Information Sciences Institute

WORKING TOGETHER TO DELIVER THE FUTURE

2020 ANNUAL REPORT

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.

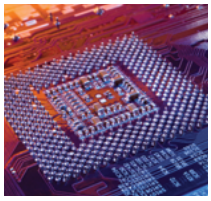
MASSACHUSETTS
890 Winter Street #115
Waltham, MA 02451
781.622.9790

VIRGINIA
3811 Fairfax Drive #200
Arlington, VA 22203
703.812.3700

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

Report Preparation Phyllis O'Neil (editor), Craig Knoblock, Yigal Arens | **Design** Russo Design, Los Angeles | **Printing** Typecraft






USC Viterbi School of Engineering

WELCOME TO ISI

I'm pleased to share our institute's accomplishments with you in our annual report.

We started the year with the institute-wide Research Day meeting last January. Research Day included everyone, from all three locations, and we came together to connect professionally and socially, to talk and listen, to visit, and to report on our research. At the time, we expected ISI's work in 2020 to continue and to progress as ever.

It was great to be together.

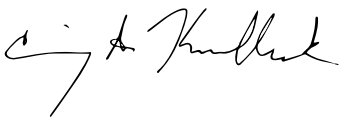
Now, in the ongoing Covid-19 pandemic, we continue to work together, in big teams and small. It's different. And it's difficult. But it's rewarding to see how ISI people collaborate and support one another. *We are still together.*

In the same way that ISI researchers in any ISI location can be part of any research group, ISI's collegiality spans home offices and Zoom meetings.

ISI leadership is working to keep our community connected. We started a "new hires buddy system" to link new researchers with an established ISI researcher, since there are no coffee conversations. In addition, senior researchers share their experience and expertise in an ongoing series of mentoring seminars. We've established a working group to focus on diversity, inclusion, and equity at ISI.

ISI's research this year includes new and ongoing work in many areas: AI modeling of the dynamics of wildfires, quantum information science, work in computer vision to predict congenital adrenal hyperplasia, and a natural language processing AI tool to detect bias in the news—these are just a few. Our research efforts in 2020 include building systems in quantum physics, application-specific intelligent integrated circuits, and a soft robotics project inspired by the anatomy of the starfish. Seven ISI researchers received faculty appointments. 19 ISI students completed their PhD degrees in 2020, and it's a pleasure to recognize these achievements.

I was honored to be appointed Keston Executive Director and Vice Dean of Engineering in June. My goals continue: to make ISI an outstanding place to work, to grow ISI's research funding, to support our researchers, faculty, and students, and to continue to pursue impactful research projects. As you look through this report, I hope you get a feel for the unique environment that is ISI.



Craig A. Knoblock
Keston Executive Director
USC Information Sciences Institute

ISI's main offices occupy nine floors of the Marina Towers South office building, due west of Los Angeles. The institute overlooks the largest small-craft harbor in the US, with some 5,000 boats.

About 325 researchers, staff, graduate research assistants, faculty, and students are based in the Marina del Rey location. Researchers in every division, plus the institute's administrative departments, are located here.

In addition to office and laboratory space—and views of the Hollywood Hills sign, the Santa Monica Mountains, and Catalina Island—the institute's 120,000 square feet of space include a lounge for collaboration or cappuccino, a yoga studio, and a machine shop. Lab space totals over 7,500 square feet.

Parks, hotels, restaurants, and shops are within walking distance from ISI, in addition to the marina. The Space Engineering Research Center and the USC-Lockheed Martin Quantum Computing Center are next door to the tower.

ISI is accessible by public transit and bike path and is a 10-minute ride from LAX.

Free shuttle service runs between ISI, the USC Institute for Creative Technologies, and USC's University Park campus.

ISI Arlington is located in Northern Virginia, near Washington, D.C. 29 researchers and staff members are based there, plus students and interns. Three Arlington researchers hold appointments as research professors at the Ming Hsieh Department of Electrical and Computing Engineering.

ISI's Secure and Robust Electronics Center and the Application Specific Intelligent Computing Lab are both located in the Arlington office.

Top areas of research carried out at ISI Arlington include:

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

In addition to collaborating with researchers in the other ISI locations, researchers at ISI Arlington collaborate with USC academic departments and centers, other universities, commercial companies, and defense contractors.

Student interns are recruited from countries around the world to work at ISI Arlington, plus local and regional universities such as George Washington University, University of Maryland, and Virginia Tech.

In February 2017, ISI opened a small office in Greater Boston with six new senior researchers and one visiting researcher. Now, ISI Boston research staff members—based in ISI Boston or reporting to ISI Boston—total 31.

Major research areas at ISI Boston are:

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

Sponsors of research 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.

ISI Boston researchers lead several projects with funded collaboration from universities, including MIT, Columbia, Carnegie Mellon, Notre Dame, UMass Amherst, and Virginia Tech, plus industry collaborators such as Northrop Grumman and Raytheon.

Within its 11,672 square feet of space, the Boston office also houses ISI's Laboratory for Quantum-Limited Information, which is dedicated to understanding and demonstrating the fundamental physical limits for extracting information from physical signals.

EXECUTIVE LEADERSHIP

Craig Knoblock, PhD
Michael Keston Executive Director

Stephen Crago, PhD
Associate Director

James Whalen, M Phil
Chief Financial Officer
Associate Director

Eileen Lu
Chief Information Officer
Director of Computing and Information Services

Yigal Arens, PhD
Senior Director for Administrative Affairs

John Damoulakis, PhD
Senior Director, Special Projects

John Wroclawski, MSEE
Senior Director, Strategic Initiatives

Elizabeth Boschee
Director, ISI Boston

Lifu Chang, PhD
Director, The MOSIS Service

Yolanda Gil, PhD
Director for Major Strategic AI and Data Science Initiatives

Derek Mikuriya, MBA
Director, Human Resources

RESEARCH LEADERSHIP

Terry Benzel, MA, MBA
Director, Networking and Cybersecurity Division

Stephen Crago, PhD
Director, Computational Systems and Technology Division

Carl Kesselman, PhD
Director, Informatics Systems Research Division

Pedro Szekely, PhD
Director, Artificial Intelligence Division

Yigal Arens, PhD
Director, Emerging Activities Group

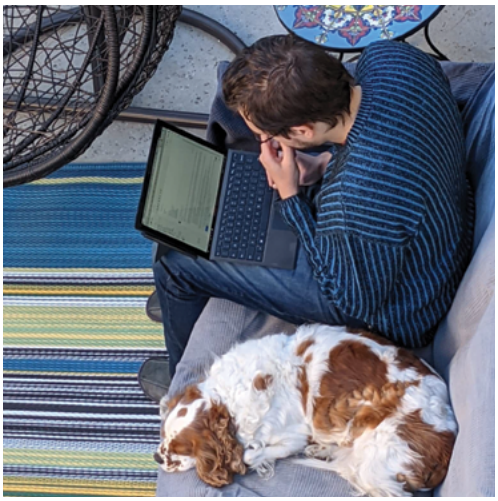
WORKING FROM HOME IN 2020

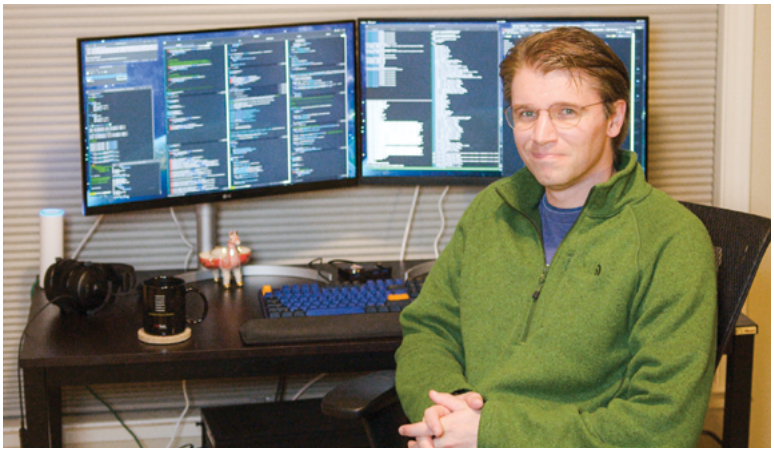
As they say, it's been a year . . . when the year started, ISI people expected to see one another in seminars, in the hallway, over coffee. Those interactions stopped abruptly in March. Almost overnight, ISI people made the switch to working from home and collaborating via Zoom meetings. New modes, new work habits, new challenges. The institute's work continued and advanced. Special thanks go to the staff members who continued to come into the office to keep operations going—in Facilities, CIS, and Administrative Services.


HOLLYWOOD


USCISI
USCISI
USCISI


Peace, Love & Cider





RESEARCH DAY 2020

On January 10, the first ISI Research Day celebrated ISI's cutting-edge research and innovation.

Scientists from all research groups at ISI, from all ISI locations, came together to present their research in interactive talks, poster sessions, and a panel discussion on future developments in research. Topics included quantum annealers, advanced electronics, spacecraft technology, artificial intelligence, and cybersecurity.

12 ISI staff members and researchers were recognized for their achievements by Director Craig Knoblock, who presented them with institute achievement awards and exploratory research grants.

With almost 300 researchers, staff, and faculty from across the country attending, ISI Research Day provided an invaluable opportunity for ISIers to meet and mingle with colleagues, in addition to learning about each other's research.

ISI administrative staff attended Research Day—meeting and mingling—and participated in a moderated workshop on improving institute business processes.

ISI Research Day 2020 was packed with great science and great people.

400+
ISI STAFF, FACULTY, AND STUDENTS

100
PHD STUDENTS

31
PROFESSORS

- Astronautical Engineering
- Civil and Environmental Engineering
- Communication
- Computer Science
- Electrical and Computer Engineering
- Industrial and Systems Engineering
- Physics and Astronomy
- Spatial Sciences Institute

ISI STAFF AND FACULTY

6 MBA degrees

47 master's degrees

63 bachelor's degrees

95 PhD degrees

RESEARCH EXPENDITURES

$85.7 million

NEW RESEARCH GRANTS IN 2020

63

QUANTUM COMPUTERS

1

POSTDOCS

8

SUMMER INTERN MENTORS

12

2020 SUMMER INTERNS

27

SEMINARS HOSTED

65

ISI LOCATIONS

- Marina del Rey, CA
- Arlington, VA
- Waltham, MA

TOTAL OFFICE SPACE

147,576 square feet

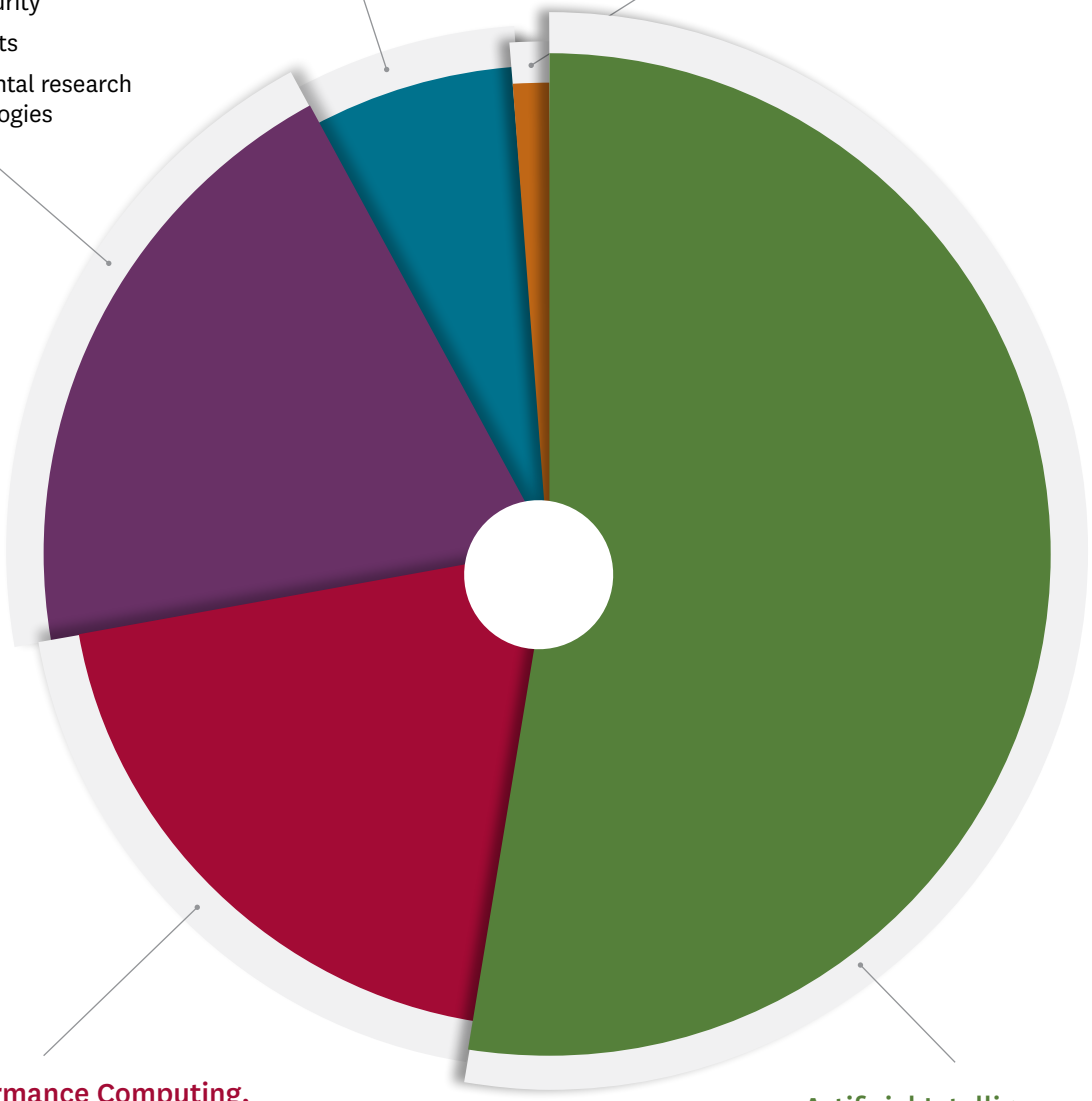
Informatics Systems Research

- Sociotechnical platforms for data-centric discovery
- Informatics cyberinfrastructure
- Collaboration infrastructure for biomedical research

Computer Networks and Cybersecurity

- Cyberphysical systems
- Internet measurement
- Cybersecurity
- Sensor nets
- Experimental research methodologies

Space Engineering Research



High Performance Computing, Quantum Computing, and Microelectronics

- Secure, robust, and trusted electronics
- Heterogeneous and cloud computing
- High performance computing architectures and software
- Quantum information science; quantum computing
- Science automation technologies
- Chip fabrication brokerage
- Multi-project wafers

Artificial Intelligence

- Social networks, analytics, and dynamics
- Bioinformatics
- Natural language processing
- Knowledge graphs
- Machine learning

PEDRO SZEKLEY | DIVISION DIRECTOR

The Artificial Intelligence Division is one of the world's largest AI research groups, with 28 researchers, three postdocs, 11 research programmers, 66 PhD students, 38 master's students, and five administrative assistants. Most researchers hold graduate degrees in computer science or related disciplines; 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 conducts fundamental and applied research in the following areas:

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; federated learning over data silos to learn joint models without sharing any subject data.

Natural language processing focusing on low resource machine translation, multilingual representation learning, transfer learning, dialogue, decision-making, question answering, summarization, ontologies, information retrieval, text decipherment.

Knowledge graphs using AI and machine learning techniques to construct and exploit large-scale knowledge bases and to induce taxonomies from data. Notable applications include probabilistic models for scientific reproducibility, incorporating extractions from scientific articles and scientific networks of citation and reference, and business knowledge graphs characterizing innovation and competition using web data and regulatory filings.

Scientific data analysis and discovery using interactive knowledge capture, intelligent user interfaces, semantic workflows, provenance, and collaboration; large-scale data integration and analysis of biomedical data, including sensor, environmental, neuroimaging, clinical, and genetic data.

Multi-modal understanding including image and video understanding for deepfake detection, visual misinformation identification and multimedia analysis, face recognition, biometric anti-spoofing, and robust AI; table understanding to automate exploitation of millions of tables on the web focusing on automatic layout detection, semantic modeling, table retrieval, table summarization, entity linking, and fact-checking.

Commonsense representation and reasoning using cognitively inspired computational paradigms for evaluating commonsense AI (including those based on large-scale language models) to create and solve new challenge tasks based on logical axioms and numeracy; human-centric dialog agents that maximize metrics of human utility alongside algorithmic utility in task-focused dialogs; game-theoretic simulators for poker, Monopoly, and wargames that enable refinement and evaluation of theories of novelty for general AI agents.

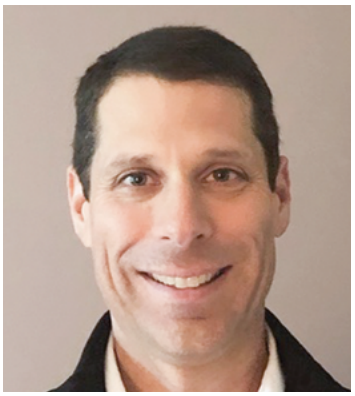
Computational social science with emphasis on structure detection and pattern matching in unusual complex systems with hidden information (e.g., human trafficking, dark money networks); large-scale, contextualized social media analysis (e.g., in the context of natural disasters) including analysis involving non-verbal tokens, such as emojis; computational social science methods for quantifying socio-demographically segmented impacts of COVID-19 on wellbeing, technological inequity, and vaccine hesitancy; applied AI in industrial applications, such as e-commerce.

AI fairness detecting and mitigating bias, robustness against adversarial attacks, identifying cultural values, polarization and misinformation, forecasting, and crowdsourcing. Notable case studies include a study of gender bias in 19th century English literature using natural language processing methods and a study of how state-of-the-art named entity recognition approaches systematically fail to identify female names.

The Artificial Intelligence Division was led by Aram Galystan from 2019 through 2020, before Pedro was appointed director.

COMPUTATIONAL SYSTEMS AND TECHNOLOGY DIVISION

STEPHEN CRAGO | DIVISION DIRECTOR



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

- **Heterogeneous cloud and embedded computing**
- **Heterogeneous integration of novel materials, devices, circuits, and architectures for advanced microelectronics, photonics, and magnetic integrated circuits; hardware architectures for artificial intelligence and sensors**
- **Numerical simulations of physical systems with high performance computing**
- **Quantum computing, communication, sensing theory, and hardware**
- **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 disciplines including electrical engineering, computer science, physics, and math.

Ongoing initiatives include theoretical adiabatic quantum computing through the **USC-Lockheed Martin Quantum Computing Center (QCC)** 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 for novel chips and field programmable gate arrays. Division researchers explore applications and algorithms for processing large-scale and real-time streaming data and solving optimization problems; 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 challenging 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 used by astronomers, physicists, and earthquake specialists.

In the **Laboratory for Quantum-Limited Information (QLIlab)** at ISI Boston, CS&T researchers are exploring how to build the most sensitive communications and sensing devices when a signal is imprinted on only a fraction of a photon—which is the fundamental quantum particle of light. One project is led by a USC undergrad examining the problem of laser range detection using a fully quantum theoretical framework, who discovered new quantum-inspired sensor architectures to improve the precision of range detection. In another project, researchers are developing hardware and software to detect the presence of a laser beam traversing the field of view of a camera, only sensing the very weak light that is scattered by atmospheric dust particulates.

