of an equal amount; therefore, the total award is \$15,000 per student over four years. These fellowships are made possible by



funding from the Rice OG-HPC Conference.

The Supplemental Fellowships are meant to help recruit the “best-of-the-best” applicants, students that Rice might otherwise lose to competitive universities without this extra incentive for students to choose Rice. For the 2017/18 Academic Year, the Ken Kennedy Institute made 19 CS&E offers to applicants. Of these 19 offers, 4 students accepted. The Institute looks forward to welcoming these exceptional students to Rice in the Fall of 2017.

K2I Sponsored and Co-Sponsored Public Lectures

2016-2017 K2I Distinguished Lectures

- **9.21.16 Kathy Yelik, University of California at Berkeley**
“How to Teach Your Exascale Computer to Do the Data Dance”
- **10.18.16 John Mitchell, Stanford University**
“Can Education Scale?”
- **10. 27.16 Moshe Vardi, Rice University**
“Humans, Machines and Work: The Future is Now”
- **11.10.17 Andrew Lo, MIT Sloan School of Management**
“Moore's Law vs. Murphy's Law: The Impact of Technology on the Financial Industry”
- **01.19.17 Sami Saydjari, Cyber Defense Agency**
“Defending Cyberspace: A New Frontier”
- **1.25.17 Alfred Spector, Two Sigma**
“Big Data: Opportunities & Unanticipated Consequences”
- **02.07.17 Ruby Mendenhall, University of Illinois at Urbana-Champaign**
“Using Big Data to Ask Sociological Questions: Recovering Black Women’s Lived Experiences From 1740 to 2014”
- **04.04.17 William Gropp, University of Illinois Urbana-Champaign**
“MPI: The Once and Future King”

K2I Sponsored Monthly Member Luncheons

FALL 2016

- **09.09.16 Adrianna Gillman**
“Fast Direct Solvers for Linear Partial Differential Equations”

- **10.07.16 Paul Hand**
“Computer Vision Algorithms Inspired by Compressed Sensing”
- **11.04.16 Anshumali Shrivastava**
“Probabilistic Algorithms for Resource - Frugal Machine Learning”

SPRING 2017

- **02.03.17 Ankit Patel**
"Understanding Convolutional Networks from First Probabilistic Principles"
- **03.03.17 Rebecca Richards-Kortum**
“Opportunities to Strengthen Collaborations with the TMC”
- **04.07.17 Lisa Spiro**
“Plan, Organize, Share: How Rice’s Research Data Management Team Helps Researchers Curate Their Data”
- **05.05.17 Paul Padley**
“Faculty IT Governance”

Miscellaneous

Master of Computational Science and Engineering Degree Program

- The Ken Kennedy Institute assisted with the application review of applicants and Rice offered admission to 23 of the 80 students that applied to the Master of Science in Computational Engineering (MCSE) Professional Masters Program, a joint program between Computer Science, Computational and Applied Mathematics and Statistics. Of the 23 admission offers made, seven students enrolled in the program.

Institute Sponsorship

- Provided monetary support for the Data Rescue Houston event hosted by the Fondren Library.
- Provided marketing and administrative support for the 2016 DeLange Conference.

Distinguished Guests

- Hosted the campus visit with Rice faculty and visiting lecture of Anastasios Kyrillidis as part of the Rice Data Science faculty recruitment process. Approximately 80 people attended the lecture: “Rethinking algorithms in Data Science: Scaling up optimization using non-convexity, provably”. Following his visit to Rice University, Anastasios Kyrillidis has accepted a position with Rice University and will be joining the faculty in FY 2018.
- Hosted the campus visit with Rice faculty for Jeannette Wing, Corporate Vice President of Microsoft Research, Consulting Professor of Computer Science at Carnegie Mellon,

and Affiliate Faculty in Computer Science and Engineering at the University of Washington.

- Hosted the campus visit with Office of Information Technology leaders for Lauren Michael, Discovery Fellow and Research Computing Facilitator at the Center for High Throughput Computing, Core Computational Technology at the University of Wisconsin, Madison