



**PennState**  
Institute for CyberScience

# 2018-19 Annual Report

## CONTENTS

2	REFLECTIONS FROM THE ICS DIRECTOR
3-6	EXPANDING SUPERCOMPUTING RESOURCES
7	IMPACTING THE LOCAL COMMUNITY
8	BROADENING INTERNATIONAL PARTNERSHIPS
9-10	LAUNCHING NEW RESEARCH
11-12	SUCCESSES FROM PAST SEED GRANTS
13-18	BUILDING THE CYBERSCIENCE COMMUNITY

### DID YOU KNOW?

Penn State researchers using the ICS-maintained supercomputer, ICS-ACI, hail from 65 academic departments and 16 campuses across Penn State. ICS is proud to support such a diverse and widespread group of scientists!

**“This has been an exciting year for the Institute for CyberScience. We have taken on a host of new challenges with zeal—building industry and international collaborations, expanding our co-hire and associate community, and diversifying our infrastructure offerings”**



Jenni Evans  
Director, Institute for CyberScience  
Professor of Meteorology and Atmospheric Science  
Centennial President, American Meteorological Society

## REFLECTIONS FROM THE ICS DIRECTOR

### ● CYBERSCIENCE: EXPANDING IN NEW DIRECTIONS

Whether it's the growth of our faculty users or the increase in attendees to our annual ICS Symposium, investment in cyberscience at Penn State is clearly on the rise. As researchers continue to harness new high-performance-computing-enabled tools — such as sophisticated modeling, complex simulations, machine learning and artificial intelligence — we will continue seeing new growth emerging in computational and data sciences. The computational capacity of the ICS-ACI computational infrastructure supports researchers probing everything from the far corners of the universe to the orbits of atoms, and from our individual brains to the complex dynamics of cultures and societies, using these and other tools in ways previously thought impractical, or even impossible.

### ● MEETING THE NEEDS OF EXPANDING RESEARCH

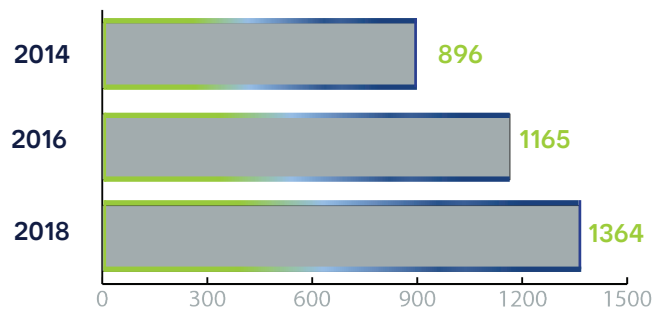
The disciplinary focus of Penn State's cyberscience community closely reflects the research diversity across Penn State. For those of us who have worked in the world of HPC research, this comes as no surprise. The intellectual strength of cyberscience is its universality: all research domains benefit from it, but none is solely defined by it. This “science of discovery” provides a unique virtual laboratory for exploring complex problems that are otherwise impossible or impractical to address.

New advances in education and availability of publicly accessible computer tools are increasing the accessibility of cyberscience, enabling scientists to look at their research from new perspectives, to tackle challenges in interdisciplinary ways, and to expedite the pace of discovery. By fostering interdisciplinary connections across the University and beyond, we at ICS will expand the capacity and reach of cyberscience at Penn State.

# EXPANDING AN ESSENTIAL RESOURCE FOR RESEARCHERS

Cyberscience methods are valuable for researchers in nearly every research discipline. ICS has seen a steady increase in the number of principal investigators on ICS-ACI, our high-performance computing system, and the disciplines of ICS-ACI users closely match the disciplines of all Penn State researchers.

## GROWTH OF PRINCIPAL INVESTIGATORS

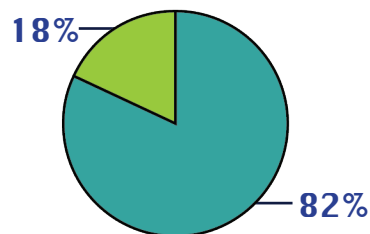


Anthropology graduate student Richard George uses ICS-ACI to analyze the DNA of scarlet macaws and glean insights into the brightly colored bird's evolution.