At the **Application Specific Intelligent Computing Lab (ASIC)** researchers apply "alternate state variables" such as electrons, photons, phonons, and magnetic spins to deliver next-generation hardware fabrics using novel monolithic and heterogeneous device integration for artificial intelligence, smart sensors, smart and secure manufacturing, and quantum information. ASIC lab researchers and students bring hardware expertise in materials, devices, circuits design, and fabrication. Key projects include:

- Novel memory technologies that allow dynamically augmenting storage capacity for data-intensive applications during runtime
- In-pixel computing
- Volatile and non-volatile optical memories for wafer-scale computer chips
- Quantum interconnects (QuIC) and quantum chip interposers (QuIP) for quantum computing inter-chip and intra-chip applications
- Deep learning for throughput improvements for microelectronics imaging technology and hardware Trojan detection during manufacturing

CARL KESSELMAN | DIVISION 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. The Informatics Systems Research Division takes a holistic, systems-oriented approach, working in areas ranging from basic network service architectures, data management abstractions, computer security, and user-interface design, to domain-specific algorithms. The division specializes in highly collaborative, user-driven research in the context of 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, 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.

Researchers work closely with ISI's artificial intelligence, networking, and distributed systems experts, and 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.

The division participates in collaborative projects with faculty in Dornsife College, Keck School of Medicine, and the Herman Ostrow School of Dentistry. Current collaborations range from basic science to clinical use cases in molecular biology, basic neuroscience, neuroimaging, stem cell research, and craniofacial dysmorphia.

The division plays a central role in three international biomedical consortiums:

The GenitoUrinary Development Molecular Anatomy Project (GUDMAP) is a consortium of laboratories working to provide the scientific and medical community with tools to facilitate research on the genitourinary tract. GUDMAP is a public resource funded by the National Institutes of Health.

(Re)Building a Kidney is a consortium led by the National Institute of Diabetes and Digestive and Kidney Diseases to optimize approaches for the isolation, expansion, and differentiation of appropriate kidney cell types and the integration of these cells into complex structures that replicate human kidney function.

FaceBase is a collaborative project, supported by the National Institute of Dental and Craniofacial Research, that houses comprehensive data in support of advancing research into craniofacial development and malformation. FaceBase serves as a resource by curating large datasets from the craniofacial research community.

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, the Common Fund Data Environment.

TERRY BENZEL | DIVISION DIRECTOR

Networking is the backbone of the interconnected world, and *cybersecurity* is its guardian. The 53 researchers, faculty, PhD students, and student workers in the Networking and Cybersecurity Division focus on understanding the Internet, analyzing human behavior through social simulation, the theory and practice of distributed computing, and analyzing vulnerabilities and attacks using scientific modeling, experimentation, and evaluation. Areas of research include:

Network and security measurement, analysis, and defenses

Viewing the Internet as the world's largest sensor, division researchers study methods of observing and collecting network and network security data and behaviors to develop novel networking capabilities and defenses.

Modeling human behavior for cybersecurity and social simulation

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

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

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

Theory and practice of distributed computing

Today all computation and communication are fundamentally distributed, involving multiple participants and their interactions; Networking and Cybersecurity Division research addresses the challenges of enforcing safety, security, and robustness in these systems.

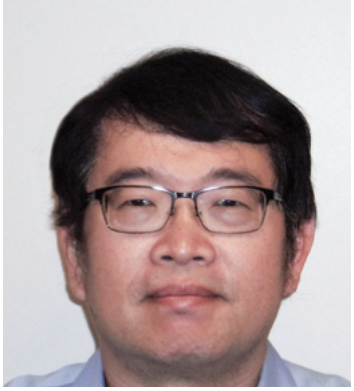
Social engineering attacks

Social engineering attacks such as phishing and impersonation continue to grow because organizations' weakest security links are the humans. Division researchers produce new methods for detecting and fingerprinting attack campaigns by leveraging metadata from communication channels and employing novel techniques to redirect attackers.

Binary program analysis and reverse engineering for vulnerability discovery

Division researchers perform reverse engineering using binary program analysis to search for vulnerabilities in software released without source code, and to assess the security of software products.

LIFU CHANG | DIRECTOR



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

Since its beginning in 1981, MOSIS ("Metal Oxide Semiconductor Implementation Service") has processed more than 60,000 IC designs, averaging five chips a week.

MOSIS enables IC designers to prototype innovative semiconductor designs within an effective cost structure, offering a range of processes: CMOS FinFET, FD-SOI, Bulk, SiGe, high-voltage BCD, and other specialty processes. MOSIS collaborates with four major foundries—TSMC, GlobalFoundries, Intel Custom Foundry, and Samsung Foundry—to offer a range of semiconductor processes from 12nm FinFET to 350nm.

MOSIS completed a significant modernization project this year:

- Moving all business and technical operations to Amazon Web Services (AWS) Gov Cloud resources
- Deploying a new secure website with enhanced functionality
- Configuring and deploying the Synopsys Lynx workflow engine to support design analysis across all foundries and technology nodes supported by MOSIS
- Launching a secure state-of-the-art Cloud Design Environment (CDE), which supports the stringent foundry IP security protections required to support advanced nodes
- Automating the design verification and fabrication sign-off process to provide a highly accurate and repeatable design analysis

These advances support access to advanced semiconductor technologies from TSMC and other advanced foundry technologies.

MOSIS supports the design enablement, design support, and fabrication sign-off for Intel Custom Foundry 22FFL process for digital, analog, and mixed-signal design flows utilizing Intel's Process Design Kit (PDK), IP, and EDA tech files. MOSIS collaborates with Intel to establish a customer design support model connecting support capabilities from MOSIS, EDA companies, and Intel Custom Foundry.

MOSIS is also planning to work with the PDK of the Samsung 28nm FD-SOI process.

MOSIS continues projects with TSMC and GlobalFoundries and is exploring new capabilities to add value to its offerings for Extreme UltraViolet (EUV) processes from TSMC.

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

- **MINSEC** (Microelectronics Needs and Innovation for National Security and Economic Competitiveness) developing a strategic plan for microelectronic innovation centers in the US
- **DoD/AFRL ATMI** (Access to Intel 22FFL Technologies through the MOSIS-Intel Alliance) facilitating and supporting silicon fabrication projects with the 22nm process of Intel Custom Foundry

The MOSIS Service performs research in silicon fabrication and Design for Manufacturability (DFM) related areas, producing both academic papers and patents. MOSIS collaborated with Xyalis on the paper *Reducing stress effects on multi-project-wafer reticles by optimizing metal densities and density gradients in an MPW placement flow*, which was accepted by SPIE Advanced Lithography, the premier conference on semiconductor manufacturing technologies. In addition to collaborating with other organizations in the semiconductor design and manufacturing community, MOSIS staff pursue independent research.

CENTER FOR VISION, IMAGE, SPEECH, AND TEXT ANALYTICS

WAEL ABDALMAGEED | CO-DIRECTOR
SCOTT MILLER | CO-DIRECTOR
SHRI NARAYANAN | CO-DIRECTOR
PREM NATARAJAN | FOUNDING DIRECTOR

In the Center for Vision, Image, Speech, and Text Analytics (VISTA), ISI researchers in various divisions work 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.

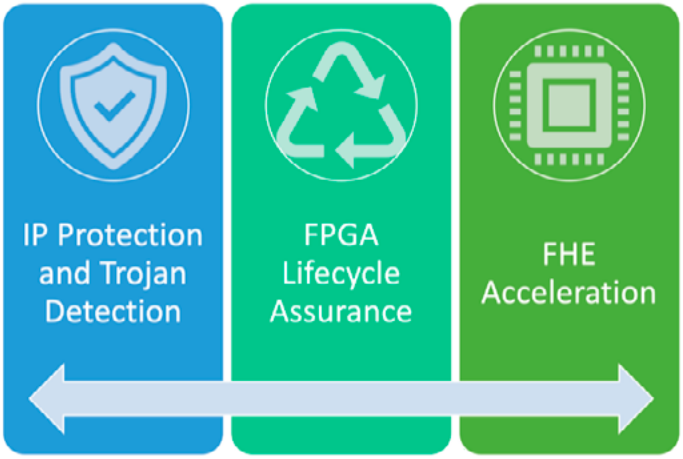
Applications of deep neural networks are spreading rapidly in health care, law enforcement, mobile phones, and self-driving vehicles. However, regardless of the sophistication and human-like accuracy of artificial intelligence models, these models are brittle and can easily be attacked or deceived into 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.

MATTHEW FRENCH | DIRECTOR

The past year has seen an enormous increase in the level of interest and need for hardware security. While there was already keen interest in protecting hardware intellectually property (IP) and re-investing in microelectronics research and startups to maintain US leadership, the COVID-19 pandemic stress-tested all our critical supply chains, including microelectronics, while also making us more reliant on them. Seemingly overnight, the workforce became distributed, education moved to online platforms, and our entertainment options were reduced to streaming video services. All of this made us dependent as never before on the underlying microelectronics that enable us to videoconference, share files, and stay connected during this time of isolation. Even ISI's Secure and Robust Electronics Center (SURE) hardware lab adapted by completing remote access and power management systems so that researchers could run experiments in the lab while largely working remotely. During this time, the SURE center persisted in making an impact in hardware IP protection, hardware Trojan detection, field programmable gate array (FPGA) assurance, and Fully Homomorphic Encryption (FHE) processing.

The hardware IP protection thrust is making significant ground and is featured later in this report in the "Quantifying Hardware Confidentiality" article. Hardware Trojans remain a concern and while much prior research has focused on detection during the development process, the SURE center is researching and developing detection techniques using in-line manufacturing metrology and deep algorithm-based detection for assurance during the manufacturing process.

FPGAs remain a strong focus as the SURE center is addressing many aspects of the FPGA design and lifecycle. This year SURE kicked off a new effort developing wrappers and bitstream linting tools that detect and prevent a user from unintentionally accessing the many undocumented states within an FPGA. The SURE team is also assisting in the development of an FPGA CAD tool that validates that the intended synthesis settings are applied. The group's work in FPGA supply chain verification also continues. The stuck-at fault detection toolset now supports all Xilinx Virtex devices (series 1 through 7), UltraScale, and UltraScale+ devices and is expanding to address timing degradation detection and Microsemi devices. At the final stage of the FPGA lifecycle, research and development on rapidly migrating FPGA designs from obsolete devices to modern devices has proven successful and is now addressing the challenges of porting previously encrypted designs.



The group is perhaps most excited by its new effort in developing a custom hardware accelerator for FHE. FHE algorithms enable end-to-end computation directly on encrypted data without the need for encryption and decryption. These algorithms come at a huge computational cost, traditionally 100,000x more than unencrypted computations. SURE researchers are bringing to bear the state of the art in hardware/software co-design, design space exploration, configurable Large Algorithmic Word Size (LAWS) multiplier units, and novel memory solutions, while teaming with external researchers at Duality Technologies in FHE algorithms, CMU in compilers, Two Six Labs in formal verification, and NYU in FHE computing architectures to realize a processor capable of bringing FHE computation to within 10x of unencrypted operations. This work has the capability to be truly transformative, with potential applications in fields such as communications, healthcare, banking, and big data.

DAVID BARNHART | DIRECTOR

The Space Engineering Research Center (SERC) is a longstanding 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 teaches space systems research and offers hands-on training to build, test, and fly actual spacecraft and satellites.

SERC hosts high school, community college, undergrad, and PhD students from all over the world. Six cadets from the French Air Force again worked and studied at SERC in 2020.

The following new projects, started in 2020, illustrate the breadth of research the center is undertaking in space science.

The STARFISH project is a unique soft crawling robotic device for spacecraft and satellite surfaces for use in anomaly detection. It uses technology from the previous year's OCTOPUS gripper for electro-adhesion, as well as the biologically inspired gait from . . . a starfish.

In conjunction with UC San Diego and UC Berkeley, SERC won a NASA research award to take the successful LEAPFROG lunar lander simulator and create a nation-wide competition using multiple LEAPFROG vehicles flying at radio-controlled flying aircraft fields across the US.

SERC was also selected to validate a new physics phenomenon that could revolutionize spacecraft propulsion using an internal laser pulsing in an enclosed cavity to generate thrust through a proposed concept in physics called "quantized inertia." USC was the single university selected to create a very high-fidelity, high-power laser lab and test this concept in vacuum at SERC, validating tests that only had been performed internationally.

2020 continued to be a growth year for SERC, despite the pandemic. Additional advanced research taking place at SERC includes:

- Starting the third satellite program in USC's history in partnership with Lockheed Martin (La Jument)
- Building a strand burner for solid propellant test grains to validate super-low temperature performance for rocket systems
- Enhancing an existing docking mechanism with new rendezvous sensors to create an all-in-one rendezvous and proximity operation (RPO) and docking mechanism that any object in space could use to "connect" to any other object
- Upgrading the USC ground station control center with command-and-control capability for planned USC satellites (including a 1.5U, a 3U, and possibly two 6U formation flight satellites) with modifications to the radio frequency (RF) front end and control systems

SERC continues to explore new concepts in space system disciplines, including 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.

USC-LOCKHEED MARTIN QUANTUM COMPUTING CENTER

DANIEL LIDAR | SCIENTIFIC AND TECHNICAL DIRECTOR
FEDERICO SPEDALIERI | OPERATIONAL DIRECTOR
ROBERT LUCAS | DIRECTOR OF RESEARCH PARTNERSHIPS

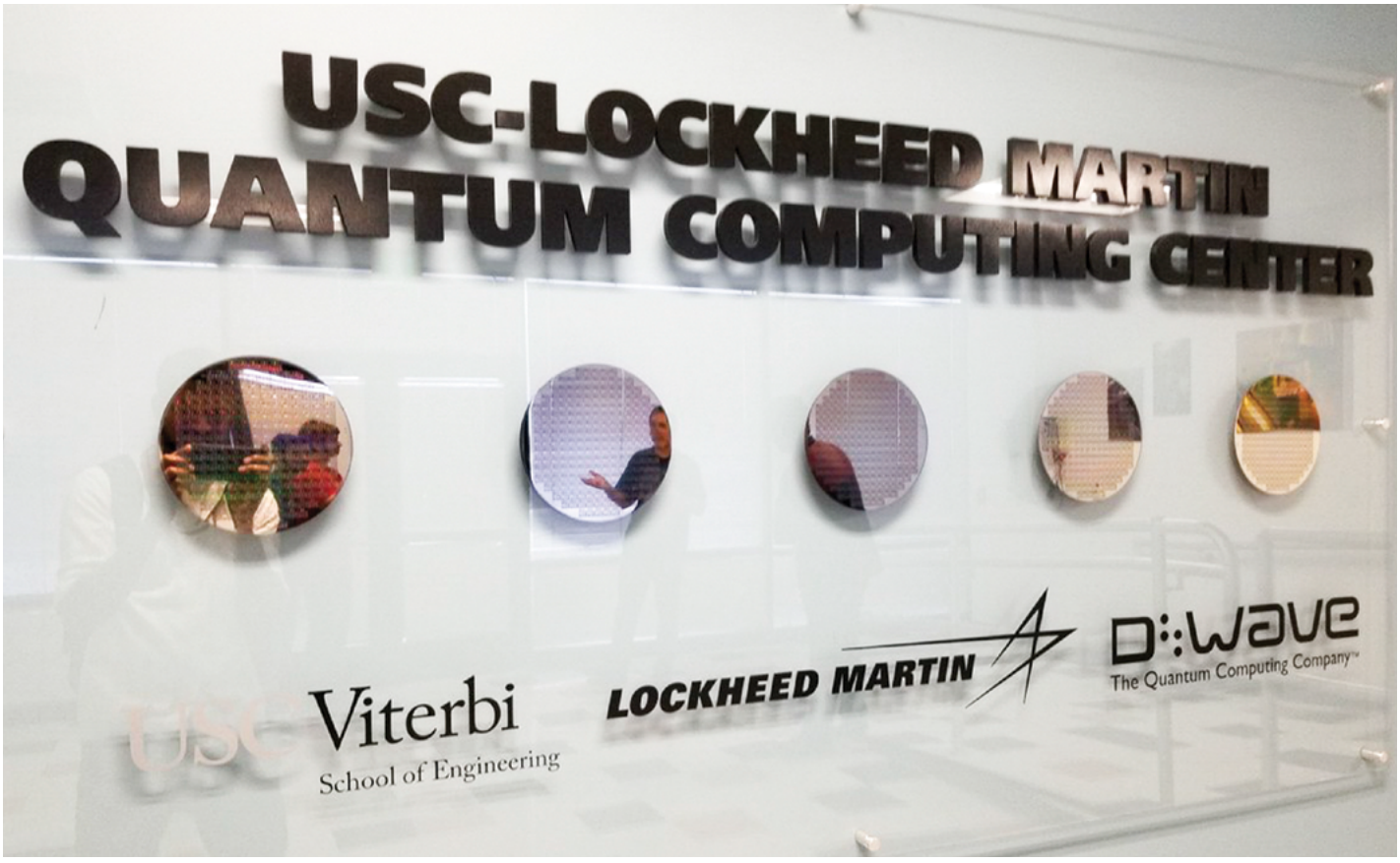
The USC-Lockheed Martin Quantum Computing Center (QCC) has housed a D-Wave quantum annealing system, manufactured by D-Wave Systems, Inc., since it was established in 2011. QCC was the first organization outside D-Wave to house and operate its own system; the center has conducted pioneering research on three different generations of these early noisy, intermediate-scale quantum (NISQ) devices.

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.

USC and ISI researchers, faculty, and students perform basic and applied research on NISQ computing devices and collaborate with researchers around the world.

Four USC doctoral students are currently engaged in research at the QCC and another received her PhD last year.

The QCC renewed its long-term relationship with D-Wave and Lockheed Martin in March 2020, leading to important upgrades at the facility. QCC will upgrade to D-Wave's forthcoming Advantage™ system with more than 5000 qubits. The upgrades will enable QCC to host a new Advantage generation of quantum annealers from D-Wave.



VITERBI DATA SCIENCE PROGRAM

Several years ago, ISI researchers, including ISI director Craig Knoblock, worked with the Computer Science Department leadership to create the Data Science Program in the USC Viterbi School of Engineering.

The Data Science Program now offers these degree programs:

- Master of Science in Applied Data Science
- Master of Science in Cyber Security Engineering

In addition, the program offers these joint degrees with other USC schools:

- Bachelor of Arts in Data Science
- Master of Science in Communication Data Science
- Master of Science in Environmental Data Science
- Master of Science in Healthcare Data Science
- Master of Science in Spatial Data Science
- Master of Science in Public Policy Data Science

LEADERSHIP

ISI's **Yolanda Gil** serves as Director of the Data Science Program. **Emilio Ferrara** is Associate Director of the Data Science Program, and **Cliff Neuman,** an ISI Research Affiliate, is Director of the Cyber Security Engineering Program.

ISI FACULTY

ISI researchers who work actively in data science research are central to the Viterbi Data Science Program, teaching along with USC faculty colleagues.

Andres Abeliuk, *Lecturer*

Jeremy Abramson, *Lecturer*

Yigal Arens, *Research Professor*

Keith Burghardt, *Lecturer*

Rafael Ferreira da Silva, *Research Assistant Professor*

Filip Ilevski, *Lecturer*

Carl Kesselman, *Professor*

Deborah Khider, *Lecturer*

Kristina Lerman, *Research Professor*

Jelena Mirkovic, *Research Associate Professor*

Fred Morstatter, *Research Assistant Professor*

Jay Pujara, *Research Assistant Professor*

Mohammad Rostami, *Research Assistant Professor*

Pedro Szekely, *Research Associate Professor*

Satish Kumar Thittamaranahalli, *Research Assistant Professor*

Each semester, the Center for Knowledge-Powered Interdisciplinary Data Science (CKIDS) and the Graduates Rising in Information and Data Science student association (GRIDS) organize the DataFest event, where 100+ data science students participate in practical data science projects proposed by ISI faculty and faculty members across USC schools.

