



2019 ANNUAL REPORT

Information Sciences Institute

DELIVERING
THE FUTURE

USC Viterbi
School of Engineering

Information Sciences Institute

Information Sciences Institute is a world leader in research
and development of advanced information processing,
computing, and communications technologies.



CALIFORNIA

4676 Admiralty Way #1001
Marina del Rey, CA 90292
310.822.1511



MASSACHUSETTS

890 Winter Street
Waltham, MA 02451
781.622.9790



VIRGINIA

3811 Fairfax Drive #200
Arlington, VA 22203
703.812.3700



2 | Welcome to ISI

4 | ISI Offices

10 | ISI Research Groups

19 | Research Funding by Area

20 | ISI in Numbers

21 | New Directors and Research Leads

23 | New Faculty Appointments

24 | New Team Members

25 | 2019 Research Grants and Awards

29 | 2019 Awards and Honors

31 | 2019 PhD Graduates

32 | Postdoctoral Scholars

32 | Visiting Scholars and PhD Students

33 | PhD Students

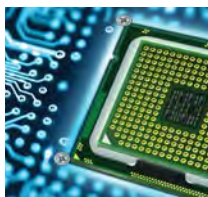
35 | Student Research

38 | 2019 Summer Internship Program

39 | 2019 Research Highlights

55 | 2019 Publications

75 | In Memoriam



WELCOME TO ISI



It's a pleasure to reflect on the growth and progress that the USC Information Sciences Institute experienced in 2019.

We've grown in many areas: the number of researchers, number of research projects, number of PhD students, and more. ISI Boston in particular has increased in size. But numbers alone don't reflect ISI's vitality.

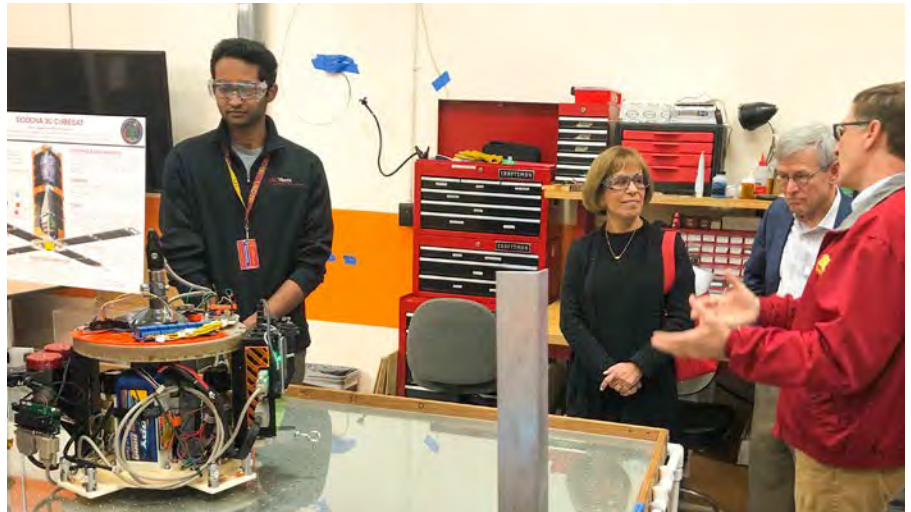
Since the institute's start, ISI researchers enabled a number of fundamental advances in computing sciences: voice-over-IP technology, the Internet domain name system (DNS), machine translation, machine learning, and experimental cybersecurity research. We continue to explore impactful technologies and drive advances in diverse fields including artificial intelligence, language understanding, quantum computing, hardware assurance, and networking technology.

ISI was founded as part of USC in 1972. Now, decades later, our connection to USC grows, strengthening us both. 28 ISI researchers are USC faculty members. These ISI faculty members not only teach at USC, but also advise over 95 PhD students. Our PhD students are in USC's computer science and electrical and computer engineering departments—and also in USC's industrial and systems engineering, physics, and astronautical engineering departments. In 2019 three USC professors were awarded joint appointments at ISI, bringing new research ideas and new students to ISI.

We're looking forward to continued synergy with USC.

When I became interim director in mid-2018, my focus was to continue providing a stimulating research environment and attracting outstanding researchers and students. And it's been quite a year. We hired a record number of researchers and staff, advised a record number of doctoral students, initiated a record number of new research projects—and we did a record amount of work.

As you go through this report, I hope you'll get a flavor of ISI's unique atmosphere and facilities, the diverse portfolio of computing research and development underway, the academic excellence of our researchers and students, and our contributions to real-world societal issues. I invite you to read about our achievements and our people.



President Carol Folt, USC's 12th president, visits ISI at Marina del Rey with Provost Charles Zukoski.



Craig Knoblock

Keston Executive Director, USC Information Sciences Institute

Director of Data Science Program

Research Professor of Computer Science and Spatial Sciences

<http://www.isi.edu/~knoblock>





ISI occupies the large building in the lower left.

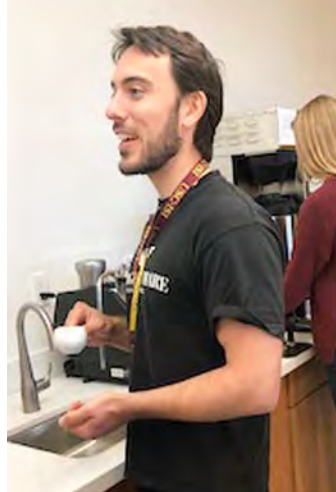


Craig Knoblock, Executive Director

ISI occupies 9 floors of the Marina Towers South office building, due west of Los Angeles, on the edge of the Pacific. In addition to office and laboratory space—and views of the Hollywood Hills sign, the Santa Monica Mountains, and Catalina Island—the institute’s 119,436 square feet of space include a lounge for collaboration or coffee, a yoga studio, and a machine shop. Lab space at ISI headquarters totals 7,578 square feet.

ISI is accessible by public transit and bikepath. In addition to the marina, beaches, and parks, it’s surrounded by diverse restaurants and shops.

The Space Engineering Research Center and the USC-Lockheed Martin Quantum Computing Center are within walking distance from the tower. Free shuttle service runs between ISI, the USC Institute for Creative Technologies, and USC’s University Park campus.





Elizabeth Boschee, Director

ISI Boston first opened in 2017 with a handful of researchers; the research staff now totals 22 and continues to grow. Major research areas at ISI Boston include:

- *Quantum information sciences*
- *Natural language and multimedia understanding*
- *Artificial intelligence*
- *Deep learning*

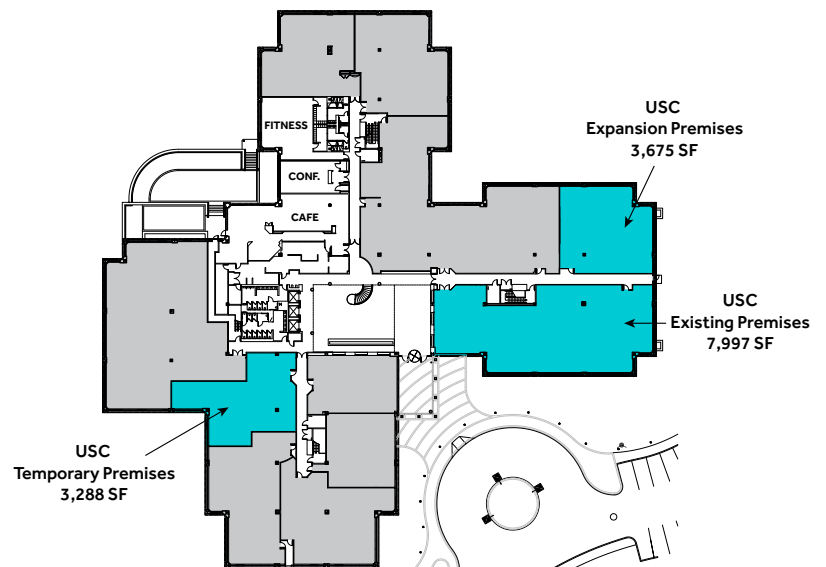


Dean Yannis Yortsos, Viterbi School of Engineering, visits ISI Boston

890 Winter Street, Floor 1

Sponsors of projects led at ISI Boston include Defense Advanced Research Projects Agency (DARPA), Intelligence Advanced Research Projects Activity (IARPA), the Air Force, and the Army Research Office.

Many projects led by ISI Boston researchers are synergized by funded collaboration with numerous universities (including MIT, Columbia, Carnegie Mellon, Notre Dame, UMass Amherst, Virginia Tech, and others) as well as industry collaborators such as Northrop Grumman and Raytheon.



ISI Boston is expanding





Stephen Crago, Director

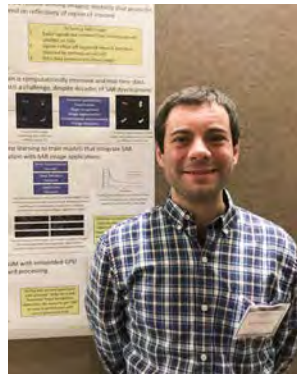
ISI Arlington is located in Northern Virginia, close to Washington, D.C. The Arlington office was started in 1997. 30 researchers and staff members are based there, plus students and interns. In addition, ISI's Secure and Robust Electronics Center (SURE) and the new Application Specific Intelligent Computing Lab are in the Arlington office. Top areas of research carried out at ISI Arlington include:

- *Heterogeneous computing*
- *Reconfigurable computing*
- *Secure electronics*
- *Networking and cybersecurity*
- *Image processing*
- *Natural language processing*
- *Machine learning*



ISI Arlington celebrates Vivek Menon's citizenship

In addition to collaborating with researchers in other ISI locations, researchers at ISI Arlington collaborate with USC academic departments and centers, commercial companies, and defense contractors. Interns are recruited from around the world to work at ISI Arlington, including local and regional universities such as the University of Maryland, Virginia Tech, and George Washington University. Three Arlington researchers hold appointments as research professors at the USC Viterbi Ming Hsieh Department of Electrical and Computing Engineering.





ARTIFICIAL INTELLIGENCE DIVISION

ARAM GALYSTAN | DIRECTOR

The Artificial Intelligence division is one of the world's largest artificial intelligence groups. Most division researchers hold graduate degrees in computer science or related disciplines and many also serve as research faculty in the USC Viterbi School of Engineering in the Computer Science, Electrical and Computer Engineering, and Industrial and Systems Engineering departments.

The division is known especially for its work in natural language processing and machine translation, machine learning, and information integration. Researchers also explore biomedical data integration and engineering, social network analysis and computational behavioral modeling, and video, image, and multimedia analysis. In addition, division researchers build working prototypes and partner with academia and industry to create commercial applications.

Research areas include:

Machine learning and data science focusing on developing efficient algorithms to analyze data from a variety of application areas, including biomedical sciences, computational social science, and cybersecurity.

Natural language processing and machine translation for which ISI's AI division is internationally renowned; this includes statistical machine translation, question answering, summarization, ontologies, information retrieval, poetry generation, text decipherment, and more.

Knowledge graphs using artificial intelligence and machine learning techniques to construct large-scale knowledge bases, with applications ranging from combatting human trafficking to predicting cyberattacks.

Knowledge technologies involving interactive knowledge capture, intelligent user interfaces, semantic workflows, provenance, and collaboration, with a focus on scientific data analysis and discovery.

Video, image, and multimedia analysis including document image processing, face recognition, and biometrics.

Biomedical data integration and knowledge engineering providing efficient access to distributed and heterogeneous biomedical data and developing biomedical informatics systems based on cutting-edge AI techniques.





COMPUTATIONAL SYSTEMS AND TECHNOLOGY DIVISION

STEPHEN CRAGO | DIRECTOR

The Computational Systems and Technology division (CS&T) focuses on:

- *Heterogeneous cloud and embedded computing*
- *Microarchitecture, integrated circuits, and advanced electronics and photonics*
- *Quantum computing, communication, and sensing*
- *Reconfigurable computing and wireless networks*
- *Science automation technologies*
- *Spaceborne and ground-based data processing*
- *System software, including operating systems, runtime systems and frameworks, and compilers*
- *Trusted and secure electronics and computing*



The division's researchers, research programmers, and graduate students represent a range of disciplines, including electrical engineering, computer science, physics, and math.

Ongoing initiatives include theoretical adiabatic quantum computing through the USC-Lockheed Martin Quantum Computing Center and hardware security through ISI's Secure and Robust Electronics Center (SURE). CS&T projects include system software for heterogeneous clouds and hardware-software design of unique chips and field-programmable gate arrays. Division researchers are exploring applications and algorithms for processing large-scale and real-time streaming data and solving optimization problems; in addition, researchers work on system engineering for space systems.

CS&T teams are creating wireless networking and edge and fog computing technologies for battlefields and other difficult environments, along with social media platforms for people who lack trustworthy Internet access. Ongoing research on scientific automation tools enables scientists to focus on conducting science—instead of managing data. CS&T automation tools are currently used by astronomers, physicists, and earthquake specialists.



INFORMATICS SYSTEMS RESEARCH DIVISION

CARL KESSELMAN | DIRECTOR

The research agenda of ISI's Informatics Systems Research division focuses on creating new types of sociotechnical systems that enable and accelerate discovery in domains of high societal impact. Launched in 2008, the division takes a holistic, systems-oriented approach, working in areas from basic network services architectures, data management abstractions, computer security, user interfaces, human factors, and domain-specific algorithms. The division specializes in highly collaborative user-driven research in the context of operational, high-impact domain science.

In earlier work, the Informatics Systems Research division developed grid computing infrastructures to support the creation and operation of "virtual organizations" as a foundation for scientific collaboration and discovery. This work, which focused on understanding methods for sharing computing and storage methods, played a role in two Nobel prizes: all the data analysis for discovering the Higgs boson was performed on a global grid infrastructure; the discovery of gravity waves took place on a data grid.



More recently, the division has focused on biomedical applications. Current collaborations range from basic science to clinical use cases in molecular biology, basic neuroscience, neuroimaging, stem cell research, and craniofacial dysmorphism.

Researchers work closely with ISI's highly regarded artificial intelligence, networking, and distributed systems experts, as well as with two of USC's nationally ranked Viterbi School of Engineering departments: Computer Science and the Daniel J. Epstein Department of Industrial and Systems Engineering.

In addition, the division participates in collaborative projects with the faculty of Dornsife College, Keck School of Medicine, and the Herman Ostrow School of Dentistry. Division researchers play a leadership role in the Michelson Center for Convergent Biosciences,

including establishing the new Center for Discovery Informatics as part of the university's convergent biosciences initiative. The division also plays a central role in four international consortiums. Most recently, the Informatics Systems Research division has been a central participant in the effort by the National Institutes of Health to define a shared data infrastructure for biomedical research.



MOSIS

LIFU CHANG | DIRECTOR

The MOSIS service offers silicon fabrication services to semiconductor integrated circuit (IC) designers at universities, research organizations, military and aerospace companies, and commercial design companies for both multiple-project wafer and dedicated wafer projects.

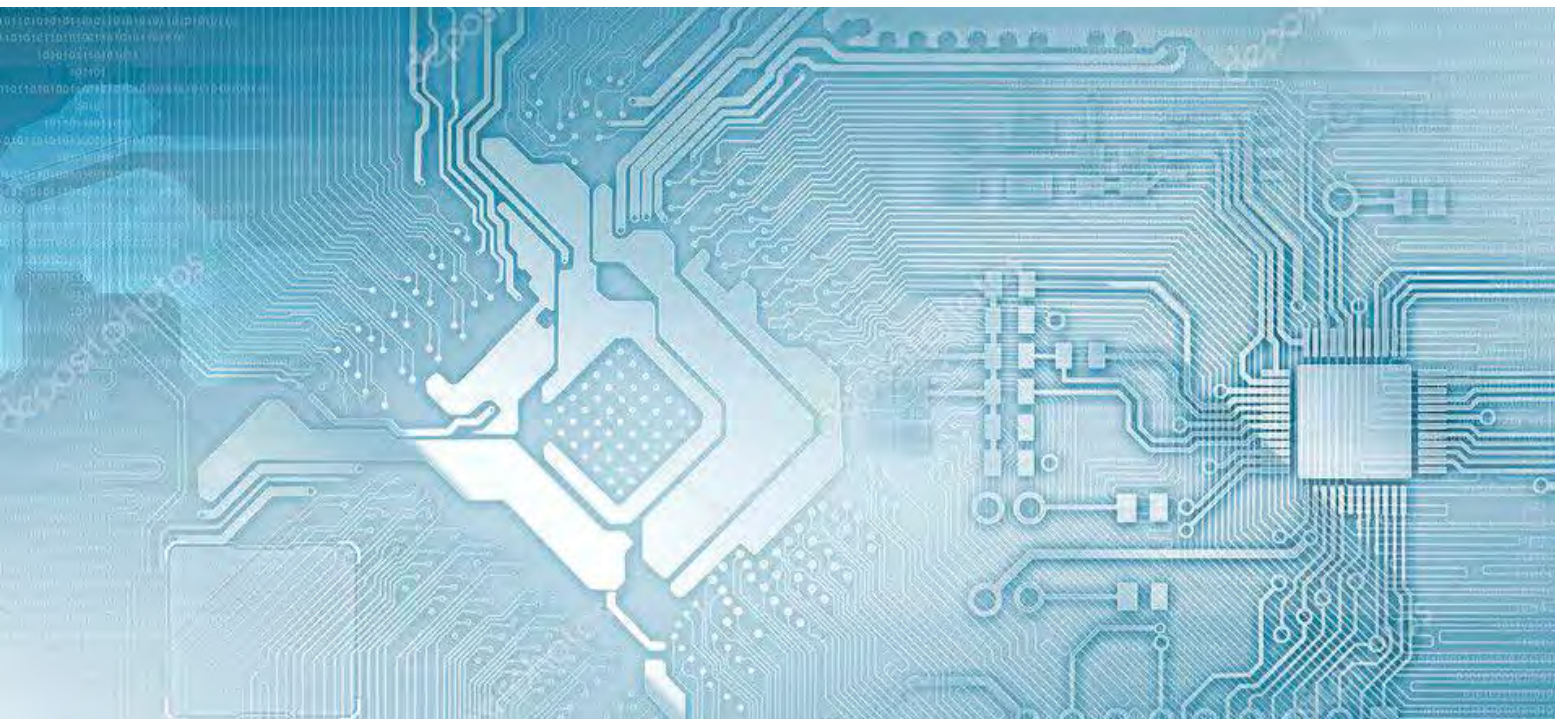
Since its beginning in 1981, MOSIS (“Metal Oxide Semiconductor Implementation Service”) has processed more than 60,000 IC designs.

MOSIS enables IC designers to prototype innovative semiconductor designs within an effective cost structure; it offers a range of processes: CMOS inFET, FD-SOI, Bulk, SiGe, high-voltage BCD, and other specialty processes. MOSIS also enables low-volume production. In 2019, the MOSIS service advanced to MOSIS 2.0, offering service flow automation, a secure cloud design environment, and business model improvements.

Lifu Chang, PhD, a design-for-manufacturability expert, joined ISI as director of MOSIS. Under Dr. Chang, MOSIS started new collaborations with state-of-the-art semiconductor foundries—Intel Custom Foundry and Samsung Foundry—to enhance its fabrication capabilities. The collaborations combine MOSIS’s IC manufacturing expertise with the foundries’ high-performance CMOS fabrication and packaging technologies.

In addition to the commercial service, ISI and MOSIS staff participate in these research programs:

- **MINSEC (Microelectronics Needs and Innovation for National Security and Economic Competitiveness)**, developing a strategic plan for microelectronic innovation centers in the US
- **DARPA CRAFT (Circuit Realization at Faster Timescales)**, developing fast-track circuit-design methods, multiple sources for integrated circuit fabrication, and a technology repository to increase reuse of integrated circuit elements





NETWORKING AND CYBERSECURITY DIVISION

TERRY BENZEL | DIRECTOR

The Networking and Cybersecurity division focuses on understanding the Internet, the theory and practice of distributed computing, approaches to analyzing vulnerabilities, and scientific methods of modeling, experimenting, and evaluating critical infrastructure systems. Areas of research include:

Network and security measurement, analysis, and defenses

Division researchers study methods of observing and collecting network and network security data and behaviors. These methods are used to develop novel networking capabilities and network defenses.

Network infrastructure supporting science and operations

The division develops infrastructure that fosters network- and cybersecurity-enabled collaborations to drive discovery in science for research, education communities, and Internet users domestically and internationally.

Cyber experimentation research, methods, and infrastructure

To evaluate and assess network systems, scientific experiments that model multiple, complex network, environmental, traffic, and behavioral effects and systems are required. Division researchers create models, experimentation frameworks, and tools to enhance the science of cyber experimentation.

Social engineering attacks

Social engineering attacks such as phishing and impersonating continue to grow because organizations' weakest security link is the human. By leveraging metadata from communication channels and using techniques to redirect attackers, division researchers produce new methods for detecting and fingerprinting attack campaigns.

Binary program analysis, vulnerability discovery, and reverse engineering

Binary program analysis is the process of analyzing software programs in their binary form ("executable"). Division researchers use reverse engineering to search for vulnerabilities in software that is released without source code and to assess the security of software products.

Modeling human behavior for cybersecurity and social simulation

Human behavior is a key factor in assessing the effectiveness of organizations' cyber defenses, including their policies. The division's current research observes and models aspects of human behavior to predict likely responses to security postures and the evolution of information in social networks.

Theory and practice of distributed computing

Understanding the foundations of distributed computing is important for designing efficient computational techniques across scientific fields. Some problems that are trivial to solve sequentially are impossible or infeasible to solve in distributed fashion, presenting researchers with problems of deep intellectual—yet practical—interest.



President Carol Folt and Provost Charles Zukoski see the DETERLab.
Photo credit: Shannon Ward



SPACE ENGINEERING RESEARCH CENTER

DAVID BARNHART | DIRECTOR

The Space Engineering Research Center (SERC) is a joint effort with USC's Department of Astronautical Engineering. SERC is dedicated to space engineering, research, and education. The center operates as an "engineering teaching hospital." Professionals and faculty are the "doctors" who work on real satellites and schedule-driven space systems for sponsors; the students are "residents" who work alongside. SERC hosts high school, community college, and undergrad/graduate/PhD students from all over the world. SERC teaches space systems research and offers hands-on training to build, test, and fly actual spacecraft and satellites.

SERC explores new concepts in a variety of space system disciplines. Research areas include small and nanosatellite build and design, rendezvous and proximity operations sensors and techniques, software algorithms for cellular morphology, swarm satellite flight techniques, advanced RF communications systems and techniques, and new concepts in laser propulsion.

In 2019, SERC built and tested a robotic gripper inspired by octopus anatomy and helped develop new standards in rendezvous and proximity operations analysis for the commercial industry. In addition, SERC students built USC's third nanosatellite "Dodona" to carry specialized payloads to test and validate advanced processing and RF capabilities in earth's low orbit.



In Marina del Rey, SERC students work with specialized equipment to create research testbeds, prototypes, and flight hardware for satellite and space activities. In addition, SERC maintains two ground-tracking antennas and a dedicated control station for tracking satellite projects.

In 2019, SERC hosted two California high school students, two Italian graduate students, and six undergraduate cadets from the French Air Force. SERC student research included:

- *Developing specialized light emitting eyeglasses for sleep cycle mitigations for astronauts*
- *Designing and building a gimbal system for a jet-engine-powered lunar lander vehicle*
- *Testing and writing code for a new Cubesat radio transceiver using open source radio code*
- *Developing and simulating a new liquid rocket engine injector system*
- *Working on a completely new physics concept for laser-based propulsion*
- *Developing and testing an IR-based visual pose/range system for a genderless docking adaptor for nanosatellites*

2019 was a growth year for SERC. Two PhD students successfully defended their dissertations and multiple master's degree students graduated and secured jobs in industry. SERC installed new capabilities for electronics design, becoming a center of excellence for US and international students to execute advanced degrees and perform research.



SECURE AND ROBUST ELECTRONICS CENTER

MATTHEW FRENCH | DIRECTOR

The need for hardware security research stems from the complexity of the developmental process and supply chain associated with microelectronics. Multi-billion-dollar ecosystems have evolved around the development of application-specific integrated circuits (ASICs) including third-party intellectual property hardware cores, CAD tools, verification tools, etc. These examine whether the device operates as intended, but they do not provide trust or “beyond functional verification” to determine if additional unwanted functionality, such as hardware-level Trojans, are present in the device. If a device is produced in a trusted manner, it increasingly encounters complex attacks in the field, where counterfeiting and reverse engineering are deployed to steal intellectual property.

Hardware security often overlaps with resiliency and reliability; the now-famous Row hammer attack illustrated how a reliability vulnerability could be exploited to circumvent security protections.



The Secure and Robust Electronics Center (SURE) recognizes that effective hardware needs to comprehensively address trust, security, resiliency, and reliability. In 2019, SURE led research in logic locking for ASICs, addressing supply chain obsolescence for programmable logic devices, and developing tools to provide counterfeit detection and reliability analysis of Field Programmable Gate Array (FPGA) devices.

There is growing concern that the globalization of the ASIC manufacturing process can lead to hardware intellectual property being “leaked” or stolen, or that hardware-level Trojans could be inserted. SURE researchers, working with faculty from USC’s Electrical and Computer Engineering department, developed the Mirage toolset, which applies a common set of security and overhead metrics to uniformly compare different logic locking techniques. The Mirage tools also include a design space search engine that determines what type of technique is best to apply to a given input circuit.

The counterfeiting of electronics is also a problem. Many government systems have lifetimes of up to 15 years and it is not uncommon for commercial electronics to sunset during a mission’s lifetime. This creates a lucrative demand for obsolete parts, incentivizing counterfeit. The SURE Center researches and develops exhaustive functional tests for FPGA devices and this year achieved a milestone, with Georgia Tech Research Institute, achieving full-scale solutions for all Xilinx parts with exhaustive coverage in realistic testing times. These tools enable the testing of all a device’s configurable elements to independently determine if the device is counterfeit.