## COMPUTATION CORRELATES WITH PRODUCTIVITY

ICS analyses revealed the high levels of productivity of ICS-ACI users, both in terms of grant dollars received and scholarly output.

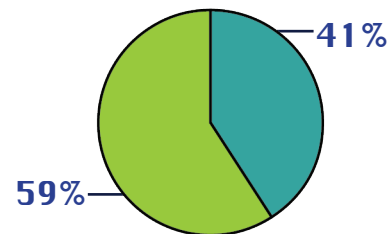
### Total Penn State Researchers



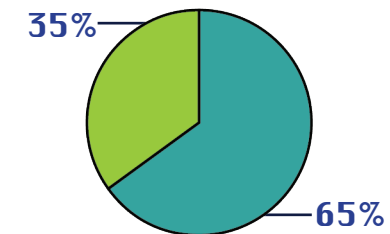
ICS-ACI USERS

OTHER PENN STATE RESEARCHERS

### Top 100 Researchers by Grant Dollars Received



### Scholarly Output



## LAYING THE GROUNDWORK FOR VIRTUAL COMPUTING

ICS has prepared ICS-ACI for a new venture that will benefit researchers by adding a new, lower cost option for high-performance computing. We have installed and begun testing a set of compute cores that, through the power of virtual computing, can each function like up to four cores. This virtual component of ICS-ACI is known as the high-performance research cloud (HPRC). HPRC provides cloud-based computing resources similar to those of commercial providers, but ICS HPRC is a fraction of the cost and has the compliant, secure environment of ICS-ACI. The current ICS-ACI compute offerings also remain, with high memory, high capacity compute, and guaranteed submission time.

Providing the cyberinfrastructure enabling researchers to run many types of experiments ensures that ICS is maximizing the capabilities of its researchers. In fall 2019, 6,000 virtual cores of HPRC will be available to researchers, scaling up to 24,000 virtual cores based upon demand.

## ICS-ACI GETS BIGGER

To meet the growing demand for ICS-ACI resources, ICS continually assesses usage of the system and increases resources appropriately. This past year, we added 3,000 compute cores to our system, bringing the total to 26,000 cores.

### ICS-ACI AT A GLANCE

**26,000 Compute Cores**

**20 Petabytes of Storage**

**650 Teraflops of Total  
Theoretical Peak Performance**



ICS-ACI equipment is located in a state of the art data center on the University Park campus.



# IBM DONATES SYSTEM TO PENN STATE TO ADVANCE AI RESEARCH



Doug Balog, general manager of client success for IBM Storage, stands with Jenni Evans.

**IBM** donated a system based on the technology in the world's most powerful supercomputers, Summit and Sierra, to Penn State, giving researchers a chance to work with the same computing architecture that handles research studies at the government's Oak Ridge National Laboratory and Lawrence Livermore National Laboratory.

The IBM POWER9 architecture of these supercomputers was built from the ground up for the most data-intensive workloads, which make them particularly well-suited for research into AI, or research that uses AI. Using Summit, researchers have already broken the exaop — or, a billion billion calculations — barrier in deep learning and genomics applications. Researchers around the world are using Summit to understand the origins of the universe, identify genetic indicators of addiction or cancer, and tackle other cutting-edge research challenges.

**“Penn State is a pioneer in interdisciplinary research, and this new system will give them a powerful tool to advance that research with AI.”**

**–Doug Balog**

## PENN STATE UPGRADES LIGO SUPERCOMPUTER CLUSTER

In 2016, an international team of scientists found definitive evidence — tiny ripples in space known as gravitational waves — to support one of the last remaining untested predictions of Einstein's theory of general relativity. The team used the Laser Interferometer Gravitational-Wave Observatory (LIGO), which has since made several gravitational wave discoveries. Each discovery was possible in part because of a global network of supercomputer clusters, one of which is housed at Penn State. Researchers use this network, known as the LIGO Data Grid, to analyze the gravitational wave data.