NEW DIRECTORS AND RESEARCH LEADS

Yolanda Gil | *Director for Major Strategic AI and Data Science Initiatives*

In this role, Yolanda will lead ISI efforts to develop new large initiatives that cross disciplinary boundaries in both artificial intelligence and data science. Yolanda is also Director of the Data Science Program in the Viterbi School of Engineering, Research Professor of Computer Science and of Spatial Sciences, and Director of the Center for Knowledge-Powered Interdisciplinary Data Science at USC. She is Past President and Awards Chair of the American Association for the Advancement of Artificial Intelligence. Yolanda joined ISI in 1992 after earning her PhD in Artificial Intelligence at Carnegie Mellon University.

Derek Mikuriya | *Director of Human Resources*

Derek came to ISI from Thermo Fisher Scientific, where he was Vice President, Human Resources, for Global IT and R&D. In addition to managing HR for the Chief Information Officer and the Chief Scientific Officer, he led the talent workstream of the company's digital transformation initiative. As Director of Human Resources at ISI, he will implement HR strategies focusing on attracting talent and developing talent—plus diversity, equity, and inclusion. Derek also worked in HR at eBay, Bausch + Lomb, and GE Capital. He holds an MBA from Saint Mary's College of California and a BA in Communication Arts and Sciences from USC.

Pedro Szekely | *Division Director, Artificial Intelligence*

Pedro has investigated knowledge representation to support the creation of flexible and contextualized applications. In his early career, he investigated the representation of human machine interactions to support generation of user interfaces tailored to different modalities and screen sizes. Today, his focus is on machine commonsense representation and reasoning, table understanding to automate the creation of large knowledge graphs, techniques to understand knowledge graphs at multiple levels of abstraction, and tools to efficiently create and exploit knowledge graphs to robust AI applications. He has extensive experience in research, fund-raising, recruiting, and building real-world systems. Pedro is also a Research Associate Professor in USC's Computer Science department.

Genevieve Bartlett | *Research Lead, Senior Computer Scientist*

Genevieve was promoted to Research Lead and Senior Computer Scientist. She is in the Networking and Cybersecurity Division. Genevieve's research interests include networking, cybersecurity, and social engineering.

Mohamed Hussein | *Research Lead*

Mohamed was promoted to Research Lead. He is in ISI's Artificial Intelligence Division, a member of the Vision, Image, Speech and Text Analytics (VISTA) group, and an Associate Professor (on leave) at Alexandria University, Egypt. He earned his PhD degree in Computer Science from the University of Maryland at College Park in 2009. Before joining ISI in 2017, Mohamed was an assistant professor at Egypt-Japan University of Science and Technology (E-JUST) and Alexandria University. Mohamed has supervised and co-supervised PhD students at Alexandria University, E-JUST, and Virginia Tech. Mohamed's research interest is the intersection of machine learning and computer vision, most recently focusing on securing biometrics systems and machine learning models against attacks.

Joshua Monson | *Research Lead, Reconfigurable Computing Researcher*

Joshua, in the Computational Systems and Technology Division, was promoted to Research Lead and Reconfigurable Computing Researcher. His research focuses on the security, reliability, and productivity of field programmable gate arrays. He is currently involved in a cooperative effort to develop a prototype wireless radio that will provide efficient communications for autonomous robotic swarms. Josh joined ISI in 2016 after completing his PhD in Electrical and Computer Engineering at Brigham Young University.

NEW TEAM MEMBERS

In 2020, ISI welcomed 29 new researchers, computer scientists, programmers, and staff members.

EXECUTIVE OFFICE

Derek Mikuriya, *Director of Human Resources*

NETWORKING AND CYBERSECURITY DIVISION

Calvin Ardi, *Computer Scientist*

Luis Garcia, *Research Computer Scientist*

Brian Kocoloski, *Research Computer Scientist*

ARTIFICIAL INTELLIGENCE DIVISION

Maura Covaci, *Administrative Assistant*

Yash Doshi, *Programmer Analyst*

Muhao Chen, *Computer Scientist*

Xuezhe Ma, *Computer Scientist*

COMPUTATIONAL SYSTEMS AND TECHNOLOGY DIVISION

Yi-Hsiang Chen, *Postdoctoral Scholar*

Ciji Davis, *Project Specialist*

Akhilesh Jaiswal, *Research Assistant Professor of Electrical and Computer Engineering*

Amir Kalev, *Quantum Information Scientist*

Ramesh Kudalippalliyalil, *Postdoctoral Scholar*

Wendy Whitcup, *Project Administrator*

MOSIS

Jaime Bravo, *Semiconductor Process Engineer*

Glen Donelson, *Application Engineer*

COMPUTING AND INFORMATION SERVICES

Brian Miller, *Systems Administrator*

Mark Schreiman, *Technical Program Manager*

Jesus Navarro, *HPC Cluster/Linux System Engineer*

EMERGING ACTIVITIES

Joe Cecil, *Programmer Analyst*

Manuel Ciosici, *Postdoctoral Scholar*

Joseph Cummings, *Programmer Analyst*

Mitchell DeHaven, *Research Programmer*

Yash Kankanampati, *Research Programmer*

Dong Ho Lee, *Research Programmer*

Baha Mansurov, *Research Programmer*

Luke Miles, *Programmer Analyst*

Daniel Napierski, *Senior Computer Systems Engineer*

Eugene Park, *Programmer Analyst*

FACULTY APPOINTMENTS

Wael AbdAlmageed | *Research Associate Professor*

Wael AbdAlmageed is based in the Arlington office, where he works in ISI's Artificial Intelligence Division. On January 1, Wael received an appointment as Research Associate Professor in the Ming Hsieh Department of Electrical and Computer Engineering.

Emilio Ferrara, *Associate Professor*

Emilio Ferrara now holds a tenured appointment as Associate Professor in the Annenberg School of Communications, with a joint appointment in Computer Science. Emilio continues at ISI as a Research Team Leader. He also continues as Associate Director of the USC Data Science program

Jonathan Habif | *Research Assistant Professor*

Jonathan Habif is an experimental physicist based in the Boston office. Jonathan's research focuses on quantum limits for photon-starved, classical communication and imaging, quantum-secured optical communications in free-space and fiber, and integrated nano-photonics. In July, he received an appointment as Research Assistant Professor in the Ming Hsieh Department of Electrical and Computer Engineering.

Akhilesh Jaiswal | *Research Assistant Professor*

Akhilesh Jaiswal is based in ISI's Arlington office. He's part of the Application Specific Intelligent Computing lab (ASIC lab), where he conducts research in alternate state variables. Akhilesh's interests include hardware for AI, smart sensors, and hardware security. He received his faculty appointment as Research Assistant Professor in August in the Ming Hsieh Department of Electrical and Computer Engineering.

Kristina Lerman | *Research Professor*

Kristina Lerman joined ISI in 1998. She is in the Artificial Intelligence Division. Kristina was promoted to Research Professor in the Computer Science Department in July. Her work focuses on social networks, social computing, and artificial intelligence.

Michael Orosz | *Research Associate Professor*

Michael Orosz received a courtesy joint appointment in the USC Dornsife Spatial Sciences Institute. He is a Research Associate Professor of Civil and Environmental Engineering. Michael started out at ISI in 2001 as a Computer Scientist undertaking research in human computer interfaces in mission critical environments. His current research centers on applying data analytic approaches to open source intelligence challenges and approving systems engineering in large-scale DoD acquisition programs.

Satish Thittamaranahalli | *Research Assistant Professor*

Satish Thittamaranahalli already held an appointment as Research Assistant Professor in the Computer Science Department, and he now also holds an appointment in the Daniel J. Epstein Department of Industrial and Systems Engineering. Satish is interested in algorithmic techniques for solving combinatorial problems in artificial intelligence and computational physics.

PROMOTIONS

Congratulations to these ISI people who earned promotions in 2020!

EXECUTIVE OFFICE

Sharon Uyeda, *Human Resources Specialist*

NETWORKING AND CYBERSECURITY DIVISION

Genevieve Bartlett, *Research Lead, Senior Computer Scientist*

Ryan Goodfellow, *Senior Computer Scientist*

Alefiya Hussain, *Senior Computer Scientist*

Erik Kline, *Senior Computer Scientist*

ARTIFICIAL INTELLIGENCE DIVISION

Yolanda Gil, *Director for Major Strategic AI and Data Science Initiatives*

Mohamed Hussein, *Research Lead*

Deborah Khider, *Research Scientist*

Kristina Lerman, *Research Professor*

Pedro Szekely, *Division Director*

COMPUTATIONAL SYSTEMS AND TECHNOLOGY DIVISION

Joshua Monson, *Research Lead, Reconfigurable Computing Researcher*

Mohammad Nikravan, *Programmer Analyst IV*

Loic Pottier, *Computer Scientist*

Mats Rynge, *Senior Computer Scientist*

Karan Vahi, *Senior Computer Scientist*

INFORMATICS SYSTEMS AND RESEARCH DIVISION

Aref Shafaeibejestan, *Research Programmer*

ISI BOSTON

Elizabeth Boschee, *Senior Supervising Computer Scientist*

SPONSORED RESEARCH AWARDS

2020 was another strong year for ISI. Institute researchers won external funding for projects that range from measuring Internet traffic during COVID-19 quarantine periods, to automatic transcription and translation systems, to culturally divergent AI. Awards range from from tens of thousands to millions of dollars. This listing of new awards in 2020 illustrates the diversity of the institute's sponsors and collaborators.

AEROSPACE CORPORATION

Build, Test and Validation of an Electro-mechanical Docking Mechanism with Electronic Radar Sensors: CLING-ERS
David Barnhart, *Principal Investigator*

Build, Test, and Validation of a Soft Transformable Adaptable Robot for In Space Inspection
David Barnhart, *Principal Investigator*

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Detecting and Mitigating Adversarial Influence Operations in Networks
Kristina Lerman, *Principal Investigator*

AIR FORCE RESEARCH LABORATORY

Bitstream Assurance Checking Engine for Undocumented Functionality
Joshua Monson, *Principal Investigator*

CATIE High Risk High Payoff Advanced Propulsion Research in 2020 (CATIE2020)
David Barnhart, *Principal Investigator*
(via California State University Long Beach)

Information Extraction for New Emerging Noisy User-generated Micro-text
Elizabeth Boschee, *Principal Investigator*
(via InferLink Corporation)

Microelectronic Innovation for National Security and Economic Competitiveness (MINSEC) Technologies Development and Engagement
John Damoulakis, *Principal Investigator*
(via The Design Knowledge Company)

Trusted Silicon Stratus (TSS) Vulnerability Risk Assessment & Mitigation (V-RAM)
Ewa Deelman, *Principal Investigator*
(via NIMBIS Services, Incorporated)

ALFRED P. SLOAN FOUNDATION

Privacy Preserving Entity Resolution (Phase 2)
Pedro Szekely, *Principal Investigator*
(via Actuate Innovation)

AMAZON.COM SERVICES LLC

Amazon Alexa Prize
Jonathan May, *Principal Investigator*

ARMY RESEARCH OFFICE

Optimal Measurements for Reaching Quantum Fisher Information in Mixed Quantum Optical States
Jonathan Habif, *Principal Investigator*

DARPA

Reference Platform Preparation for Automatic Implementation of Secure Silicon
John Paul Walters, *Principal Investigator*
(via Boeing Company)

Culturally Divergent Artificial Intelligence
Fred Morstatter, *Principal Investigator*
(via Honeywell International, Inc.)

Direct and Coherent Off-axis Detection of Near Infrared Scattered Laser Light
Jonathan Habif, *Principal Investigator*

Exploiting Language Information for Situational Awareness
Jonathan May, *Principal Investigator*

High-assurance Advanced SW-based HW-independent Tagged Architecture Game-changer (HASHTAG)
Joshua Monson, *Principal Investigator*
(via Greensight)

Secure, Adaptive, roBust, Resilient, and Efficient Slices
Erik Kline, *Principal Investigator*

Scalable All-Optical Processing - Remotely
Jonathan Habif, *Principal Investigator*

Searchlight Testbed
Stephen Schwab, *Principal Investigator*
(via Sandia National Laboratories)

Secure Heterogeneous Learning Federation with Information-theoretic Guarantees
José Luis Ambite, *Principal Investigator*

BAE VENOM
James Blythe, *Principal Investigator*
(via BAE Systems)

VENICE: VErifyiNg Implicit Cultural modEls
Fred Morstatter, *Principal Investigator*
(via US Army)

Media Forensics Integrity Analytics - Scientific Integrity Project
Wael AbdAlmageed, *Principal Investigator*
(via Purdue University)

KGTK: Knowledge Graph Toolkit
Pedro Szekely, *Principal Investigator*
(via Air Force Research Laboratory)

LR2: Learning Robust Representations
Wael AbdAlmageed, *Principal Investigator*

SMELLCPS: Symbolic Math Expressions from Low-level Logic in Cyber-physical Systems
Luis Garcia, *Principal Investigator*

DEPARTMENT OF ENERGY

Replacing Aging Programmable Electronics Rapidly (REAPER) - Phase 2
Andrew Schmidt, *Principal Investigator*
(via Honeywell Federal Manufacturing & Technologies LLC)

Replacing Aging Programmable Electronics Rapidly (REAPER) - Phase 3
Andrew Schmidt, *Principal Investigator*
(via Honeywell Federal Manufacturing & Technologies LLC)

GOOGLE

Google Focused Research Award
Jay Pujara et al., *Principal Investigator*

Google Focused Research Award, Machine Perception Category
Shri Narayan, *Principal Investigator*

JP MORGAN

JP Morgan Artificial Intelligence Research Program
2020 JP Morgan Faculty Research Award
Jay Pujara, Craig Knoblock, *Principal Investigators*

MISSILE DEFENSE AGENCY

Extensible Virtual Processing Hardware Environment
John Paul Walters, *Principal Investigator*
(via Intelligent Automation, Inc.)

NASA

Deep Learning in Deep Space (DLDS): A HPSC-compatible Deep Neutral Network Coprocessor
John Paul Walters, *Principal Investigator*
(via Intelligent Automation, Inc.)

LEAPFROG: Next Generation Lunar Lander Prototype for National STEM Competition
David Barnhart, *Principal Investigator*
(via UC San Diego)

StereoBit: Advanced Onboard Science Data Processing to Enable Future Earth Science
Matthew French, *Principal Investigator*
(via Carr Astronautics)

Integrating TAT-C, STARS, and VCE for New Observing Strategies Mission Design
Matthew French, *Principal Investigator*
(via Stevens Institute of Technology)

NATIONAL SCIENCE FOUNDATION

Collaborative Research: A Big Data Approach to Three Fundamental Paleoclimate Questions
Deborah Khider, *Principal Investigator*

CC* Compute: A Customizable, Reproducible, and Secure Cloud Infrastructure as a Service for Scientific Research in Southern California
Carl Kesselman, *Principal Investigator*

CCRI: Planning: Collaborative Research: Infrastructure for Enabling Systematic Development and Research of Scientific Workflow Management Systems
Rafael Ferreira Da Silva, *Principal Investigator*

NSF Convergence Accelerator Track C: Chiral-Based Quantum Interconnect Technologies (CirQuITs)
Ajey Jacob, *Principal Investigator*
(via UCLA)

CNS Core: Small: Event Identification in Evaluation of Internet Outages
John Heidemann, *Principal Investigator*

FET: Small: Collaborative Research: Efficient and Robust Characterization of Quantum Systems
Amir Kalev, *Principal Investigator*

Collaborative Research: EAGER: Leveraging Advanced Cyberinfrastructure and Developing Organizational Resilience for NSF Large Facilities in the Pandemic Era
Ewa Deelman, *Principal Investigator*

Collaborative Research: EAGER: Advancing Reproducibility in Multi-Messenger Astronomy
Ewa Deelman, *Principal Investigator*

CCRI: ENS: Modernizing and Streamlining DeterLab Testbed Experimentation
Jelena Mirkovic, *Principal Investigator*

CC* Integration-Large: An "On-the-fly" Deeply Programmable End-to-end Network-centric Platform for Edge-to-Core Workflows
Ewa Deelman, *Principal Investigator*
(via University of Massachusetts)

EAGER: QSA: Solving Optimization Problems on NISQ Computers
Amir Kalev, *Principal Investigator*

Artificial Intelligence and Community Driven Wildland Fire via a WIFIRE Commons Infrastructure for Data and Model Sharing
Yolanda Gil, *Principal Investigator*

Partnership to Advance Throughput Computing (PATh)
Ewa Deelman, *Principal Investigator*
(via University of Wisconsin-Madison)

PReSto: A Paleoclimate Reconstruction Storehouse to Broaden Access and Accelerate Scientific Inference
Deborah Khider, *Principal Investigator*

RAPID-MINCEQ: Measuring the Internet During Novel Coronavirus to Evaluate Quarantine
John Heidemann, *Principal Investigator*

Collaborative Research: PPoSS: Planning: Performance Scalability, Trust, and Reproducibility: A Community Roadmap to Robust Science in High-throughput Applications
Ewa Deelman, *Principal Investigator*

RAPID: Supply Chain Portal to Serve Entrepreneurs Producing Critical Items in Response to COVID-19 (SupCovid)
Jay Pujara, *Principal Investigator*
(via University of Maryland)

CC*DNI DIBBS: Merging Science and Cyberinfrastructure Pathways: The Whole Tale
Carl Kesselman, *Principal Investigator*
(via University of Illinois)

NATIONAL INSTITUTES OF HEALTH

Limited Competition: Continuation of the Center for Genomic Studies on Mental Disorders (U24)
Yigal Arens, *Principal Investigator*
(via Rutgers University)

NOVARTIS PHARMACEUTICALS CORPORATION

Novartis-USC Research Collaboration: Developing FAIR Data Platform to Scaling Up Data FAIR-ification
José Luis Ambite, *Principal Investigator*

NAVAL SURFACE WARFARE CENTER

Rapid Assured Microelectronics Prototypes Using Advanced Commercial Capabilities
Devanand Krishna Shenoy, *Principal Investigator*
(via IBM Thomas J. Watson Research Center)

OFFICE OF NAVAL RESEARCH

FogSys: A System and Framework for Fault Tolerant Real-time Fog Computing
John Paul Walters, *Principal Investigator*

SONY

Sony Faculty Research Award
Xiang Ren, *Principal Investigator*

US ARMY

Learning Best Practices for Building Automatic Transcription and Translation Systems
Marjorie Freedman, *Principal Investigator*
(via Office of Naval Research)

REDSTAR - RE-engineering Deterlab for Scalability robusTness And Reliability
Jelena Mirkovic, *Principal Investigator*
(via Office of Naval Research)

US DEPARTMENT OF STATE

Backpack: Content Bucket Viewer and Distributor
Michael Orosz, *Principal Investigator*

US GOVERNMENT – OTHER

Cosmic Canary
Andrew Schmidt, *Principal Investigator*
(via Georgia Institute of Technology)

HONORS AND AWARDS

2020 PROFESSIONAL AND ACADEMIC RECOGNITION

Yigal Arens
Senior Member, American Association for Artificial Intelligence (AAAI)

Terry Benzel
Senior Member, Institute of Electrical and Electronics Engineers (IEEE)
Board of Governors, IEEE Computer Society

Muhao Chen, Junheng Hao, Chelsea Ju, Yizhou Sun, Carlo Zaniolo, Wei Wang
ACM BCB 2020 Best Student Paper Award, Association for Computing Machinery Bioinformatics, Computational Biology and Biomedical Informatics Special Interest Group (ACM SIGBio)

Daniel Garijo
Best Reviewer Award, Extended Semantic Web Conference (ESWC 2020)
Spotlight Paper Runner Up, International Semantic Web Conference

Yolanda Gil
Fellow, American Association for the Advancement of Science (AAAS)
Fellow, IEEE

John Heidemann and Yuri Pradkin
2020 Networking Systems Award, ACM Special Interest Group on Data Communication (SIGCOMM)

Mohamed Hussein
Senior Member, IEEE
Senior Member, ACM

Carl Kesselman
2020 Harry H. Goode Memorial Award, IEEE Computer Society

Craig Knoblock
Fellow, IEEE
Outstanding Engineering Merit Award, Orange County Engineering Council

Jonathan May
Advisor Support Award, Bloomberg

George Papadimitriou
Best Student Paper Award, Practice & Experience in Advanced Research Computing (PEARC20) Conference

Daniel Napierski
Best Demo Paper, 2020 Association for Computational Linguistics Conference

Shri Narayanan
exploreCSR Award, Google Research
University Professor, USC
Senior Research Award, USC Viterbi School of Engineering
Sustained Accomplishment Award, ACM International Conference on Multimodal Interaction (ICMI)

Yesenia Ornelas
Staff Early Career Award, USC Viterbi School of Engineering

Xiang Ren
Best Paper Award Runner Up, The Web Conference

Alexander Spangher
Bloomberg Data Science PhD Fellowship, Bloomberg LP

Peng Xie
DSN 2020 Test-of-Time Award, DSN-2020 IEEE/IFIP International Conference on Dependable Systems and Networks

2020 ISI ACHIEVEMENT AWARDS

These awards recognize contributions to ISI with significant impact on the reputation, research productivity, or administrative efficiency of the institute.