For some systems, parts are no longer available. SURE researchers address the challenge of replacing deprecated in-system programmable hardware components when source code, programming files, and schematics are no longer available. The team is developing a methodology to lift binaries from old programmable device types, extract their contents, remap them to modern technology, and verify the resulting system.



USC-LOCKHEED MARTIN QUANTUM COMPUTING CENTER

DANIEL LIDAR | SCIENTIFIC AND TECHNICAL DIRECTOR

FEDERICO SPEDALIERI | OPERATIONAL DIRECTOR

ROBERT LUCAS | DIRECTOR OF RESEARCH PARTNERSHIPS

The USC-Lockheed Martin Quantum Computing Center (QCC) houses a D-Wave 2X quantum annealing system, manufactured by D-Wave Systems, Inc.

QCC was the first organization outside D-Wave to house and operate its own system; the center has conducted pioneering research on three different generations of these early noisy, intermediate-scale quantum NISQ processors.

Operating quantum computing systems is demanding: the temperature of these systems needs to be kept near absolute zero (-273 degrees Celsius) and the devices must be electromagnetically shielded to protect the fragile quantum states from degradation by external noise. The main thrust of the research conducted at QCC has been to understand how this noise can adversely affect the computational power of these devices.

Faculty, researchers, and students perform basic and applied research to NISQ computing devices and collaborate with researchers around the world.

Five USC doctoral students are currently engaged in quantum computing research at the QCC.



President Carol Folt and Provost Charles Zukoski visit the USC-Lockheed Martin Quantum Computing Center

Photo credit: Shannon Ward





CENTER FOR VISION, IMAGE, SPEECH, AND TEXT ANALYTICS

Wael AbdAlmageed | CO-DIRECTOR

Scott Miller | CO-DIRECTOR

Prem Natarajan | FOUNDING DIRECTOR

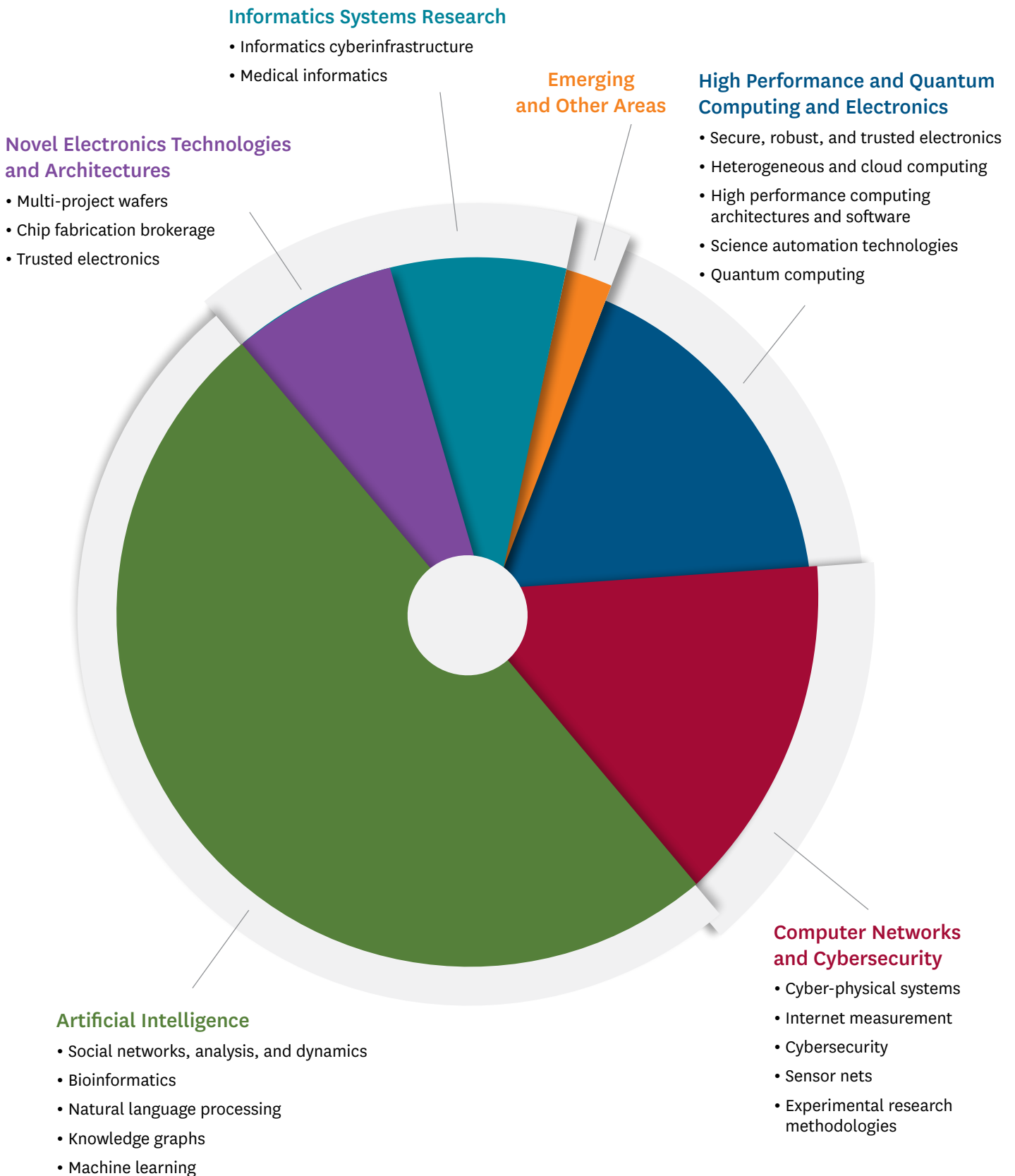
The Center for Vision, Image, Speech, and Text Analytics (VISTA) works in natural language processing, computer vision, biometrics, optical character recognition, face recognition, speech and text analytics, and multimedia forensics. The center's research addresses pressing challenges, including improving the security of biometric systems, making deep learning algorithms more robust against adversarial attacks, and identifying deepfake videos and fake news.

Democratized artificial intelligence technology, where anyone can generate Hollywood-like special effects using off-the-shelf applications on a personal computer, has increased the spread of malevolent image manipulation. A deepfake video is a modified video in which the original face is replaced with a victim's face shown saying something he or she never said. The deepfake video is usually very realistic, so viewers believe that the swapped subject is the actual person. Deepfakes represent a potential threat to national security, stock markets, and to child and personal safety. Under the DARPA MediFor program, VISTA researchers have developed an accurate deepfake detection system, which has been transitioned to news and entertainment agencies.

Applications of deep neural networks (DNN) are spreading rapidly in areas including health care, law enforcement, mobile phones, and self-driving vehicles. However, regardless of the sophistication and human-like accuracy of artificial intelligence models, these models are brittle and can easily be attacked or deceived into producing erroneous outputs. Funded by the DARPA GARD program, VISTA researchers are developing modality-agnostic defenses, including a unified framework using adversarial learning that incorporates extracting disentangled learned representations and various classification regularization techniques. In addition, the center is developing multiple modality-specific defenses for images, video, audio, and text.



RESEARCH FUNDING BY AREA



ISI IN NUMBERS

ISI staff, faculty, and students 400+		96 USC PhD Students	23 USC PhD advisors	75+ USC master's degree students
28 USC professors		Computer Science Spatial Sciences Physics and Astronomy Astronautical Engineering Electrical and Computer Engineering Industrial and Systems Engineering Civil and Environmental Engineering		
ISI staff and faculty				Research funding in 2019
7 MBA degrees	64 master's degrees	70 bachelor's degrees	107 PhDs degrees	\$108.8 million
New research grants in 2019		Quantum computers		Summer intern mentors
44		1		20
2019 visiting summer interns				
10 undergraduate students		5 master's degree students		24 PhD students
ISI locations				Total office space
 Marina del Rey, CA  Arlington, VA  Waltham, MA				147,576 square feet

NEW DIRECTORS

ISI was privileged to bring three outstanding individuals to key roles in 2019. They manage vital ISI operations and a new lab. The operation and growth of the MOSIS service, the Application Specific Intelligent Computing Lab, and the forward-looking computing environment of ISI continue to increase ISI's impact and support the institute as a world-class research organization.



LIFU CHANG | DIRECTOR, MOSIS

Lifu Chang leads MOSIS, which he envisions as the semiconductor manufacturing and yield enablement provider for advanced chip designs. He has more than 20 years of work experience in semiconductor device, process technology, Electronic Design Automation (EDA), design enablement, and Design for Manufacturability (DFM). His work resulted in patents that laid critical foundations of multiple DFM EDA tool flows in use worldwide today. Lifu obtained his PhD degree from Purdue University in 1996 in the area of solid-state electronics. He has published nearly 50 journal and conference papers and holds multiple US patents. He received the Elected Membership Award from American Association for the Advancement of Science in 2001 and is a Senior Member of IEEE.



AJEY JACOB | DIRECTOR, APPLICATION SPECIFIC INTELLIGENT COMPUTING LAB

Ajey Jacob brings 16 years of research and development experience from Intel and GLOBALFOUNDRIES (GF). At Intel, Ajey worked as a senior research scientist for the component research group on beyond-CMOS materials and device technology. At GF, Ajey managed research for 14, 10, 7 nm technology nodes and More than Moore differentiating and disruptive elements. Ajey received his PhD in physics from the Chalmers University of Technology/Gothenburg University, Sweden, in 2002. He holds over 135 issued USPTO patents. In 2013 and 2016 the Semiconductor Research Corporation awarded him the Mahboob Khan Outstanding Industry Liaison Award. Ajey joined ISI's Computational Systems and Technology division to build a new research group, the Application Specific Intelligent Computing Lab, to focus on hardware development in areas including neuromorphic computing, quantum information processing, and smart lab on a chip using alternate state variables such as electronics, photonics, and spintronics.



EILEEN LU | CHIEF INFORMATION OFFICER AND DIRECTOR OF COMPUTING AND INFORMATION SERVICES

As CIO, Eileen Lu provides leadership in information technology to advance ISI's research, including strategic planning, policy and process development, and implementation of innovative and secure IT services. Eileen is also the Director of Computing and Information Services, which oversees and executes the institute's IT operations, including infrastructure, network security, client services, communications/messaging, and project management. Eileen has more than 20 years of experience in the IT industry. She has led secure public cloud implementations for scientific research, large-scale enterprise cloud email migrations, enterprise-level user-support, and process improvement and automation.

NEW RESEARCH LEADS

Three ISI researchers were promoted to new roles in 2019.



MARJORIE FREEDMAN | RESEARCH TEAM LEADER

Marjorie Freedman holds degrees in linguistics and computer science from Cornell University. At ISI, she is the principal investigator of DARPA's Active Interpretation of Disparate Alternatives (AIDA), Active Social Engineering Defense (ASED), and Knowledge-directed Artificial Intelligence Reasoning Over Schemas (KAIROS) efforts. Marjorie is also contributing to the MICS (Common Sense) and ADAM (grounded language learning) efforts. Her work includes tailoring speech recognition and optical character recognition systems for use in an information extraction pipeline. In addition, she is exploring the impact of uncertainty in anaphora resolution to downstream tasks and working with vision researchers to understand and address the challenges of mapping the output of vision analytics to classic information extraction ontologies. Her work in information extraction has explored how to address limited training data, including fusing rule-based and learned systems, exploring alternative approaches to annotation, and measuring the impact of coreference in bootstrap learning for information extraction.



MAYANK KEJRIWAL | RESEARCH LEAD

Mayank Kejriwal is a Research Assistant Professor of Industrial and Systems Engineering. His research has been funded by DARPA and private endowments; it has been covered in the press for its success in spawning real-world systems for tackling human trafficking. Before joining ISI in 2016, he earned his PhD from the University of Texas at Austin. His dissertation *Populating a Linked Data Entity Name System* was awarded the Best Dissertation Award by the Semantic Web Science Association in 2017. He is also the author of *Domain-specific Knowledge Graph Construction* (Springer), which has been downloaded thousands of times in the last year and is available internationally. Mayank is a passionate advocate of using AI technology for social good and collaborates with domain experts to build such systems.



FRED MORSTATTER | RESEARCH LEAD

Fred Morstatter is a Research Assistant Professor of Computer Science. He graduated from Arizona State University with a PhD in Computer Science in 2017; he joined ISI shortly after as a computer scientist. Fred's research interests include computational social science, data science, data mining, and machine learning. At ISI, Fred is deeply involved in the SAGE project, Synergistic Anticipation of Geopolitical Events. SAGE attempts to predict future events based on forecasts from a set of human participants, non-expert "forecasters," who interact with unique AI data tools and machine learning methods on the SAGE platform. Fred has also been involved in DARPA's Understanding Group Biases project where he identifies the biases inherent to cultural groups based upon their online behavior.

NEW FACULTY APPOINTMENTS

31 ISI researchers held faculty appointments at USC in 2019. Many are in the Viterbi School of Engineering Department of Computer Science. ISI researchers also hold faculty appointments in these departments:

- *Ming Hsieh Department of Electrical and Computer Engineering*
 - *Daniel J. Epstein Department of Industrial and Systems Engineering*
 - *Sonny Astani Department of Civil and Environmental Engineering*
- *Department of Astronautical Engineering*
 - *Department of Physics and Astronomy*
 - *Spatial Sciences Institute*

DEPARTMENT OF COMPUTER SCIENCE



Fred Morstatter



Nanyun Peng



Jay Pujara



Srivatsan Ravi



Satish Kumar Thittamaranahalli

DANIEL J. EPSTEIN DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING



Mayank Kejriwal

MING HSIEH DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE



Peter Beerel



Shri Narayanan



Xiang Ren

JOINT APPOINTMENTS WITH ISI



NEW ISI TEAM MEMBERS

In 2019, 42 computer scientists, researchers, directors, programmers, and staff members joined ISI.

MARINA DEL REY

Joseph Barnes, Jr., *Research Systems Administrator*

Rishbha Bhagi, *Public Communications Specialist*

Sahib Bhai, *Computer Services Consultant*

Leticia Calleros, *Contracts & Grants Administrator*

Phil Chandler, *Senior Computer Systems Engineer*

Lifu Chang, *Director of MOSIS Services*

Craig Charlton, *Senior Systems Engineer*

Winsome Chee, *Post Award Administrator*

Justin Cho, *Programmer Analyst*

Nalin Gujuluva, *Budget/Business Analyst*

Xiao Guo, *Programmer Analyst*

Thomas Halverson, *Postdoctoral Scholar*

Harun Ibrahim, *Research Programmer*

Filip Ilievski, *Computer Scientist*

Jun Liu, *Research Programmer*

Eileen Lu, *Chief Information Officer, Director of CIS*

Arjun Mani, *Postdoctoral Scholar*

Kimberly Olmedo, *Paralegal*

Phyllis O'Neil, *Senior Grant Writer*

Maximiliano Osorio, *Research Programmer*

Loic Pottier, *Postdoctoral Scholar*

Mohammad Rostami, *Computer Scientist*

Ahmed Shabbir, *Windows System Administrator*

Ryan Tanaka, *Programmer Analyst*

Sharon Tang, *Infrastructure Operations Manager*

Rob Younes, *Facilities Project Manager*

BOSTON

Shantanu Agarwal, *Research Programmer*

Deniz Beser, *Research Programmer*

Michelle Bonner, *Administrative Assistant*

Valentino Crespi, *Senior Computer Scientist*

Steven Fincke, *Computer Scientist*

Alex Hedges, *Research Programmer*

Christopher Jenkins, *Research Programmer*

Elizabeth Lee, *Research Programmer*

Michael Spector, *Research Programmer*

ARLINGTON

Darrell Best, *Research Programmer*

Priyatam Chiliki, *VLSI Engineer*

Michael Collins, *Senior Computer Scientist*

Ajey Jacob, *Director, Application Specific Intelligent Computing Lab*

Will Richardson, *Research Programmer*

Ting-Yuan Sung, *Research Programmer*

Peng Xie, *Computer Scientist*

2019 RESEARCH GRANTS AND AWARDS

This listing of ISI's externally funded research projects illustrates the diversity of the institute's research sponsors, including federal agencies, industry, and state government. Likewise, the scope of research spans computing technology, electrical engineering, artificial intelligence, and other fields. The projects here range from small satellite science to the study of craniofacial deformities, from verifying the integrity of field-programmable gate arrays to automatic language translation. Project funding ranges from tens of thousands of dollars to millions; some projects started this year while others continue decades of research effort. 2019 was another strong year for the institute.

AIR FORCE RESEARCH LABORATORY

ISI subcontract for InferLink STTR on Autonomous Cyber Defense

(via InferLink Corp)

Principal Investigator: Christophe Hauser

ALFRED P. SLOAN FOUNDATION

Preliminary study of semi-automated methods for privacy-preserving linkage discovery

(via Actuate Innovation)

Principal Investigator: Pedro Szekely

ARAMCO SERVICES COMPANY

Deep Learning Architectures for Characterization, Performance,
Prediction and Optimization of Unconventional Reservoirs

Principal Investigator: Young Cho

ARMY RESEARCH OFFICE

Structured Optical Receiver Design for Active Sensing at the Quantum

Principal Investigator: Jonathan Habif

CALIFORNIA DEPARTMENT OF PUBLIC HEALTH

California Cancer Registry

Principal Investigator: Pedro Szekely

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY (DARPA)

Expert Comprehensive Experimental Evaluation of Defenses (EXCEED) for XD3

Principal Investigator: Erik Kline

Accurate and Precise Recognition of Obscured Payloads in Operational Systems

Principal Investigator: Erik Kline

BESPOKE: Learning Individualized Interventions for Human Performance

Principal Investigator: Kristina Lerman

Discovering Common Sense from Video, Images, Text and Knowledge Bases

Principal Investigator: Ralph Weischedel

CASPER: Compiler Abstractions Supporting High Performance on Extreme-scale Resources

Principal Investigator: John Paul Walters

Macroscopic Models for Reproducibility and Replicability of Research Claims
in the Social and Behavioral Sciences (MACRO-SCORE)

Principal Investigator: Jay Pujara



DEFENSE ADVANCED RESEARCH PROJECTS AGENCY (DARPA)

Generating Novelty in Open-world Multi-agent Environments (GNOME)

Principal Investigator: Mayank Kejriwal

Abduction to Demonstrate an Articulate Machine

Principal Investigator: Ryan Gabbard

CORAL: Combined Representations for Adept Learning

Principal Investigator: Aram Galstyan

LESTAT: Discovering Schemas from Diverse Data

Principal Investigator: Marjorie Freedman

Multi-modal Open World Grounded Learning and Inference

Principal Investigator: Pedro Szekely

DEPARTMENT OF DEFENSE

WRT-1012: Global Positioning Systems-Mission Engineering and Integration of Emerging Technologies

(via Stevens Institute of Technology)

Principal Investigator: Michael Orosz

AI-based Forensic Object Classification and Identification System

(via SRI International)

Principal Investigator: Craig Knoblock

DEPARTMENT OF ENERGY

Replacing Aging Programmable Electronics Rapidly

(via Honeywell Federal Manufacturing & Technologies LLC)

Principal Investigator: Andrew Schmidt

Resource Efficient Quantum Simulations on NISQ Devices: Advancing the State of the Art

Principal Investigator: Itay Hen

FUJITSU LIMITED

Benchmarking the Fujitsu Digital Annealer

Principal Investigator: Itay Hen

INTELLIGENCE ADVANCED RESEARCH PROJECTS ACTIVITY (IARPA)

CLEAR: Cross-Lingual Event & Argument Retrieval

Principal Investigator: Elizabeth Boschee

2019 RESEARCH GRANTS AND AWARDS (continued)

NASA

SWARMS: JPL Student Undergraduate Research Proposal (JPLSURP 2020)

(via Jet Propulsion Laboratory)

Principal Investigator: David Barnhart

NATIONAL SCIENCE FOUNDATION

GREPSEC IV: Underrepresented Groups in Security Research

Principal Investigator: Terry Benzel

ABI Development: Building a Pipeline for Validation, Curation and Archiving of Integrative/Hybrid Models

(via Rutgers University)

Principal Investigator: Carl Kesselman

2019 NSF Workshop on Connecting Large Facilities and Cyberinfrastructure

Principal Investigator: Ewa Deelman

Convergence Accelerator Phase I (RAISE): Leveraging Financial and Economic Data Business OKN

Principal Investigator: Jay Pujara

CCRI: Medium: DNS, Identity, and Internet Naming for Experimentation and Research (DIINER)

Principal Investigator: John Heidemann

Collaborative Research: CyberTraining: Implementation: Small: Integrating core CI literacy and skills into university curricula via simulation-driven activities

Principal Investigator: Rafael Ferreira Da Silva

SaTC-CCRI: Collaborative Research: Sharing Expertise and Artifacts for Reuse through Cybersecurity CommunityHub (SEARCCH)

Principal Investigator: Terry Benzel

NATIONAL INSTITUTES OF HEALTH

Center for Genomic Studies on Mental Disorders (Year 12)

(via Rutgers University)

Principal Investigator: Yigal Arens

Page III: Population Architecture using Genomics and Epidemiology

(via Rutgers University)

Principal Investigator: Jose Luis Ambite

University of Maryland NIH Data Commons Facilitation Center

(via University of Maryland)

Principal Investigator: Carl Kesselman

USC FaceBase III Craniofacial Development and Dysmorphology Data Management and Integration Hub

Principal Investigator: Carl Kesselman

NAVAL SURFACE WARFARE CENTER CRANE

Comprehensive Independent FPGA Functional Testing Tool

Principal Investigator: Matthew French

SANDIA NATIONAL LABORATORIES

Subject Matter Expert Support to Sandia National Laboratories for SMC/SY Programs

Principal Investigator: Michael Orosz

US ARMY

ALFRED: Tensor Factorization for Dynamic Recommendation for Alfred the Butler

(via Aptima, Inc.)

Principal Investigator: Kristina Lerman

US GOVERNMENT – OTHER

Embedded Processor Software Architecture Trade Studies

(via Boeing Company)

Principal Investigator: John Paul Walters

BIGFOOT

(via Peraton Inc.)

Principal Investigator: John Paul Walters

BreadFan

(via Boeing Company)

Principal Investigator: Matthew French

CORPORATE AWARDS

Google AI Faculty Research Award

Recipient: Xiang Ren

JP Morgan AI Research Award

Recipient: Xiang Ren

Yahoo! Faculty Research Engagement Program (FREP)

Recipient: Mayank Kejriwal

2019 AWARDS AND HONORS

ISI researchers earned honors throughout the year. Recognition for their contributions came from professional associations, foundations, government agencies, and universities.

ISI ACHIEVEMENT AWARDS

Wael AbdAlmageed

ISI Institute Achievement Award for outstanding academic achievements and research leadership that established his group as a leader in biometrics, face recognition, and visual misinformation detection and characterization

James Whalen and John Damoulakis

ISI Institute Achievement Award for outstanding work developing and executing a strategic plan to propel MOSIS to the center stage of the semiconductor industry



Craig Knoblock presents the ISI Institute Achievement Award to Wael AbdAlmageed

ISI RESEARCH AWARDS

ISI Research Awards and the **Keston Research Awards**, totaling five this year, provide funding for research in novel areas of research. Keston Research Awards support intellectually intriguing research with societal impact, with the potential to produce a tangible result within a year or so. ISI Research Awards support new research, in areas where outside sources of funding are not yet available.

Iacopo Masi and Wael AbdAlmageed

Keston Research Award to support their new project, Fighting Misinformation: An Internet System for Detecting Fake Face Videos

Pedro Szekely, with Childrens Hospital LA and LA Cancer Surveillance Program collaborators

Keston Research Award to support their new project, Mining Cancer Registry Data to Improve Long Term Effects of Pediatric Cancer Treatment

Jonathan May

ISI Research Award to support the Universal Translators for Asylum Seekers at the Border project

Andrew Rittenbach, Priyatam Chiliki, and Dev Shenoy

ISI Research Award to support their new work in automating programmability of hybrid digital-analog hardware for stochastic cell simulation in biological systems

Joel Mathew and Ulf Hermjakob

ISI Research Award to support their new Advancing No-Resource Languages project

PROFESSIONAL AND ACADEMIC RECOGNITION

Ewa Deelman

American Association for the Advancement of Science (AAAS) Fellow

Emilio Ferrara

ACM Senior Member

Yolanda Gil

Inaugural EarthCube Legacy Award

Yolanda Gil, Shrikanth Narayanan et al.

Best Paper Award, 10th International Conference on Indoor Air Quality, Ventilation and Energy Conservation in Buildings

Wes Hardaker

Distinguished Paper, Proceedings of the Internet Measurement Conference

Connor Imes

Karsten Schwan Best Paper Award, 16th IEEE International Conference on Autonomic Computing (ICAC)

Mayank Kejriwal

Mensa Copper Black Award for Creative Achievement for scientific work fighting human trafficking and broader impacts of technology

Mensa Foundation's Intellectual Benefits to Society Award for research and continued outreach on human trafficking

Red Judge, IBM Watson AI XPRIZE for serving as a mentor and auditing judge for Project Spock

Deborah Khider, Daniel Garijo, Yolanda Gil et al.

Paper selected as one of 15 contributions to the Grand Challenges in the Earth and Space Sciences for the Centennial Collection of the Review of Geophysics Journal of the American Geophysical Union

Iacopo Masi

2019 Computer Vision and Pattern Recognition Conference Outstanding Reviewer special mention

Shrikanth Narayanan

USC Associates Award for Creativity in Research and Scholarship

Distinguished Alumni Award, College of Engineering—Guindy (India)

Best Paper Award (co-author), 5th International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESC2018)

American Institute for Medical and Biological Engineering (AIMBE) Fellow

Xiang Ren

Forbes Asia 30 Under 30: Asia – Healthcare and Science

Mats Rynge, Karan Vahi, Ewa Deelman et al.