A recent Penn State investment to upgrade our portion of the data grid roughly quadruples the capacity of the LIGO cluster for cutting-edge astronomy and astrophysics research. The new cluster, 192 servers working in tandem, is administered by ICS. The upgrade boosts the cluster from 1,152 compute cores to 4,608 cores, allowing more researchers to use the system simultaneously.

Bangalore Sathyaprakash, professor of astronomy and astrophysics and Elsbach Professor of Physics, and Chad Hanna, associate professor of physics and astronomy and astrophysics, and ICS co-hired faculty member, are the primary researchers who will be using the new system with their research team and collaborators.

**“Without this new cluster, we would be very severely hampered from doing the science that we want to do.”**

**- BANGALORE SATHYAPRAKASH**



## BUILDING A CYBERSCIENCE COMMUNITY

A supercomputer cluster donated by ICS and interdisciplinary partnerships across the University helped turn a group of Bald Eagle Area High School physics students into climate scientists. The class presented a poster session detailing some of their findings in their investigation of global radiation. According to Max Besong, a physics teacher at Bald Eagle Area High School, the equipment and his training at the College of Education's Center for Science and the School's (CSATS) Research Experience for Teachers program were critical in providing this immersive science experience for his students.

"I was interested in seeing how I could implement the research practices that I learned during a Research Experience for Teachers into my classroom," said Besong. "I also wanted the class to experience the culture of an actual lab: having group meetings, sharing each other's work and learning how to communicate research between one another and the community through presentations and posters."

**“I felt like the smartest kid in the school when we first started to code. It’s a project you might think of doing in college, but nobody really thinks about doing something like this in high school.”**  
– Matthew Reese, student



Students from Bald Eagle Area High School presented the results of their semester of research to a group that included Penn State faculty and staff.



## BROADENING INTERNATIONAL PARTNERSHIPS

To strengthen collaborations between Penn State and major institutions from around the world, ICS has partnered with key universities to create interdisciplinary opportunities for faculty, students, and staff.

### ICS'S KEY INTERNATIONAL PARTNERS INCLUDE:

- INDIA INSTITUTE OF SCIENCE, INDIA
- MONASH UNIVERSITY, AUSTRALIA
- NATIONAL TAIWAN NORMAL UNIVERSITY
- TOHOKU UNIVERSITY, JAPAN
- UNIVERSITY OF SALERNO, ITALY



Monash University faculty and staff pose for a photo Penn State's Iconic Lion Shrine.

## CONFERENCES DOWN UNDER AND AT HOME

This past year, ICS and Monash University partnered on multiple research-focused trips that brought Australian researchers to the U.S., and transported American researchers to Australia. In January, ICS hosted a competitive travel grant competition, awarding funds to five faculty members from five Penn State academic colleges. The faculty, along with several ICS staff members, traveled to Australia for a symposium, “Data Science & Artificial Intelligence for Humanities, Arts & Social Sciences.”

In April, Monash University brought a cohort of researchers and staff members to University Park for the annual ICS Symposium, this year themed “Artificial Intelligence and Machine Learning in Science and Society” (see p. 17 for more information about this event). After both symposia, ICS and the Monash eResearch Centre partnered on a seed grant fund to allow researchers to continue their intercontinental collaborations.

## SEED GRANTS HELP GROW NEW RESEARCH

After a fourth consecutive annual request for seed grant proposals in fall 2018, ICS has awarded more than \$1.3 million in funding to launch research. ICS promotes interdisciplinary research in these grants by increasing funds available for projects involving researchers from multiple colleges.