Matt Binkley for his passionate work on process improvement, information sharing, and creating a positive and efficient work environment at ISI

Ryan Goodfellow for his leadership in formulating the architectural vision, implementing and testing the software, and designing, standing up, and operating working testbeds

Rajiv Mayani for his outstanding work on the data analysis and management tools, and the complete redesign of the website for the NIMH Repository and Genomics Resource Center

2020 KESTON AND ISI EXPLORATORY RESEARCH AWARDS

ISI received a generous endowment gift in 2015 from Los Angeles entrepreneur, philanthropist, and engineer-at-heart Michael Keston, and his wife and philanthropic partner, Linda Keston. Under the terms of this endowment, a portion of the income generated is dedicated each year to sponsoring the Keston Exploratory Research Awards, intended to foster and support exploratory early-stage ISI research not yet funded by outside sponsors.

ISI has contributed additional funds to expand the program, creating the ISI Exploratory Research Awards. Exploratory Research Awards support significant, intellectually intriguing projects with the potential to produce a tangible, meaningful result, such as a tool or artifact, within about a year. Craig Knoblock presented these awards in January 2020 during ISI Research Day. The award winners and summaries of the ensuing research follow.

Wael AbdAlmageed | Keston Exploratory Research Award

Fighting Misinformation: An Internet System for Detecting Fake Face Video

Deepfakes are a new class of AI-generated videos showing a person saying or doing something they did not say or do. Deepfakes can have significant adverse impacts on national security, stock markets, criminal justice, and child safety. The research team developed a deep learning algorithm that produces compact representations (i.e., features) for real videos while producing dispersed representations for the fake videos. This approach provides robustness against new methods of generating deepfakes. During the 2020 elections, the team detected multiple deepfake videos out-of-the-box without retraining the deep neural networks. They also developed a web-based application that enables organizations and users to upload and examine suspect videos. The system has been transitioned to government and private sector partners.

Pedro Szekely | Keston Exploratory Research Award

Mining Cancer Registry Data to Improve Long Term Effects of Pediatric Cancer Treatment

Although the majority of children diagnosed with cancer in the US survive, they often live with adverse consequences from their treatment. In collaboration with researchers from Children's Hospital of Los Angeles and the USC Keck School of Medicine, this project developed data extraction techniques to identify and characterize patients at high risk of adverse effects—from the existing data resources of the Los Angeles County Cancer Surveillance Program, a database of over 40,000 records of treatment. Future work will integrate treatment data with effects data to derive correlations and produce clinically actionable findings.

Jonathan May | ISI Exploratory Research Award

Universal Translators for Asylum Seekers at the Border

This AI project is developing machine translation techniques for translating personal narratives, spoken in a variety of languages such as K'iche', Kanjobal, and languages from the Mixtec family into English. The goal is to assist asylum seekers at the US Southern border who do not speak Spanish or English. Because of the lack of translators for these languages, the asylum seekers cannot get crucial legal assistance to prepare for their interviews with immigration judges. The researchers collected a corpus of translated text, plus a corpus of translated text from trauma and displacement narratives, to build machine translation models using a transfer learning approach. Although collaboration with the School of Social Work to obtain asylum interview recordings was delayed by COVID restrictions, the work on expanding, cleaning, and analyzing the corpora is continuing.

Andrew Rittenbach, Priyatam Chiliki, and Dev Shenoy | ISI Exploratory Research Award

Automating Programmability of Hybrid Digital-analog Hardware for Stochastic Cell Simulation in Biological Systems

Biological system modeling enables biologists to quantitatively explain observations made in a laboratory to predict how biological processes respond. One of the "holy grails" of biological system modeling is a fully functional model of the entire human cell. Today, biological systems are modeled in software, which is not viable for large-scale gene-protein networks. In this project, the researchers investigated an alternative approach: a cytomorphic computing platform that consists of programmable hybrid analog/digital circuitry designed to model biochemical reactions. One challenge, however, is to configure the platform's analog circuit parameters to accurately model the reactions. The research team developed a reinforcement learning-based approach for automatic configuration of a Simulink-based model of a cytomorphic circuit without the need for expertise in analog circuit design. Results showed that the biochemical reaction simulated using the cytomorphic circuit-based simulation data closely matched the simulation data generated by software.

Joel Mathew and Ulf Hermjakob | ISI Exploratory Research Award

Advancing No-Resource Languages

This project aims to provide speakers of languages with few or no digital text resources with natural language processing tools for translation and original authoring, thereby also supporting literacy and language preservation. The researchers sentence-aligned New Testament Bible translations of 12 minority languages in Northern India to Hindi, built a tool for data cleaning and spell checking, and are working on automatically predicting translation target words with previously unseen morphological forms, employing both traditional natural language processing and neural network approaches.

2020 US ISSUED PATENTS

Young Cho et al.

Subcircuit Physical Level Power Monitoring Technology for Real-Time Hardware Systems and Simulators (patent number 10,830,800)

Ming Ting Wu, **Adam Cohen** et al.

Multi-Layer, Multi-Material Fabrication Methods for Producing Micro-Scale and Millimeter-Scale Devices with Enhanced Electrical and/or Mechanical Properties (patent number 10,641,792)

Ajey Jacob et al.

Transverse-electric (TE) *pass polarizer* (patent number 10,871,614)
Semiconductor detectors integrated with Bragg reflectors (patent number 10,818,807)
Tunable grating couplers (patent number 10,816,872)
Polarizers with confinement cladding (patent number 10,816,728)
Multimode waveguide bends with features to reduce bending loss (patent number 10,816,727)
Edge couplers for photonics applications (patent number 10,816,726)
Waveguide intersections incorporating a waveguide crossing (patent number 10,816,725)
Integrated graphene detectors with waveguides (patent number 10,804,416)
Heterogeneous directional couplers for photonics chips (patent number 10,795,083)
Bragg gratings with airgap cladding (patent number 10,795,082)
FinFET with multilayer fins for multi-value logic (MVL) applications (patent number 10,756,213)
Electro-optic modulators with stacked layers (patent number 10,747,030)
Silicon nitride grating couplers (patent number 10,746,925)
Stacked waveguide arrangements providing field confinement (patent number 10,746,921)
Grating couplers with cladding layer(s) (patent number 10,746,907)
Resistive nonvolatile memory structure employing a statistical sensing scheme and method (patent number 10,726,896)
Waveguide crossings with a non-contacting arrangement (patent number 10,718,903)

Ajey Jacob et al. (continued)

Multiple-layer arrangements including one or more dielectric layers over a waveguide (patent number 10,698,159)
Three dimensional optical interconnects (patent number 10,690,845)
Vertical nanowires formed on upper fin surface (patent number 10,685,847)
Electro-optic modulators with layered arrangements (patent number 10,684,530)
Neuromorphic memory devices (patent number 10,672,465)
Composite waveguiding structures including semiconductor fins (patent number 10,670,804)
Resistive nonvolatile memory cells with shared access transistors (patent number 10,665,281)
Electro-optic modulators with stacked metal, dielectric, and active layers (patent number 10,649,245)
Back-end-of-line blocking structures arranged over a waveguide core (patent number 10,649,140)
Polarizers and polarization splitters phase-matched with a back-end-of-line layer (patent number 10,641,956)
Multiple-layer arrangements using tunable materials to provide switchable optical components (patent number 10,585,245)
Grating couplers with multiple configurations (patent number 10,585,219)
Slot assisted grating based transverse magnetic (TM) transmission mode pass polarizer (patent number 10,557,989)

Akhilesh Jaiswal et al.

Backup and/or restore of a memory circuit (patent number 10,854,291)
Multi-bit dot product engine (patent number 10,825,510)
Memory device having in-situ in-memory stateful vector logic operation (patent number 10,802,827)
Read and logic operation methods for voltage-divider bit-cell memory devices (patent number 10,783,957)
MRAM read and write methods using an incubation delay interval (patent number 10,593,397)

Michael J. Giering, **Vivek Menon** et al.

Deep convolutional neural networks for crack detection from image data (patent number 10,860,879)

PHD GRADUATES

ISI is proud to congratulate these terrific USC graduates for completing their PhD degrees in 2020.

David Barnhart, Advisor

William Bezouska
Dissertation: *Cooperative Localization of a Compact Spacecraft Group Using Computer Vision*
Current position: Senior Researcher, Aerospace Corporation

Stephen Crago, Advisor

Geoffrey Tran
Dissertation: *Data and Computation Redundancy in Stream Processing Applications for Improved Fault Resiliency and Real-Time Performance*
Current position: Software Engineer, Google

Emilio Ferrara, Advisor

Palash Goyal
Dissertation: *Graph Embedding Algorithms for Attributed and Temporal Networks*
Current position: Senior Research Scientist & Project Lead in Core Machine Learning, Samsung Research America

Adam Badawy
Dissertation: *Analysis and Prediction of Malicious Users on Online Social Networks*
Current position: Principal Data Scientist, Capital One

Ashok Deb
Dissertation: *Modeling Information Operations and Diffusion on Social Media Networks*
Current position: Lieutenant Colonel, US Army

John Heidemann, Advisor

Lan Wei
Dissertation: *Anycast Stability, Security and Latency in The Domain Name System (DNS) and Content Deliver Networks (CDNs)*
Current position: Member of Technical Staff, Roblox

Calvin Ardi
Dissertation: *Improving Network Security through Collaborative Sharing*
Current position: Computer Scientist, USC Information Sciences Institute

Hang Guo
Dissertation: *Detecting and Characterizing Network Devices Using Signatures of Traffic About End-Points*
Current position: Member of Technical Staff, Microsoft

Shri Narayanan, Advisor

Brandon Booth
Dissertation: *Improving Modeling of Human Experience and Behavior: Methodologies for Enhancing the Quality of Human-produced Data and Annotations of Subjective Constructs*
Current position: Postdoc Faculty Researcher, Colorado University, Boulder

Prashanth Gurunath Shivakumar
Dissertation: *Speech Recognition Error Modeling for Robust Speech Processing and Natural Language Understanding Applications*
Current position: Applied Scientist, Amazon

Haoqi Li
Dissertation: *Behavior Understanding from Speech under Constrained Conditions – Exploring Sparse Networks, Transfer and Unsupervised Learning*
Current position: Applied Scientist, Amazon Web Services

Yongwan Lim
Dissertation: *Improved Real-Time MRI of Speech Production*
Current position: Staff Research Scientist, University of Southern California

Shri Narayanan, Advisor (continued)

Pavlos Papadopoulos
Dissertation: *Noise Aware Methods for Robust Speech Processing Applications*
Current position: Applied Scientist, Amazon

Manojkumar Prabakaran
Dissertation: *Context-aware Models for Understanding and Supporting Spoken Interactions with Children*
Current position: Applied Scientist, Amazon

Tanner Sorensen
Dissertation: *Dynamics of Speech Tasks and Articulator Synergies*
Current position: Research Scientist, Cobalt Speech & Language

Shao-yen Tseng
Dissertation: *Human Behavior Understanding from Language through Unsupervised Modeling*
Current position: Deep Learning Data Scientist, Intel Labs

Prem Natarajan, Advisor
Wael AbdAlmageed, Co-advisor

Ayush Jaiswal
Dissertation: *Invariant Representation Learning for Robust and Fair Predictions*
Current position: Applied Scientist, Amazon

Cliff Neuman, Advisor

Abdulla Alwabel
Dissertation: *Enabling Symbolic Execution String Comparison During Code-analysis of Malicious Binaries*
Current position: Software Engineer, Google

Pedro Szekely, Advisor

Majid Ghasemi Goll
Dissertation: *Learning Distributed Representations of Cells in Tables*
Current position: Software Engineer, Google

POSTDOCTORAL SCHOLARS AND VISITORS

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

POSTDOCTORAL SCHOLARS

Emilio Ferrara, Advisor

Goran Muric | *Artificial intelligence and social networks*

Itay Hen, Advisor

Emre Akaturk | *Quantum many-body simulations*

Yi-Hsiang Chen | *Quantum computing and algorithms*

Tom Halverson | *Quantum many-body simulations*

Arjun Mani | *Quantum algorithms*

Ajey Jacob, Advisor

Ramesh Kudalippalliyalil | *Photonics-based quantum chip photonics interconnects (QuIC) and quantum chip photonics interposers (QuIP) for chip-to-chip and on-chip quantum information processing*

Fred Morstatter, Advisor

Daniel Benjamin | *Decision science*

Ralph Weischedel, Advisor

Manuel Ciosici | *Natural language processing, using deep neural nets and pre-trained language models to learn both common sense knowledge and typical event sequences from the web*

VISITING SCHOLARS

Craig Knoblock, Host

Marianna Bucchi (Politecnico of Milan) | *Exploiting the semantic web for the automatic extraction of Los Angeles city data*

Pedro Szekely, Host

Jingwei Cheng (Northeastern University, China) | *Knowledge graphs, natural language processing, semantic web*

Daniel Schwabe (Pontificia Universidade Católica, Brazil) | *Semantic web, knowledge graphs*

Fu Zhang (Northeastern University, China) | *Knowledge graphs*

VISITING PHD STUDENT

Mohamed Hussein, Host

Ehsan Kazemi (University of Central Florida) | *Adversarial machine learning*

PHD STUDENTS

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

Wael AbdAlmageed, Advisor

Jiazhi Li | *Debiasing artificial intelligence algorithms*

Ranga Sai Shreyas Manchikanti | *Multimodal machine common sense*

Hanchen Xie | *Self- and un-supervised learning*

Jiageng Zhu | *Learning robust representations*

José Luis Ambite, Advisor

Ruhollah Shemirani | *Scaling genetic analysis, identity-by-descent (IBD) mapping*

Dimitrios Stripelis | *Federated learning*

David Barnhart, Advisor

Rahul Rughani | *Genetic algorithms to create non-conjunctive trajectories for multiple spacecraft or satellites (swarms) relative to a single area or another object*

Stephen Crago, Advisor

Yang Hu | *Fault tolerant real-time edge computing*

Haonan Wang | *Fault tolerant real-time edge computing*

Ewa Deelman, Advisor

Tu Mai Anh Do | *Data analytics, high-performance computing, distributed systems*

Patrycja Krawczuk | *Anomaly detection, machine learning workflows*

George Papadimitriou | *Dynamic resource provisioning and workflow management for scientific applications*

Emilio Ferrara, Advisor

Alexander Bisberg | *Machine learning*

Emily Chen | *Machine learning*

Ashok Deb | *Machine learning*

Julie Jiang | *Machine learning*

Hsien-Te Kao | *Machine learning*

Akira Matsui | *Machine learning*

Shen Yan | *Machine learning*

Yilei Zeng | *Machine learning*

Rafael Ferreira Da Silva, Advisor

Taina Coleman | *Workflow graph structures and benchmarking*

Aram Galstyan, Advisor

Sami Abu-El-Haija | *Machine learning, graph analysis*

Shushan Arakelyan | *Machine learning*

Robert Brekelmans | *Machine learning, information theory*

Sarik Ghazarian Ghalemaleki | *Machine learning, natural language processing*

Rujun Han | *Machine learning, information extraction*

Hrayr Harutyunyan | *Machine learning, information theory*

Neal Lawton | *Machine learning*

Myrl Marmarelis | *Machine learning, time series*

Ninareh Mehrabi | *Machine learning, fair AI*

Mehrnoosh Sadat Mirtaheri Feijani | *Machine learning, knowledge graphs*

Kyle Reing | *Machine learning, information theory*

Serban Stan | *Machine learning*

John Heidemann, Advisor

Guillermo Baltra Elorriaga | *Internet outage detection with active methods*

Asma Enayet | *Internet outage detection with passive data*

Basileal Yoseph Imana | *Privacy and fairness*

Abdul Qadeer | *Big data processing for Internet data*

ASM Rizvi | *Anycast, DNS, and DDoS*

Itay Hen, Advisor

Yi Hsiang Chen | *Quantum computing, quantum algorithms*

Lalit Gupta | *Computational physics, quantum Monte Carlo simulations*

Craig Knoblock, Advisor

Minh Tran Xuan Pham | *Data cleaning for building knowledge graphs*

Basel Shbita | *Knowledge graph construction using auxiliary supervision*

Binh Vu | *Learning semantic models of data sources for building knowledge graphs*

Kristina Lerman, Advisor

Nazanin Alipourfard | *Biases in social networks*

Nathan Bartley | *Algorithmic bias*

Fiona Guo | *Causal modeling of human performance*

Yuzi He | *AI fairness*

Zihao He | *Modeling complex data*

Ashwin Shreyas Mohan Rao | *Social media analysis*

Negar Mokhberian | *Semantic text mining*

Nazgol Tavabi | *Semantic text mining*

Xuezhe Ma, Advisor

Jiao Sun | *Controlled language generation, natural language processing fairness*

Jonathan May, Advisor

Nada Aldarrab | *Natural language processing, decipherment*

Hyundong Cho | *Natural language processing, dialogue*

Katy Felkner | *Natural language processing, machine translation*

Mozhdeh Gheini | *Natural language processing, machine translation, transfer learning*

Thamme Gowda | *Natural language processing, machine translation*

Meryem M'Hamdi | *Natural language processing, multilingual representations, information extraction*

Alexander Spangher | *Natural language processing, computational journalism (Emilio Ferrara, co-advisor)*

Xusen Yin | *Natural language processing, reinforcement learning, text game playing*

Xiyang Zhang | *Natural language processing, event sequence extraction*

Jelena Mirkovic, Advisor

Sima Arasteh | *Binary analysis*

Sivaramakrishnan Satyamangalam Ramanathan | *Network security, network management*

Rajat Tandon | *Network security, denial of service*

Nicolaas Weideman | *Binary analysis*

Wei-Cheng Wu | *Binary analysis*

Fred Morstatter, Advisor

Yuzhong Huang | *Graph embedding and hybrid forecasting*

Shri Narayanan, Advisor

Victor Ardulov | *Human-centered AI, behavioral machine intelligence*

Sabyasachee Baruah | *Natural language processing*

Zhuohao Chen | *Behavioral machine intelligence*

Ming-Chang Chiu | *Robust audio processing*

Tiantian Feng | *Human-centered sensing and modeling*

Nikolaos Flemotomos | *Speech, natural language processing, behavioral machine intelligence*

Victor Martinez Palacios | *Natural language processing, media machine intelligence*

Karel Bogomir Mundnich Batic | *Human-centered sensing and modeling*

Amrutha Nadarajan | *Robust audio processing*

Pavlos Papadopoulos | *Robust audio processing*

Raghuveer Peri | *Robust audio processing*

Karan Singla | *Natural language processing, behavioral machine intelligence*

Prem Natarajan, Advisor

Jiaxin Cheng | *Computer vision*

I-Hung Hsu | *Information extraction (Nanyun Peng, co-advisor)*

Soumyaroop Nandi | *Computer vision*

Ekraam Sabir | *Detection of image manipulation in scientific documents*

Emily Sheng | *Fairness in natural language processing (Nanyun Peng, co-advisor)*

Jay Pujara, Advisor

Pegah Jandaghimeibodi | *Human-centric dialog, quantitative verification of text*

Lee Kezar | *Psycholinguistic models, machine common sense, translating American Sign Language*

Kexuan Sun | *Table understanding, knowledge graphs, explainable AI*

Avijit Thawani | *Natural language processing, representation learning, commonsense reasoning*

Pei Zhou | *Commonsense reasoning, human-centered natural language processing (Xiang Ren, co-advisor)*

Srivatsan Ravi, Advisor

Weizhao Jin | *Multi-party computation for edge compute systems*

Xiang Ren, Advisor

Yuchen Lin | *Commonsense reasoning, neural-symbolic reasoning*

Jun Yan | *Question answering, neural-symbolic reasoning, information extraction*

Qinyuan Ye | *Question answering, learning from human explanations*

Pedro Szekely, Advisor

Majid Ghasemi Gol | *Machine learning, table understanding*

Ehsan Qasemi | *Commonsense reasoning*

Avijit Thawani | *Commonsense reasoning*

Peifeng Wang | *Commonsense reasoning*

Satish Thittamaranahalli Ka, Advisor

Ang Li | *Computational physics and data science*

Greg Ver Steeg, Advisor

Umang Gupta | *Machine learning*

Hannes Leipold | *Quantum computing*

2020 GRADUATE STUDENT SYMPOSIUM

The institute's 14th annual Graduate Student Symposium took place—remotely this year—on July 24. The symposium is an opportunity for students who are working, visiting, or volunteering at ISI to present their new or recently published research as short papers or one-page posters.