Best Technical Paper, PEARC19 Conference

Phil Andrews Most Transformative Contribution Award, PEARC19 Conference

Pedro Szekely

DARPA AI Colloquium: Advances in Natural Language Understanding invited presentation

Satish Kumar Thittamaranahalli

USC Stevens Technology Commercialization Award for work on multi-agent coordination in warehouses and fulfillment centers

2019 PHD GRADUATES

ISI proudly congratulates these 2019 graduates for completing their PhD degrees.

EMILIO FERRARA, ADVISOR

Palash Goyal

Affiliation: Samsung

Position: Senior AI Scientist

Dissertation: *Graph Embedding Algorithms for Attributed and Temporal Graphs*

ARAM GALSTYAN, ADVISOR

Sahil Garg

Affiliation: Icahn School of Medicine at Mount Sinai

Position: Postdoctoral Scholar

Dissertation: *Hashcode Representations of Natural Language for Relation Extraction*

CRAIG KNOBLOCK, ADVISOR

Yuan Shi

Affiliation: Turing Video

Position: Research Scientist

Dissertation: *Learning to Adapt to Sensor Changes and Failure*

KRISTINA LERMAN, ADVISOR

Shin-Chieng Ngo (Xin-Zeng Wu)

Affiliation: Blume Global, Inc.

Position: Data Scientist

Dissertation: *Global Consequences of Local Information Biases in Complex Networks*

SHRIKANTH NARAYANAN, ADVISOR

Nikolaos Malandrakis

Affiliation: Amazon

Position: Research Scientist

Dissertation: *Generating Psycholinguistic Norms and Applications*

Md Nasir

Affiliation: Microsoft

Position: Researcher/Data Scientist

Dissertation: *Interaction Dynamics and Coordination for Behavioral Analysis in Dyadic Conversations*

Anil Ramakrishna

Affiliation: Amazon

Position: Applied Scientist

Dissertation: *Computational Models for Multidimensional Annotations of Affect*

Ruchir Travadi

Affiliation: Apple

Position: Research Scientist

Dissertation: *Efficient Estimation and Discriminative Training for the Total Variability Model*

Colin Vaz

Affiliation: Behavioral Signal Technologies, Inc.

Position: Senior Machine Learning Engineer

Dissertation: *Matrix Factorization for Noise-Robust Representation of Speech Data*

CLIFF NEUMAN, ADVISOR

Yatin Wadhawan

Affiliation: Microsoft

Position: Software Engineer

Dissertation: *Defending Industrial Control Systems: An End-to-End Approach for Managing Cyber-Physical Risk*

2019 POSTDOCTORAL SCHOLARS

ISI welcomes postdocs who come to work with experienced mentors and develop their independent and collaborative research skills while contributing to ongoing research projects.

EWA DEELMAN, ADVISOR

Loic Pottier | *Parallel and scheduling algorithms*

EMILIO FERRARA, ADVISOR

Goran Muric | *Artificial intelligence and social networks*

Homa Hosseinmardi | *Machine learning*

Anna Sapienza | *Machine learning*

ARAM GALSTYAN, ADVISOR

Daniel Benjamin | *Behavioral decision science*

K.S.M. Tozammel Hossain | *Machine learning*

ITAY HEN, ADVISOR

Thomas Halverson | *Adiabatic quantum computing and quantum simulations*

Arjun Mani | *Quantum computing*

Brendan Reid | *Adiabatic quantum computing*

SHRIKANTH NARAYANAN, ADVISOR

Michelle L'Hommedieu | *Behavioral modeling*

Benjamin Girault | *Behavioral modeling*

2019 VISITING SCHOLARS

ITAY HEN, HOST

Milad Marvian (MIT) | *Quantum computing*

Lev Barash (Landau Institute of Theoretical Physics) | *Computational physics*

Elizabeth Crosson (University of New Mexico) | *Quantum computing*

Michael Jarrett (Perimeter) | *Quantum computing*

Amir Kalev (University of Maryland) | *Quantum computing*

Jack Raymond (D-Wave Systems) | *Quantum computing*

KRISTINA LERMAN, HOST

Allon Percus (Claremont Graduate University) | *Complex networks*

2019 VISITING PHD STUDENTS

EMILIO FERRARA, HOST

Hanyu Chwe (Northeastern University) | *Machine learning*

Boli Fang (Indiana University) | *Machine learning*

Desheng Hu (Northeastern University) | *Machine learning*

Leonardo Nizzoli (Illinois Institute of Technology, University of Pisa) | *Machine learning*

Mark Heimann (University of Michigan) | *Machine learning*

Mariam Nouh (Oxford University) | *Machine learning*

Hao Peng (University of Michigan) | *Machine learning*

Miley Yao (SUNY Albany) | *Machine learning*

KRISTINA LERMAN, HOST

Tiago Santos (Graz University, Austria) | *Online communities*

MATTHEW FRENCH, HOST

Pantea Kiaei (Virginia Tech) | *FPGA assurance*

Gaurav Kolhe (George Mason University) | *Logic obfuscation, malware detection*

Sanil Rao (Carnegie Mellon University) | *High performance computing, machine learning*

2019 PHD STUDENTS

The institute's cohort of PhD students exceeds passes those of previous years, not only in number, but also in the breadth of their research interests.

JOSE LUIS AMBITE, ADVISOR

Ruhollah Shemirani | *Computational algorithms for genetics*
Dimitrios Stripelis | *Large-scale data integration and federated learning*

DAVID BARNHART, ADVISOR

Rahul Rughani | *Genetic algorithms applied to satellite swarms*

STEPHEN CRAGO, ADVISOR

Yu-An (Victor) Chen | *Dynamic resource allocation in real-time systems*
Geoffrey Tran | *Fault tolerance in cloud-based analytics*

EWA DEELMAN, ADVISOR

Tu Mai Anh Do | *Workflow systems in high performance computing environments*
Patrycja Krawczuk | *Applying artificial to scientific workflow systems*
Georgios Papadimitriou | *Execution and anomaly detection for distributed workflow systems*

EMILIO FERRARA, ADVISOR

Emily Chen | *Machine learning*
Di (Diana) Huang | *Machine learning*
"Julie" Yi Ou Jiang | *Machine learning*
Hsien-Te Kao | *Machine learning*
Akira Matsui | *Machine learning*
Alex Spangher | *Machine learning*
Shen-Yan | *Machine learning*
Yilei Zeng | *Machine learning*

ARAM GALSTYAN, ADVISOR

Sami Abu-El-Haija | *Machine learning*
Shushan Arakelyan | *Machine learning*
Robert Brekelmans | *Machine learning*
Sarik Ghazarian Ghalemaleki* | *Natural language processing and dialog modeling*
Rujun Han* | *Artificial intelligence*
Hrayr Harutyunyan | *Machine learning*
Neal Lawton | *Machine learning*
Ninareh Mehrabi | *Artificial intelligence*
Mehrnoosh Mirtaheri Feijani | *Machine learning*
Kyle Reing | *Machine learning*
Serban Stan | *Machine learning*

JOHN HEIDEMANN, ADVISOR

Calvin Ardi | *Network measurement and data sharing for anti-phishing*
Guillermo Baltra | *Internet outage detection*
Asma Enayet | *Internet outage detection through analysis of passive sources*
Hang Guo | *Network measurement and IoT security*
Basileal Imana | *Network measurement, privacy, and fairness*
Abdul Qadeer | *Large-scale data processing of network data*
A. S. M. Rizvi | *Network measurement and anycast for anti-distributed denial of service attacks*
Lan Wei | *Network measurement and anycast reliability*

ITAY HEN, ADVISOR

Zoe Gonzalez Izquierdo | *Physics/quantum computing*
Lalit Gupta | *Physics/computational computing*
Xiaofan Zhang | *Physics/quantum computing*

CRAIG KNOBLOCK, ADVISOR

Minh Tran Xuan Pham | *Unsupervised data cleaning for building knowledge graphs*
Basel Shbita | *Building knowledge graphs from historical graphs*
Binh Vu | *Semantic data representation and knowledge graphs*

KRISTINA LERMAN, ADVISOR

Nazanin Alipourfard | *Machine learning, data science*
Nathan Bartley | *Artificial intelligence*
Yuzi He | *Data science, artificial intelligence fairness*
Negar Mokherian | *Artificial intelligence*
Nazgol Tavabi | *Artificial intelligence*

JONATHAN MAY, ADVISOR

Nada Aldarrab | *Machine translation, information retrieval*
Mozhdeh Gheini | *Machine translation, transfer learning, visualization*
Thamme Gowda | *Machine translation, information representation*
Meryem M'Hamdi | *Information extraction, event recognition, multilingual representations*
Xusen Yin | *Dialogue and interaction*
Xiyang Zhang | *Information extraction, event sequence recognition*

JELENA MIRKOVIC, ADVISOR

Sima Arasteh | *Binary analysis*
Xiyue Deng | *Malware analysis*
Sivaramkrishnan Ramanathan | *Network management and cybersecurity*
Rajat Tandon | *Distributed denial of service attacks*
Nicolaas Weideman | *Binary analysis*
Wei-Cheng Wu | *Binary analysis*

FRED MORSTATTER, ADVISOR

Yuzhong Huang | *Representation learning, forecasting, crowdsourcing*

SHRIKANTH NARAYANAN, ADVISOR

Victor Ardulov | *Human-centered machine learning*
Sabyasachee Baruah | *Computational media intelligence*
Brandon Booth | *Human-centered machine learning*
Tiantian Feng | *Human-centered machine learning*
Nikolaos Malandrakis | *Low-resource natural language processing*
Victor Martinez Palacios | *Natural language processing*
Karel Bogomir Mundnich Batic | *Human-centered machine learning*
Amrutha Nadarajan | *Human-centered machine learning*
Pavlos Papadopoulos | *Human-centered machine learning*
Raghuv eer Peri | *Signal processing, audio, speech*
Karan Singla | *Natural language processing*

PREM NATARAJAN, ADVISOR

Jiaxin Cheng | *Computer vision*
I-Hung Hsu* | *Natural language processing*
Ayush Jaiswal | *Representational learning, adversarial learning, multimedia integrity analysis*
Soumyaroop Nandi | *Computer vision*
Ekraam Sabir | *Multimedia Integrity*
Emily Sheng* | *Towards Fairness in Natural Language Processing*

CLIFF NEUMAN, ADVISOR

Abdulla Alwabel | *Enabling symbolic execution string comparison during code-analysis of malicious binaries*
Yatin Wadhawan | *Risk assessment, mitigation of cyber-physical attacks in industrial control systems*

NANYUN PENG, ADVISOR

Sidi Lu | *Natural language generation, creative generation*
Mingyu Ma | *Event schema and event ontology*
Te-Lin Wu | *Machine commonsense reasoning*
Johnny Wei | *Low-resource machine translation and evaluation metrics*

JAY PUJARA, ADVISOR

Pegah Jandaghimeibodi | *Artificial intelligence*
Alexander Kezar | *Artificial intelligence*
Pei Zhou | *Artificial intelligence*

XIANG REN, ADVISOR

Woojeong Jin | *Machine learning, temporal reasoning, knowledge graphs*
Yuchen Lin | *Natural language processing, commonsense reasoning*
Jun Yan | *Natural language processing, commonsense reasoning*
Qinyuan Ye | *Natural language processing, learning with explanation*

FEDERICO SPEDALIERI, ADVISOR

Hannes Leipold | *Quantum computing*

PEDRO SZEKELY, ADVISOR

Majid Ghasemi Gol | *Knowledge graphs, table understanding*
Ehsan Qasemi | *Knowledge graphs, machine commonsense*
Avijit Thawani | *Machine commonsense*
Peifeng Wang* | *Commonsense reasoning*

SATISH KUMAR THITTAMARANAHALLI, ADVISOR

Ang Li | *Computational physics, materials science*

GREG VER STEEG, ADVISOR

Umang Gupta | *Representation learning*

KE-THIA YAO, ADVISOR

Jeremy Liu | *Quantum computing*

*Nanyun Peng, Co-advisor

STUDENT RESEARCH

GRADUATE STUDENT SYMPOSIUM 2019

The annual ISI Graduate Student Symposium invited graduate students to submit their ongoing or recent work, in short papers or posters, in these areas:

Networks

Advanced electronics

Intelligent systems

Health informatics

Computational systems

Cybersecurity systems

Machine learning

Natural language processing



ISI Graduate Student Symposium 2019 award winners, pictured with Harkeerat Bedi (far right) from Verizon, the event's sponsor

At the 13th annual symposium in March 2019, 13 projects and 11 posters were presented. They were judged by students, ISI researchers, and the broader USC community. Topics included observing the deep web to anticipate cybercrimes, analyzing social bot behavior in the mid-term elections, and detecting cryptocurrency manipulation. At the close of the symposium, held at ISI's Marina del Rey location, Verizon presented awards to these students:

BEST PAPERS

Minh Pham and Craig Knoblock:

“From Instances to Transformations: Minimizing User Effort”

Mehrnoosh Mirtaheri, Sami Abu-El Haija, Fred Morstatter, Greg Ver Steeg, Aram Galstyan:

“Fraudulent Cryptocurrency Manipulations on Social Media”

Nazanin Alipourfard, Buddhika Nettasinghe, Andres Abeliuk, Vikram Krishnamurthy, Kristina Lerman:

“Friendship Paradox Skews Perceptions of Popularity in Directed Networks”

BEST POSTERS

Ninareh Mehrabi, Fred Morstatter, Nanyun Peng, Aram Galstyan:

“Debiasing Community Detection: The Importance of Lowly-Connected Nodes”

Geoffrey Tran:

“Resilient Real-time Performance in Stream Processing Systems”

BEST PRESENTATION

Nicolaas Weideman:

“Binary Analysis for Low-rate Denial of Service Attack Vulnerability Detection”

NSF RESEARCH EXPERIENCES FOR UNDERGRADUATES

10 terrific undergrads from colleges and universities across the country participated in 10 weeks of research experience at ISI headquarters. The Research Experiences for Undergraduates (REU) program is funded by the National Science Foundation's Directorate for Computer and Information Sciences and Engineering.

- **Alexandra Fernandez** (Loyola University, Maryland)
- **Anne Lin** (University of Michigan)
- **Caleb Ziems** (Emory University)
- **Katherine B. Schroeder** (Marymount University)
- **Ryan Patrick Bogutz** (College of New Jersey)
- **Allen Shang-Long Mao** (Ohlone College)
- **Bryan Hayes** (San Diego State University)
- **Claire Cannati** (Wellesley College)
- **Mirabela Medallon** (University of Hawaii at Manoa)
- **Sunshine Chong** (City College of San Francisco)

In addition to working on current ISI research projects, the students also led their own smaller research projects, with guidance from ISI researchers who are experts in their fields.

Four of the students, working with ISI researchers, wrote papers that were accepted for presentation at the IEEE BigData 2019 Conference in December:

Ryan Bogutz

Mentor: John Heidemann

Paper: "Identifying Important Network Outages"

Sunshine Chong

Mentor: Andres Abeliuk

Paper: "Quantifying the Effects of Recommendation Systems"

Anne Lin

Mentor: Emilio Ferrara

Paper: "Effects of Network Structure on Subjective Preference Diversity"

Allen Mao

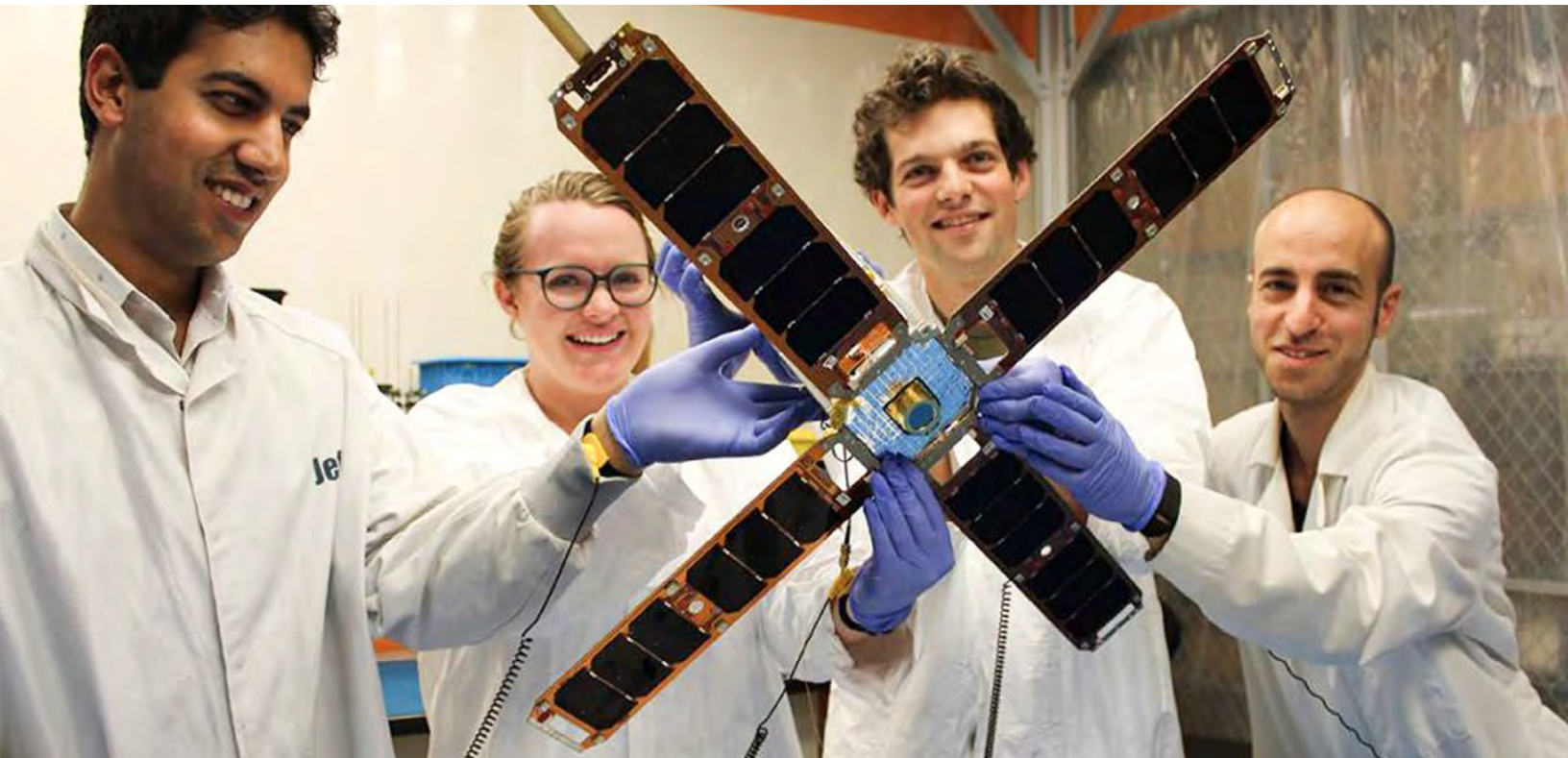
Mentor: Daniel Garijo

Paper: "SoMEF: A Framework for Capturing Scientific Software Metadata from its Documentation"



2019 REU undergraduate students and ISI researchers

STUDENT RESEARCH (continued)



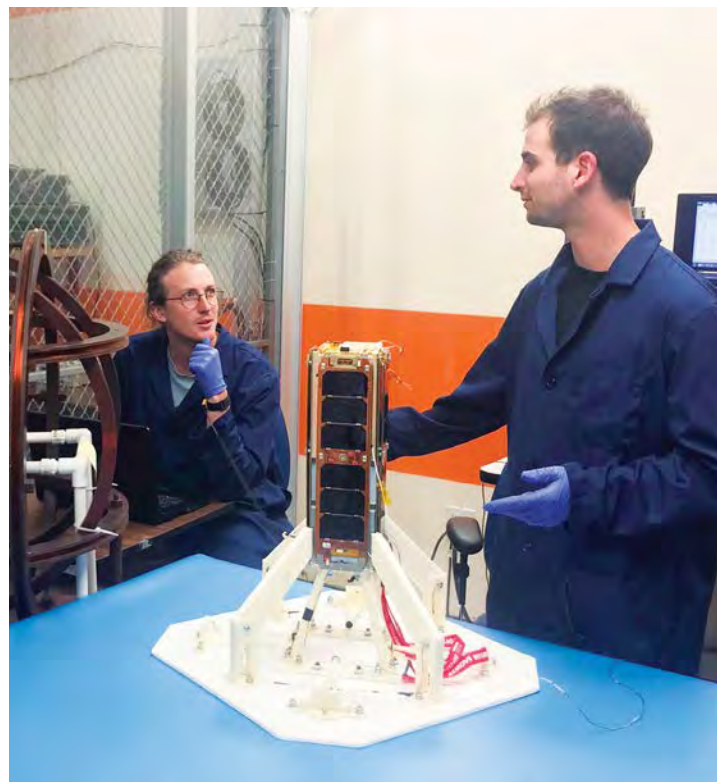
Members of the USC student team with GalacticSky-1 Photo credit: David Barnhart

STUDENTS BUILDING SATELLITES

Students at USC's Space Engineering Research Center (SERC) built a miniature satellite, or CubeSat, USC's third. (A CubeSat is a type of nanosatellite defined by the CubeSat standard.) The GalacticSky-1 satellite carries specialized payloads to test and validate advanced processing capabilities in Earth's low orbit. This CubeSat includes a set of new boards and components for flight research projects designed for operation in low Earth orbit's thermal and radiation environment.

GalacticSky-1 will also host and support new flight payload components and test out a new B-Dot controller for the satellite's magnetic torque guidance and control system. This new algorithm and controller will de-spin the satellite using the earth's magnetic field vectors saving power, which is in short supply on nanosatellites.

When launched, the satellite will be monitored by SERC students and staff from a ground antenna at the USC campus and from SERC at ISI's office in Marina del Rey.



2019 SUMMER INTERNSHIP PROGRAM

Every summer, ISI welcomes undergraduate, master's, and PhD degree students as paid interns to work with senior research leaders and their teams. Internships are available at all three ISI locations—Marina del Rey, CA, Arlington, VA, and Waltham, MA.

In 2019, five master's degree students and 24 PhD students, along with 10 undergraduates, were mentored at ISI. One student, from the School of Information at the University of Michigan, interned at ISI as a master's student in 2017 and then returned as a PhD student in 2019.

ISI interns worked in research areas including machine learning, stochastic processes, logic obfuscation, image processing, hardware security and FPGA assurance, networking protocols, natural language processing, and high performance computing.

ISI welcomed students from all over the world in the summer of 2019. These are the universities where they study:

Arizona State University

Carnegie Mellon University

CEA-List/University of Rennes 1

Columbia University

Florida State University

Fudan University

George Mason University

Indiana University

Kansas State University

Northeastern University

Oxford University

Ruhr University Bochum

SUNY Albany

Texas A & M

UMass Amherst

UC Santa Barbara

University of Alexandria

University of Edinburgh

University of Kentucky

University of Liverpool

University of Maryland

University of Michigan

University of Pisa

University of Southern California

University of Toronto

University of Virginia

University of Washington

Virginia Tech



Emilio Ferrara (bottom right) and ISI summer interns at a lab lunch in 2019

RESEARCH HIGHLIGHTS

Advancing the State of the Art in Computational Physics: Efficiently Simulating Quantum Many-body Systems

One of main challenges in physics and chemistry is gaining a true understanding of the properties of quantum materials—materials whose traits are governed by the laws of quantum mechanics. This task is, however, demanding. Optimally parallelized methods of simulation are being sought to address this challenge as supercomputers increase in power.

ISI computational physics researchers are designing computational algorithms to simulate complex physical models. In collaboration with researchers from the Landau Institute for Theoretical Physics in Moscow, the ISI researchers are devising and testing novel Monte Carlo (stochastic) techniques that allow the study of quantum many-body systems. The goal is to uncover new physics that have been unreachable with existing techniques.

Results from the studies are expected to advance the scientific understanding of a range of unique physical phenomena and impact technological innovation. As these simulations succeed, novel materials will become better understood.

This work is funded by DARPA and the Department of Energy.

Improving Systems Engineering in Mission-critical Environments

In 2019 ISI's Decision Systems group continued expanding its research into the space sciences—particularly in the area of improving systems engineering and operations, both space and ground segments. A particularly challenging aspect of the space domain is the ability to produce robust solutions in a timely matter. Traditional approaches to acquisition rely on defining requirements up front, followed by design, development, integration, testing, and then deployment—the Waterfall method. Such approaches are rigid and cannot adapt to changing system requirements and the availability of new technologies. This rigidity often results in cost overruns and delayed product releases that can result in obsolete and error-prone systems.



GPS III SV02 Photo credit: Breaking Defense

The alternative approach is to apply an Agile framework that produces frequent deliverables to the customer. These “code fast, fail fast” approaches benefit from getting the product quickly to the customer to elicit feedback that feeds into the next development cycle. Unfortunately, this rapid turn-around approach presents challenges in a mission-critical environment where solutions *cannot fail*. The Decision Systems group is exploring how to adapt Agile methods to traditional Waterfall development efforts while also addressing the mission-critical nature of the space domain.

In 2019, the team worked closely with the US Space Force to introduce novel Agile and DevSecOps approaches to current systems acquisition efforts, including the introduction of new metrics for measuring development performance and new approaches to system engineering combining Waterfall and Agile to handle acquisition efforts.

Much of the effort was in support of the successful launch of the second GPS III Satellite (SV02) in August 2019. In addition, members of the Decision Systems group participated as subject matter experts in ISI's multi-division SpaceAware project. The SpaceAware project developed techniques for data fusion and analysis of satellites in earth orbit. The team assisted with Agile feature development, space domain expertise, and demonstrations to the US Air Force's Combined Space Operations Center at Vandenberg Air Force Base.