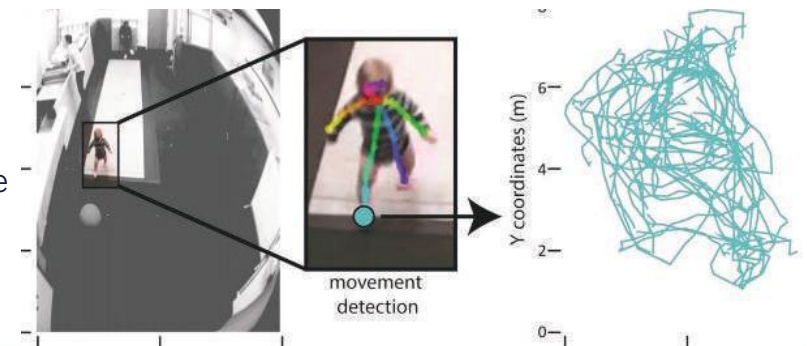
## 2018-19 FOCUS: ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

In 2018-19, ICS awarded 16 seed grants, most of which focused on applications of artificial intelligence and machine learning in a variety of fields.

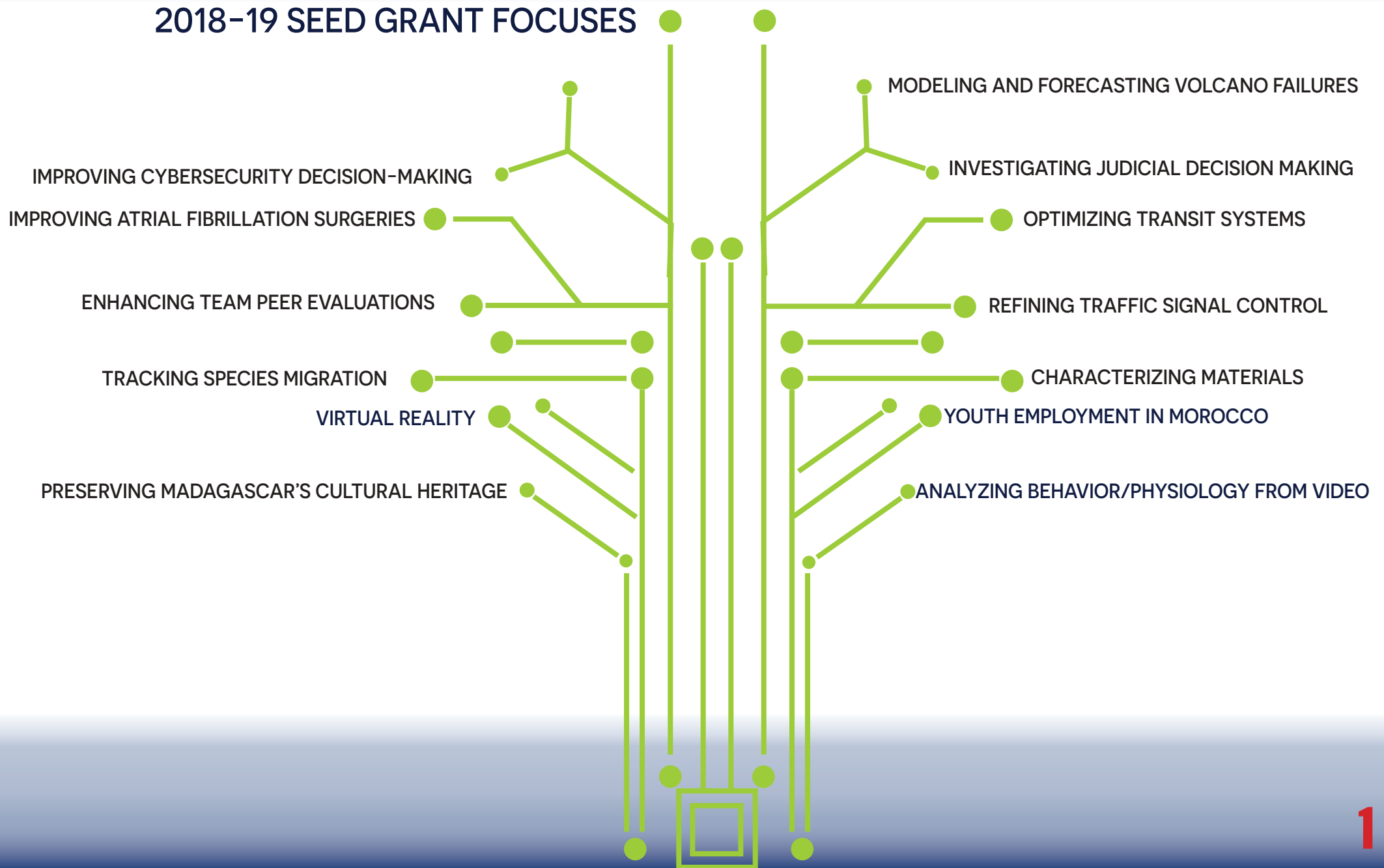
**\$367,632.20** AWARDED — **22** DEPARTMENTS — **10** COLLEGES — **3** CAMPUSES