The symposium was opened by Craig Knoblock, ISI's director. Dr Ying Zhang, Software Engineering Manager and Research Scientist at Facebook, then spoke on "Top-Down Network Management at Facebook." Dr. Zhang is an expert on network management, and her talk discussed how to plan, evaluate, and implement route changes at Facebook scale.

The symposium program featured nine paper presentations in the areas of natural language processing, quantum computing, and cybersecurity. In addition, students gave lightning talks in two poster sessions, with 14 posters from graduate students, plus nine posters submitted by undergraduate researchers from ISI's Research Experiences for Undergraduates site.

The symposium audience voted for the best paper, the best graduate student poster, and the best REU student poster.

BEST PAPER

Zoe Gonzalez Izquierdo, ISI
Ruilin Zhou, Northwestern University
Klas Markström, Umeå University
Itay Hen, ISI

Discriminating Non-Isomorphic Graphs with an Experimental Quantum Annealer

BEST POSTER

Avijit Thawani, ISI
Jay Pujara, ISI
Pedro Szkeley, ISI

BERT Takes a Math Class

BEST REU POSTER

Marlowe McCraney, Birmingham-Southern College
Yuzi He, ISI
Kristina Lerman, ISI

Educational Disparities in Attainability

2020 RESEARCH EXPERIENCES FOR UNDERGRADUATES

ISI sponsored eight undergrads—from colleges and universities across the country—to spend eight weeks working with ISI research faculty and their teams on problems related to improving human communication in a connected world. The NSF Research Experiences for Undergraduates (REU) program offers an intellectually and socially stimulating research experience to undergrads, while providing training in research and working in research teams, information about research careers, and guidance on writing for publication. Each student concentrates on a specific research project, and he or she works closely with ISI research faculty and other ISI researchers. The 2020 cohort included four male and four female students.

ISI was selected to host an REU site during the summers of 2018 – 2020 with Jelena Mirkovic, an ISI cybersecurity researcher and project leader, as the site supervisor.

Although 2020 was the third year of ISI's REU site, it was the first year that the undergrads worked with ISI researchers remotely instead of coming to ISI. While each student worked with their own mentor, they attended twice-weekly seminars. The seminars explored a variety of research areas and tools. In addition to the seminars, students took part in two online social events organized by the site administrators.

At the end of the program, the students presented their research as poster sessions at two virtual events: an REU symposium attended by undergrads and their mentors from six other REU sites, and the annual ISI Graduate Student Symposium, where they presented to ISI researchers and students.

Thanks for having me this summer, I really enjoyed my time and I hope that I will have the opportunity to work at ISI again in the future!

I had an excellent experience both times I've been an REU mentor. The students have been awesome, and I've been able to learn from them while they were doing their internship.

Andres Abeliuk, Mentor
Artificial Intelligence Division
Antonio Camara, Columbia University
Research topic: *Identifying and Quantifying Polarization in Categorized Corpora*

Jeremy Abramson, Mentor
Networking and Cybersecurity Division
An Nguyen, Harvey Mudd College
Research topic: *Record Linkage for Learning Introductory Data Science Using American Football Data*

Emilio Ferrara, Mentor
Artificial Intelligence Division
Shawn Huesman, Northern Kentucky University
Research topic: *Detecting Similar Image Groups on Instagram During #BlackLivesMatter Protests*

Rafael Ferreira da Silva, Mentor
Computational Systems and Technology Division
Tabitha See Ya Lee, Vanderbilt University
Research topic: *Building Simulation-Driven Pedagogic Activities for Cyberinfrastructure Learning*

Daniel Garijo, Mentor
Artificial Intelligence Division
Aidan Kelley, Washington University in St. Louis
Research topic: *SOSEN: A Framework for Generating Scientific Software Knowledge Graphs*

Kristina Lerman, Mentor
Artificial Intelligence Division
Marlowe McCraney, Birmingham-Southern College
Research topic: *Educational Disparities in Attainability*

Jelena Mirkovic, Mentor
Networking and Cybersecurity Division
Carson Rupp, Purdue University
Research topic: *Identifying Kid-Friendly YouTube Content*

Fred Morstatter, Mentor
Artificial Intelligence Division
Emilee Daniel, Washington College
Research topic: *California County Twitter Practices and Voter Turnout*

2020 SUMMER INTERNSHIP PROGRAM

Every summer, ISI welcomes PhD students, master's degree students, and undergraduates as paid interns to work with senior research leaders and their teams. ISI researchers at all three locations mentor summer interns. Research areas for summer interns typically include:

- **Natural language processing**
- **Knowledge technologies**
- **Machine learning and data science**
- **Computing systems and technologies**
- **Internet and networked systems**
- **Cybersecurity research and experimentation**

Instead of coming to an ISI location to work, the 27 interns worked remotely with their ISI mentors and research teams. They participated in seminars, reading groups, and social activities. Although most summer interns work on PhD degrees, this year, the ISI interns included nine undergrads.

In 2020, 12 senior ISI researchers served as mentors to interns:

Wael AbdAlmageed
Artificial Intelligence Division

Ehsan Kazemi, University of Central Florida
Research area: *Robustifying deep learning against adversarial attacks*

David Barnhart
Emerging Activities Group

Christine Zhou, Brown University
Research Area: *Deposition of silver on complex shapes*

Emilio Ferrara
Artificial Intelligence Division

Aparna Ananthasubramaniam, University of Michigan
Research area: *Machine learning*

Danaja Maldeniya, University of Michigan
Research area: *Machine learning*

Prerna Jueja, Virginia Tech University
Research area: *Machine learning*

Sophie Wang, Northeastern University
Research area: *Machine learning*

Xinyi Zhou, Syracuse University
Research area: *Machine learning*

Marjorie Freedman
ISI Boston

Jacob Litchfeld, University of Louisville
Research area: *Natural language processing*

Justin Martin, University of Louisville
Research area: *Natural language processing*

Sarah Payne, University of Pennsylvania
Research area: *Natural language processing*

Ashwin Kumar, University of Louisville
Research area: *Natural language processing* (spring 2020)

Jonathan Habif
Computational Systems and Technology Division

Marina Furbush, *Wellesley College*
Research area: Materials emissivity modulation measurements for applications in steganography

Zi Chua, Wellesley College
Research area: *Quantum-limited estimation of coherence under thermal noise in photon-starved states*

Travis Haroldsen
Computational Systems and Technology Division

Monica Pottipati, UC Davis
Research area: *VLSI and Hardware Design*

Yuanzhe Jin, Northwestern University
Research area: *Embedded systems*

Jonathan May
Artificial Intelligence Division

Weiqiu You, University of Pennsylvania
Research area: *Natural language processing*

Naitian Zhou, University of Michigan
Research area: *Natural language processing*

Omar Shaikh, Georgia Tech
Research area: *Natural language processing*

Ugur Yavuz, Dartmouth College
Research area: *Natural language processing*

Vivek Menon
Computational Systems and Technology Division

Sanket Shukla, George Mason University
Research area: *Machine learning*

Shervin Roshanisefat, George Mason University
Research area: *Logic locking*

Fred Morstatter
Artificial Intelligence Division

Parisa Kaghazgaran, Texas A&M
Research area: *Mining cultural divergence in source code*

Andrew Schmidt
Computational Systems and Technology Division

Abhijit Dhavlle, George Mason University
Research area: *Electrical engineering*

Mohamed Hassan, Virginia Tech
Research area: *Electrical engineering*

Richard Becker, University of Arkansas
Research area: *Electrical engineering*

Michael Spector
Emerging Activities Group

Tanner Willis, University of Kentucky
Research area: *Non-named entity coreference*

John Paul Walters
Computational Systems and Technology Division

Sudhanshu Agarwal, Georgia Tech
Research area: *Heterogeneous fog and edge computing*

During the summer and fall of 2020, high school student Kelsie Lam interned under the guidance of Patrycja Krawczuk in the Computational Systems and Technology Division.

RISING STAR MS INTERNS

In a new program, ISI offers Rising Star Internships to students applying to the Viterbi School of Engineering Master of Science program in Computer Science, Data Science, Electrical and Computer Engineering, and similar disciplines.

The interns work on ISI teams on ongoing research projects—and work directly with individual mentors. The internships last throughout the course of the master's degree.

- **Research software engineer** track – engaging with emerging software technologies, developing advanced software engineering skills
- **Research systems/hardware engineer** track – contributing to the development of a novel hardware system or prototype, electronic and electronic fabrication technologies, developing engineering skills
- **Research data scientist** track – applying data science theory and tools to significant real-world problems, contributing to the development of new data analytics tools, approaches, and software
- **Researcher development** track – for students interested in a career in research, developing and publishing peer-reviewed research papers in highly-ranked conferences and technical journals

ISI's first Rising Star interns began fall semester 2020, working remotely with research teams and their individual mentors.

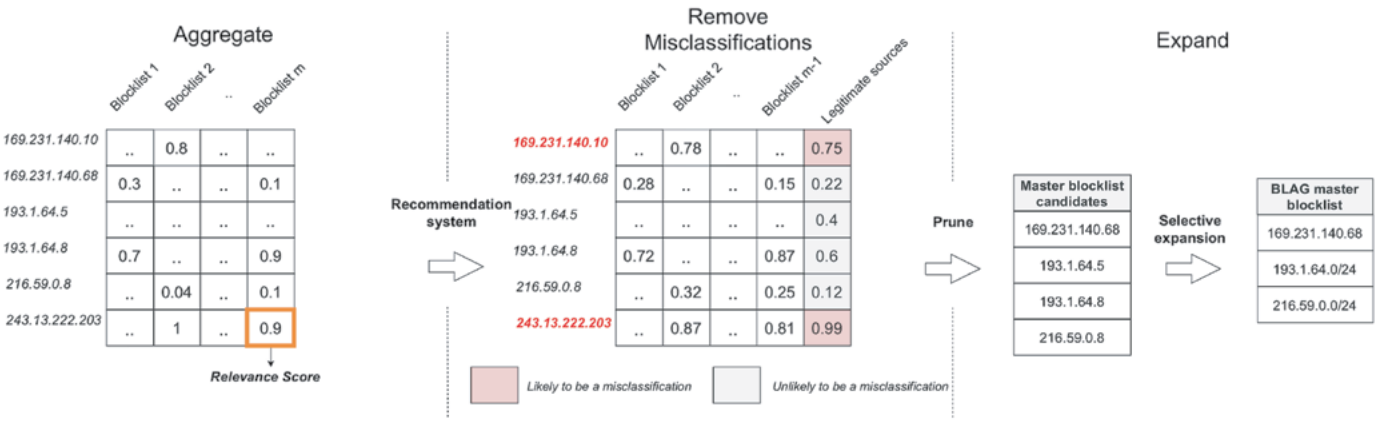
Wes Hardaker, in the Networking and Cybersecurity Division, is advising **Erin Szeto**. Erin is pursuing the research data scientist track and participates in the Global Analysis of Weak Signals for Enterprise Event Detection (GAWSEED) project.

Filip Ilievski, in the Artificial Intelligence Division, is advising **Nicholas Klein**. Nic is interested in AI research, and he too is pursuing the research data scientist track. Together they are working on a project about knowledge graph profiling.

Blocking Past Online Offenders Accurately and Without Bias

Many companies use blocklists—public or commercial feeds of IP addresses of known offenders—to scrutinize or even drop traffic from the blocklisted sources. Researchers in the Networking and Cybersecurity Division address two drawbacks related to blocklists: inaccuracy and bias.

Existing blocklists may have limited coverage of attacks and offenders. Such inaccuracy leads to missing offenders or dropping legitimate network traffic. The team of researchers developed the Blocklist Aggregator system (**https://steel.isi.edu/Projects/BLAG/**) to address this. The Blocklist Aggregator system aggregates and expands hundreds of blocklists to improve their coverage and deploys a recommendation system to customize the final blocklist to the customer network and improve its accuracy, as shown below.



This work improves existing blocklisting approaches by increasing recall (offender identification) from 0.1 - 18.4% to 6.4 - 69.7% while maintaining high specificity (legitimate source identification) of 95 - 99.5%. The system also improves detection delay, discovering offenders 8.8 - 13.4 days sooner than competing approaches.

Blocklists can also exhibit bias due to their use of IP addresses as identifiers. Legitimate users might share an IP address with an offender, for example, because their shared ISP uses network address translation or dynamic addressing. The legitimate users have a higher risk of being blocked at sites that employ blocklists, with no recourse. The Networking and Cybersecurity Division researchers quantified the impact of this phenomenon, called *shared blocking*. They showed that users from smaller countries with fewer Internet resources have a higher chance of suffering from shared blocking. The team published a list of shared IP prefixes that should be treated carefully by blocklists and other mechanisms to reduce the negative impact on legitimate users.

Contact: Jelena Mirkovic, sunshine@isi.edu

Enhancing National Cyberinfrastructure

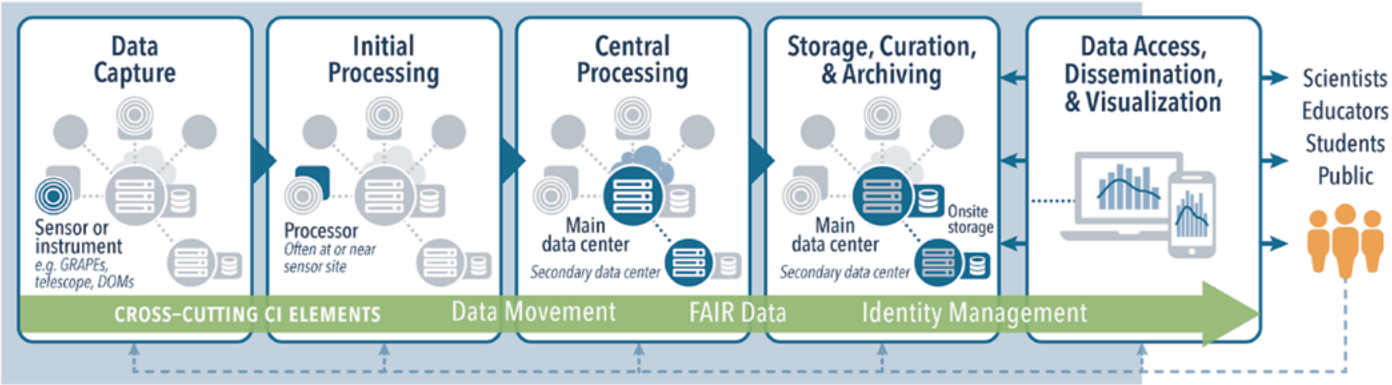
The NSF's major facilities are large, multi-user research facilities that operate and manage sophisticated and diverse research instruments, including large telescopes, interferometers, and distributed sensor arrays. They serve a variety of scientific disciplines—from astronomy and physics to geology and biology.

The Cyberinfrastructure Center of Excellence (CI CoE) Pilot project (**https://cicoe-pilot.org**) supports cyberinfrastructure practitioners at NSF major facilities to accelerate the data lifecycle and ensure the integrity and effectiveness of the cyberinfrastructure that research and discovery rely on.

The project developed a data lifecycle model that identifies the specific stages of the data flow within major facilities, from scientific data collection to dissemination to scientists' use of the data. The team conducted a number of activities, including one-on-one multi-month engagements, interviews with selected facilities, and community meetings. It helped establish joint work between the National Ecological Observatory Network (NEON) and the National Center for Atmospheric Research (NCAR) to develop a strategy and a project proposal for cloud-based data processing using NCAR models and NEON data.

At the same time, the project engaged with two other facilities, the Seismological Facility for the Advancement of GEoscience (SAGE) and the Geodetic Facility for the Advancement of Geoscience (GAGE), to help them plan for a common cloud platform capable of hosting, processing, and disseminating their collective data. More recently, the project began exploring the value of cyberinfrastructure in the broader context of how people and the facilities perceive and respond to major disruptive events, such as a global pandemic. (**https://ci4resilience.org**).

Contact: Ewa Deelman, deelman@isi.edu



CI CoE Pilot's data lifecycle model within NSF major facilities

Integrating Cyberinfrastructure Literacy into University Curricula

Teaching parallel and distributed computing topics in a hands-on manner is challenging, especially at the introductory undergraduate level, because appropriate compute platforms are generally not available. Even when all students have access to a platform, not all relevant learning objectives can be achieved. In particular, the platforms do not represent emerging and future cyberinfrastructure, specifically, highly distributed, heterogeneous platforms with large numbers of high-end compute nodes.

An ISI research team developed a set of pedagogic modules (**https://eduwrench.org**) that can be integrated piecemeal into university courses to address these challenges. They focus on quantitative reasoning and performance. The modules include simulation-driven activities so that students can experience relevant application and platform scenarios hands-on. The modules do not require any programming. They feature simulation-based pedagogic activities (supported by simulators developed using the WRENCH simulation framework, **https://wrench-project.org**) through which students learn by experimenting with various application and platform scenarios in simulation. The simulations provide metrics and visualizations of simulated application executions, so that students can empirically verify their answers. Students can also use simulations to explore complex design spaces to learn independently and with instructor-provided scaffolding. Several modules include a "capstone" activity, case studies in which students apply what they learned to solve problems inspired by real-world scenarios. This work is a collaboration with the University of Hawaii at Manoa.

Contact: Rafael Ferreira da Silva, rafsilva@isi.edu

A distributed execution pedagogic module for running scientific workflow applications

Exacerbating Bias in Algorithms through Fairness Attacks

Algorithmic fairness is attracting significant attention in the scientific and popular press. AI researchers are working to design algorithms that are robust to biases in the data. In recent years, researchers—and the public—have recognized that algorithms have the potential to propagate human error. This is largely because the datasets used to build the algorithms contain human biases. The effects can be tragic. For instance, an AI designed to predict recidivism was found to overpredict that Black offenders would re-commit their crimes. In another example, a study of AI systems that attempt to identify names in text was found to systematically ignore female names.

It is of utmost importance for researchers to identify and address these issues in algorithms.