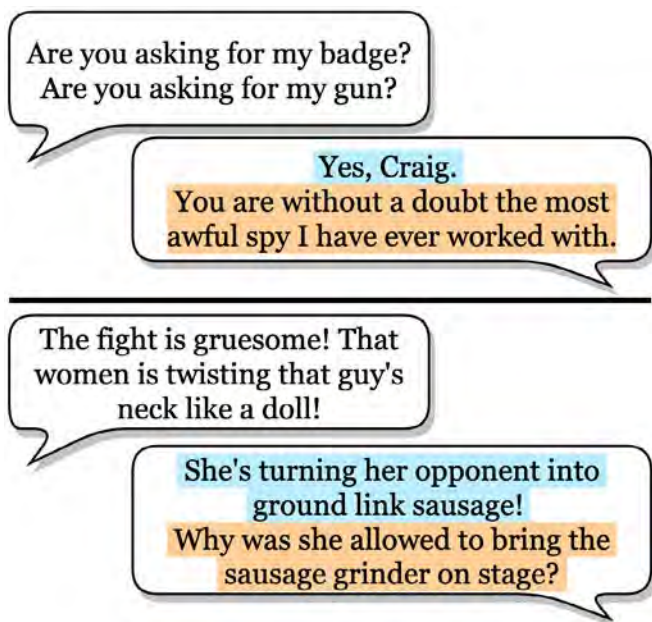
Creating AI Dialogue Partners with Improvisational Theater

When you interact with a dialogue agent such as Siri, you probably don't regard your interaction as one among equals. Treating dialogue systems of the future as partners, not servants, will be crucial in order to gain the power of trust, coordinated activity, and brainstorming; these are benefits of today's human-to-human interaction. ISI researchers work toward this goal by designing dialogue agents that can engage in natural, open-ended, meaningful conversations.

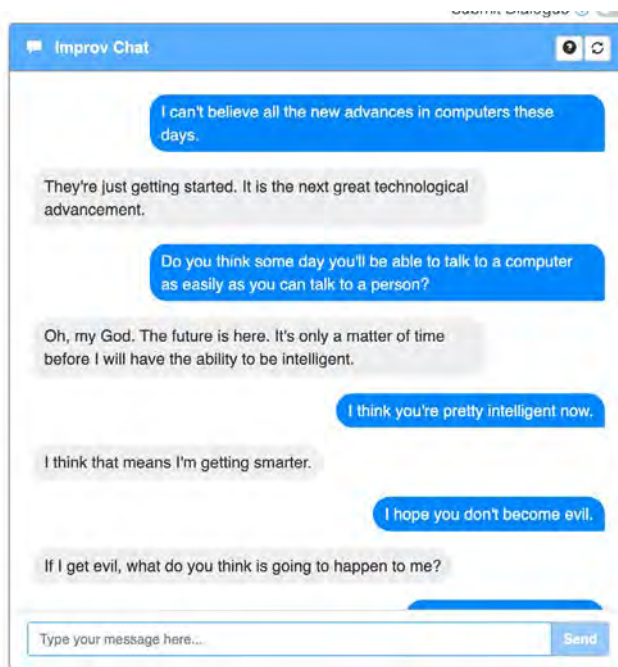
A key dialogue act is *grounding*, the establishment of "mutual knowledge, mutual beliefs, and mutual assumptions." In improvisational theater, this *grounding* act is known as a *yes-and*. It is a cornerstone of improv. ISI researchers built a *yes-and* corpus by annotating an improv theater podcast and using bootstrapping techniques to mine more than 64,000 examples from movie scripts and subtitle collections. They then adapted a neural dialogue model on this dataset, resulting in an improved conversational bot that is able to propose creative and relevant information in more satisfying ways than previous conversation bots.

This work, which is sponsored by DARPA, was principally carried out by a PhD student who first came to ISI as part of an undergraduate study abroad program. The student knocked on doors of improv comedy clubs in Hollywood, recorded improv theater classes at USC, and listened to every relevant podcast he could find until he discovered the *Spontaneanation* podcast, run by veteran actor and improviser Paul F. Tompkins. When told that ISI was trying to use his podcast to create an AI with the power to improvise, Tompkins remarked "Hahaha this is TERRIFYING!" and then agreed to let the team use snippets from the show in a publicly released dataset. The dataset is the largest, and possibly the only, corpus of *yes-and*s known to exist.

To try talking to the bot, visit spolin.isi.edu.



Examples of *yes-and* dialogue turn pairs included in the corpus. Each response reflects (in blue) acceptance of the context established in the prompt ("yes") and (in orange) a new relevant contribution to the dialogue ("and").



Example conversation between a human and a dialogue agent trained with the corpus. The agent's generated turns are on the left.

RESEARCH HIGHLIGHTS

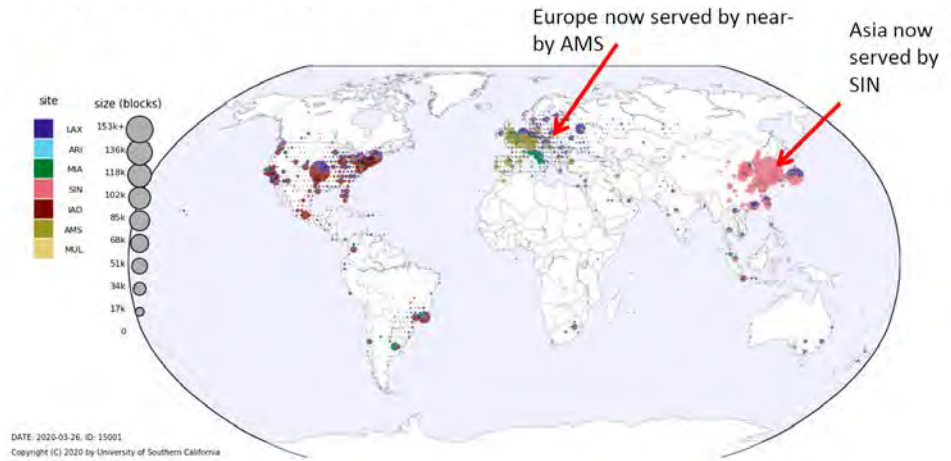
B-Root Expands Its Global Service Footprint

The initial step in nearly all Internet communication is querying the Domain Name System (DNS). Since the creation of the DNS at ISI in the early days of the Internet, ISI has operated a DNS “root server.” ISI’s root server, B-Root, is still a core piece of Internet critical infrastructure.

In 2019, B-Root tripled its physical footprint by adding new servers in Chile, Singapore, the Netherlands, and Virginia—extending its existing sites in Los Angeles and Florida. In addition to reducing latency for B-Root’s service around the world, these sites allow ISI to advance its research in Anycast, the Domain Name System, and routing.

The B-Root team used the new Verfploeter tool, which was developed in collaboration with the University of Twente, to map geographic regions to their respective B-Root servers.

This image shows which regions communicate with which deployed B-Root instance. The newly deployed sites appropriately attract traffic from neighboring regions.

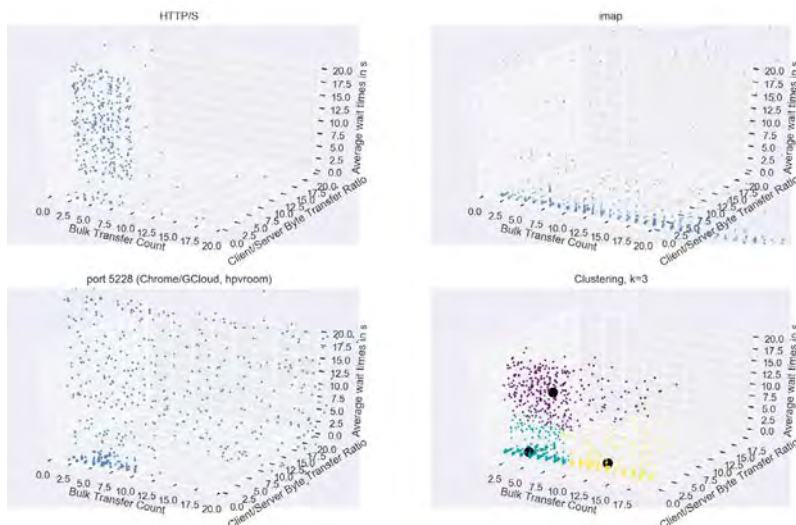


Detecting and Tracking Encrypted Application Instances and Entities

The increasing use of encryption hampers the ability of network operators to give important low-latency streams, such as video conferencing, higher bandwidth with faster throughput.

The Accurate and Precise Recognition of Obscured Payloads in Operational Systems (APROPOS) project aims to identify the links in a mesh network that contain important application traffic requiring a faster class of service. APROPOS rapidly clusters encrypted traffic into application classes and uses novel techniques to identify and track the communicating entities behind the communication.

ISI researchers are developing technology to analyze encrypted application streams using machine learning techniques and pass the results to software being developed by other researchers in DARPA’s SEARCHLIGHT program.



By estimating round trip times (RTTs) from TCP network flows, researchers are able to remove network timing effects in order to extract and model application data directly. Features in the network data sent and received by applications are then used to do k-means clustering as a first pass toward classifying traffic. The network data and the results of the initial clustering are then used to drive more expensive techniques that will identify flows of interest.

Deploying an Octopus to Grab a Passing Satellite

As the commercial space sector evolves to perform servicing activities in space, ubiquitous, easy-to-use methods of safe grasping will be required.

Services in space involve rendezvous and eventual docking, or contact, with another spacecraft. That spacecraft might be cooperative; or it could be non-cooperative, that is, non-operational or of limited operation. Contact between two disparate entities in space has a high level of potential inadvertent mechanical contact risk due to the micro-gravity environment. (By contrast, in ocean or air there is a medium to provide friction and thus a corresponding backing force.) The question is: how can contact be made between objects in space with a safe and highly compliant mechanism? To date, rendezvous and docking mechanisms are mechanical and very proscribed.



On the left, an 8-armed octopus.

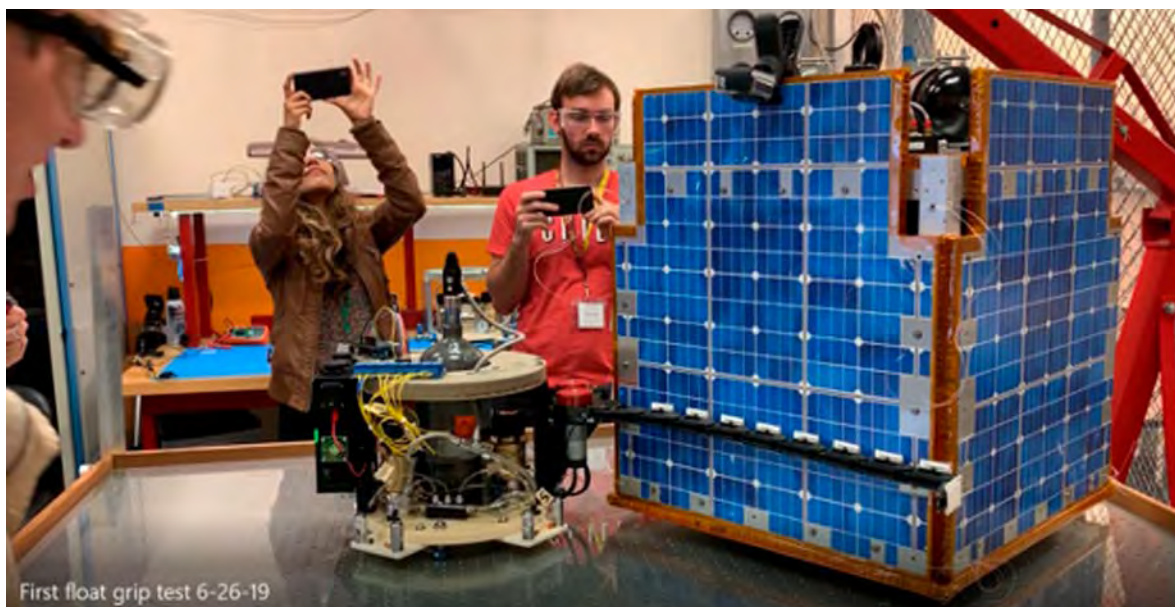
On the right, a space-rated version to capture satellites.



The octopus, with its flexible, soft body and tentacles that can sense surfaces and grip, formed the basis of the REACCH project.

Sponsored by DARPA, REACCH (Reactive ElectroAdhesive Capture Cloth) uses the geometry and methodology of an octopus, coupled with new grasping technologies. The USC Space Engineering Research Center (SERC), in partnership with the Jet Propulsion Laboratory (JPL), created working prototypes of various “tentacles” and put them on a floating platform to test against a candidate space object. The tentacles are 3D-printed joints with mechanical cables. Unique electro-adhesive/gecko adhesion “pads” are placed along the length of the cables (similar to the suckers on octopus tentacles).

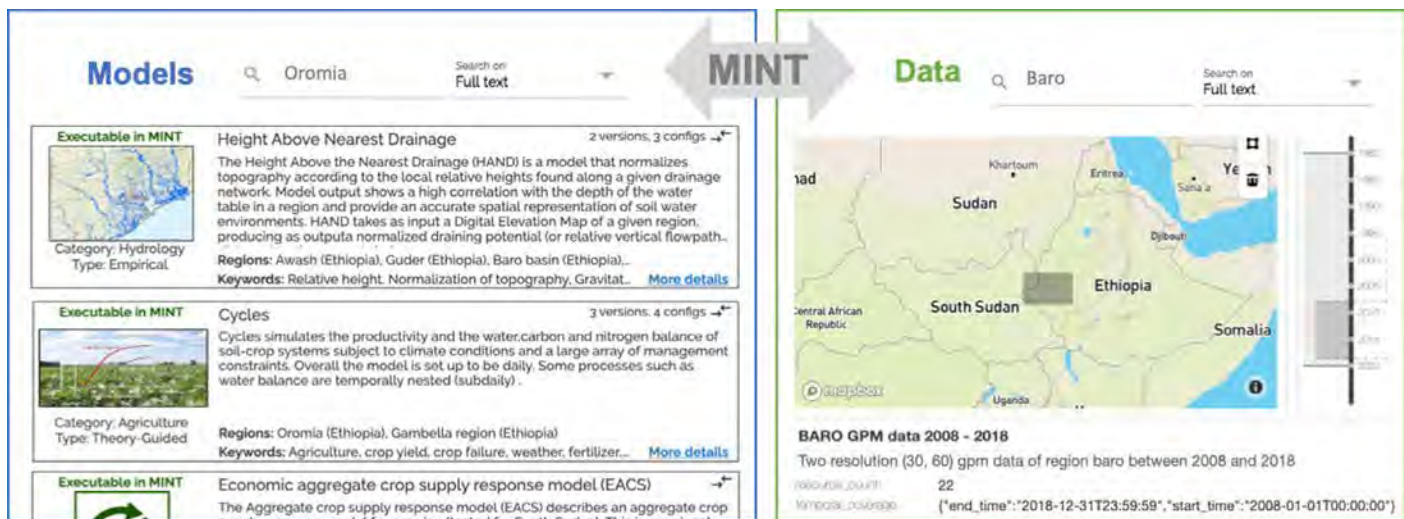
The team ran a number of tests on different geometries, surfaces, and types of materials. In almost all cases, REACCH performed as expected in the gravity of the earth by “capturing” an object and proving it can grip using opposing shear forces. The next step for REACCH is a space flight.



Two of the prototype device’s REACCH tentacles have attached to a “client” satellite (the blue solar panel). Both the device and the satellite are floating on an air-bearing table to simulate the space environment in three dimensions.

RESEARCH HIGHLIGHTS

AI for Modeling the Impact of Climate and Human Activities in Water and Food Resources



MINT is an integrated environment to model complex systems. MINT uses AI techniques to assemble relevant models and data efficiently.

Major societal and environmental challenges require understanding the interactions between natural phenomena and human activities. For example, consider how droughts and market prices affect crop production or how agriculture or industrial needs affect water quality and availability.

Integrated modeling enables better understanding of how natural and human phenomena affect one another, helping to forecast the effects of undesirable situations and design preventive measures. But the diversity of modeling approaches across disciplines and the wide variety of data sources available in different formats requiring complex conversions impede modeling efforts.

Integrated modeling requires resolving semantic, spatial-temporal, and execution mismatches, which can take months or years of largely manual effort.

ISI is leading the development of an integrated modeling environment that uses a variety of artificial intelligence techniques to significantly reduce modeling effort while ensuring utility for decision making. The Model INtegration (MINT) framework makes several innovative contributions:

- An intelligent user interface that guides analysts to frame their modeling problems and assists them by suggesting relevant choices and automating steps along the way
- Semantic metadata for models, including their modeling variables and constraints, that ensures model relevance and proper use for a given decision-making problem
- Semantic representations of datasets that enable automated data selection and data transformations
- An interactive visualization environment to support analysis of future scenarios and potential interventions to support decision making

MINT includes data and models to analyze the effects of climate on food insecurity for select regions in Sub-Saharan Africa and on water resources in the South-Central region of the US.

AI for Systematic Scientific Data Analysis

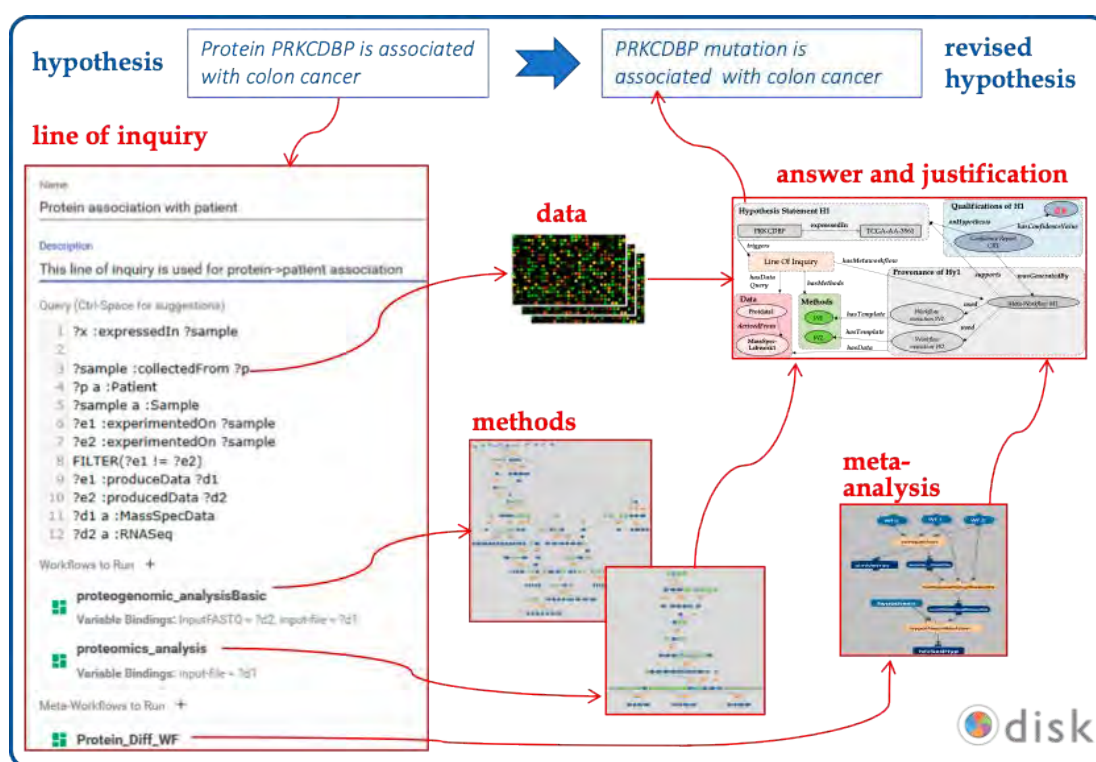
The quantity of complex scientific data that labs are collecting has surpassed their ability to analyze it.

The DISK system (automated Discovery of Scientific Knowledge) automates data analysis by systematically exploring the methods available for complex cancer omics data. In collaboration with researchers at Stanford University, AI researchers at ISI are capturing scientists' knowledge of data analysis methods for various questions and datasets.

DISK contains methods for transcriptomics, proteomics, and genomics that come from scientists with different specialties and diverse expertise. Given a hypothesis, such as whether a protein is expressed in a certain type of cancer, a scientist knows the kind of data and analysis that is appropriate, which DISK represents as a line of inquiry. DISK uses this information to query community repositories of cancer data, execute all available methods for each type of data found, and run meta-workflows to combine the results. It then shows the scientist the evidence it found, using the top-performing methods. DISK may suggest a revised hypothesis, for example if the protein found was a mutant form of the original hypothesis.

DISK can reason systematically about all the data, explore the use of all available methods with different parameters, and use different reference databases. In experiments to date, researchers have found that using equivalent methods leads to substantial differences in protein and peptide identifications, as well as in single amino acid variants and variant peptides. In a more recent project, DISK is being extended with knowledge about neuroscience data analysis.

AI systems such as DISK have the potential to make science more systematic and accelerate scientific discovery.



The DISK AI system contains extensive knowledge about how to analyze diverse types of cancer omics data and uses it to do systematic hypothesis testing. When a scientist poses a hypothesis, DISK uses its knowledge to find relevant data, identify methods appropriate for each data type, and perform meta-analysis to combine the results. DISK provides an answer with a justification of the methods that produced the best results.

RESEARCH HIGHLIGHTS

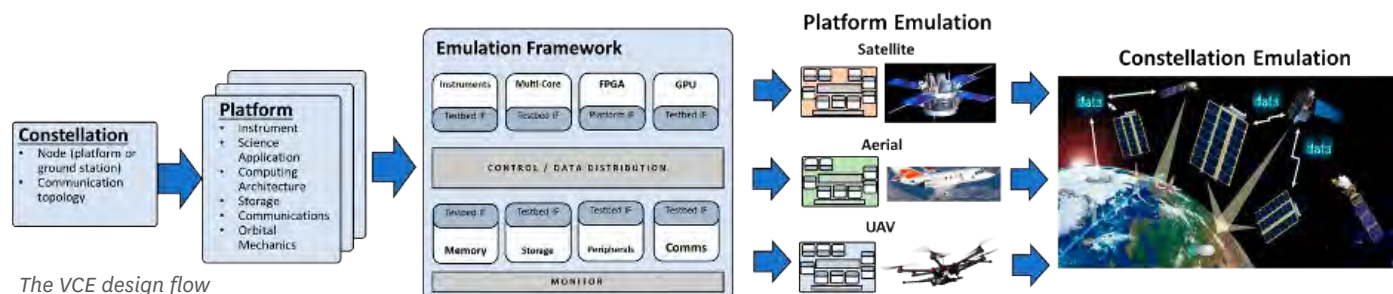
Constellations in the Cloud: Virtualizing Distributed Remote Sensing Systems

As scientists seek to answer fundamental questions about how the earth works—carbon cycle, coral reef erosion, severe weather—they find it critical, and more economical, to have a variety of data produced by multiple small sensors across many satellites, as opposed to fewer high-fidelity sensors on one large satellite.

A group of such sensors is a distributed remote sensing system (DRSS), also called a constellation or a sensor web. A DRSS may be a mix of homogenous sensing nodes, such as identical remote ground sensors or satellites, or it can be a heterogeneous mix of ground nodes, unmanned aerial vehicles (UAVs), and space-based nodes. A heterogeneous DRSS can provide a mix of instrument fidelities and types, plus a mix of look angles and revisit times, making for richer data.

But DRSSs present a challenge for mission planners. Instead of optimizing a single satellite for a particular mission, the mission planner has the complex task of designing a heterogeneous system.

Researchers at ISI teamed with NASA GSFC, NASA ARC, and NASA JPL under sponsorship from the NASA Earth Science Technology Office. The team developed the Virtual Constellation Engine (VCE), a high-fidelity emulator that enables mission planners to experiment with different DRSS concepts to prove out their designs before committing to build. VCE also enables mission teams to start application development and verification before hardware resources are available.



VCE uses Amazon Web Services (AWS) to spawn a node for each node in the DRSS, such as a ground sensor, unmanned aerial vehicle (UAV), ground station, or satellite. Each AWS node then emulates the exact instrument type, onboard processor, operating system, and application that the actual DRSS node would contain.

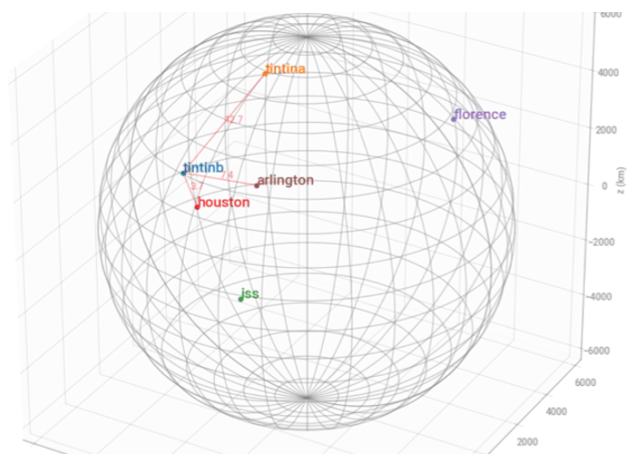
A control AWS node calculates the orbitology, instrument observation, and communication link bandwidth for each node. As a mission scenario runs, VCE tracks which nodes observe which locations on earth, as well as which nodes communicate with one other.

VCE can support systems ranging from just a few nodes to systems with thousands of nodes being conceived of by Google and other commercial companies.

Using VCE, the ISI team demonstrated that by using new onboard computing hardware being developed by NASA GSFC named SpaceCube 3, the processing rate of a new Multispectral, Imaging, Detection, and Active Reflectance (MiDAR) application developed by NASA ARC could be increased by 39,137x, sufficient to achieve 13,552 frames per second and support real time operation.

VCE has been open sourced (<https://isi-rcg.github.io/vce/>) and is now transitioning to be a part of NASA ESTO's New Observation Strategies Testbed, where it will be integrated with Stevens Institute's Trade-space Analysis Tool for designing Constellations (TAT-C) and Ohio State's Simulation Toolset for Adaptive Remote Sensing (STARS) to provide a suite of tools to help mission planners conceive of and validate DRSS design concepts.