## SEED GRANT SPOTLIGHT: COMPUTER VISION TO STUDY HUMAN BEHAVIOR

Led by Rick Gilmore, professor of psychology, a team of researchers will use computer vision algorithms to analyze videos from an NIH-funded project focusing on mother-infant interactions, characterizing body movements, respiration rates, and other aspects of physiology and human behavior. The team will make their extensible computational framework available for researchers who are not specialists in computer vision or AI so that they can conduct similar analyses on their own.



## 2018-19 SEED GRANT FOCUSES





## SEED GRANT SUCCESS: PREDICTING INFECTIOUS DISEASE IN UGANDA

A team covering an array of disciplines — including neuroscience, meteorology, and biology — is leveraging predictive models to improve personalized approaches to treating infectious disease in Uganda.

“Personalized medicine today is reactive, often expensive, and in many places unavailable. Complicating matters further, many infectious epidemic diseases are strongly dependent on environmental factors and weather/climate which are factors not addressable in hospital laboratories. We here propose a radically different approach to the treatment of infectious diseases,” said Steven Schiff, Harvey F. Brush Chair in the College of Engineering, Engineering Science and Mechanics and Physics, and a professor of Neurosurgery at Penn State College of Medicine in Hershey.

Schiff and his collaborators received an \$ 8.1 million NIH grant, in part due to work they completed as part of a past ICS seed grant. The team is taking a multi-pronged approach involving disciplines as disparate as epidemiology, meteorology and genomics, and using a variety of tools including machine learning, statistics, and engineering control theory. The Ugandan National Planning Authority, Meteorological Authority, and Ministry of Health are among the critical departments that have offered their support, and Schiff noted that the work of these agencies within Uganda will be key to successful implementation trials.

“

The **ICS seed grant** enabled us to pull our team together and start generating data. This helped lead to tremendous support. ”

–Steven Schiff





## SEED GRANT SUCCESS: TAPPING THE RESEARCH POTENTIAL OF TWITTER DATA

Each day, hundreds of millions of messages flow across the social media platform Twitter. For scientists, who are always looking for new sources of data, those real-time messages are an information gold mine. This could be used for everything from tracking diseases to studying social movements. But there's a problem. There's no way to estimate whether Twitter users think, act, and react the same way as the rest of the world. It's a problem that researchers usually refer to as representativeness, or generalizability, of data.

Funded in part through an ICS seed grant, Guangqing Chi, associate professor of rural sociology and demography and public health sciences in agricultural economics, sociology and education, is exploring ways to assess the representativeness of Twitter data. According to Chi, who is an ICS associate, the seed grant helped pave the way to a \$500,000 grant from the National Science Foundation.

Chi said that the seed grant provided initial funds to hire staff and begin the study. But the seed grant award process had unexpected benefits for Chi. "The process, itself, was really helpful in preparing our proposal for the NSF grant," said Chi. "The feedback we received from the ICS reviewers was invaluable, as well."

The NSF grant will fund the three-year project to collect geotagged tweets from 2014 and 2017 and compare that information to county census data in the U.S. The research team will then investigate methods to determine demographics and use those values to predict county-wide characteristics.

## BUILDING THE CYBERSCIENCE COMMUNITY

Researchers are the core of the Penn State cyberscience community. Several ongoing ICS initiatives are aimed at increasing the number of cyberscience-focused researchers throughout the University, and also increasing collaborations among scientists in different fields.