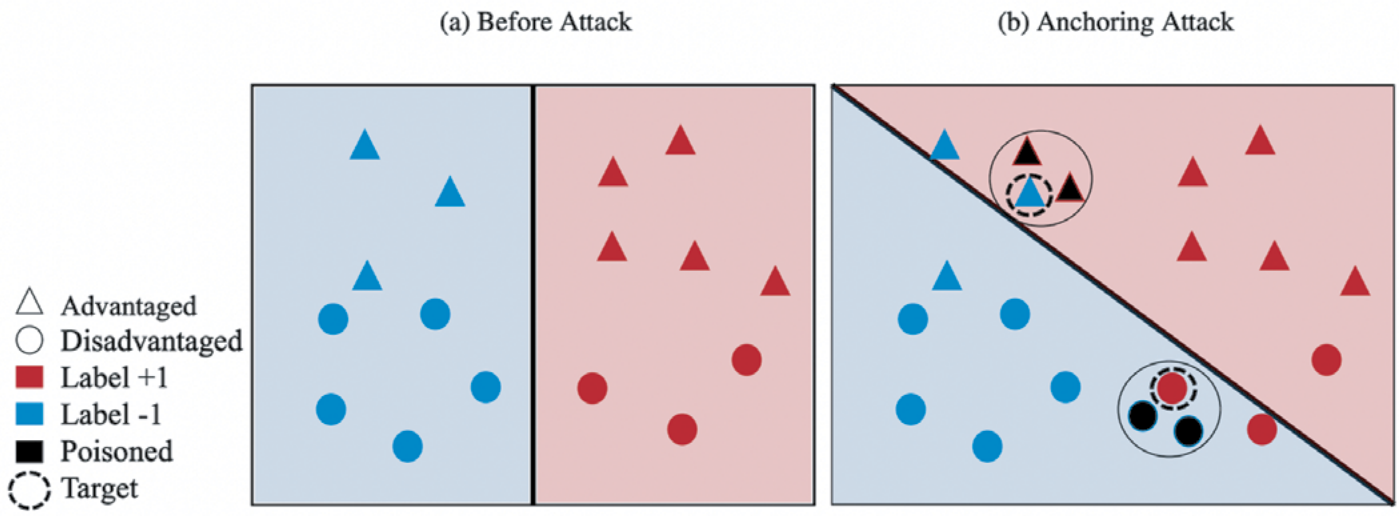
Recent research at ISI identified the potential for malign actors to poison AI systems. The research identified two families of attacks that can drive an AI system to make unfair, biased decisions:

> In the anchoring attack, the researchers skew the decision boundary by placing poisoned points near selected target points to bias the outcome. The adversary places points strategically near these points to move the decision boundary and cause unfair predictions, as demonstrated in the figure.
>
> In the influence attack, the adversary aims to maximize the covariance between the sensitive attributes and the decision outcome in order to decrease the fairness of the model.

Extensive experimentation showed that an adversary can generate poisoned data that affects the fairness of the algorithm. Interestingly, these experiments showed that the adversary can do so with little impact on the model's predictive accuracy. This magnifies the concern, because it demonstrates that these attacks could easily go unnoticed unless audited explicitly for bias. The research leads to a way for researchers to combat such attacks, either through more robust algorithms or through novel auditing approaches that can identify such poisoned data.

The figure depicts an anchoring attack. The left represents the data before the attack, and the right represents the anchoring attack in which poisoned points are positioned close to the selected points to skew the data and bias the algorithm.

Contact: Ninareh Mehrabi, ninarehm@isi.edu



Collecting Malware Samples from Honeypot Logs to Measure Industry Lag

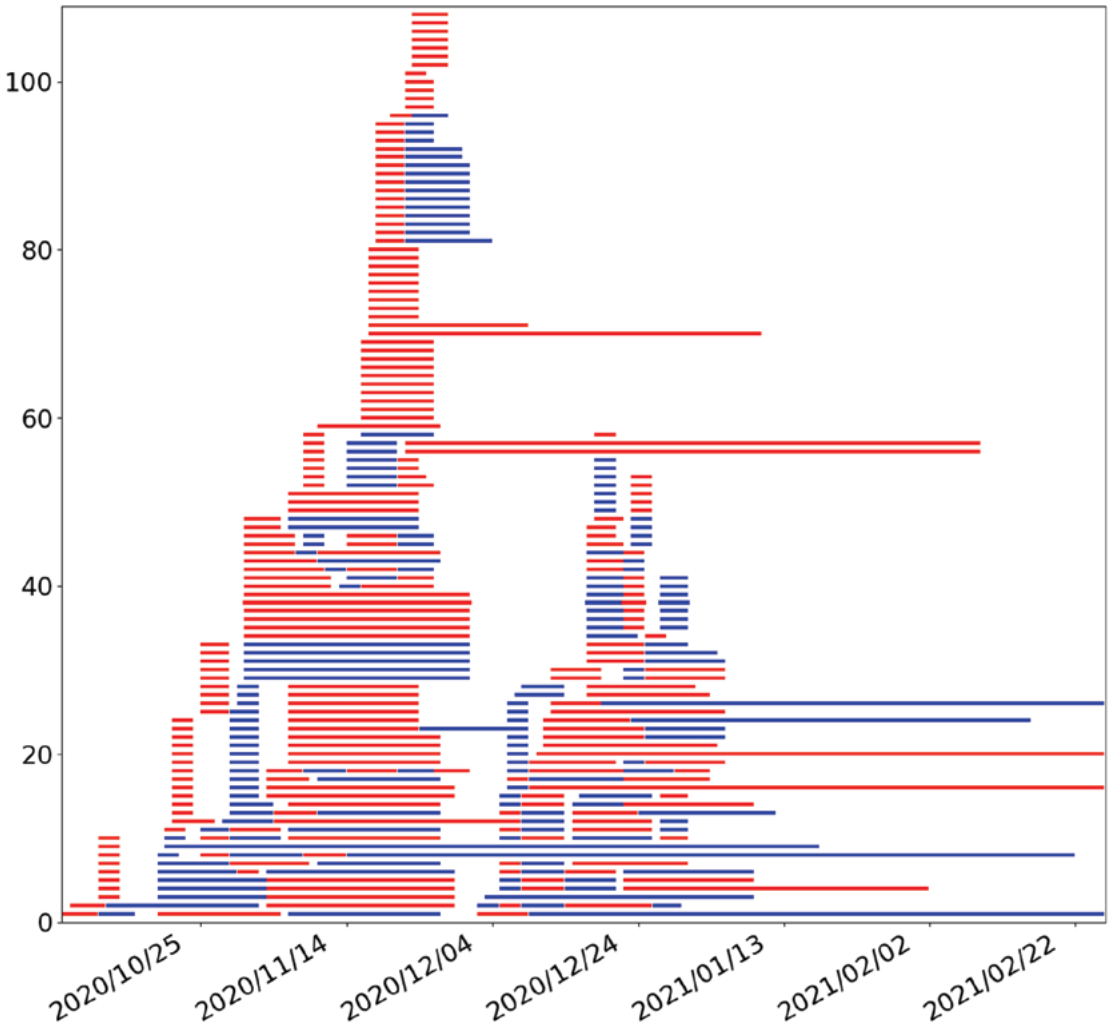
Delay in identifying malware leads to late discovery of intrusion; this provides attackers more time to infect and further penetrate their target systems and networks.

To estimate the lag time in the security industry, a team of Networking and Cybersecurity Division researchers gathered malware samples from web URLs extracted from SSH honeypot network logs from October 1 to December 31, 2020. (Internet honeypots are systems designed with intentional security vulnerabilities used to lure hackers in and log their behaviors for study.)

The researchers compared their findings against the leading VirusTotal malware catalog. Of the 3,785 unique files they discovered, 1,722 had never appeared in the VirusTotal malware catalog. 368 malware files were discovered first by the research team and only later identified in the industry database.

To visualize the time lag until the security industry identified these 368 malware files, the researchers plotted each one. In the graph below, each bar is one of the previously unidentified malware files. The leftmost point of each bar represents the date when the ISI researchers found the malware; the rightmost point represents the date when the malware file was identified in the VirusTotal malware catalog.

Contact: Wes Hardaker, hardaker@isi.edu

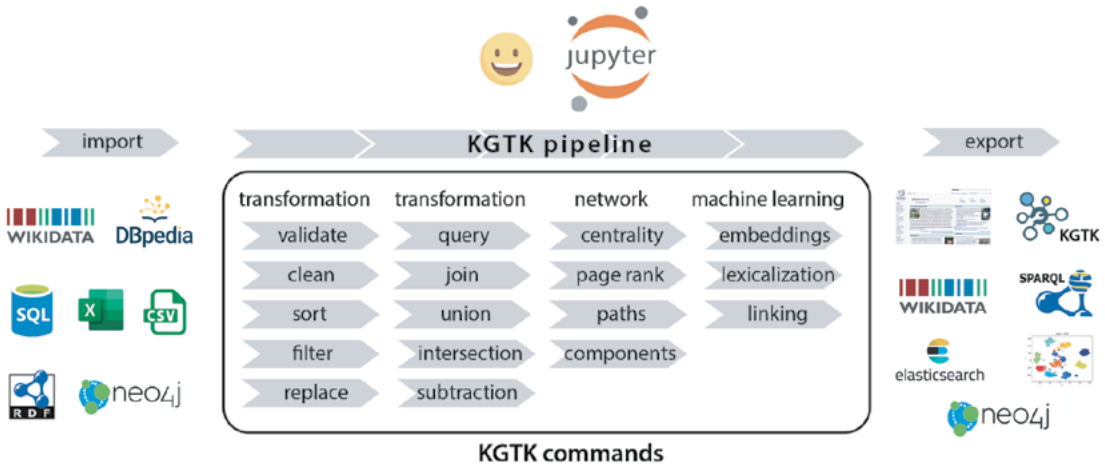


KGTK: A Knowledge Graph ToolKit

In the last decade, knowledge graphs have emerged as a preferred technology for representing and cataloging organizations' mission-critical knowledge. The goal is to enable rapid data assembly and analysis from multiple perspectives by connecting data that was previously siloed. In practice, it is costly and challenging to construct and exploit knowledge graphs because existing tools are deficient: they use specialized technologies unfamiliar to most developers; they are slow and unscalable; and they do not integrate with modern machine learning tools.

ISI's knowledge graph group is working on KGTK, a new Knowledge Graph ToolKit, to address deficiencies in existing tools. KGTK is built from the ground up to provide state-of-the-art components to address construction and exploitation requirements, while being easy to use, scalable, and fast.

KGTK defines a new simple knowledge interchange format that represents hyper-relational knowledge graphs using tab separated (TSV) files and can import and export knowledge graphs in many popular formats including Wikidata (the large publicly maintained knowledge graph), the Resource Description Framework (RDF), property graphs, and others. All KGTK commands consume and produce knowledge graphs represented in this simple format, so they can be composed into pipelines to perform complex transformations and analyses. KGTK provides a rich suite of commands to build, transform, combine, query, and analyze knowledge graphs, including commands to seamlessly integrate knowledge graphs with language models such as BERT and graph embeddings such as TransE and ComplEx. KGTK is open source (**https://github.com/usc-isi-i2/kgtk**), well-documented, actively used and developed, and released under the MIT License.



KGTK is a tool for both practitioners and researchers. Notable use cases include the creation of knowledge graphs for food security and the pharmaceutical industry, built as extensions of Wikidata. KGTK pipelines extract relevant subsets of Wikidata, augmenting them with data from spreadsheets and CSV files. These knowledge graphs are enhanced with a variety of network analytics and embeddings computed using KGTK.

ISI researchers also use KGTK to study knowledge graphs. In one use case, KGTK is being used to study the evolution of Wikidata at an unprecedented scale by computing differences between more than 300 Wikidata dumps released since the inception of the Wikidata project. The resulting study shows the impact of communities of interest and how their contributions shaped the present-day Wikidata.

The knowledge graph group is also researching new tools to address challenges with real-world knowledge graphs, including knowledge graph abstraction, identification and correction of errors, knowledge graph summarization and profiling, automated ingestion of structured sources, entity linking, creation of multi-lingual knowledge graphs, computation of knowledge graph embeddings, and intuitive browsing and visualization tools.

Contact: Pedro Szekeley, pszekely@isi.edu; Jay Pujara, jpujara@isi.edu

Reducing Stress Effects on Multi-project-wafer Reticles by Optimizing Metal Densities and Density Gradients

The manufacture, packaging, and testing of semiconductors create stress effects that limit yield, performance, and reliability. This has led to work in back-end-of-line (BEOL) layer density enhancement for Chip-Packaging Interaction (CPI). The CPI effect largely consists of thermal-mechanical stress effects resulting from wafer bumping and chip packaging processes.

The same stress effects also occur at the wafer level, and they are as significant as the chip-level CPI effect. Wafer manufacturing processes, such as wafer bumping and wafer thinning, induce thermo-mechanical stress effects that greatly limit yield. Cracks and delamination resulting from the stress can be severe, even across dies on a wafer. Wafer warpage also limits yield.

One design solution to mitigate the stress effects is to enhance the BEOL layer density and uniformity, namely, the lowest density gradient. However, it has been difficult to achieve in a silicon fabrication project, especially on a multi-project wafer (MPW) reticle where multiple chips are being assembled. Each chip has completed metal routes and different densities from other chips on the same reticle. The completed routes usually do not leave enough space for inserting additional dummy metals. The spaces (scribe lines) between the chips are not accounted for and not checked for density rule requirements; however, metal densities in those spaces still cause stress effects and other process variations from the chemical mechanical polish process. To date, electronic design automation (EDA) methods for reducing stress effects in MPW reticle chip assemble flows have not been found.

Researchers in the MOSIS group are pursuing a new method to reduce thermo-mechanical stress effects on MPW reticles. The method optimizes metal densities and density gradients on metal layers. This is achieved through a new MPW chip placement method and a new inter-die dummy metal fill method. The team has demonstrated intentional stress reduction structures in the form of a guard-ring type of crack-stopping structures; they also demonstrated effective crack-stop cell placement around a chip in MPW placement. Future work is planned in TCAD stress effect simulations, direct silicon measurement, and intentional on-chip stress reducing dummy fill methods.

Contact: Lifu Chang, lchang@mosis.com

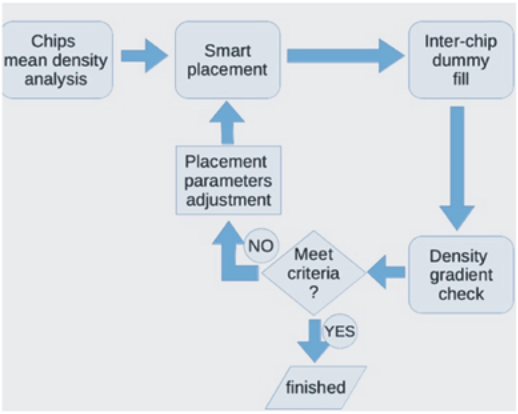


Figure 1. MPW chip placement flow optimizing density and density gradients between chips on the MPW reticle.

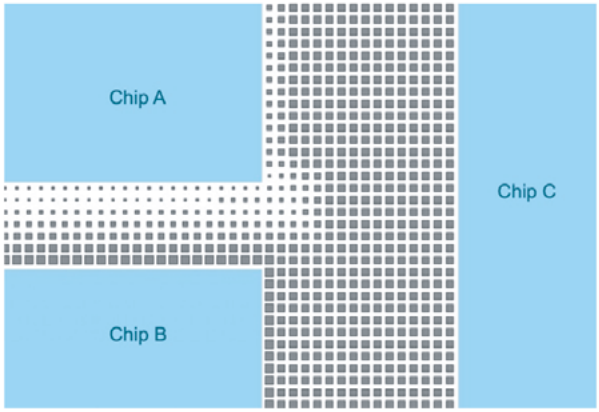


Figure 2. Progressive dummy fill near three adjacent chips after density gradient is optimized.

Figure 3. Dummy fill with ring structures surrounding each chip on an MPW reticle.

(a) Placement of 10 different chips with each average density of a metal layer noted.

(b) Zoom-in of corners of four chips, as marked by the yellow zoom-in window, showing the dummy fill ring cell structures surrounding each chip and the dummy fills in the scribe line space.

DERIVA Scientific Asset Management Platform: Supporting USC Research on Seniors and Exercise

Fewer than 10% of people over 65 meet national guidelines for physical activity. Researchers in the USC Chan Division of Occupational Science and Occupational Therapy are researching this issue; they are conducting experiments with a mobile phone app and a cohort of older adults who serve as experiment participants. Dr. Stacey Schepens Niemiec, Associate Professor of Research, designed the Moving Up app, which offers a suite of features specifically optimized to address inactivity in older individuals.

But Moving Up nearly didn't get off the ground until a chance encounter with a researcher in the Informatics Systems Research Division.

Professor Niemiec's team struggled to find collaborators capable of developing the intricate technological functionality of Moving Up. She was introduced to a division Research Lead with expertise in scientific data management who researches and develops innovative data management technology as part of the DERIVA platform. DERIVA is the underlying data management system for numerous scientific projects in labs, core facilities, multi-site collaborations, and NIH-supported data coordinating centers.

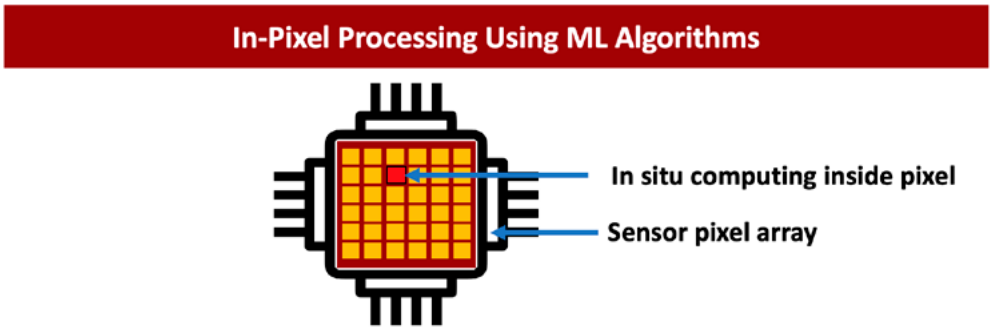
DERIVA is similar to a content management system but specialized to help scientists work with data in a streamlined, efficient way. DERIVA enables scientists and researchers to manage the unprecedented volume and diversity of data generated by today's scientific instruments. The Moving Up app utilizes the DERIVA platform to manage all its research data, including real-time events obtained by the app from the study participants.

This collaboration delivered a unique platform to empower USC Chan researchers in the study of intervention and motivational protocols for increasing activity levels in older adults.

Contact: Rob Schuler, schuler@isi.edu

X-PIXEL Computing: Designing an Imaging Sensor Chip with Built-in Compute Power

X-PIXEL is a novel computing model for edge intelligence using in-pixel, massively parallel analog computing for AI-enabled sensors. Examples of imaging sensor integrated circuits include "smart glasses" like the Google Glass Enterprise Edition and LIDAR remote sensing. The idea is to design a chip that is not only an imaging sensor but also processes the image data using built-in neural network logic.



X-PIXEL uses novel memory-embedded pixels (using CMOS or emerging non-volatile memory technologies) to support analog vector-matrix multiplication operations, enabling researchers in ISI's Application Specific Intelligent Computing lab to map a neural network computation directly into individual pixels. The in-pixel computing concept, called *X-PIXEL: Xcelerated Processing In-Pixel for Lightweight Neural Networks*, provides significant energy-delay improvements compared to existing near-sensor and in-sensor compute solutions.

Further, the researchers ran algorithmic simulations to determine if analog in-pixel computing can be useful for real-life applications. Analysis showed close to state-of-the-art accuracy for well-known Tiny Machine Learning (TinyML) datasets, including the German Traffic Signs Recognition Benchmark (GTSRB) and Visual Wake Words (VWW), while achieving significant energy-delay improvements.

The research team also took an initial set of data to analyze Hyperspectral Image Recognition using in-pixel processing. Results show close to state-of-the-art accuracy for pixel-based classification for hyperspectral cameras. In-pixel computing could find broad applicability for intelligent imaging on platforms such as autonomous drones, providing significant improvements in size, weight, and power.