VCE was also recently selected to be a part of the team lead by Carr Astronautics Corporation to support development of StereoBit, a multi-satellite concept investigating atmospheric dynamics with 3D stereo tracking of cloud and moisture features.



VCE visualization of satellite nodes and communication links

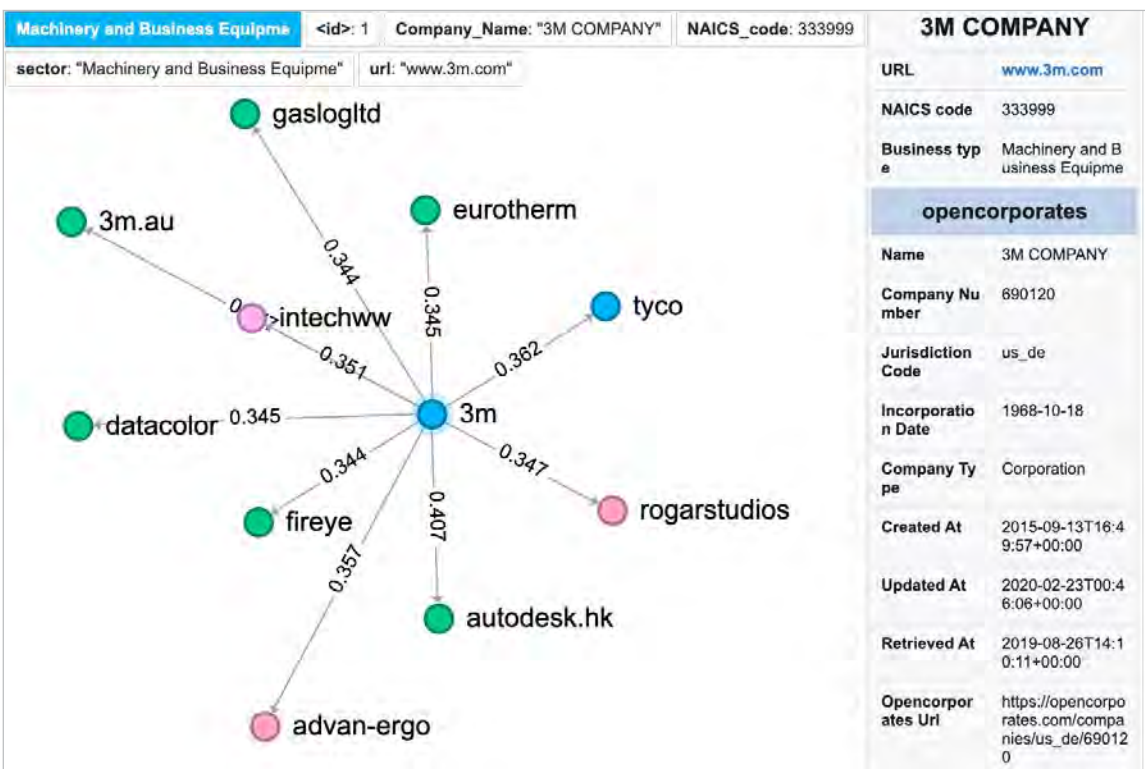
Using Knowledge Graphs to Help Entrepreneurs Navigate Their Competitive Environment

Starting a successful business is an essential component of what many consider the American dream. But entrepreneurial activity in the US is declining. Beyond hard work and innovative ideas, entrepreneurs face information-critical challenges: finding customers, differentiating from their competitors, securing suppliers and partners, and commercializing innovations. Spending hours on search engines can yield meager results and commercial tools are costly.

The Business Open Knowledge Network (BOKN) is designed to fill these informational gaps. BOKN is one of the first projects funded through the NSF's Convergence Accelerator program, which focuses on multidisciplinary collaboration and real-world impact. The BOKN team combines AI techniques developed by ISI researchers and domain expertise from professors at the Robert H. Smith School of Business (University of Maryland), the Tuck School of Business (Dartmouth), and the Marshall School of Business (USC).

Using a vast, diverse set of data from websites, social media, regulatory filings, and patents, the project team is creating a knowledge graph of nearly a million public and private US companies and modeling the relationships between them. BOKN provides a hybrid knowledge graph, where the information and relationships in the knowledge graph are a combination of traditional data queries for factual information (such as a firm's address or CEO) and results from machine learning to predict the most likely competitors and customers.

The knowledge in BOKN provides the foundation for the entrepreneurship portal (E-Portal). E-Portal offers an accessible interface for entrepreneurs to describe their business plans and instantly receive guidance. Entrepreneurs can use E-Portal to find market intelligence to identify close competitors and to differentiate their business plans. Using models of patents, the E-Portal suggests potential commercialization strategies and licensors of intellectual property. Customer discovery features allow entrepreneurs to engage with other businesses who may wish to purchase their products and services; the portal's supply chain analysis can help find materials or components. Together, these features help enable the fledgling entrepreneur to achieve the American dream.



Using the entrepreneurship portal, entrepreneurs can explore the BOKN business knowledge graph to learn about related businesses and understand the relationships between them.

RESEARCH HIGHLIGHTS

Multitask Learning and Benchmarking with Clinical Time Series Data

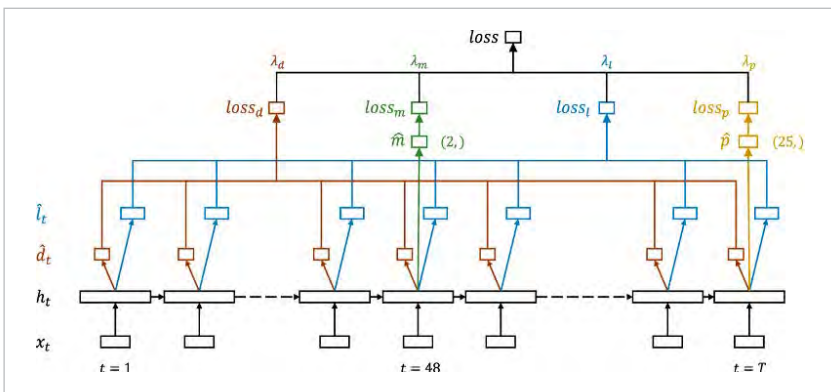
In the United States alone, each year over 30 million patients visit hospitals, 83% of which use electronic health record systems. Although there has been steady growth in machine learning research for health care, obstacles remain. First, in the absence of widely accepted benchmarks, it is hard to evaluate competing predictive models. Secondly, the prevalent approach has been to develop new methods for one single clinical prediction task at a time (e.g., mortality prediction or condition monitoring). This is contrary to the reality of clinical decision-making, where clinical staff perform several prediction tasks simultaneously. There is accumulating evidence that such prediction tasks are interrelated; studying them separately is suboptimal.

ISI researchers in the AI division took a comprehensive approach to addressing these challenges, proposing a public benchmark suite that includes four clinical prediction tasks:

- In-hospital mortality
- Physiologic decompensation
- Length of stay (LOS)
- Phenotype classification

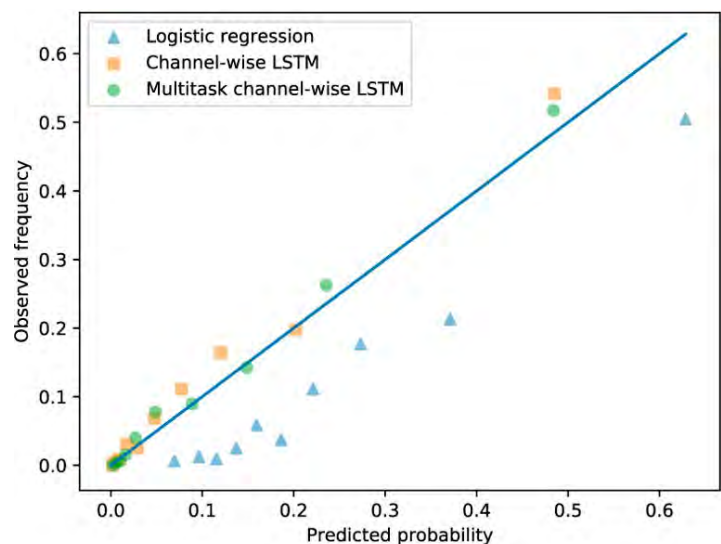
The new benchmark suite is derived from the publicly available Medical Information Mart for Intensive Care database (MIMIC-III), which contains rich multivariate time series data from over 40,000 intensive care unit (ICU) stays.

In addition, the research team also formulated a heterogeneous multitask learning problem that involves jointly learning all four prediction tasks simultaneously. It demonstrates that carefully designed recurrent neural networks are able to exploit these correlations to improve performance for multiple tasks.



Long Short-Term Memory-based (LSTM) network architecture for multitask learning

Calibration of in-hospital mortality using the best 1) linear model; 2) non-multitask model; and 3) multitask LSTM-based model. The plot shows that the multi-task LSTM-based model has better calibration.



Transforming Science through Data Collaboration

Any time scientists disagree, it's because we have insufficient data. Then we can agree on what kind of data to get; we get the data; and the data solves the problem. Either I'm right, or you're right, or we're both wrong. And we move on. That kind of conflict resolution does not exist in politics or religion.

Neil deGrasse Tyson

Non-reproducible single occurrences are of no significance to science.

Karl Popper, 1959

The Logic of Scientific Discovery

Reproducible data is at the heart of modern scientific discovery. Yet the reproducibility of published scientific results is shockingly low, impeding scientific collaboration.

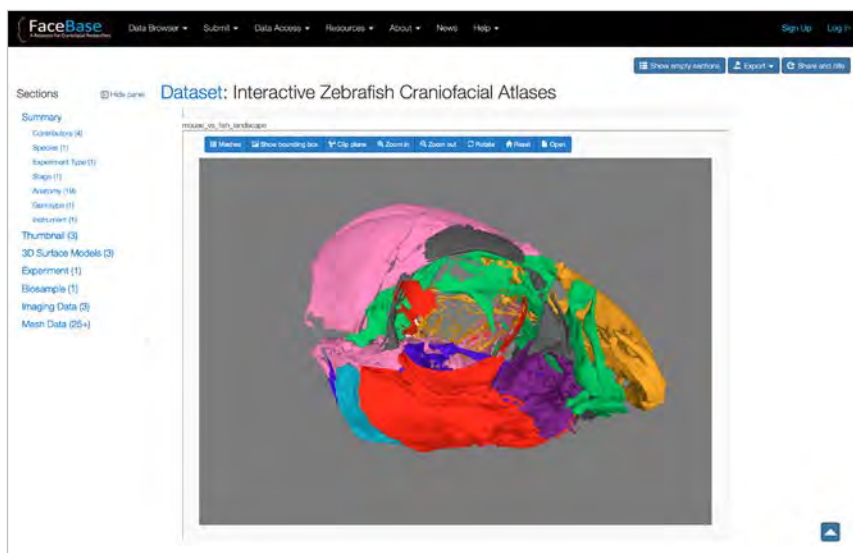
This is the case across all branches of science, from basic biology to machine learning. In today's data driven-world, scientists and engineers lack the methods, tools, training, and infrastructure required to make irreproducible results a thing of the past.

The Informatics Systems Research division (ISRD) is tackling this problem by creating innovative sociotechnical ecosystems that can radically improve the reproducibility of scientific results with the benefits of increasing scientific data sharing and decreasing the time impact.

The division's research is dedicated not only to technology, but to using this technology to transform the research of global scientific communities in areas of societal importance. One example is the FaceBase repository for the collection and dissemination of diverse data types related to dental and craniofacial development and dysmorphia (www.facebase.org).

Nearly half of all birth defects involve the face and skull, yet scientists don't know why most occur. To better serve families at risk for these conditions, scientists need a comprehensive and systematic understanding of how the faces of healthy children form and what goes wrong to cause common malformations such as cleft lip and palate.

Partnering with scientists at USC's Herman Ostrow School of Dentistry, ISRD developed and operates the Facebase data repository with funding from the NIH National Institute of Dental and Craniofacial Research. Used by a global community of thousands of developmental biologists, biomedical researchers, and clinicians, FaceBase provides the laboratory by which ISRD research is advanced, while contributing to finding solutions to pressing medical problems.



The zebrafish is a model for human craniofacial development. This figure shows processed surfaces extracted from micro-CT images of zebrafish annotated with the detailed anatomy of the skull. These images are linked against all other data in FaceBase related to the same parts of the craniofacial anatomy.

RESEARCH HIGHLIGHTS

Efficiently Finding Genetic Relatives and the Genetic Basis of Disease

The ability to identify segments of the genome shared across individuals in a population due to common inheritance, a problem known as Identity-By-Descent (IBD), is an important component of analysis pipelines in statistical, medical, and population genetics.

Population-based linkage extends ideas from genetic linkage of family trees to large populations. The increasing growth of genetic datasets, such as the 500,000 individuals in the UK Biobank or the 1,000,000 individuals expected in the NIH All-of-US project, offers new opportunities to interrogate the human genome for associations between genetic variation and diseases or traits.

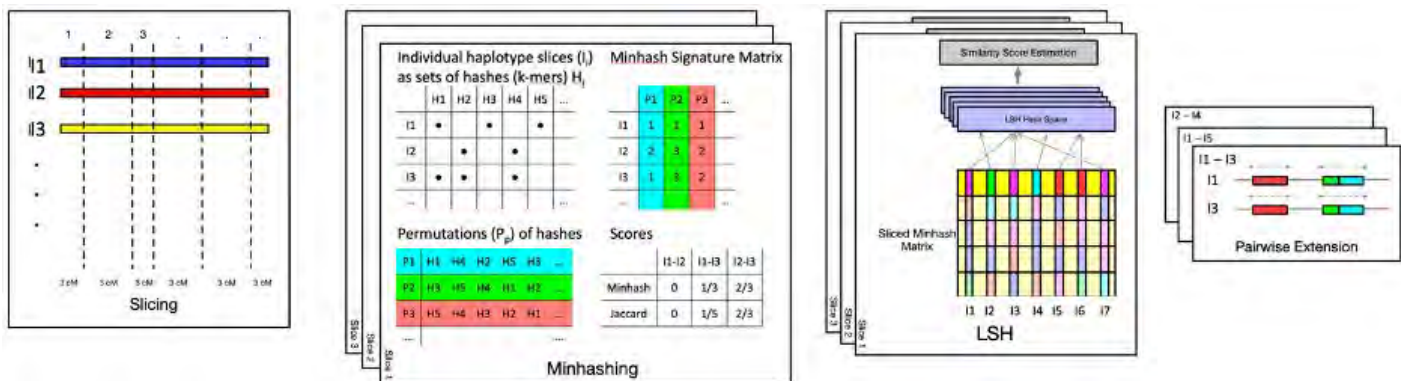
Using IBD algorithms, groups of individuals that share (small) segments of their DNA through a common distant ancestor can be identified. If some of these “distant families” have a significantly higher prevalence of any given trait or disease, then the DNA segment shared by the “family” may contain the genetic basis for the disease. Moreover, accurate, large-scale IBD data can help with other tasks, such as understanding population and disease structure, improving imputation of genetic data, and estimating disease risks on an individual level.

However, traditional methods for finding local IBD across all pairs of individuals scale poorly.

To scale IBD for very large datasets, two AI division researchers developed iLASH (IBD by LocAlity-Sensitive Hashing) an algorithm based on similarity detection techniques. iLASH shows improved accuracy compared to current methods; it speeds analysis by several orders of magnitude, making IBD identification tractable for hundreds of thousands of individuals, or even millions.

The figure below shows a schematic of the adaptation of locality-sensitive hashing to the IBD problem, including breaking up the genome into slices adaptively (close to the desired minimum IBD threshold) and minhashing. The research team applied iLASH to the Population Architecture using Genomics and Epidemiology (PAGE) dataset of ~52,000 multi-ethnic participants, including several founder populations with elevated IBD sharing. iLASH identified IBD segments with high accuracy in one hour on a single machine. This is roughly 3 minutes per chromosome—compared to over 6 days per chromosome for a state-of-the-art algorithm. The team also applied iLASH to the UK Biobank, detecting nearly 13 billion pairwise IBD connections shared among ~11% of participants.

iLASH enables fast and accurate detection of identity-by-descent, an important method for population genetics and trait mapping. The researchers are now in the process of using the IBD results for population-based linkage in the UK Biobank.



Schematic of the iLASH algorithm. Starting from the left with the **Slicing** step where haplotypes are broken into slices (segments of uniform or variable length). The **Minhashing** step creates minhash signatures by generating a table of random permutations. The **LSH** step bands together minhash values to create an integrated LSH hash table where candidate matches are grouped together. Finally, in the **Pairwise Extension** step, these candidates are further analyzed to be extended in the (likely) case that an IBD tract spans multiple slices.

Cyberinfrastructure Influencers: Connecting Large Scientific Facilities to Experts, Knowledge, and Each Other

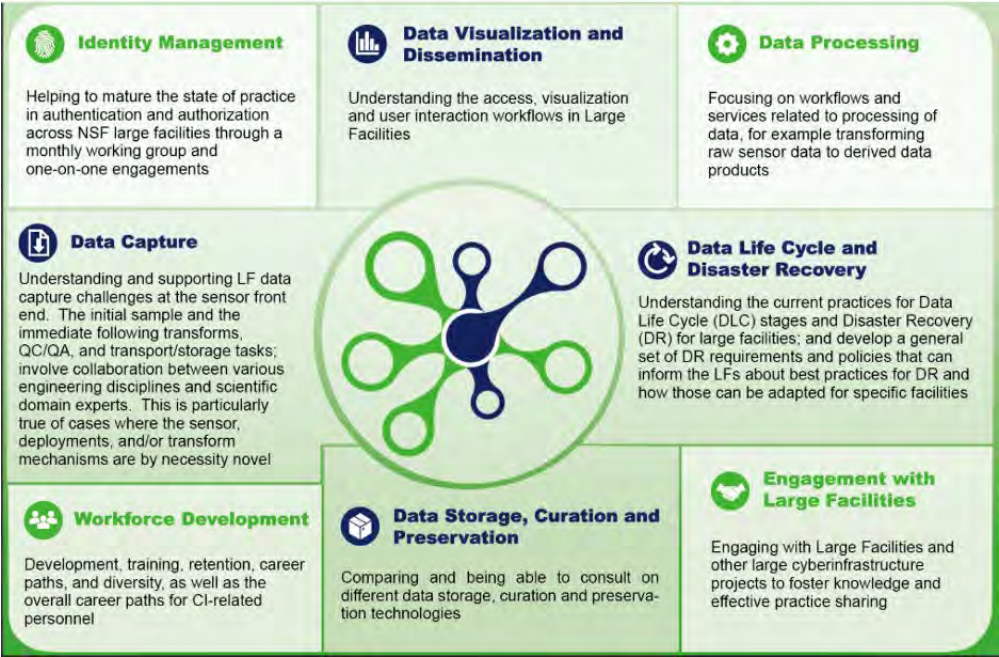
The CI CoE had four types of profound influence on NEON developers. First, as we transitioned from construction to operations, our developers benefited from greater awareness of the wider NSF CI community practices. Second, deep engagement with CI CoE experts produced three major technologies insertions into NEON CI, remarkably within 6 months. Third, open dialog and prototyping with CI CoE experts affirmed our workflow-based sensor message handling strategy and built our confidence to invest in this novel method. NEON’s CI and Data Sciences team mission includes advancing methods and ecological science; interaction with CI CoE nudged our efforts ahead significantly through community workgroup involvement, presentations and publications.

Tom Gulbransen, NEON

The Cyberinfrastructure Center of Excellence pilot (CI CoE) provides leadership, expertise, and support to cyberinfrastructure practitioners at the National Science Foundation (NSF) major facilities. (<https://cicoe-pilot.org>)

To better understand the specific cyberinfrastructure (CI) challenges faced by NSF major facilities, the pilot identified the National Ecological Observatory Network (NEON, <https://www.neonscience.org>) for the first collaboration. NEON is an ecological observation facility that collects and provides open data about the changes in North America’s ecosystems. NEON’s capture, processing, and dissemination of ecological data improves the understanding of the environment to forecast how human activities impact ecology. NEON builds and operates ecological sensors at a number of geographic sites and collects a rich set of data. The collection sites are strategically located across 20 ecoclimatic domains in the US and represent regions of distinct landforms, vegetation, climate, and ecosystem dynamics.

During the first year, the CI CoE pilot produced CI plans, templates for disaster recovery, and prototypes for architectures. NEON and the pilot produced a joint research publication. The pilot project promoted awareness of the wider NSF CI community practice and served as the link between NEON and larger community efforts such as Earth Science Information Partners and schema.org for semantic technologies and data management and then extended those community efforts to include NEON data schemas.



CI CoE working groups

RESEARCH HIGHLIGHTS

Assembling a Live Cybersecurity Threat-feed

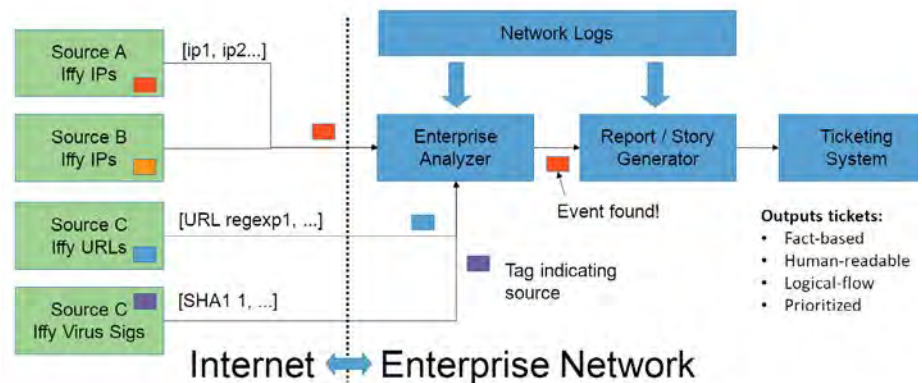
ISI researchers are searching for global signals of emerging cybersecurity threats in the Global Analysis of Weak Signals for Enterprise Event Detection (GAWSEED) project. The project is part of DARPA's Cyber Hunting at Scale (CHASE) program.

Project members are assembling a live threat-feed of detected cybersecurity threats and events, combining the results of ISI's own threat research with those from external sources into a live stream entitled "GiFT" (GAWSEED Feed of Threats) to identify global cyber-security threat signatures.

When GiFT is combined with ISI's flexible log analyzing toolset, it provides cybersecurity analysts with a prioritized list of active cybersecurity events (threat analysis) within an enterprise network.

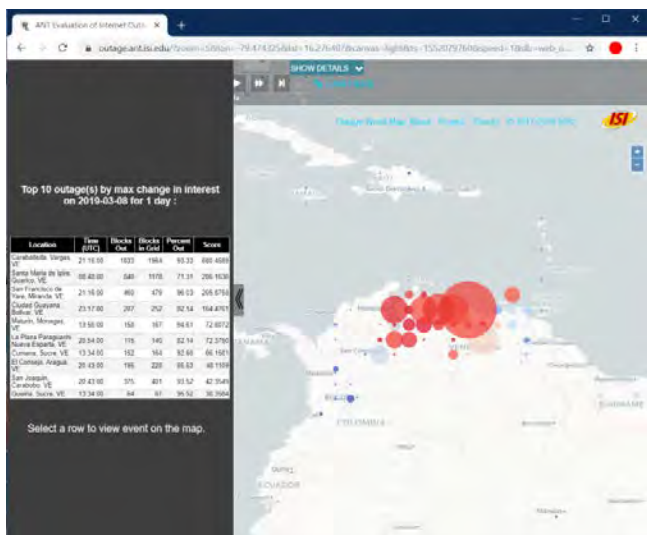
founder populations with elevated IBD sharing. iLASH identified IBD segments with high accuracy in one hour on a single machine. This is ~3 minutes per chromosome—compared to over 6 days per chromosome for a state-of-the-art algorithm. The team also applied iLASH to the UK Biobank, detecting nearly 13 billion pairwise IBD connections shared among ~11% of participants.

iLASH enables fast and accurate detection of identity-by-descent, an important method for population genetics and trait mapping. The researchers are now in the process of using the IBD results for population-based linkage in the UK Biobank.



The GAWSEED threat-searching tools pull in data from internal and external analysis sources and create an actionable report to be automatically submitted to a network operator's ticketing system.

Undergraduate Research Helps Understand Internet Outages



The sidebar shows the important outages on March 8, 2019, including this outage in Venezuela.

ISI has been observing Internet outages—globally, with 24/7 coverage—for more than five years. With all that data, it can be hard to discern what is *important*.

A student participating in the NSF Research Experiences for Undergraduates program at ISI during the summer of 2019 developed an outage reporting system that helps reveal large events in the global Internet outage data collected by ISI's ANT Lab: Analysis of Network Traffic (<https://ant.isi.edu/>).

The student's project examined Internet outage data from Trinocular (ISI's probing system that tracks outages across the Internet) and developed an outage report that summarizes the most "interesting" outages each day. He worked with ISI network researchers to integrate the report into the ISI ANT Internet Outage Maps website <https://outage.ant.isi.edu/>.

The undergrad, with two ISI researchers, wrote up the project in the paper "Identifying Important Network Outages." It was accepted at the 6th National Symposium for NSF REU Research in Data Science, Systems, and Security at the IEEE BigData 2019 Conference.

Improving Computing Performance and Efficiency—with Introspection

In the first decade of this century, clock speeds for microprocessors stopped getting faster and single-threaded performance stopped increasing. Since then, parallelism, heterogeneity, and specialization have become the predominant methods of increasing performance in computing architectures. The adoption of these techniques comes at a cost: increased complexity for programmers. This complexity leads to decreased productivity and increased cost.