## ICS CO-HIRED FACULTY

ICS partners with academic colleges to recruit energetic, talented researchers who are recognized experts in their respective fields and who employ cyberscience tools to pursue interdisciplinary research.

### NEW CO-HIRE POSITIONS FUNDED IN 2018–19

Unit(s) Involved	Research focus/theme
College of Earth and Mineral Sciences, Earth and Environmental Systems Institute	Water and Earth systems
College of Health and Human Development	Individualized health Research and Intervention
Eberly College of Science (Astronomy and Astrophysics), College of Engineering, College of Information Sciences and Technology	Data science related to large scale synoptic telescopic data
College of Education	Big data analytics and methodology, educational psychology
College of Earth and Mineral Sciences, Eberly College of Science	Big data and remote sensing
Penn State Law, School of International Affairs, College of Engineering	A combination of energy law and policy, cyber privacy and security issues, and big data analysis
College of Information Sciences and Technology, College of Medicine, College of Health and Human Development	Health-related science

## CO-HIRE SNAPSHOT: 2019-20



### AFFILIATES AND ASSOCIATES

ICS includes its Affiliates and Associates in collaborative projects and seeks feedback on cyberscience needs in every discipline. ICS has 326 Affiliates and Associates from 14 Colleges and 13 Campuses.



Want to join our ranks of affiliates and associates?

Learn more about our program and apply at [WWW.ICS.PSU.EDU/AFFILIATES](http://WWW.ICS.PSU.EDU/AFFILIATES)

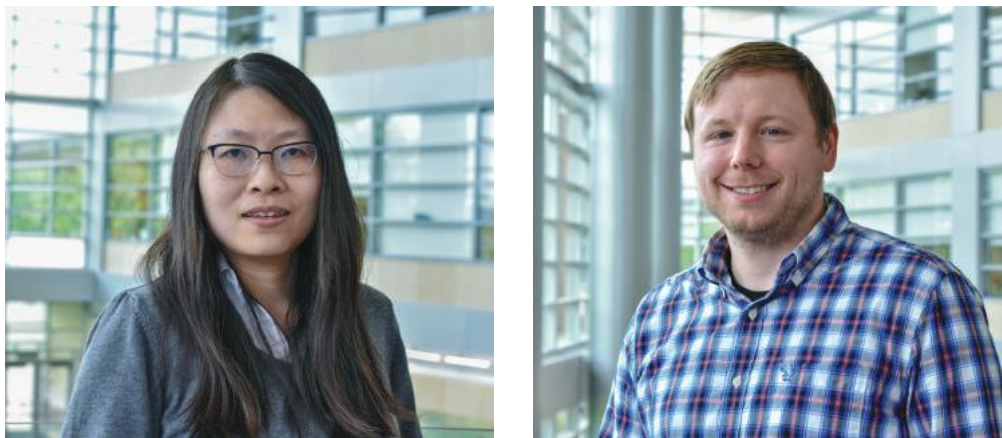
### HITTING THE GROUND RUNNING WITH STARTUP PACKAGES

Working with academic colleges, ICS provides computational resources on ICS-ACI to select incoming researchers as part of startup packages. Since 2015, ICS has contributed nearly **\$1.75 MILLION** worth of ICS-ACI high-performance computing resources to new faculty.

# BUILDING SKILLS AND CAPACITY FOR THE RESEARCH COMMUNITY

## COLLABORATING WITH RISE

Formerly known as Advanced Technical Services, ICS's **Research Innovations with Scientists and Engineers (RISE)** team is a group of R&D engineers who specialize in various applications of HPC tools. They are available to form close-knit collaborations with researchers across the University, assisting with a range of research applications, from setting up science workflows to providing guidance on the best tools to use in specific domain areas.