Contact: Ajey Jacob, ajey@isi.edu; Akhilesh Jaiswal, akjaiswal@isi.edu; Peter Beerel, pabeerel@usc.edu

Open World Commonsense Inference

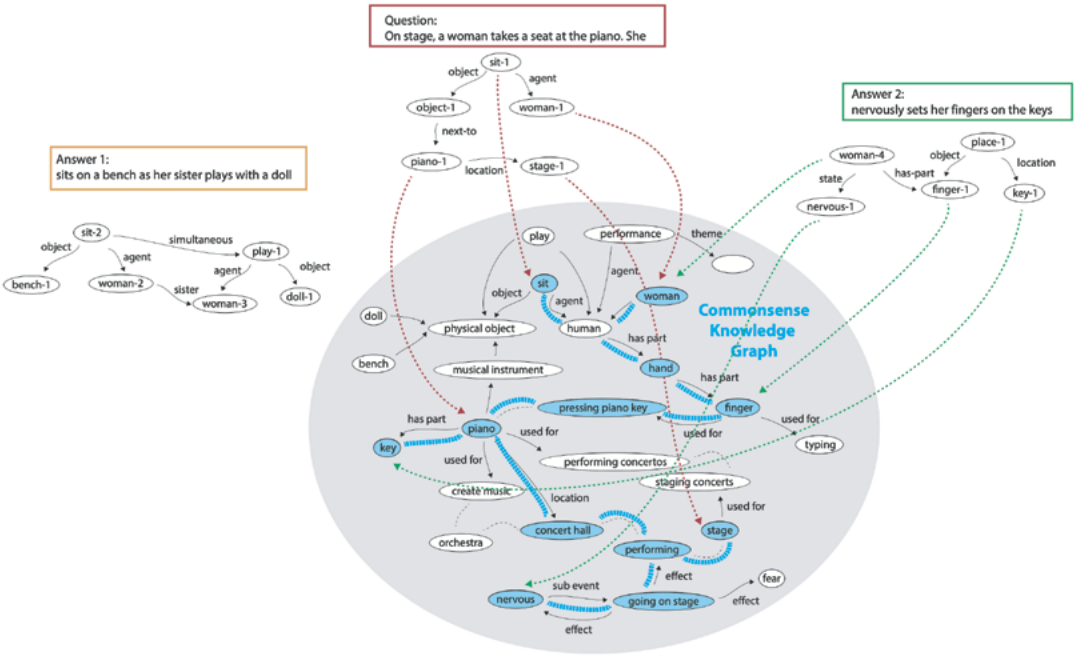
The Multi-modal Open World Grounded Learning and Inference (MOWGLI) team is developing open world AI techniques that perform physical, social, numerical, and event-based commonsense reasoning with the goal of enabling systems that can seamlessly understand—and interact with—humans in a range of scenarios.

The work starts with large-scale language representation models and extends them with a more intelligent approach by explicitly incorporating knowledge sources with varying degrees of structure, explicitness, and conceptual coverage.

The illustration presents an example. The task is to select the most probable follow-up to the sentence: "On stage a woman takes a seat at the piano. She . . . "

Which answer is best supported by the commonsense knowledge?

A commonsense knowledge source is used to find implicit connections between the question and each of the possible answers. The MOWGLI software then picks the answer that is best supported by the commonsense knowledge. In the illustration, this is "nervously sets her fingers on the keys."



The research team developed methods of answering such commonsense questions in natural language and explaining the reasoning process by leveraging commonsense knowledge graphs and other resources through advanced machine learning techniques. The techniques include graph neural networks, open-ended contextualized retrieval, and auto-prompting using gradient-guided search. MOWGLI focuses on generalizable zero/few-shot reasoning methods that typically perform commonsense inference using consolidated knowledge sources instead of specialized training data. A key contribution is the Commonsense Knowledge Graph (CSKG), which consolidates seven disjoint sources of commonsense knowledge.

The team developed a dozen benchmarks to systematically evaluate reasoning progress for both multiple-choice and generative automated evaluation of open-ended reasoning. The benchmarks test the ability of algorithms to:

- Perform numeric reasoning
- Remain consistent regardless of linguistic variations
- Compose words into commonsensical situations
- Understand preconditions of assumed facts
- Solve riddles

Contact: Filip Ilievski, ilievski@isi.edu; Mayank Kejriwal, kejriwal@isi.edu

Quantifying Hardware Confidentiality

Metrics are critical in all fields of science as they allow us to measure results and progress objectively. Some fields utilize well-established metrics; for example, power, performance, and area are used to measure a microprocessor's capability. However, security poses challenges because it is a human-derived concept, not bound by the laws of physics. Security metrics are deeply entwined with assumptions about what is being protected and what the adversary can do. This is especially true in the nascent field of hardware security. One challenge is quantitatively measuring the confidentiality level of a circuit to assess if that circuit might leak critical intellectual property (IP). IP leakage can allow an adversary to steal or reproduce a design that cost millions of dollars to develop.

Researchers in the Computational Systems and Technology Division have developed the Mirage platform, an electronic design automation platform that quantitatively compares tradeoffs in security against traditional power, performance, and area metrics.

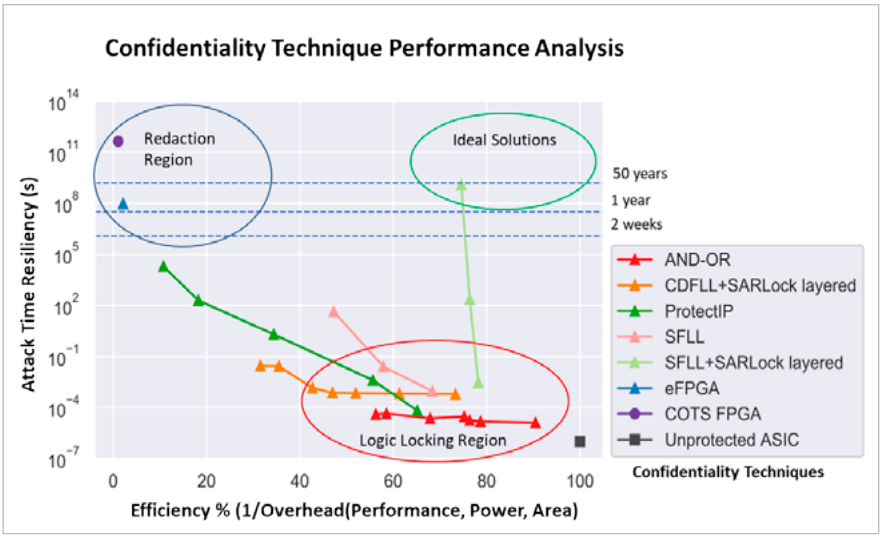
Mirage security metrics currently include:

- Attack Time Resiliency, the time a confidentiality technique can withstand its most challenging attack
- Attack Robustness, the ability of a confidentiality technique to withstand combinations of attacks or future attacks
- Structural Entropy, a measure of the non-uniformity of a confidentiality implementation that indicates a potential vulnerability

The Mirage framework has now been used on multiple projects, evaluating over 10 confidentiality techniques, and is advancing the state-of-the-art in circuit confidentiality.

The community had debated the efficacy of redaction techniques (wherein the critical circuitry is completely removed from a circuit and replaced with programmable elements) versus logic-locking techniques (which add locking mechanisms to an existing circuit). The Mirage tools revealed that redaction techniques provided high security but with severe cost to efficiency; logic-locking techniques were highly efficient but provided lower security. The ideal implementation would offer high security and high efficiency. The Mirage data provides insights into potential solutions, such as support-set expansion, state machine redaction, and additional ways to add more programmability to logic-locking techniques or efficiency to redaction techniques.

The data produced from the Mirage platform has led to insights that are spurring the evolution of confidentiality metrics. Currently, most hardware security metrics are reactive; they measure a confidentiality technique's effectiveness after attacks against it are developed. The analysis of thousands of Mirage data sets reveals that certain physical signatures of a confidentiality technique result in vulnerabilities and that an adversary can develop an attack leveraging these vulnerabilities. Because these signatures are known at design time, a new metric assessing a given technique's risk of being broken is in development. Assessing the potential of a confidentiality technique to be broken is critical in hardware—because unlike software, hardware design cannot be patched after production.



Contact: Matthew French, mfrench@isi.edu

Grasping the Information Contained in Tables on the Web

The web contains hundreds of millions of tables embedded in web pages, PDF documents, spreadsheets, and delimited files. Today, the valuable knowledge in tables on the web goes largely unused because of the cost to extract, normalize, and understand the data. The table understanding group at ISI is exploring this problem, including finding tables relevant to a search or topic, identifying the structure of the tables, determining semantic relationships, linking the tables to knowledge graphs, and fact-checking using tabular data. Their work will allow users to quickly transform tabular data into useful knowledge using an integrated set of tools.

The table understanding researchers have developed solutions to several problems:

- Retrieving relevant tables in a large corpus using a natural language query
- Detecting table layout to uncover the relationships between the elements in complex tables including multi-level headers, metadata, and sub-tables
- Linking data in tables to entities in knowledge graphs to relate new data to existing knowledge
- Summarizing tables to create concise natural language summaries of the table data
- Fact-checking using tables

One of the key techniques for these tasks is representation learning for tables, using millions of tables to learn patterns in structure and content. Like language models, these pretrained table models can be used in a variety of downstream applications. The group has explored transformer architectures and graph neural networks, and it has produced state-of-the-art results on table understanding benchmarks for cell classification, layout detection, and table retrieval.

The group also builds practical tools and applications. The T2WML tool maps tables to Wikidata with a simple user interface. T2WML can ingest an arbitrary table downloaded from the web, apply the automated layout detection, semantic mapping, and entity linking algorithms, and then present the results as simple, colorized overlays over the original data. The user interface enables users to quickly correct imperfections in the AI algorithms before uploading the data to a knowledge graph or downloading it in a canonical form ready for processing in popular tools such as the Python Data Analysis Library (pandas) and Jupyter Notebook.

Contact: Pedro Szekely, pszekely@isi.edu

Candidate	Party/alliance	First round		Second round	
		Votes	%	Votes	%
Iván Duque Márquez	Grand Alliance for Colombia	7,569,693	39.14	10,373,080	53.98
Gustavo Petro	List of Decency	4,851,254	25.09	8,034,189	41.81
Sergio Fajardo	Colombia Coalition	4,589,696	23.73		
Germán Vargas Lleras	Mejor Vargas Lleras	1,407,840	7.28		
Humberto De la Calle	PLC–ASI	399,180	2.06		
Jorge Antonio Trujillo	We Are All Colombia	75,614	0.39		
Promotores Voto En Blanco	Party of Ethnic Reclamation "PRE"	60,312	0.31		
Viviane Morales Hoyos	Somos Región Colombia	41,458	0.21		
Blank votes		341,087	1.76	808,104	4.21
Invalid votes		300,080	–	295,499	–
Total		**19,336,134**	**100**	**19,510,684**	**100**
Registered voters/turnout		36,227,267	53.37	36,783,940	53.04
Source: El Tiempo Government					

T2WML

Goal: To add data from a Wikipedia table to Wikidata

2018 Colombian presidential election.csv [R... Output Wikify Annotate

	A	B	C	D	E	F	G
1	Candidate		Party/alli...	First round		Second r...	
2				Votes	%	Votes	%
3		Iván Duq...	Grand Alli...	7569693	39.14	10373080	53.98
4		Gustavo ...	List of De...	4851254	25.09	8034189	41.81
5		Sergio Fa...	Colombia...	4589696	23.73		
6		Germán ...	Mejor Var...	1407840	7.28		
7		Humbert...	PLC–ASI	399180	2.06		
8		Jorge Ant...	We Are Al...	75614	0.39		
9		Promotor...	Party of E...	60312	0.31		
10		Viviane M...	Somos R...	41458	0.21		
11	Blank vot...			341087	1.76	808104	4.21
12	Invalid vo...			300080	–	295499	–
13	Total			19336134	100	19510684	100
14	Registere...			36227267	53.37	36783940	53.04
15	Source: E...						
16	URL	https://en...					
17							

The user annotates the data.

WIKIDATA

T2WML transforms the data into Wikidata statements.

Teaching AI to Play Text-based Adventure Games

Although AI models can hold enormous quantities of knowledge and learn from past mistakes, they lack a general understanding of implicit information and the common sense that informs human decision-making. To test AI's ability to master decision-making skills in diverse settings and contexts, a team of researchers taught models to master text-based games that follow a "choose-your-own-adventure" structure, similar to the '70s classic "Zork."

For example, in a game based on cooking, the player might read a recipe that says, "Dice and grill a yellow bell pepper." To win the game, the player must select an appropriate sequence of actions: the player must find a pepper, then use a knife to dice it, and then use a stove to grill it. Knowing these sub-steps and the order to perform them is obvious for most adult humans, but AI models need to be taught how to take a logical sequence of steps to complete a task. Humans are also good at generalizing—once we learn how to take logical sequences of steps, we can do this for new domains, such as following a "recipe" for sewing or furniture assembly.

The researchers employed Deep Reinforcement Learning, a process in which deep neural networks help models learn from their mistakes and make correct decisions. They developed models that can play a series of games based on the theme of cooking recipes. At first, the models constantly made mistakes, but they eventually learned how to search for ingredients and how to combine them appropriately based on a recipe. The team then integrated BERT, a large language model developed at Google, to provide large amounts of common background knowledge to the models. Equipped with this knowledge and the general ability to make decisions, the models were able to transition to novel games in a completely new treasure-hunting domain. This work was presented in the *Findings of the Association for Computational Linguistics: EMNLP 2020*.

The development of sequential decision-making skills will prove important in artificial intelligence models because it enables more contextually flexible interaction. If modern dialogue and assistant AI bots can adopt complex decision-making skills, our interactions with them will be much more efficient and helpful. With the success of research studies like this, AI moves closer to resembling human characteristics previously exclusive to our kind. This study, and others like it, will propel the artificial intelligence field into one that truly understands the ins-and-outs of being human.

Contact: Jonathan May, jonmay@isi.edu

COMPUTING ARCHITECTURES AND SYSTEMS

Evaluation of Remote-sensing Architectures Using the Virtual Constellation Engine
M. Paolieri, V. Menon, A. Schmidt, M. French
American Geophysical Union Fall Meeting 2020

FPGA Virtualization for Deprecated Devices
I. Taras and A. Schmidt
2020 IEEE 28th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)

Independent Testing of Untrusted FPGAs for Faulty Interconnect
T. Haroldsen, M. French, T. Sung, D. Glick, J. Danner, L. Lerner
Government Microcircuit Applications & Critical Technology Conference (GOMAC Tech 2020)

Logic Obfuscation: Modeling Attack Resiliency
V. Menon, G. Kolhe, J. Fifty, A. Schmidt, J. Monson, M. French, Y. Hu, P. Beerel, P. Nuzzo
Government Microcircuit Applications & Critical Technology Conference (GOMACTech)

Scalable Parallel File Write from a Large NUMA System
D. Kang, J. P. Walters, S. Crago
2020 IEEE High Performance Extreme Computing Conference (HPEC)

StereoBit: An Innovative SpaceCube Application for Atmospheric Science
J. Carr, C. Wilson, D. Wu, M. French, M. Kelly
IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium

Emulating and Verifying Sensing, Computation, and Communication in Distributed Remote Sensing Systems
M. French, M. Paolieri, V. Menon, A. Schmdit
IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium

CYBERSECURITY

An Experimental Approach for Estimating Cyber Risk: A Proposal Building upon Cyber Ranges and Capture the Flags
G. Di Tizio, F. Massacci, L. Allodi, S. Dashevskyi, J. Mirkovic
Workshop on Cyber Range Technologies and Applications (CACOE)

Be the Phisher - Understanding Users' Perception of Malicious Domains
F. Quinkert, M. Degeling, J. Blythe, T. Holz
15th ACM ASIA Conference on Computer and Communications Security

Behavioral Determinants of Target Shifting and Deterrence in an Analog Cyber-attack Game
S. Kusumastuti, J. Blythe, H. Rosoff, R. John
Risk Analysis

BLAG: Improving the Accuracy of Blocklists
S. Ramanathan, J. Mirkovic, M. Yu
27th Annual Network and Distributed System Security Symposium, NDSS

Challenges in Forecasting Malicious Events from Incomplete Data
N. Tavabi, A. Abeliuk, N. Mokhberian, J. Abramson, K. Lerman
Companion Proceedings of the Web Conference 2020

Eyes on URLs: Relating Visual Behavior to Safety Decisions
V. Kothari, C. Mills, N. Ramkumar, J. Blythe, R. Koppel, S. Smith, A. Kun
ACM Symposium on Eye Tracking Research and Applications

Human-computability Boundaries
V. Kothari, P. Anantharaman, I. Jenkins, M. Millian, S. Bratus, J. Blythe, R. Koppel, S. Smith
Security Protocols XXVII, 27th International Workshop, Revised Selected Papers

Multidisciplinary publications appear in all relevant areas.