As these trends of parallelism, heterogeneity, and specialization continue, computing will need to exploit the parallelism of hundreds of cores per chip, thousands per rack, and millions per server room. This is further complicated by heterogeneity and specialization, which put constraints on which hardware best executes which software tasks. In order to efficiently exploit these new computing architectures, especially in the context of applications that interact dynamically with the non-virtual world, systems will need to be able to dynamically optimize their behavior without programmer intervention.

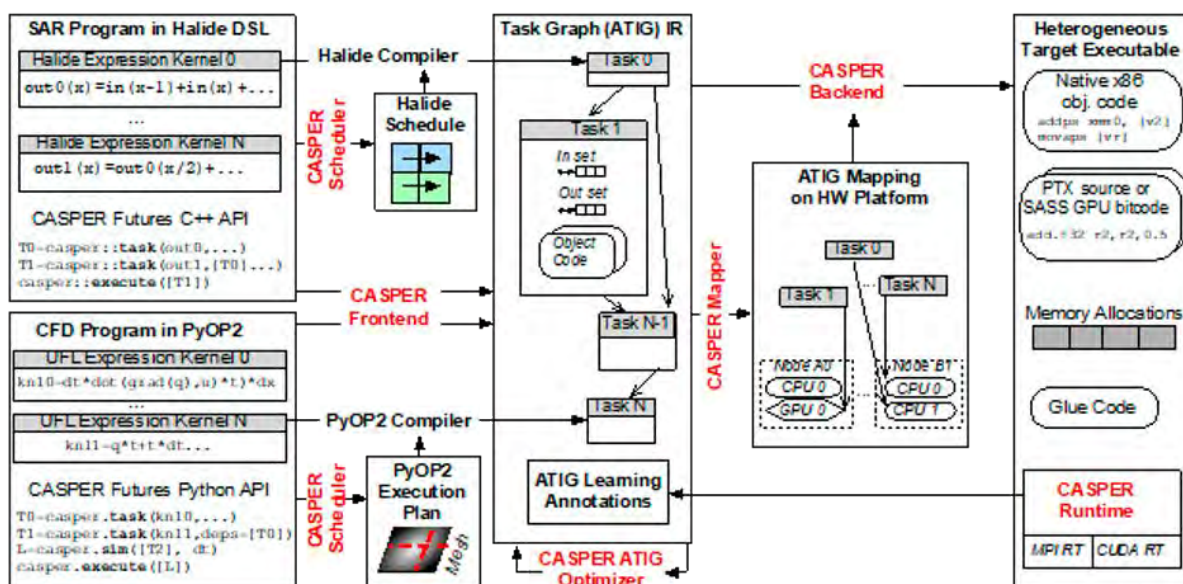
Researchers in the Computational Systems and Technology division (CS&T) are developing ways to manage resources in computing systems using introspection, where the computing systems monitor their own behavior and performance and can make adjustments to improve performance and efficiency. Several projects are pursuing this.

In the FogSys project, sponsored by the Office of Naval Research, researchers are monitoring real-time performance of fog and edge computing systems to reallocate resources to meet real-time performance goals and provide fault tolerance, so that embedded computing systems can use compute resources efficiently while reliably processing sensor data to protect naval systems and sailors.

In the BigFoot project, the team is developing run-time provisioning systems to allow the shifting of resources between the real-time processing of sensor data with batch jobs that are less time sensitive.

In the ARPA-E sponsored STEAM project, in collaboration with Networking and Cybersecurity division researchers, CS&T researchers are building models of large-scale data centers, with a focus on interconnect that could be used to project performance when making decisions in introspective systems.

In the CASPER project, sponsored by DARPA, CS&T researchers are working with faculty in USC's Ming Hsieh Department of Electrical and Computer Engineering to develop a domain-specific software environment that targets heterogeneous architectures with up to a million processing elements. The software environment will perform dynamic optimizations with input from the application programmer and compiler to achieve orders of magnitude improvement in performance, portability, and productivity.



The CASPER software stack enables the exploitation of parallel, heterogeneous compute elements while increasing programmer productivity with domain-specific languages.

RESEARCH HIGHLIGHTS

Multilingual Representation Spaces for Zero-Shot and Few-Shot Natural Language Processing

Nearly all human knowledge is expressed via human language. Natural language processing (NLP) analysis, such as syntactic parsing and information extraction, is employed to distill and organize that knowledge efficiently. NLP tools enable researchers to ingest a large body of text and use analysis to rapidly determine, for instance, all the names of people and organizations, all the activities they were involved in, and who did what to whom.

With sufficient annotated training data, researchers can also build tools that perform this analysis in many languages. However, the languages with annotated training data are often in one language, usually English, and do not enable analysis in other languages such as Uyghur, Chinese, Arabic, or Lithuanian.

To combat this data mismatch problem, NLP researchers at ISI are pursuing strategies based on the notion of a *multilingual representation space*. A sentence is encoded as a many-dimensional vector that represents the ideas it contains but is insensitive to the language of the test data/inference data. This allows the translation and linguistic analysis of data in a language without training data.

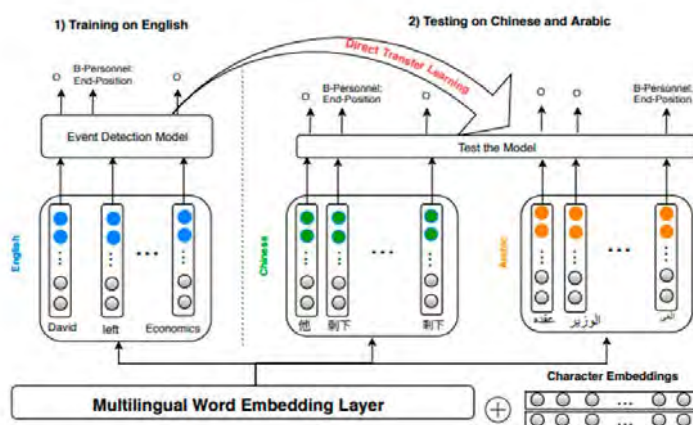
The research team uses techniques centered around *transfer learning* to create these representations. Broadly speaking, transfer learning leverages data from the “wrong” task or language and is used, often with a smaller amount of “correct” data, to fine-tune general purpose architectures. Language-independence is encouraged by:

- Introducing specialized training methodologies such as adversarial learning
- Using common multidimensional representation spaces as a merge point for multitask training
- Explicitly harmonizing heterogeneous data into a common symbol set

Many types of data are used: data annotated for the task of interest but in a different language, data annotated for another task but for the right language, data that is neither for the task nor the language of interest, and *community annotated corpora* like Wikipedia.

When there is no data relevant to both the training language and the analysis task, the models operate in a *zero-shot* manner. When there is a small amount of data relevant to both the task and the language, it is included; the models operate in a *few-shot* manner. These techniques are applied to machine translation, information retrieval, named entity prediction, relation prediction, event prediction, part-of-speech tagging, and syntactic parsing.

This work encompasses nine papers in 2019, involving colleagues from the University of Illinois, Facebook AI, and the Idiap Research Institute. Two of the papers were primarily the work of ISI summer interns. The work is sponsored by DARPA and IARPA.



Schematic of transfer learning across languages; a model trained on English event detection and English/Chinese/Arabic masked word prediction is transferred to enable Arabic and Chinese event detection.

Principal Component Analysis visualization of English entities and English and Chinese words in a multilingual representation space trained on multilingual data harmonized by projecting common entity identification cross-lingually, with the aid of Wikipedia links. Courtesy of UIUC.

PIRANHA Turns the Tables on Spearfishers

Targeted email attacks (“spearphishing”) increasingly plague businesses and individuals.

According to a 2019 Symantec report, 65% of attacker groups doing targeted attacks used spearphishing as the primary infection vector. Spearphishing attacks are difficult to detect, for both end users and anti-spam devices, because they “look and feel” similar to legitimate email. In addition, many spearphishing attacks come from inside an enterprise via a compromised email account, which makes malicious email messages indistinguishable from valid email.

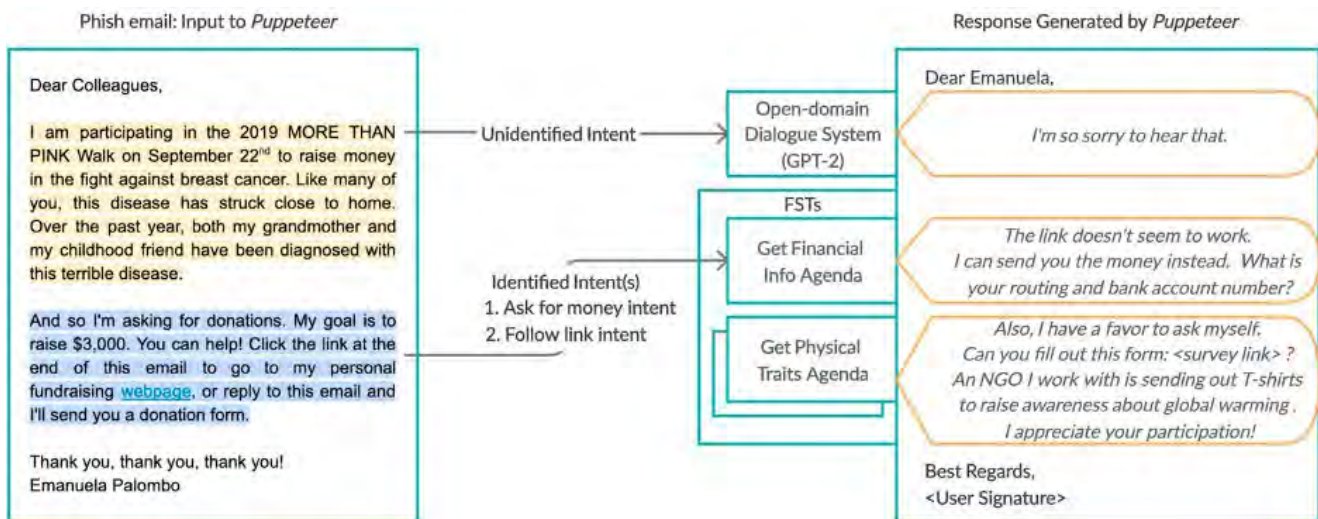
Past research in this area focused on detection and avoidance. However, the PIRANHA effort

is developing automated systems to 1) automatically detect phishing attacks; 2) engage the attackers in realistic dialogue; and 3) surface intelligence about the attackers. If successful, this effort will change the paradigm for one of the most virulent and frequently used cyberattack vectors.

The detection focuses on persona vetting, that is, comparing who the sender claims to be with details in their message. To engage with the attacker, automatic dialog pushes the scammers toward actions, such as directing the scammer to a honeypot or eliciting specific information from them. The actions can then be used by law enforcement or other agencies to help identify the threat actors.

The ISI team is developing a hybrid dialogue system, Puppeteer, which uses customizable probabilistic finite state transducers to orchestrate pushing agendas, coupled with an open-domain dialogue system that generates responses to unexpected prompts. This hybrid system helps meet the unique challenges of the problem space, enabling the PIRANHA system to push scripted agendas and respond naturally to open-ended dialogue.

ISI is working with subcontractors Northrup Grumman, UC Davis, and Next Century Corporation as part of the DARPA Active Social Engineering Defense program (ASED).



CYBERSECURITY

Mismorphism: The Heart of the Weird Machine

P. Anantharaman, V. Kothari, I. Jenkins, M. Millian, S. Bratus, J. Blythe, R. Koppel and S. Smith
International Workshop on Security Protocols

Optimization of Large-scale Agent-based Simulations through Automated Abstraction and Simplification

J. Blythe and A. Tregubov
International Joint Conference on Artificial Intelligence

The DARPA SocialSim Challenge: Massive Multi-Agent Simulations of the Github Ecosystem

J. Blythe, E. Ferrara, D. Huang, K. Lerman, G. Muric, A. Sapienza, A. Tregubov, D. Pacheco, J. Bollenbacher, A. Flammini, P. Hui and F. Menczer
International Conference on Practical Applications of Agents and Multi-Agent Systems AAMAS

Massive Multi-Agent Data-Driven Simulations of the GitHub Ecosystem

J. Blythe, J. Bollenbacher, D. Huang, P. Hui, R. Krohn, D. Pacheco, G. Muric, A. Sapienza, A. Tregubov, Y. Ahn, A. Flammini, K. Lerman, F. Menczer, T. Weninger and E. Ferrara
International Conference on Practical Applications of Agents and Multi-Agent Systems

BootKeeper: Validating Software Integrity Properties on Boot Firmware Images

R. Chevalier, S. Cristalli, C. Hauser, Y. Shoshitaishvili, R. Wang, C. Kruegel, G. Vigna, D. Bruschi and A. Lanzi
Proceedings of the Ninth ACM Conference on Data and Application Security and Privacy (CODASPY)

Learning Adversarial Interactions in Stackelberg Security Games with Limited Data

B. Fang, M. Jiang, A. Tregubov, J. Blythe and E. Ferrara
NeurIPS workshop on Bridging Game Theory and Deep Learning

Sleak: Automating Address Space Layout Derandomization

C. Hauser, J. Menon, Y. Shoshitaishvili, R. Wang, G. Vigna and C. Kruegel
Proceedings of the 35th Annual Computer Security Applications Conference

Security-driven Metrics and Models for Efficient Evaluation of Logic Encryption Schemes

Y. Hu, V. V. Menon, A. Schmidt, J. Monson, M. French and P. Nuzzo
Proceedings of the 17th ACM-IEEE International Conference on Formal Methods and Models for System Design

Human-Computability Boundaries

V. Kothari, P. Anantharaman, I. Jenkins, M. Millian, S. Bratus, J. Blythe, R. Koppel and S. Smith
International Workshop on Security Protocols

Thin Wing EMA-System on Chip with Synchronized TRust and Assurance (SaSTRA)

P. Kshirsagar and V. Venugopalan
Government Microcircuit Applications & Critical Technology Conference (GOMACTech)

Measuring Student Learning on Network Testbeds

P. Lepe, A. Aggarwal, J. Mirkovic, J. Mache, R. Weiss and D. Weinmann
ICNP MERIT Workshop

System-Level Framework for Logic Obfuscation with Quantified Metrics for Evaluation

V. V. Menon, G. Kolhe, A. Schmidt, J. Monson, M. French, Y. Hu, P. A. Beerel and P. Nuzzo
IEEE Cybersecurity Development (SecDev)

Designing for Fallible Humans

J. Mirkovic and S. S. Woo
Proceedings of Humans and CyberSecurity Workshop (HACS) - Invited Paper

Defending Web Servers Against Flash Crowd Attacks

R. Tandon, A. Palia, J. Ramani, B. Paulsen, G. Bartlett and J. Mirkovic
ICNP MERIT Workshop

Optimization of Large-scale Agent-based Simulations through Automated Abstraction and Simplification

A. Tregubov and J. Blythe

AAMAS International Workshop on Multi-Agent-Based Simulation

Quantifying Security and Overheads for Obfuscation of Integrated Circuits

V. Venugopalan, G. Kolhe, A. Schmidt, J. Monson, M. French, Y. Hu, P. A. Beerel and P. Nuzzo

Government Microcircuit Applications & Critical Technology Conference (GOMACTech)

Using Episodic Memory for User Authentication

S. S. Woo, R. Artstein, E. Kaiser, X. Le and J. Mirkovic

ACM Transactions on Privacy and Security

KNOWLEDGE TECHNOLOGIES

Automating Ontology Engineering Support Activities with OnToology

A. Alobaid, D. Garijo, M. Poveda-Villalón, I. Santana-Perez, A. Fernández-Izquierdo and O. Corcho

Journal of Web Semantics

Optimal and Bounded Suboptimal Multi-Agent Motion Planning

L. Cohen, T. Uras, S. Thittamaranahalli and S. Koenig

Proceedings of the Twelfth International Symposium on Combinatorial Search (SoCS-2019)

Automatic Alignment of Contemporary Vector Data and Georeferenced Historical Maps Using Reinforcement Learning

W. Duan, Y. Chiang, S. Leyk, J. H. Uhl and C. Knoblock

International Journal of Geographical Information Science

Semantic Workflows and Machine Learning for the Assessment of Carbon Storage by Urban Trees

J. M. C. Garcia, D. Garijo, M. Crowley, R. Carrillo, Y. Gil and K. Borda

Proceedings of the Third Workshop on Capturing Scientific Knowledge (SciKnow 2019), in conjunction with the 2019 ACM International Conference on Knowledge Capture (K-CAP)

Towards Automated Hypothesis Testing in Neuroscience

D. F. Garijo

Heterogeneous Data Management, Polystores, and Analytics for Healthcare

WDPlus: Leveraging Wikidata to Link and Extend Tabular Data

D. Garijo and P. Szekely

Proceedings of the Third Workshop on Capturing Scientific Knowledge (SciKnow 2019), in conjunction with the 2019 ACM International Conference on Knowledge Capture (K-CAP)

An Intelligent Interface for Integrating Climate, Hydrology, Agriculture, and Socioeconomic Models

D. Garijo, D. Khider, V. Ratnakar, Y. Gil, E. Deelman, R. F. da Silva, C. Knoblock, Y. Chiang, M. Pham,

J. Pujara, B. Vu, D. Feldman, R. Mayani, K. Cobourn, C. Duffy, A. Kemanian and others

Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion

OKG-Soft: An Open Knowledge Graph with Machine Readable Scientific Software Metadata

D. Garijo, M. Osorio, D. Khider, V. Ratnakar and Y. Gil

Proceedings of the Fifteenth IEEE International Conference on eScience

Towards Human-Guided Machine Learning

Y. Gil, J. Honaker, S. Gupta, Y. Ma, V. D. Orazio, D. Garijo, S. Gadewar, Q. Yang and N. Jahanshad

Proceedings of the 24th ACM International Conference on Intelligent User Interfaces (IUI)

PUBLICATIONS | 2019

Intelligent Systems for Geosciences: an Essential Research Agenda

Y. Gil, S. Pierce, H. Babaie, A. B., K. Borne, G. Bust, M. Cheatham, I. E., C. Gomes, M. Hill, J. Horel, L. Hsu, J. K., C. Knoblock, D. Krum, V. Kumar and others
Communications of the ACM

PaCTS 1.0: A Crowdsourced Reporting Standard for Paleoclimate Data

D. Khider, J. Emile-Geay, N. P. McKay, Y. Gil, D. Garijo, V. Ratnakar, M. Alonso-Garcia, S. Bertrand, O. Bothe, P. Brewer, A. Bunn, M. Chevalier, L. Comas-Bru, A. Csank, E. Dassié, K. DeLong, T. Felis, P. Francus and others
Paleoceanography and Paleoclimatology

MINT: An intelligent Interface for Understanding the Impacts of Climate Change on Hydrological, Agricultural and Economic Systems

D. Khider, Y. Gil, K. M. Cobourn, E. Deelman, C. Duffy, R. F. da Silva, A. Kemanian, C. Knoblock, V. Kumar, S. D. Peckham and others
AGU Fall Meeting 2019

Solving Graph Problems in a Euclidean Space Using FastMap

J. Li, A. Felner, S. Koenig and S. Thittamaranahalli
Proceedings of the Twenty-Ninth International Conference on Automated Planning and Scheduling (ICAPS-2019)

Multi-Agent Path Finding for Large Agents

J. Li, P. Surynek, A. Felner, H. Ma, S. Thittamaranahalli and S. Koenig
Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-2019)

GAIA at SM-KBP 2019 - A Multi-Media Multi-Lingual Knowledge Extraction and Hypothesis Generation System

M. Li, Y. Lin, A. Subburathinam, S. Whitehead, X. Pan, D. Lu, Q. Wang, T. Zhang, L. Huang, H. Ji, A. Zareian, H. Akbari, B. Chen, B. Wu, E. Allaway, S. Chang, K. McKeown, Y. Yao and others
Proceedings of the 2019 Text Analysis Conference (TAC-2019)

Lifelong Path Planning with Kinematic Constraints for Multi-Agent Pickup and Delivery

H. Ma, W. Hoenig, S. Thittamaranahalli, N. Ayanian and S. Koenig
Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-2019)

SoMEF: A Framework for Capturing Scientific Software Metadata from its Documentation

A. Mao, D. Garijo and S. Fakhraei
REU Symposium, co-located with 2019 IEEE International Conference on Big Data (IEEE BigData 2019)

Linking Abstract Plans of Scientific Experiments to their Corresponding Execution Traces

M. Markovic, D. Garijo and P. Edwards
Proceedings of the Third Workshop on Capturing Scientific Knowledge (SciKnow 2019), in conjunction with the 2019 ACM International Conference on Knowledge Capture (K-CAP)

Semantic Modelling of Plans and Execution Traces for Enhancing Transparency of IoT Systems

M. Markovic, D. Garijo, P. Edwards and W. Vasconcelos
2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)

Software Adaptation for an Unmanned Undersea Vehicle

A. Pfeffer, C. Wu, G. Fry, K. Lu, S. Marotta, M. Repos, Y. Shi, T. S. Kumar, C. Knoblock, D. Parker, I. Muhammad and C. Novakovic
IEEE Software

Learning Data Transformations with Minimal User Effort

M. Pham, C. Knoblock and J. Pujara
2019 IEEE International Conference on Big Data (IEEE BigData 2019)

PRIMAL: Pathfinding via Reinforcement and Imitation Multi-Agent Learning

G. Sartoretti, J. Kerr, Y. Shi, G. Wagner, S. Thittamaranahalli, S. Koenig and H. Choset
IEEE Robotics and Automation Letters (RA-L-2019)

Parsing, Representing and Transforming Units of Measure

B. Shbita, A. Rajendran, J. Pujara and C. Knoblock
Proceedings of the Conference on Modeling the World's Systems

Automatic Adaptation to Sensor Replacements

Y. Shi, T. S. Kumar and C. Knoblock
The Thirty-Second International Flairs Conference

Semantic Workflows for Benchmark Challenges: Enhancing Comparability, Reusability and Reproducibility

A. Srivastava, R. Adusumilli, H. Boyce, D. Garijo, V. Ratnakar, R. Mayani, T. Yu, R. Machiraju, Y. Gil and P. Mallick
Proceedings of the Pacific Symposium on Biocomputing (PSB)

Multi-Agent Path Finding: Definitions, Variants and Benchmarks

R. Stern, N. Sturtevant, D. Atzmon, T. Walker, J. Li, L. Cohen, H. Ma, S. Thittamaranahalli, A. Felner and S. Koenig
Proceedings of the Twelfth International Symposium on Combinatorial Search (SoCS-2019)

Top K Hypotheses Selection on a Knowledge Graph

K. Sun, K. Maddali, S. Salian and S. Thittamaranahalli
Proceedings of the Thirty-Second International FLAIRS Conference (FLAIRS-2019)

Multi-Agent Path Finding with Capacity Constraints

P. Surynek, S. Thittamaranahalli and S. Koenig
Proceedings of the Eighteenth International Conference of the Italian Association for Artificial Intelligence (AIIA-2019)

T2WML: A Cell-Based Language to Map Tables into Wikidata Records

P. A. Szekely, D. Garijo, J. Pujara, D. Bhatia and J. Wu
ISWC Satellites

T2WML: Table To Wikidata Mapping Language

P. Szekely, D. Garijo, D. Bhatia, J. Wu, Y. Yao and J. Pujara
Proceedings of the 10th International Conference on Knowledge Capture

ENIGMA and Global Neuroscience: A Decade of Large-Scale Studies of the Brain in Health and Disease across more than 40 Countries

P. Thompson, N. Jahanshad, C. R. K. Ching, L. Salminen, S. I. Thomopoulos, J. Bright, B. T. Baune, S. Bertolí, J. Bralten, W. B. Bruin, et. al.
PsyArXiv

Automated Extraction of Human Settlement Patterns from Historical Topographic Map Series Using Weakly Supervised Convolutional Neural Networks

J. H. Uhl, S. Leyk, Y. Chiang, W. Duan and C. Knoblock
IEEE Access

Learning Semantic Models of Data Sources Using Probabilistic Graphical Models

B. Vu, C. Knoblock and J. Pujara
Proceedings of the 2019 World Wide Web Conference (WWW '19)

D-REPR: A Language for Describing and Mapping Diversely-Structured Data Sources to RDF

B. Vu, J. Pujara and C. Knoblock
Proceedings of the 10th International Conference on Knowledge Capture

A New Constraint Satisfaction Perspective on Multi-Agent Path Finding: Preliminary Results

J. Wang, J. Li, H. Ma, S. Koenig and S. Thittamaranahalli
Proceedings of the Eighteenth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-2019)

PUBLICATIONS | 2019

Quadratic Reformulation of Nonlinear Pseudo-Boolean Functions via the Constraint Composite Graph

K. W. Yip, H. Xu, S. Koenig and S. Thittamaranahalli

Proceedings of the Sixteenth International Conference on the Integration of Constraint Programming, Artificial Intelligence and Operations Research (CPAIOR-2019)

Extracting Geographic Features from the Internet: A Geographic Information Mining Framework

Y. Zhang, Q. Ma, Y. Chiang, C. Knoblock, X. Zhang, P. Yang, M. Gao and X. Hu

Knowledge-Based Systems

MACHINE LEARNING AND DATA SCIENCE

SAGE: A Hybrid Geopolitical Event Forecasting System

F. Morstatter, A. Galstyan, G. Satyukov, D. Benjamin, A. Abeliuk, M. Mirtaheeri, K. T. Hossain, P. A. Szekely, E. Ferrara, A. Matsui, M. Steyvers, S. Bennett, D. V. Budescu and M. A. Himmelstein

Proceedings of the 28th International Joint Conference on Artificial Intelligence

MixHop: Higher-Order Graph Convolution Architectures via Sparsified Neighborhood Mixing

S. Abu-El-Haija, B. Perozzi, A. Kapoor, H. Harutyunyan, N. Alipourfard, K. Lerman, G. V. Steeg and A. Galstyan

International Conference on Machine Learning (ICML)