ICS'S two newest RISE team members, Danying Shao and Justin Petucci

## PYTHON WORKSHOPS

Patrick Dudas and Adam Lavelly of ICS's RISE team hosted a new, four-workshop series on Python training for the Penn State community. Topics covered in the workshops:

- HOW TO USE THE WEB APPLICATION JUPYTER WITH PYTHON
- READING DATA AND MAKING PLOTS
- SPEEDING UP PYTHON
- SCRAPING PUBLICLY AVAILABLE DATA FROM THE WEB

4 workshops

316 attendees

116 academic departments and administrative units represented





## MACHINE LEARNING

ICS RISE team member Christopher Blanton launched an introductory course on machine learning for the Penn State community. The session covered topics ranging from how researchers use machine learning to basic applications using Tensor Flow technology.

## XSEDE TRAINING

The Extreme Science and Engineering Discovery Environment (XSEDE) is an NSF-funded virtual organization that integrates and coordinates the sharing of advanced digital services with researchers nationally to support science. Diego Menéndez is Penn State's XSEDE Campus Champion, and he coordinates Penn State's participation in monthly HPC training sessions and a summer boot camp. The trainings are led by Pittsburgh Supercomputing Center and videocast to sites around the U.S.

## SYMPOSIUM FOCUSES ON AI OPPORTUNITIES AND CHALLENGES

This year's annual ICS Symposium brought in a record number of students, faculty, staff and visitors (310) to explore the use, as well as the legal and ethical implications, of artificial intelligence and machine learning in science and research. The symposium, held on April 1 at Penn State's Nittany Lion Inn, featured panel discussions on artificial intelligence and machine learning, including topics such as security, fairness, privacy and law. Other panel discussions focused on the science applications of big data and machine learning and the benefits of leveraging AI to help society. Michael J. Franklin, Liew Family Chair of Computer Science and senior adviser to the provost for Computation and Data Science at the University of Chicago, offered the symposium's keynote address.



**“The ICS Symposium is designed to be a forum that helps researchers learn new ways to explore science and society’s biggest questions.”**

**–Jenni Evans**


## HACKING BACK: SIMULATION HELPS TRAIN INTERDISCIPLINARY CYBERSECURITY TEAMS

Cyber hacks cost the U.S. economy more than \$100 billion a year, or about .64 percent of the nation's GDP, according to the White House. Training the next generation of cybersecurity experts is now a national priority. A group of 70 Penn State students, faculty, and staff from law, business, IST, and other disciplines participated in the 2019 Penn State Hack Response Simulation Competition. The simulation was a hack to a large company, with teams challenged to account for the computational, public relations, and legal aspects of the situation. Anne Toomey McKenna, distinguished scholar of cyber law at Dickinson Law, and ICS Co-Hire and professor of practice, spearheaded the Hack Response Simulation Competition to help train students to understand the scope of a cyber-attack and how to respond as part of an interdisciplinary team.



**“These students are gaining an unmatched interdisciplinary experience with some of the brightest minds.”**  
–Shelly Curling,  
Instructor in  
Accounting



The background is a deep blue with a complex pattern of glowing, interconnected lines and circles. The lines are of varying thickness and orientation, creating a sense of depth and movement. Some lines are straight, while others are curved or bent at angles. The circles are also of different sizes and are scattered throughout the composition. Many of these elements have a bright, glowing center, giving them a three-dimensional, almost ethereal appearance. The overall effect is one of a futuristic or technological environment.

This publication is available in alternative media on request. The University is committed to equal access to programs, facilities, admission, and employment for all persons. It is the policy of the University to maintain an environment free of harassment and free of discrimination against any person because of age, race, color, ancestry, national origin, religion, creed, service in the uniformed services (as defined in state and federal law), veteran status, sex, sexual orientation, marital or family status, pregnancy, pregnancy-related conditions, physical or mental disability, gender, perceived gender, gender identity, genetic information, or political ideas. Discriminatory conduct and harassment, as well as sexual misconduct and relationship violence, violates the dignity of individuals, impedes the realization of the University's educational mission, and will not be tolerated. Direct all inquiries regarding nondiscrimination policy to the Affirmative Action Office, The Pennsylvania State University, 328 Boucke Building, University Park, PA 16802-5901; Email: [aao@psu.edu](mailto:aao@psu.edu); Tel: 814-863-0471. U.Ed. RES RES 19-107