Linguistic Cues to Deception: Identifying Political Trolls on Social Media

A. Addawood, A. Badawy, K. Lerman and E. Ferrara

Proceedings of the International AAAI Conference on Web and Social Media

A Large-Scale Study of ISIS Social Media Strategy: Community Size, Collective Influence, and Behavioral Impact

M. Alfifi, P. Kaghazgaran, J. Caverlee and F. Morstatter

Proceedings of the International AAAI Conference on Web and Social Media

Linking Educational Resources on Data Science

J. L. Ambite, J. Gordon, L. Fierro, G. Burns and J. Matthew

Proceedings of the 31st Innovative Applications of Artificial Intelligence Conference (IAAI)

BD2K Training Coordinating Centers ERuDIte: the Educational Resource Discovery Index for Data Science

J. L. Ambite, L. Fierro, J. Gordon, G. A. Burns, F. Geigl, K. Lerman and J. D. V. Horn

IEEE Transactions on Emerging Topics in Computing

Characterizing the 2016 Russian IRA Influence Campaign

A. Badawy, A. Addawood, K. Lerman and E. Ferrara

Social Network Analysis and Mining

Who Falls for Online Political Manipulation

A. Badawy, K. Lerman and E. Ferrara

Companion Proceedings of the 2019 World Wide Web Conference: CyberSafety Workshop

The DARPA SocialSim Challenge: Massive Multi-Agent Simulations of the Github Ecosystem

J. Blythe, E. Ferrara, D. Huang, K. Lerman, G. Muric, A. Sapienza, A. Tregubov, D. Pacheco, J. Bollenbacher, A. Flammini and others

Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems

Massive Multi-Agent Data-Driven Simulations of the GitHub Ecosystem

J. Blythe, J. Bollenbacher, D. Huang, P. Hui, R. Krohn, D. Pacheco, G. Muric, A. Sapienza, A. Tregubov, Y. Ahn and others

International Conference on Practical Applications of Agents and Multi-Agent Systems

Massive Multi-agent Data-Driven Simulations of the GitHub Ecosystem

J. Blythe, J. Bollenbacher, D. Huang, P. Hui, R. Krohn, D. Pacheco, G. Muric, A. Sapienza, A. Tregubov, Y. Ahn, A. Flammini, K. Lerman, F. Menczer, T. Weninger and E. Ferrara

Advances in Practical Applications of Survivable Agents and Multi-Agent Systems: The PAAMS Collection

Multimodal Human and Environmental Sensing for Longitudinal Behavioral Studies in Naturalistic Settings: Framework for Sensor Selection, Deployment, and Management

B. M. Booth, K. Mundnich, T. Feng, A. Nadarajan, T. H. Falk, J. L. Villatte, E. Ferrara and S. Narayanan

Journal of Medical Internet Research

Exact Rate-Distortion in Autoencoders via Echo Noise

R. Brekelmans, D. Moyer, A. Galstyan and G. V. Steeg

Advances in Neural Information Processing Systems (NeurIPS)

Inferring Models of Opinion Dynamics from Aggregated Jury Data

K. Burghardt, W. Rand and M. Girvan

PLOS ONE

Building Explainable Predictive Analytics for Location-Dependent Time-Series Data

Y. Chiang, Y. Lin, M. Franklin, S. P. Eckel, J. L. Ambite and W. Ku

The First IEEE International Conference on Cognitive Machine Intelligence

Quantifying the Effects of Recommendation Systems

S. Chong and A. Abeliuk

2019 IEEE International Conference on Big Data (IEEE BigData 2019)

Impact of fine-scale population structure in the UK Biobank on Mendelian disease variants

S. Cullina, G. Belbin, R. Shemirani, J. Ambite, C. Gignoux and E. Kenny

Annual Meeting of the American Society of Human Genetics

Perils and Challenges of Social Media and Election Manipulation Analysis: The 2018 US Midterms

A. Deb, L. Luceri, A. Badawy and E. Ferrara

Companion Proceedings of the 2019 World Wide Web Conference

Statistical and Machine Learning Methods for Using Sensor Monitoring Systems to Predict Asthma Exacerbations

S. P. Eckel, K. Li, H. Deng, S. Xu, R. Habre, M. Girguis, A. Bui, K. Sward, R. Gouripeddi, S. Collingwood, R. Urman, J. Morrison, M. Franklin, J. L. Ambite, Y. Chiang and D. Stripe

Military Health System Research Symposium

Collective Alignment of Large-scale Ontologies

V. Embar, J. Pujara and L. Getoor

AKBC Workshop on Federated KBs and the Open Knowledge Network

NSEEN: Neural Semantic Embedding for Entity Normalization

S. Fakhraei, J. Mathew and J. L. Ambite

European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)

Learning Adversarial Interactions in Stackelberg Security Games with Limited Data

B. Fang, M. Jiang, A. Tregubov, J. Blythe and E. Ferrara

NeurIPS 2019

Dynamics of Attention and Public Opinion in Social Media

E. Ferrara

The Oxford Handbook of Networked Communication

PUBLICATIONS | 2019

The History of Digital Spam

E. Ferrara

Communications of the ACM

Auto-Encoding Total Correlation Explanation

S. Gao, R. Brekelmans, G. V. Steeg and A. Galstyan

AI & Statistics (AISTATS)

Nearly-Unsupervised Hashcode Representations for Biomedical Relation Extraction

S. Garg, A. Galstyan, G. V. Steeg and G. A. Cecchi

Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)

Kernelized Hashcode Representations for Relation Extraction

S. Garg, A. Galstyan, G. V. Steeg, I. Rish, G. Cecchi and S. Gao

Proceedings of the AAAI Conference on Artificial Intelligence

An Intelligent Interface for Integrating Climate, Hydrology, Agriculture, and Socioeconomic Models

D. Garijo, D. Khider, V. Ratnakar, Y. Gil, E. Deelman, R. F. da Silva, C. Knoblock, Y. Chiang, M. Pham, J. Pujara, B. Vu, D. Feldman, R. Mayani, K. Cobourn, C. Duffy, A. Kemanian and others

Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion

OKG - Soft: An Open Knowledge Graph with Machine Readable Scientific Software Metadata

D. Garijo, M. Osorio, D. Khider, V. Ratnakar and Y. Gil

Proceedings of the Fifteenth IEEE International Conference on eScience

Tabular Cell Classification Using Pre-Trained Cell Embeddings

M. Ghasemi-Gol, J. Pujara and P. Szekely

International Conference on Data Mining

Multitask Learning and Benchmarking with Clinical Time Series Data

H. Harutyunyan, H. Khachatrian, D. C. Kale, G. V. Steeg and A. Galstyan

Nature Scientific Data

Hidden in Plain Sight: A Machine Learning Approach for Detecting Prostitution Activity in Phoenix, Arizona

E. Helderop, J. Huff, F. Morstatter, A. Grubestic and D. Wallace

Applied Spatial Analysis and Policy

Advancing the International Data Science Workforce Through Shared Training and Education

J. V. Horn, S. Abe, J. L. Ambite, T. K. A. a. Beard, L. Bellis, A. Bhattra, A. Bui, G. B. a. Fierro, J. Gordon, J. Grethe, J. K. a. Lei, K. Lerman, A. McGrath and others

F1000 Research

A Gentle Introduction to Graph Neural Network (Basics, DeepWalk, and GraphSage)

K. Huang

Towards Data Science

Hands-on Graph Neural Networks with PyTorch & PyTorch Geometric

K. Huang

Towards Data Science

Understanding Cyberbullying on Instagram and Ask.fm via Social Role Detection

H. Kao, S. Yan, D. Huang, N. Bartley, H. Hosseinmardi and E. Ferrara

Companion Proceedings of the 2019 World Wide Web Conference

PaCTS 1.0: A Crowdsourced Reporting Standard for Paleoclimate Data

D. Khider, J. Emile-Geay, N. McKay, Y. Gil, D. Garijo, V. Ratnakar, M. Alonso-Garcia, S. Bertrand, O. Bothe, P. Brewer, A. Bunn, M. Chevalier, L. Comas-Bru, A. Csank and E. Dassié
Paleoceanography and Paleoclimatology

Collective Entity Resolution in Multi-relational Familial Networks

P. Kouki, J. Pujara, C. Marcum, L. Koehly and L. Getoor
Knowledge and Information Systems

Personalized Explanations for Hybrid Recommender Systems

P. Kouki, J. Schaffer, J. Pujara, J. O. Donovan and L. Getoor
ACM International Conference on Intelligent User Interfaces

Lessons Learned: Recommendations For Implementing a Longitudinal Study Using Wearable and Environmental Sensors in a Health Care Organization

M. L. Hommedieu, J. L. Hommedieu, C. Begay, A. Schenone, L. Dimitropoulou, G. Margolin, T. Falk, E. Ferrara, K. Lerman and S. Narayanan
JMIR mHealth and uHealth

SoReC: A Social-Relation Based Centrality Measure in Mobile Social Networks

B. Li, Z. Gao, X. Shan, W. Zhou and E. Ferrara
26th International Conference on Telecommunications

Applying Multivariate Segmentation Methods to Human Activity Recognition from Wearable Sensors Data

K. Li, R. Habre, H. Deng, R. Urman, J. Morrison, F. D. Gilliland, J. Ambite, D. Stripelis, Y. Chiang, Y. Lin, A. A. T. Bui, C. King, A. Hosseini, E. V. Vliet, M. Sarrafzadeh and S. P. Eckel
JMIR MHealth and UHealth

Effects of Network Structure on Subjective Preference Diversity

A. Lin, A. Abeliuk and E. Ferrara
2019 IEEE International Conference on Big Data (IEEE BigData 2019)

Genetics of Chronic Kidney Disease Stages Across Ancestries: The PAGE Study

B. M. Lin, G. N. Nadkarni, R. Tao, M. Graff, M. Fornage, S. Buyske, T. C. Matise, H. M. Highland, L. R. Wilkens, C. S. Carlson, S. L. Park, V. W. Setiawan, J. L. Ambite, G. Heiss and others
Frontiers in Genetics

Red Bots Do It Better: Comparative Analysis of Social Bot Partisan Behavior

L. Luceri, A. Deb, A. Badawy and E. Ferrara
Companion Proceedings of the 2019 World Wide Web Conference

Evolution of Bot and Human Behavior During Elections

L. Luceri, A. Deb, S. Giordano and E. Ferrara
First Monday

Biomedical Named Entity Recognition via Reference-Set Augmented Bootstrapping

J. Mathew, S. Fakhraei and J. L. Ambite
International Conference on Machine Learning (ICML), Workshop on Computational Biology

Automated Processing of Phenotypic Data Submissions

R. Mayani, K. Vahi, J. Ambite, S. Sharma, M. Azaro, S. Wilson, B. Ruocco, G. Davis, M. Romanella, L. Brzustowicz, E. Deelman and Y. Arens
Annual Meeting of the American Society of Human Genetics

Debiasing Community Detection: the Importance of Lowly Connected Nodes

N. Mehrabi, F. Morstatter, N. Peng and A. Galstyan
Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

PUBLICATIONS | 2019

Collaboration Drives Individual Productivity

G. Muric, A. Abeliuk, K. Lerman and E. Ferrara

Proceedings of the 2019 ACM Conference on Human-Computer Interaction

A phenome-wide Association Study (PheWAS) in the Population Architecture using Genomics and Epidemiology (PAGE) Study Reveals Potential Pleiotropy in African Americans

S. Pendergrass, S. Buyske, J. Jeff, A. Frase, S. Dudek, Y. Bradford, J. Ambite, C. Avery, P. Buzkova, E. Deelman, M. Fesinmeyer, C. Haiman, G. Heiss, L. Hindorff, C. Hsu, R. Jackson, Y. Lin, L. L. Marchand, T. Matise, K. Monroe and others

PLOS One

Analyzing Tumor Gene Expression Factors with the CorExplorer Web Portal

S. Pepke, W. M. Nelson and G. V. Steeg

Journal of Visualized Experiments

Learning Data Transformations with Minimal User Effort

M. Pham, C. Knoblock and J. Pujara

2019 IEEE International Conference on Big Data (IEEE BigData 2019)

A Common Framework for Developing Table Understanding Models

J. Pujara, A. Rajendran, M. Ghasemi-Gol and P. Szekely

International Semantic Web Conference Posters

Maximizing Multivariate Information with Error-Correcting Codes

K. Reing, G. V. Steeg and A. Galstyan

IEEE Transactions on Information Theory

Considerations for Globigerinoides Ruber (White and Pink) Paleoceanography: Comprehensive Insights from a Long-Running Sediment Trap

J. N. Richey, K. Thirumalai, D. Khider, C. E. Reynolds, J. W. Partin and T. M. Quinn

Paleoceanography and Paleoclimatology

Enhancing Access to Data at the NIMH Repository and Genomics Resource

B. Ruocco, R. Mayani, S. Sharma, S. Wilson, K. Vahi, S. Voinea, G. Davis, J. Valentine-Cooper, W. Valentine-Cooper, J. Mathew, M. Romanella, Y. Arens, E. Deelman, M. A., V. Vieland, J. Ambite and L. Brzustowicz

Annual Meeting of the American Society of Human Genetics

Deep Neural Networks for Optimal Team Composition

A. Sapienza, P. Goyal and E. Ferrara

Frontiers in Big Data

Finding Prerequisite Relations Using the Wikipedia Clickstream

M. Sayyadiharikandeh, J. Gordon, J. L. Ambite and K. Lerman

Companion Proceedings of the 2019 World Wide Web Conference: CyberSafety Workshop

Parsing, Representing, and Transforming Units of Measure

B. Shbita, A. Rajendran, J. Pujara and C. Knoblock

Modeling the World's Systems

Efficient and Rapid Clustering of Identity-by-descent Tracts in Biobank-scale Datasets

R. Shemirani, G. Belbin, E. Kenny, C. Gignoux and J. Ambite

Annual Meeting of the American Society of Human Genetics

Fast Structure Learning with Modular Regularization

G. V. Steeg, H. Harutyunyan, D. Moyer and A. Galstyan

Advances in Neural Information Processing Systems

A Pipeline for Large-Scale Simulation of Population-Based Linkage

E. Sticca, G. Belbin, J. Soto, R. Shemirani, D. Nelson, S. Gravel, J. Ambite, G. Wojcik, E. Kenny and C. Gignoux
Annual Meeting of the American Society of Human Genetics

T2WML: Table to Wikidata Mapping Language

P. Szekely, D. Garijo, D. Bhatia, J. Wu, Y. Yao and J. Pujara
ACM International Conference on Knowledge Capture (KCAP)

Characterizing Activity on the Deep and Dark Web

N. Tavabi, N. Bartley, A. Abeliuk, S. Soni, E. Ferrara and K. Lerman
Companion Workshop Proceedings of the 2019 World Wide Web Conference: CyberSafety Workshop

D-REPR: A Language for Describing and Mapping Diversely-Structured Data Sources to RDF

B. Vu, C. Knoblock and J. Pujara
ACM International Conference on Knowledge Capture (KCAP)

Learning Semantic Models of Data Sources Using Probabilistic Graphical Models

B. Vu, C. Knoblock and J. Pujara
Proceedings of the 2019 World Wide Web Conference (WWW '19)

Personality and Behavior in Role-based Online Games

Z. Wang, A. Sapienza, A. Culotta and E. Ferrara
IEEE Conference on Games

Comparison of Heart Rate Measurement Between the Fitbit Charge 2 and Omsignal Smart Garments: A free-living study

M. Wildman, R. Eulogio, R. Singh, E. Ramirez, L. Foschini, A. Nadarajan, B. M. Booth, K. Mundnich, E. Ferrara, K. Lerman and others
Annals of Behavioral Medicine

Genetic Analyses of Diverse Populations Improves Discovery for Complex Traits

G. L. Wojcik, M. Graff, K. K. Nishimura, R. Tao, J. Haessler, C. R. Gignoux, H. M. Highland, Y. M. Patel, E. P. Sorokin, C. L. Avery, G. M. Belbin, S. A. Bien, I. Cheng and S. Cullina
Nature

Misinformation in Social Media: Definition, Manipulation, and Detection

L. Wu, F. Morstatter, K. M. Carley and H. Liu
ACM SIGKDD Explorations Newsletter

Estimating Individualized Daily Self-Reported Affect with Wearable Sensors

S. Yan, H. Hosseinmardi, H. Kao, S. Narayanan, K. Lerman and E. Ferrara
2019 IEEE International Conference on Healthcare Informatics (ICHI)

Arming the Public with Artificial Intelligence to Counter Social Bots

K. Yang, O. Varol, C. A. Davis, E. Ferrara, A. Flammini and F. Menczer
Human Behavior and Emerging Technologies

Extensible and Scalable Entity Resolution for Financial Datasets Using RLTK

Y. Yao, P. Szekely and J. Pujara
SIGMOD Workshop on Data Science for Macromodeling with Financial and Economic Datasets

The Influence of Social Ties on Performance in Team-based Online Games

Y. Zeng, A. Sapienza and E. Ferrara
IEEE Transactions on Games

Climate Models can Correctly Simulate the Continuum of Global-average Temperature Variability

F. Zhu, J. Emile-Geay, N. P. McKay, G. J. Hakim, D. Khider, T. R. Ault, E. J. Steig, S. Dee and J. W. Kirchner
Proceedings of the National Academy of Science

NATURAL LANGUAGE PROCESSING

SARAL: A Low-Resource Cross-Lingual Domain-Focused Information Retrieval System for Effective Rapid Document Triage

E. Boschee, J. Barry, J. Billa, M. Freedman, T. Gowda, C. Lignos, C. Palen-Michel, M. Pust, B. K. Khonglah, S. Madikeri and J. May
Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics - System Demonstrations

Cross-lingual Dependency Parsing with Unlabeled Auxiliary Languages

W. U. Ahmad, Z. Zhang, X. Ma, K. Chang and N. Peng
2019 SIGNLL Conference on Computational Natural Language Learning (CoNLL)

Building Deep Learning Models for Evidence Classification From the Open Access Biomedical Literature

G. A. Burns, X. Li and N. Peng
Database

A Grounded Unsupervised Universal Part-of-Speech Tagger for Low-Resource Languages

R. Cardenas, Y. Lin, H. Ji and J. May
Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)

Better Automatic Evaluation of Open-Domain Dialogue Systems with Contextualized Embeddings

S. Ghazarian, J. T. Wei, A. Galstyan and N. Peng
2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2019), NeuralGen Workshop

Plan, Write, and Revise: An Interactive System for Open-Domain Story Generation

S. Goldfarb-Tarrant, H. Feng and N. Peng
2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2019), Demonstrations Track

Deep Structured Neural Network for Event Temporal Relation Extraction

R. Han, I. Hsu, M. Yang, A. Galstyan, R. Weischedel and N. Peng
2019 SIGNLL Conference on Computational Natural Language Learning (CoNLL)

Contextualized Word Embeddings Enhanced Event Temporal Relation Extraction for Story Understanding

R. Han, M. Liang, B. Alhafni and N. Peng
2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2019), Workshop on Narrative Understanding

Joint Event and Temporal Relation Extraction with Shared Representations and Structured Prediction

R. Han, Q. Ning and N. Peng
2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Pun Generation with Surprise

H. He, N. Peng and P. Liang
2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2019)

Cross-lingual Multi-Level Adversarial Transfer to Enhance Low-Resource Name Tagging

L. Huang, H. Ji and J. May
Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)

What Matters for Neural Cross-Lingual Named Entity Recognition: An Empirical Analysis

X. Huang, J. May and N. Peng
2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Learning A Unified Named Entity Tagger from Multiple Partially Annotated Corpora For Efficient Adaptation

X. Huang, L. Dong, E. Boschee and N. Peng

2019 SIGNLL Conference on Computational Natural Language Learning (CoNLL)

GAIA at SM-KBP 2019 - A Multi-media Multi-lingual Knowledge Extraction and Hypothesis Generation System

M. Li, Y. Lin, A. Subburathinam, S. Whitehead, X. Pan, D. Lu, Q. Wang, T. Zhang, L. Huang, H. Ji, A. Zareian, H. Akbari,

B. Chen, B. Wu, E. Allaway, S. Chang, K. McKeown, Y. Yao and others

Text Analysis Conference TAC 2019 Proceedings Papers

Real-World Causal Relationship Discovery from Text

C. Lignos, C. Palen-Michel, O. Singer, P. A. Szekely and E. Boschee

ISWC Satellites

Contextualized Cross-Lingual Event Trigger Extraction with Minimal Resources

M. M'hamdi, M. Freedman and J. May

Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)

Debiasing Community Detection: The Importance of Lowly-Connected Nodes

N. Mehrabi, F. Morstatter, N. Peng and A. Galstyan

The 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2019)

Target Language-Aware Constrained Inference for Cross-lingual Dependency Parsing

T. Meng, N. Peng and K. Chang

2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Do Nuclear Submarines Have Nuclear Captains? A Challenge Dataset for Commonsense Reasoning over Adjectives and Objects

J. Mullenbach, J. Gordon, N. Peng and J. May

Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing

Task-offload Tools Improve Productivity and Performance in Geopolitical Forecasting

I. Juvina, O. Larue, C. Widmer, S. Ganapathy, S. Nadella, B. Minnery, L. Ramshaw, E. Servan-Schreiber, M. Balick, and R. Weischedel

Short Paper Proceedings of the Workshop on Cognitive Systems for Anticipatory Thinking (COGSAT 2019),

co-located with the AAAI Fall Symposium 2019

Cross-lingual Joint Entity and Word Embedding to Improve Entity Linking and Parallel Sentence Mining

X. Pan, T. Gowda, H. Ji, J. May and S. Miller

Proceedings of the 2nd Workshop on Deep Learning Approaches for Low Resource NLP (DeepLo2019)

Translating Translationese: A Two-Step Approach to Unsupervised Machine Translation

N. Pourdamghani, N. Aldarrab, M. Ghazvininejad, K. Knight and J. May

Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics

The Woman Worked as a Babysitter: On Biases in Language Generation

E. Sheng, K. Chang, P. Natarajan and N. Peng

2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Cross-lingual Structure Transfer for Relation and Event Extraction

A. Subburathinam, D. Lu, H. Ji, J. May, S. Chang, A. Sil and C. Voss

Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing

PUBLICATIONS | 2019

Espresso: A Fast End-to-end Neural Speech Recognition Toolkit

Y. Wang, T. Chen, H. Xu, S. Ding, H. Lv, Y. Shao, N. Peng, L. Xie, S. Watanabe and S. Khudanpur
The 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)

Plan-and-Write: Towards Better Automatic Storytelling

L. Yao, N. Peng, R. Weischedel, K. Knight, D. Zhao and R. Yan
Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI-19)

Comprehensible Context-driven Text Game Playing

X. Yin and J. May
Proceedings of the 2019 IEEE Conference on Games (CoG)

NETWORKING

Precise Detection of Content Reuse in the Web

C. Ardi and J. Heidemann
ACM Computer Communication Review

Mitigating Soft Failures Using Network Analytics And Sdn to Support Distributed Bandwidth-Intensive Scientific Instruments Over International Networks

J. Bezerra, J. Ibarra, D. Boertjes, F. Santillo, L. Williford, H. Morgan, C. Cox and L. Lopez
SubOptic Conference 2019

Identifying Important Internet Outages

R. Bogutz, Y. Pradkin and J. Heidemann
Sixth National Symposium for NSF REU Research in DataScience

Factors Influencing the Adoption of Software Defined Networking by Research and Educational Networks

V. Chergarova, J. Bezerra, J. Ibarra and H. Morgan
25th Annual Americas Conference on Information Systems (AMCIS) 2019

Novel Network Services for Supporting Big Data Science Research

J. Chung, S. Donovan, J. Bezerra, H. Morgan, J. Ibarra, R. Clark and H. Owen
Future Generation Computer Systems

The DComp Testbed

R. Goodfellow, S. Schwab, E. Kline, L. Thurlow and G. Lawler
12th USENIX Workshop on Cyber Security Experimentation and Test (CSET 19)

Cybersecurity Experimentation at Program Scale: Guidelines and Principles for Future Testbeds

E. Kline and S. Schwab
Cyber Range Applications and Technologies (CACOE)

NRE-011: Big Data Express: A Scalable and High-performance Data Transfer Platform

Q. Lu, L. Zhang, W. Wu, P. Demar, J. Chen, J. Mambretti, C. Guok, I. Monga, T. Lehman, X. Yang, J. Kim, B. Cho, J. Ibarra, H. Morgan, J. Bezerra, V. Chergarova, A. Zahir and G. v. Malenstein
2019 Supercomputing Conference (SC19)

Roll, Roll, Roll Your Root: A Comprehensive Analysis of the First Ever DNSSEC Root KSK Rollover

M. Muller, M. Thomas, D. Wessels, W. Hardaker, T. Chung, W. Toorop and R. v. Rijswijk-Deij
Proceedings of the 2019 Internet Measurement Conference

Collaborating to leverage R&E Network Infrastructures between Africa, Brazil, and the U.S.

H. Morgan, J. Ibarra, J. Bezerra, L. F. Lopez, V. Chergarova, D. A. Cox, M. Stanton, A. Hazin, L. Lotz and S. Mammen
UbuntuNet-Connect 2019

Cache Me If You Can: Effects of DNS Time-to-Live

G. C. M. Moura, J. Heidemann, R. d. O. Schmidt and W. Hardaker
Proceedings of the ACM Internet Measurement Conference

NRE-019: Global Petascale to Exascale Workflows for Data Intensive Science Accelerated by Next Generation Programmable SDN Architectures and Machine Learning Applications

H. Newman, J. Balcas, R. Sirvinskas, C. Iordache, J. Chiu, S. Anderson, J. Barayoga, J. Chang, D. Boyd, L. Watanabe, D. S. Williams, A. Mughal, J. Kantor, M. Kollross, J. Ibarra and others
2019 Supercomputing Conference (SC19)

NRE-023: International Data Transfer over AmLight Express and Protect (ExP)

H. Newman, T. Lehman, J. Ibarra, J. Bezerra, A. Zahir, V. Chergarova, H. Morgan, R. Clark, S. Donovan, S. Novaes, R. Iope, B. Leal, L. Lopez, M. Stanton, R. F. d. Lima and A. Astudillo
2019 Supercomputing Conference (SC19)

Sled: System-Loader for Ephemeral Devices

L. Thurlow and R. Goodfellow
5th IEEE INFOCOM Workshop on Computer and Networking Experimental Research using Testbeds 2019 (IEEE CNERT 2019)

NON-TRADITIONAL AND HETEROGENEOUS COMPUTER ARCHITECTURE

Design Considerations for Mapping FPGA High-Level Synthesis Algorithms to Next-Generation Memory Devices

A. D. Rajagopala., R. Sass and A. G. Schmidt
International Symposium on Memory Systems (MEMSYS)

Impact of Off-Chip Memories on HLS-Generated Circuits

A. D. Rajagopala., R. Sass and A. G. Schmidt
International Workshop on FPGAs for Software Programmers (FSP)

Volcan: System Integration of HLS and HMC with FPGAs

A. D. Rajagopala., R. Sass and A. G. Schmidt
International Conference on Reconfigurable Computing and FPGAs (ReConFig 2019)

Computational Requirements for Real-Time Ptychographic Image Reconstruction

K. Datta, A. Rittenbach, D. Kang, J. P. Walters, S. P. Crago and J. Damoulakis
Applied Optics

Enhanced Independent Functional Testing of Xilinx FPGAs

T. Haroldsen, M. French, A. Schmidt and D. Khamar
Government Microcircuit Applications & Critical Technology Conference (GOMACTech)

Security-driven Metrics and Models for Efficient Evaluation of Logic Encryption Schemes

Y. Hu, V. V. Menon, A. G. Schmidt, J. Monson, M. French and P. Nuzzo
Proceedings of the 17th ACM-IEEE International Conference on Formal Methods and Models for System Design (MEMOCODE)

CoPPer: Soft Real-Time Application Performance Using Hardware Power Capping

C. Imes, H. Zhang, K. Zhao and H. Hoffmann
2019 IEEE International Conference on Autonomic Computing (ICAC)

PUBLICATIONS | 2019

Thin Wing EMA-System on Chip with Synchronized TRust and Assurance (SaSTRA)

P. Kshirsagar and V. Venugopalan

Government Microcircuit Applications & Critical Technology Conference (GOMACTech)

System-Level Framework for Logic Obfuscation with Quantified Metrics for Evaluation

V. V. Menon, G. Kolhe, A. Schmidt, J. Monson, M. French, Y. Hu, P. A. Beerel and P. Nuzzo

IEEE Cybersecurity Development (SecDev)

Constellations in the Cloud: Virtualizing Remote Sensing Systems

A. G. Schmidt, V. Venugopalan, M. Paolieri and M. French

2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019)

Increased Fault-Tolerance and Real-Time Performance Resiliency for Stream Processing Workloads through Redundancy

G. P. Tran, J. P. Walters and S. Crago

2019 IEEE International Conference on Services Computing (SCC)

Quantifying Security and Overheads for Obfuscation of Integrated Circuits

V. Venugopalan, G. Kolhe, A. Schmidt, J. Monson, M. French, Y. Hu, P. A. Beerel and P. Nuzzo

Government Microcircuit Applications & Critical Technology Conference (GOMACTech)

QUANTUM INFORMATION PROCESSING

Role of Nonstoquastic Catalysts In Quantum Adiabatic Optimization

T. Albash

Physical Review A

Future of Physical Quantum Annealers: Impediments and Hopes

T. Albash and I. Hen

Science and Culture

Analog Errors in Ising Machines

T. Albash, V. Martin-Mayor and I. Hen

Quantum Science and Technology

Estimating the Density Of States of Frustrated Spin Systems

L. Barash, J. Marshall, M. Weigel and I. Hen

New Journal of Physics

Demonstration of the Generalized-Kennedy Receiver for Quantum-Limited Discrimination of Photon-Starved Classical Light

S. Gartenstein, J. L. Habif and A. Jagannathan

Laser Science

Demonstration of the Generalized Kennedy Receiver as a Near Quantum-Optimal Measurement for the Discrimination of Weak Classical Optical States

J. L. Habif and A. Jagannathan

APS Meeting Abstracts

Harvesting Planck Radiation for Free-Space Optical Communications in the Long-Wave Infrared Band

J. L. Habib and A. Jagannathan

Laser Science

Equation Planting: A Tool for Benchmarking Ising Machines

I. Hen

Physical Review Applied

How Quantum is the Speedup in Adiabatic Unstructured Search?

I. Hen

Quantum Information Processing

Resolution of the Sign Problem for a Frustrated Triplet of Spins

I. Hen

Physical Review E

Power of Pausing: Advancing Understanding of Thermalization in Experimental Quantum Annealers

J. Marshall, D. Venturelli, I. Hen and E. G. Rieffel

Physical Review Applied

On the Computational Complexity of Curing Non-Stoquastic Hamiltonians

M. Marvian, D. A. Lidar and I. Hen

Nature Communications

Sensitivity of Quantum Speedup by Quantum Annealing to a Noisy Oracle

S. Muthukrishnan, T. Albash and D. A. Lidar

Physical Review Applied

Analog Errors in Quantum Annealing: Doom and Hope

A. Pearson, A. Mishra, I. Hen and D. A. Lidar

NPJ Quantum Information

Analog Nature of Quantum Adiabatic Unstructured Search

M. Slutskii, T. Albash, L. Barash and I. Hen

New Journal of Physics

Quantitative Laser Speckle Contrast Imaging for Presentation Attack Detection in Biometric Authentication Systems

C. Sun, A. Jagannathan, J. L. Habib, M. Hussein, L. Spinoulas and W. AbdAlmageed

Smart Biomedical and Physiological Sensor Technology XV

SCIENTIFIC COMPUTING

A Secure Gateway for Enabling ASIC Design Collaborations

S. Bogol, P. Brenner, A. Brinckman, E. Deelman, R. F. da Silva, S. Gupta, J. Nabrzyski, S. Park, D. Perez, M. Rynge, I. Taylor, K. Vahi, M. V. Werf, R. Sarah and S. Wyngaard

11th International Workshop on Science Gateways (IWSG 2019)

Collaborative Circuit Designs using the CRAFT Repository

A. Brinckman, E. Deelman, S. Gupta, J. Nabrzyski, S. Park, R. F. da Silva, I. J. Taylor and K. Vahi

Future Generation Computer Systems

Applicability Study of the PRIMAD Model to LIGO Gravitational Wave Search Workflows

D. Chapp, D. Rorabaugh, D. A. Brown, E. Deelman, K. Vahi, V. Welch and M. Tauber

Proceedings of the 2nd International Workshop on Practical Reproducible Evaluation of Computer Systems

PUBLICATIONS | 2019

Computational Requirements for Real-Time Ptychographic Image Reconstruction

K. Datta, A. Rittenbach, D. Kang, J. P. Walters, S. P. Crago and J. Damoulakis
Applied Optics

The Role of Machine Learning in Scientific Workflows

E. Deelman, A. Mandal, M. Jiang and R. Sakellariou
The International Journal of High Performance Computing Applications

Cyberinfrastructure Center of Excellence Pilot: Connecting Large Facilities Cyberinfrastructure

E. Deelman, A. Mandal, V. Pascucci, S. Sons, J. Wyngaard, C. F. V. II, S. Petruzza, I. Baldin, L. Christopherson, R. Mitchell, L. Pottier, M. Rynge, E. Scott, K. Vahi, M. Kogank and J. Mann
15th International Conference on eScience (eScience)

The Evolution of the Pegasus Workflow Management Software

E. Deelman, K. Vahi, M. Rynge, R. Mayani, R. F. da Silva, G. Papadimitriou and M. Livny
Computing in Science Engineering

Search SRA Gateway for Metagenomics Data

R. Edwards, E. Abeysinghe, M. Pierce and M. Rynge
Gateways 2019

An Intelligent Interface for Integrating Climate, Hydrology, Agriculture, and Socioeconomic Models

D. Garijo, D. Khider, V. Ratnakar, Y. Gil, E. Deelman, R. F. da Silva, C. Knoblock, Y. Chiang, M. Pham, J. Pujara, B. Vu, D. Feldman, R. Mayani, K. Cobourn, C. Duffy, A. Kemanian and others
ACM 24th International Conference on Intelligent User Interfaces

MCEM: Multi-Level Cooperative Exception Model for HPC Workflows

S. Herbein, D. Domyancic, P. Minner, I. Laguna, R. F. da Silva and D. H. Ahn
9th International Workshop on Runtime and Operating Systems for Supercomputers (ROSS 2019)

Toward a Dynamic Network-centric Distributed Cloud Platform for Scientific Workflows: A Case Study for Adaptive Weather Sensing

E. Lyons, G. Papadimitriou, C. Wang, K. Thareja, P. Ruth, J. Villalobos, I. Rodero, E. Deelman, M. Zink and A. Mandal
15th International Conference on eScience (eScience)

Exploration of Workflow Management Systems Emerging Features from Users Perspectives

R. Mitchell, L. Pottier, S. Jacobs, R. F. da Silva, M. Rynge, K. Vahi and E. Deelman
First International Workshop on Big Data Tools, Methods, and Use Cases for Innovative Scientific Discovery (BTSD)

Training Classifiers to Identify TCP Signatures in Scientific Workflows

G. Papadimitriou, M. Kiran, C. Wang, A. Mandal and E. Deelman
Innovating the Network for Data Intensive Science (INDIS)

A phenome-wide association study (PheWAS) in the Population Architecture using Genomics and Epidemiology (PAGE) study reveals potential pleiotropy in African Americans

S. A. Pendergrass, S. Buyske, J. M. Jeff, A. Frase, S. Dudek, Y. Bradford, J. Ambite, C. L. Avery, P. Buzkova, E. Deelman and others
PLOS One

Enabling Data Streaming-based Science Gateway through Federated Cyberinfrastructure

I. Rodero, Y. Qin, J. Valls, A. Simonet, J. Villalobos, M. Parashar, C. Youn, C. Wang, K. Thareja, P. Ruth, G. Papadimitriou, E. Lyons and M. Zink
Gateways 2019

Integrity Protection for Scientific Workflow Data: Motivation and Initial Experiences

M. Rynge, K. Vahi, E. Deelman, A. Mandal, I. Baldin, O. Bhide, R. Heiland, V. Welch, R. Hill, W. L. Poehlman and F. A. Feltus
Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (Learning)

FaceBase 3.0: A Comprehensive Resource on Craniofacial Development

B. Samuels, T. Ho, C. Kesselman, J. Li, R. Schuler, C. Williams, Y. Yuan and Y. Chai
IADR/AADR/CADR General Session

A High-level User-oriented Framework for Database Evolution

R. E. Schuler and C. Kesselman
31st International Conference on Scientific and Statistical Database Management (SSDBM 19)

Toward FAIR Knowledge Turns in Bioinformatics

R. Schuler, A. Bugacov, M. Blow and C. Kesselman
2019 IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM 2019) StashCache: A Distributed Caching

StashCache: A Distributed Caching Federation for the Open Science Grid

D. Weitzel, M. Zvada, I. Vukotic, R. W. Gardner, B. Bockelman, M. Rynge, E. F. Hernandez, B. Lin and M. Selmececi
Practice and Experience in Advanced Research Computing (PEARC 19)

Federation for the Open Science Grid

D. Weitzel, M. Zvada, I. Vukotic, R. W. Gardner, B. Bockelman, M. Rynge, E. F. Hernandez, B. Lin and M. Selmececi
Practice and Experience in Advanced Research Computing (PEARC 19)

Accurately Simulating Energy Consumption of I/O-intensive Scientific Workflows

R. F. da Silva, A. Orgerie, H. Casanova, R. Tanaka, E. Deelman and F. Suter
Computational Science -- ICCS 2019

Bridging Concepts and Practice in eScience via Simulation-driven Engineering

R. F. da Silva, H. Casanova, R. Tanaka and F. Suter
Workshop on Bridging from Concepts to Data and Computation for eScience, 15th International Conference on eScience (eScience)

Using Simple PID-inspired Controllers for Online Resilient Resource Management of Distributed Scientific Workflows

R. F. da Silva, R. Filgueira, E. Deelman, E. Pairo-Castineira, I. M. Overton and M. Atkinson
Future Generation Computer Systems

Empowering Agroecosystem Modeling with HTC Scientific Workflows: The Cycles Model Use Case

R. F. da Silva, R. Mayani, Y. Shi, A. R. Kemanian, M. Rynge and E. Deelman
First International Workshop on Big Data Tools, Methods, and Use Cases for Innovative Scientific Discovery (BTSD)

Measuring the Impact of Burst Buffers on Data-Intensive Scientific Workflows

R. F. da Silva, S. Callaghan, T. M. A. Do, G. Papadimitriou and E. Deelman
Future Generation Computer Systems

Teaching Parallel and Distributed Computing Concepts in Simulation with WRENCH

R. Tanaka, H. Casanova and R. F. da Silva
Workshop on Education for High-Performance Computing (EduHPC)

Characterization of In Situ and In Transit Analytics of Molecular Dynamics Simulations for Next-generation Supercomputers

S. Thomas, M. Wyatt, T. M. A. Do, L. Pottier, R. F. da Silva, H. Weinstein, M. A. Cuendet, T. Estrada, E. Deelman and M. Taufer
15th International Conference on eScience (eScience)

PUBLICATIONS | 2019

Workflows using Pegasus: Enabling Dark Energy Survey Pipelines

K. Vahi, M. H. Wang, C. Chang, S. Dodelson, M. Rynge and E. Deelman
Astronomical Data Analysis Software and Systems XXVIII

Custom Execution Environments with Containers in Pegasus-enabled Scientific Workflows

K. Vahi, M. Rynge, G. Papadimitriou, D. Brown, R. Mayani, R. F. da Silva, E. Deelman, A. Mandal, E. Lyons and M. Zink
15th International Conference on eScience (eScience)

Genetic Analyses of Diverse Populations Improves Discovery for Complex Traits

G. L. Wojcik, M. Graff, K. K. Nishimura, R. Tao, J. Haessler, C. R. Gignoux, H. M. Highland, Y. M. Patel, E. P. Sorokin, C. L. Avery, G. M. Belbin, S. A. Bien, I. Cheng, S. Cullina and others
Nature

SPACE SYSTEMS AND TECHNOLOGIES

Initial Safety Posture Investigations for Earth Regime Rendezvous and Proximity Operations

D. A. Barnhart, R. Rughani, J. J. Allam and K. W. Clarke
10th International Association for the Advancement of Space Safety (IAASS) Conference

Spacecraft Pose Estimation and Swarm Localization performance under Varying Illumination and Viewing Conditions

W. Bezouska and D. Barnhart
AAS/AIAA Space Flight Mechanics Meeting

Visual Sensor Selection for Satellite Swarm Cooperative Localization

W. Bezouska and D. Barnhart
SPIE Defense and Security Conference

Mission Dodona: Electronic Power System Design, Analysis and Integration

S. Narayanan, R. Rughani, R. Rogers, K. Clarke and J. Allam
33rd Annual Conference on Small Satellites

Swarm RPO and Docking Simulation on a 3DOF Air Bearing Platform

R. Rughani, L. Villafana and D. A. Barnhart
70th International Astronautical Congress (IAC)

Improved CubeSat Mission Reliability Using a Rigorous Top-Down Systems-Level Approach

R. Rughani, R. A. Rogers, J. J. Allam, S. Narayanan, P. Patil, K. Clarke, M. Lariviere, J. D. Plessis, L. Villafana, D. Healy and others
70th International Astronautical Congress (IAC)

Constellations in the Cloud: Virtualizing Remote Sensing Systems

A. G. Schmidt, V. Venugopalan, M. Paolieri and M. French
2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019)

VISION, IMAGE, SPEECH, AND TEXT ANALYTICS

Deep, Landmark-Free Face: Face Alignment, Modeling, and Expression Estimation

F. Chang, A. T. Tran, T. Hassner, I. Masi, R. Nevatia and G. Medioni

International Journal of Computer Vision

Finding Structure in Point Cloud Data with the Robust Isoperimetric Loss

S. Deutsch, I. Masi and S. Soatto

International Conference on Scale Space and Variational Methods in Computer Vision

Efficient and Robust Skeleton-Based Quality Assessment and Abnormality Detection in Human Action Performance

A. Elkholy, M. Hussein, W. Gomaa, D. Damen and E. Saba

IEEE Journal of Biomedical and Health Informatics

Quantitative Laser Speckle Contrast Imaging for Presentation Attack Detection in Biometric Authentication Systems

A. Jagannathan, C. Sun, L. Spinoulas, M. Hussein, J. Habif and W. AbdAlmageed

SPIE Defense and Commercial Systems

RoPAD: Robust Presentation Attack Detection through Unsupervised Adversarial Invariance

A. Jaiswal, S. Xia, I. Masi and W. AbdAlmageed

International Conference on Biometrics (ICB)

AIRD: Adversarial Learning Framework for Image Repurposing Detection

A. Jaiswal, Y. Wu, W. AbdAlmageed, I. Masi and P. Natarajan

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition

Face-Specific Data Augmentation for Unconstrained Face Recognition

I. Masi, A. T. Tran, T. Hassner, G. Sahin and G. Medioni

International Journal of Computer Vision

Does Generative Face Completion Help Face Recognition?

J. Mathai, I. Masi and W. AbdAlmageed

International Conference on Biometrics (ICB)

On the Effectiveness of Laser Speckle Contrast Imaging and Deep Neural Networks for Detecting Known and Unknown Fingerprint Presentation Attacks

H. Mirzaalian, M. Hussein and W. AbdAlmageed

International Conference on Biometrics 2019 (ICB)

Age-Invariant Face Recognition Using Gender Specific 3d Aging Modeling

S. Riaz, Z. Ali, U. Park, J. Choi, I. Masi and P. Natarajan

Multimedia Tools and Applications

Recurrent Convolutional Strategies for Face Manipulation Detection in Videos

E. Sabir, J. Cheng, A. Jaiswal, W. AbdAlmageed, I. Masi and P. Natarajan

CVPR Workshop

IN MEMORIAM | DANNY COHEN

Danny Cohen, a distinguished computer scientist, died in August 2019.

Cohen joined ISI in 1973, recruited from Harvard with encouragement from ARPA and Keith Uncapher, ISI's founder and director. For 20 years, in the early days of ISI, Danny Cohen was an amazing force.

One of his first projects at ISI was the packet-voice project to allow interactive, real-time speech over the ARPANET (and early Internet). It was a forerunner of Voice over Internet Protocol (VoIP). He later implemented "packet-video" when he adapted his visual flight simulator to run over the ARPANET.

In 1974, Cohen implemented online voice communications, the first use of packet switching for real-time voice applications.

He published "RFC 0141: Specifications for the Network Voice Protocol (NPV)" in 1977. Cohen and a group of scientists from all over the US devised a forerunner of Internet telephony and teleconferencing, paving the way for voice and video across the Internet. In 1978, they conducted the first conference call over the ARPANET.

In the late '70s, Cohen spent a year as Caltech computer science faculty teaching graphics, on leave of absence from ISI. In his Computer History Museum oral history, he explains that he was recruited by Ivan Sutherland, who had been his advisor at Harvard; he says also that he learned about VLSI that year. Back at ISI, he and his team developed the "Metal Oxide Semiconductor Implementation Service" (hence, MOSIS), which started as a low-cost, small-volume system for silicon chip prototyping and production. The service combined multiple customers' jobs into shared-project wafers, greatly reducing each participant's manufacturing costs. Moreover, it helped train a generation of students in VLSI integrated circuit design. The first day run was in August 1980. The MOSIS fabrication service—now MOSIS 2.0—continues to this day, having processed some 60,000 chip designs for businesses, universities, and government.

1980 was also the year of his "On Holy Wars and a Plea for Peace" paper. This seminal work examined issues of byte order in computer memory with regard to messaging. He coined the computer terms Big Endian and Little Endian, employing Jonathan Swift's *Gulliver's Travels* as an engaging and insightful metaphor. (In the book, the Big Endians were a political faction that broke their eggs at the large end and rebelled against the Lilliputian King, who required his subjects the Little Endians to break their eggs at the small end.)



In 1985, Cohen and Steve Casner introduced packet video transmitted via satellite. (The photo to the left is the satellite dish that was on the roof of ISI.) Video, of course, is now everywhere.

He intuited that the method of running MOSIS, where customers transacted electronically with the fab service, could be used for e-commerce. In 1989, his paper "Computerized Commerce" was published in the proceedings of IFIP 11th World Computing Congress. He started the FAST project, Fast Acquisition of Standard Parts, during the first Gulf War. The FASTXchange brokering system was a pioneering online business where the Department of Defense ordered parts and the service located the best source. In his

oral history, he mentions the letter of appreciation that the team received from the Department of Defense.

Cohen was also central to the USC/ISI ATOMIC LAN, based on the Caltech Mosaic components. This was the research prototype for Myrinet, a high-speed local area networking system used as an interconnect between multiple machines to form computer clusters. He and colleagues went on to found Myricom, which commercialized Myrinet.

Danny Cohen left ISI in 1993 and went on to continue his long career, with many additional accomplishments.

From Steve Crocker, an ISI colleague:

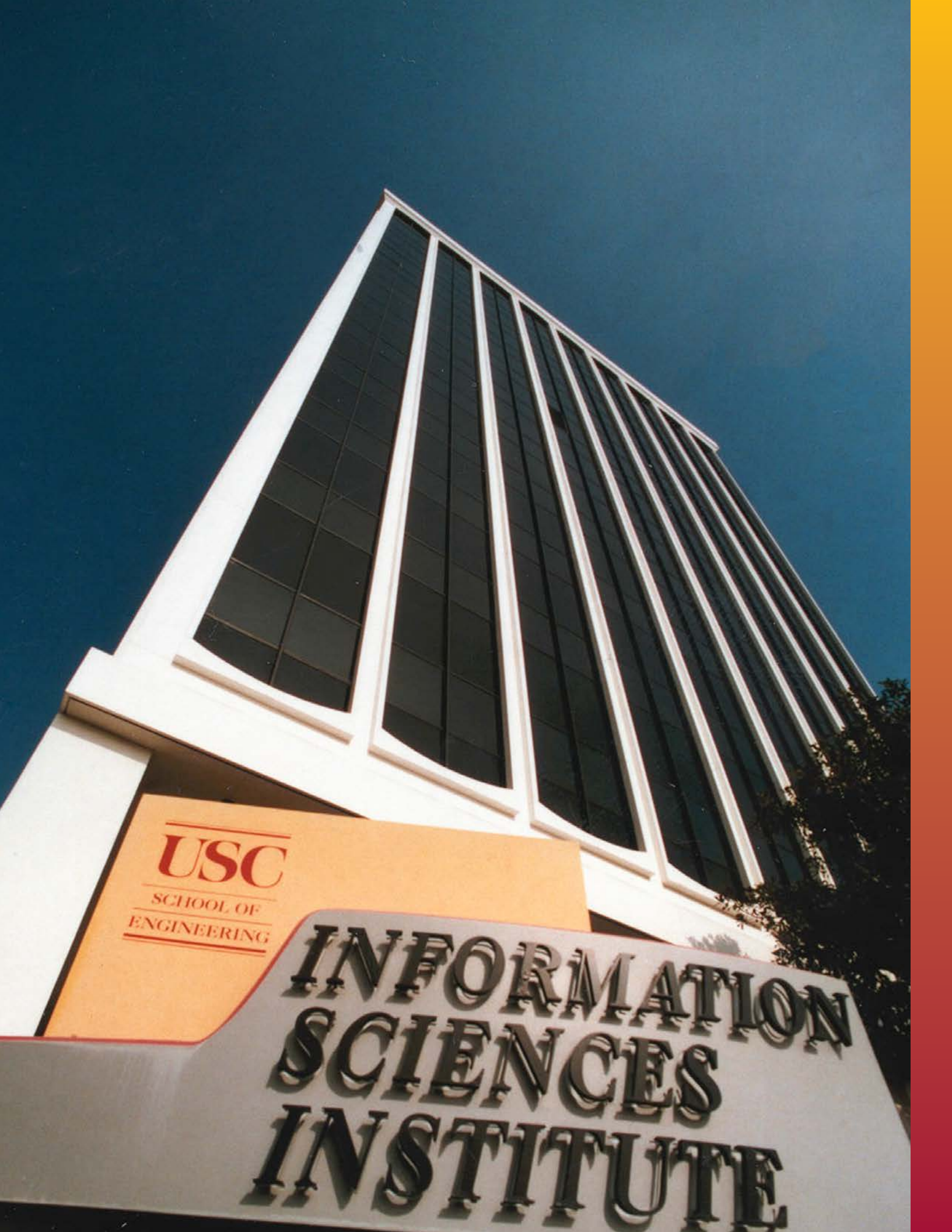
Danny's thick Israeli accent couldn't hide his quick mind, playful sense of humor, and his total command of English. His tales from Oceanview, Kansas, conveyed deep insights about protocol design in disarmingly simple language. And every once in a while, he'd challenge us with a devious math problem. With depth, personality, and leadership, he was one of the formative forces who created ISI.

RIP, Danny Cohen.

Footnote: For a broader view of Danny Cohen's work and life, see the New York Times obituary at <https://www.nytimes.com/2019/08/16/obituaries/danny-cohen-who-helped-set-the-stage-for-a-digital-era-dies-at-81.html>.



Danny Cohen at ISI's 40th anniversary celebration



USC
SCHOOL OF
ENGINEERING

**INFORMATION
SCIENCES
INSTITUTE**

Information Sciences Institute

4676 Admiralty Way #1001; Marina del Rey, CA 90292

310.822.1511 | isi.edu | media@isi.edu

USC Viterbi
School of Engineering