Malware Behavior Through Network Trace Analysis
X. Deng and J. Mirkovic
Proceedings of International Networking Conference (INC)

Massive Cross-platform Simulations of Online Social Networks
G. Muric, A. Tregubov, J. Blythe, A. Abeliuk, D. Choudhary, K. Lerman, E. Ferrara
Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems

Mismorphism: The Heart of the Weird Machine
P. Anantharaman, V. Kothari, I. Jenkins, M. Millian, S. Bratus, J. Blythe, R. Koppel, S. Smith
Security Protocols XXVII, 27th International Workshop, Revised Selected Papers

Optimization of Large-scale Agent-based Simulations Through Automated Abstraction and Simplification
A. Tregubov and J. Blythe
AAMAS Workshop on Multi-Agent Systems and Agent-Based Simulation

Quantifying Cloud Misbehavior
R. Tandon, P. Charnsetikul, J. Mirkovic
Proceedings of the IEEE CloudNet

Quantifying the Impact of Blocklisting in the Age of Address Reuse
S. Ramanathan, A. Hossain, J. Mirkovic, M. Yu, S. Afroz
Proceedings of the ACM Internet Measurement Conference (IMC 2020)

Replicated Testbed Experiments for the Evaluation of a Wide-range of DDoS Defenses
A. Hussain, D. DeAngelis, E. Kline, S. Schwab
2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)

The DARPA SocialSim Challenge: Cross-platform Multi-agent Simulations
G. Muric, A. Tregubov, J. Blythe, A. Abeliuk, D. Choudhary, K. Lerman, E. Ferrara
International Conference on Practical Applications of Agents and Multi-Agent Systems AAMAS

Toward Orchestration of Complex Networking Experiments
A. Hussain, P. Jaipuria, G. Lawler, S. Schwab, T. Benzel
13th USENIX Workshop on Cyber Security Experimentation and Test (CSET 20)

User-centered Risk Communication for Safer Browsing
S. Das, J. Abbott, S. Gopavaram, J. Blythe, L. J. Camp
Proceedings of the First Asia USEC - Workshop on Usable Security, In Conjunction with the Twenty-Fourth International Conference International Conference on Financial Cryptography and Data Security 2020

Using Terminal Histories to Monitor Student Progress on Hands-on Exercises
J. Mirkovic, A. Aggarwal, D. Weinman, P. Lepe, J. Mache, R. Weiss
Proceedings of Special Interest Group on Computer Science Education Symposium (SIGCSE)

KNOWLEDGE TECHNOLOGIES

A Benchmarking Study of Embedding-based Entity Alignment for Knowledge Graphs
Z. Sun, Q. Zhang, W. Hu, C. Wang, M. Chen, F. Akrami, C. Li
Proceedings of the VLDB Endowment, Vol. 13

A Combinatorial Perspective on Ising Model Hysteresis
Y. Guan, A. Li, S. Koenig, S. Haas, S. Thittamaranahalli
American Physical Society, March Meeting (APS-2020)

A Meta-engine for Building Domain-specific Search Engines
M. Kejriwal
Software Impacts

A Template-based Approach for Annotating Long-tail Datasets
D. Garijo, K-T. Yao, A. Singh, P. Szekely
Joint Proceedings of Workshops AI4LEGAL2020, NLIWOD, PROFILES 2020, QuWeDa 2020, and SEMIFORM2020, Co-located with the 19th International Semantic Web Conference (ISWC 2020) Virtual Conference

An Automatic Approach for Generating Rich, Linked Geo-metadata from Historical Map Images
Z. Li, Y. Chiang, S. Tavakkol, B. Shbita, J. Uhl, S. Leyk, C. A. Knoblock
Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining

Analogous Process Structure Induction for Sub-event Sequence Prediction
H. Zhang, M. Chen, H. Wang, Y. Song, D. Roth
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Automatic Alignment of Contemporary Vector Data and Georeferenced Historical Maps Using Reinforcement Learning
W. Duan, Y. Chiang, S. Leyk, J. Uhl, C. A. Knoblock
International Journal of Geographical Information Science

Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web
D. Garijo and M. Poveda
Applications and Practices in Ontology Design, Extraction, and Reasoning

Bio-joie: Joint Representation Learning of Biological Knowledge Bases
J. Hao, C. Ju, M. Chen, Y. Sun, C. Zaniolo, W. Wang
Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics

Building Linked Spatio-temporal Data from Vectorized Historical Maps
B. Shbita, C. A. Knoblock, W. Duan, Y. Chiang, J. Uhl, S. Leyk
European Semantic Web Conference

Coming to Terms with FAIR Ontologies: A Position Paper
M. Poveda-Villalón, P. Espinoza-Arias, D. Garijo, O. Corcho
Proceedings of EKAW 2020 - 22nd International Conference on Knowledge Engineering and Knowledge Management

Decision Tree Learning-inspired Dynamic Variable Ordering for the Weighted CSP
H. Xu, K. Sun, S. Koenig, S. Thittamaranahalli
Proceedings of the Thirteenth International Symposium on Combinatorial Search (SOCS-2020)

Denoising Autoencoders for High-qubit Quantum Dynamics Simulations on Quantum Computers
C. Powers, L. Bassman, Y. Geng, R. Kalia, S. Thittamaranahalli, T. Linker, K. Liu, A. Nakano, PJ Rajak, P. Vashishta
Proceedings of the NeurIPS-2020 Workshop on Machine Learning and the Physical Sciences

Diagnostic Prediction with Sequence-of-sets Representation Learning for Clinical Events
T. Zhang
Artificial Intelligence in Medicine: 18th International Conference on Artificial Intelligence in Medicine (AIME 2020)

Domain-specific Compilers for Dynamic Simulations of Quantum Materials on Quantum Computers
L. Bassman, S. Gulania, C. Powers, R. Li, T. Linker, K. Liu, S. Thittamaranahalli, R. Kalia, A. Nakano, P. Vashishta
Special Issue Article in Journal of Quantum Science and Technology (QST-2020)

Embedding Directed Graphs in Potential Fields Using FastMap-D
S. Gopalakrishnan, L. Cohen, S. Koenig, S. Thittamaranahalli
Proceedings of the Thirteenth International Symposium on Combinatorial Search (SOCS-2020)

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 et al.
Nature Translational Psychiatry

Exact Approaches to the Multi-agent Collective Construction Problem
E. Lam, P. Stuckey, S. Koenig, S. Thittamaranahalli
Proceedings of the Fourth International Workshop on Multi-Agent Path Finding (WoMAPF-2020)

FAIR Computational Workflows
C. Goble, S. Cohen-Boulakia, S. Soiland-Reyes, D. Garijo, Y. Gil, M. Crusoe, K. Peters, D. Schoberhidden
Data Intelligence, Special Issue on FAIR (Findable, Accessible, Interoperable and Reusable) Principles

Multidisciplinary publications appear in all relevant areas.

Generating the Top K Solutions to Weighted CSPs: A Comparison of Different Approaches
A. Li, Y. Guan, S. Koenig, S. Haas, S. Thittamaranahalli
Proceedings of the Thirty-Second International Conference on Tools with Artificial Intelligence (ICTAI-2020)

Hybrid Quantum-classical Algorithms for Solving the Weighted CSP
H. Xu, K. Sun, S. Koenig, I. Hen, S. Thittamaranahalli
Proceedings of the Sixteenth International Symposium on Artificial Intelligence and Mathematics (ISAIM-2020)

Idle Time Optimization for Target Assignment and Path Finding in Sortation Centers
N. M. Kou, C. Peng, H. Ma, S. Thittamaranahalli, S. Koenig
Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-2020)

Joint Constrained Learning for Event-event Relation Extraction
H. Wang, M. Chen, H. Zhang, D. Roth
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Knowledge Association with Hyperbolic Knowledge Graph Embeddings
Z. Sun, M. Chen, W. Hu, C. Wang, J. Dai, W. Zhang
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Knowledge Graph Alignment Network with Gated Multi-hop Neighborhood Aggregation
Z. Sun, C. Wang, W. Hu, M. Chen, J. Dai, W. Zhang, Y. Qu
Proceedings of the AAAI Conference on Artificial Intelligence

Knowledge Graphs and COVID-19: Opportunities, Challenges, and Implementation
M. Kejriwal
Harvard Data Science Review - Special Issue 1 - COVID-19: Unprecedented Challenges and Changes

Lifelong Multi-agent Path Finding in Large-Scale Warehouses
J. Li, A. Tinka, S. Kiesel, J. Durham, S. Thittamaranahalli, S. Koenig
Proceedings of the Nineteenth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-2020)

Locally Constructing Product Taxonomies from Scratch Using Representation Learning
M. Kejriwal, R. Selvam, C-C. Ni, N. Torzec
Proceedings of the 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

Mapping the Web Ontology Language to the OpenAPI Specification
P. Espinoza-Arias, D. Garijo, O. Corcho
2nd Workshop on Conceptual Modeling, Ontologies and (Meta)data Management for Findable, Accessible, Interoperable and Reusable (FAIR) Data (CMOMM4FAIR), Colocated with the 2020 Conference on Advances in Conceptual Modeling

Moving Agents in Formation in Congested Environments
J. Li, K. Sun, H. Ma, A. Felner, S. Thittamaranahalli, S. Koenig
Proceedings of the Nineteenth International Conference on Autonomous Agents, Multi-Agent Systems (AAMAS-2020)

Multi-agent Path Finding with Mutex Propagation
H. Zhang, J. Li, P. Surynek, S. Koenig, S. Thittamaranahalli
Proceedings of the Thirtieth International Conference on Automated Planning and Scheduling (ICAPS-2020)

Multilingual Knowledge Graph Completion via Ensemble Knowledge Transfer
X. Chen, M. Chen, C. Fan, A. Uppunda, Y. Sun, C. Zaniolo
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings

Mutation Effect Estimation on Protein-protein Interactions Using Deep Contextualized Representation Learning
G. Zhou, M. Chen, C. Ju, Z. Wang, J-Y. Jiang, W. Wang
NAR Genomics and Bioinformatics

Mutex Propagation for SAT-based Multi-agent Path Finding
P. Surynek, J. Li, H. Zhang, S. Thittamaranahalli, S. Koenig
Proceedings of the Twenty-Third International Conference on Principles and Practice of Multi-Agent Systems (PRIMA-2020)

Network-theoretic Modeling of Complex Activity Using Online Sex Advertisements
M. Kejriwal and Y. Gu
Applied Network Science

On Detecting Urgency in Short Crisis Messages Using Minimal Supervision and Transfer Learning
M. Kejriwal and P. Zhou
Social Network Analysis and Mining (Springer)

On Using Product-specific Schema.org from Web Data Commons: An Empirical Set of Best Practices
R. K. Selvam and M. Kejriwal
Workshop on Knowledge Graphs and E-Commerce, in conjunction with KDD'20

ReadNet: A Hierarchical Transformer Framework for Web Article Readability Analysis
C. Meng, M. Chen, J. Mao, J. Neville
European Conference on Information Retrieval

Structural Studies of the Global Networks Exposed in the Panama Papers
M. Kejriwal and A. Dang
Applied Network Science

The Genetic Architecture of the Human Cerebral Cortex
K. Grasby, N. Jahanshad et al.
Science

Using Historical Maps in Scientific Studies: Challenges and Best Practices
Y. Chiang, W. Duan, S. Leyk, J. Uhl, C. A. Knoblock
SpringerBriefs in Geography Series

What Are You Trying to Do? Semantic Typing of Event Processes
M. Chen, H. Zhang, H. Wang, D. Roth
Proceedings of the 24th Conference on Computational Natural Language Learning

KGTK: A Toolkit for Large Knowledge Graph Manipulation and Analysis
F. Ilievski, D. Garijo, H. Chalupsky, N. Teja Divvala, Y. Yao, C. Rogers, R. Li, J. Liu, A. Singh, D. Schwabe, P. Szekely
Proceedings of The Semantic Web – ISWC 2020

OBA: An Ontology-based Framework for Creating REST APIs for Knowledge Graphs
D. Garijo and M. Osorio
Proceedings of The Semantic Web – ISWC 2020

MACHINE LEARNING AND DATA SCIENCE

A Geometric Solution to Fair Representations
Y. He, K. Burghardt, K. Lerman
Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society

A Global Database of Holocene Paleotemperature Records
D. Kaufman et al.
Scientific Data

Affect Estimation with Wearable Sensors
S. Yan, H. Hosseinmardi, H-T. Kao, S. Narayanan, K. Lerman, E. Ferrara
Journal of Healthcare Informatics Research

Aggressive, Repetitive, Intentional, Visible, and Imbalanced: Refining Representations for Cyberbullying Classification
C. Ziems, Y. Vigfusson, F. Morstatter
Proceedings of the International AAAI Conference on Web and Social Media

All in the Exponential Family: Bregman Duality in Thermodynamic Variational Inference
R. Brekelmans, V. Masrani, F. Wood, G. Ver Steeg, A. Galstyan
Proceedings of ICML 2020

Anchor Attention for Hybrid Crowd Forecasts Aggregation
Y. Huang, A. Abeliuk, F. Morstatter, P. Atanasov, A. Galstyan
Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems

Multidisciplinary publications appear in all relevant areas.

Annealed Importance Sampling with q-Paths
R. Brekelmans, V. Masrani, T. Bui, F. Wood, A. Galstyan, G. Ver Steeg, F. Nielsen
NeurIPS Workshop on Deep Learning through Information Geometry

Bots, Elections, and Social Media: A Brief Overview
E. Ferrara
Disinformation, Misinformation, and Fake News in Social Media

Building Autocorrelation-aware Representations for Fine-scale Spatiotemporal Prediction
Y. Lin, Y. Chiang, M. Franklin, S. Eckel, J. L. Ambite
20th IEEE International Conference on Data Mining (ICDM)

Can Badges Foster a More Welcoming Culture on Q&A Boards?
T. Santos, K. Burghardt, K. Lerman, D. Helic
Proceedings of the International AAAI Conference on Web and Social Media

Can Oncologists Predict the Efficacy of Treatments in Randomized Trials?
D. Benjamin, D. Mandel, T. Barnes, M. Krzyzanowska, N. Leighl, I. Tannock, J. Kimmelman
The Oncologist

Challenges in Forecasting Malicious Events from Incomplete Data
N. Tavabi, A. Abeliuk, N. Mokhberian, J. Abramson, K. Lerman
Companion Proceedings of the Web Conference 2020

Characterizing Social Media Manipulation in the 2020 US Presidential Election
E. Ferrara, H. Chang, E. Chen, G. Muric, J. Patel
First Monday

Charting the Landscape of Online Cryptocurrency Manipulation
L. Nizzoli, S. Tardelli, M. Avvenuti, S. Cresci, M. Tesconi, E. Ferrara
IEEE Access

Comparison of Patient and Expert Perceptions of the Attainment of Research Milestones in Parkinson's Disease
P. B. Kane, D. Benjamin, R. Barker, A. Lang, T. Sherer, J. Kimmelman
Movement Disorders

Detecting Multi-timescale Consumption Patterns from Receipt Data: A Non-Negative Tensor Factorization Approach
A. Matsui, T. Kobayashi, D. Moriwaki, E. Ferrara
Journal of Computational Social Science

Detecting Troll Behavior Via Inverse Reinforcement Learning: A Case Study of Russian Trolls in the 2016 US Election
L. Luceri, S. Giordano, E. Ferrara
Proceedings of the International AAAI Conference on Web and Social Media

Discovering Patterns of Online Popularity from Time Series
M. Ozer, A. Sapienza, A. Abeliuk, G. Muric, E. Ferrara
Expert Systems with Applications

Does Streaming Esports Affect Players' Behavior and Performance?
A. Matsui, A. Sapienza, E. Ferrara
Games and Culture

Enabling Low-resource Transfer Learning Across COVID-19 Corpora by Combining Event-extraction and Co-training
A. Spangher, N. Peng, J. May, E. Ferrara
ACL 2020 Workshop on Natural Language Processing for COVID-19 (NLP-COVID)

End-to-end Learning of Compressible Features
S. Singh, S. Abu-El-Haija, N. Johnston, J. Balle, A. Shrivastava, G. Toderici
IEEE International Conference on Image Processing

Fair Class Balancing: Enhancing Model Fairness without Observing Sensitive Attributes
S. Yan, H. Kao, E. Ferrara
Proceedings of the 29th ACM International Conference on Information & Knowledge Management

Forecasts for the Attainment of Major Research Milestones in Parkinson's Disease
P. B. Kane, D. Benjamin, R. Barker, A. Lang, T. Sherer, J. Kimmelman
Journal of Parkinson's Disease

Friendship Paradox Biases Perceptions in Directed Networks
N. Alipourfard, B. Nettasinghe, A. Abeliuk, V. Krishnamurthy, K. Lerman
Nature Communications

Generating and Understanding Personalized Explanations in Hybrid Recommender Systems
P. Kouki, J. Schaffer, J. Pujara, J. O'Donovan, L. Getoor
ACM Transactions on Interactive Intelligent Systems

Generative Continual Concept Learning
M. Rostami, S. Kolouri, P. Pilly, J. McClelland
Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI-20)

Graph Embedding with Personalized Context Distribution
D. Huang, Z. He, Y. Huang, K. Sun, S. Abu-El-Haija, B. Perozzi, K. Lerman, F. Morstatter, A. Galstyan
Companion Proceedings of the Web Conference

Human Languages in Source Code: Auto-translation for Localized Instruction
C. Piech and S. Abu-El-Haija
ACM Conference on Learning @ Scale

Human-like Time Series Summaries Via Trend Utility Estimation
P. Jandaghi and J. Pujara
Ninth International Workshop on Statistical Relational AI

Improving Generalization by Controlling Label-noise Information in Neural Network Weights
H. Harutyunyan, K. Reing, G. Ver Steeg, A. Galstyan
International Conference on Machine Learning (ICML)

Invariant Representations Through Adversarial Forgetting
A. Jaiswal, D. Moyer, G. Ver Steeg, W. AbdAlmageed, P. Natarajan
Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI-20)

Learning Behavioral Representations from Wearable Sensors
N. Tavabi, H. Hosseinmardi, J. Villatte, A. Abeliuk, S. Narayanan, E. Ferrara, K. Lerman
International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation

Learning Cell Embeddings for Understanding Table Layouts
J. Pujara, M. Ghasemi-Gol, P. Szekely
Knowledge and Information Systems

Learning to Reason in Round-based Games: Multi-task Sequence Generation for Purchasing Decision Making in First-person Shooters
Y. Zeng, D. Lei, B. Li, G. Jiang, E. Ferrara, M. Zyda
Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment

Likelihood Ratio Exponential Families
R. Brekelmans, F. Nielsen, A. Makhzani, A. Galstyan, G. Ver Steeg
NeurIPS Workshop on Deep Learning through Information Geometry

Low-rank Regularization and Solution Uniqueness in Over-parameterized Matrix Sensing
K. Geyer, A. Kyrillidis, A. Kalev
Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics

Man is to Person as Woman is to Location: Measuring Gender Bias in Named Entity Recognition
N. Mehrabi, T. Gowda, F. Morstatter, N. Peng, A. Galstyan
Proceedings of the 31st ACM Conference on Hypertext and Social Media

Multidisciplinary publications appear in all relevant areas.

Massive Cross-Platform Simulations of Online Social Networks
G. Muric, A. Tregubov, J. Blythe, A. Abeliuk, D. Choudhary, K. Lerman, E. Ferrara
Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems

Maximizing Multivariate Information with Error-correcting Codes
K. Reing, G. Ver Steeg, A. Galstyan
IEEE Transactions on Information Theory

Measuring Bot and Human Behavioral Dynamics
I. Pozzana and E. Ferrara
Frontiers in Physics

Minority-centric Meta-analyses of Blood Lipid Levels Identify Novel Loci in the Population Architecture Using Genomics and Epidemiology (PAGE) Study
Y. Huh, M. Graff et al.
PLoS Genetics

Misinformation, Manipulation, and Abuse on Social Media in the Era of COVID-19
E. Ferrara, S. Cresci, L. Luceri
Journal of Computational Social Science

Modeling "Newsworthiness" for Lead-generation Across Corpora
A. Spangher, N. Peng, E. Ferrara
Computation + Journalism Symposium

Modeling Dialogues with Hashcode Representations: A Nonparametric Approach
S. Garg, I. Rish, G. Cecchi, P. Goyal, S. Ghazarian, S. Gao, G. Ver Steeg, A. Galstyan
Proceedings of The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-2020), The Thirty-Second Innovative Applications of Artificial Intelligence Conference (IAAI-2020), The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-2020)

Moral Framing and Ideological Bias of News
N. Mokhberian, A. Abeliuk, P. Cummings, K. Lerman
International Conference on Social Informatics

Network Modularity Controls the Speed of Information Diffusion
H. Peng, A. Nematzadeh, D. Romero, E. Ferrara
Physical Review E

Origins of Algorithmic Instabilities in Crowdsourced Ranking
K. Burghardt, T. Hogg, R. D'Souza, K. Lerman, M. Posfai
Proceedings of the ACM on Human-Computer Interaction

Overview of Scanner Invariant Representations
D. Moyer, G. Ver Steeg, P. Thompson
Medical Imaging with Deep Learning

Political Polarization Drives Online Conversations About COVID-19 in the United States
J. Jiang, E. Chen, S. Yan, K. Lerman, E. Ferrara
Human Behavior and Emerging Technology

Predictability Limit of Partially Observed Systems
A. Abeliuk, Z. Huang, E. Ferrara, K. Lerman
Scientific Reports

Predictive Engagement: An Efficient Metric for Automatic Evaluation of Open-Domain Dialogue Systems
S. Ghazarian, R. Weischedel, A. Galstyan, N. Peng
Proceedings of The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-2020), The Thirty-Second Innovative Applications of Artificial Intelligence Conference (IAAI-2020), The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-2020)

Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020
K. Verspoor, K. B. Cohen, M. Dredze, E. Ferrara, J. May, R. Munro, C. Paris, B. Wallace
1st Workshop on NLP for COVID-19 at ACL 2020

Quantifying Machine Influence Over Human Forecasters
A. Abeliuk, D. Benjamin, F. Morstatter, A. Galstyan
Scientific Reports

ReCOVery: A Multimodal Repository for COVID-19 News Credibility Research
X. Zhou, A. Mulay, E. Ferrara, R. Zafarani
CIKM '20: Proceedings of the 29th ACM International Conference on Information & Knowledge Management

Scanner Invariant Representations for Diffusion MRI Harmonization
D. Moyer, G. Ver Steeg, C. Tax, P. Thompson
Magnetic Resonance in Medicine

Stacking Models for Nearly Optimal Link Prediction in Complex Networks
A. Ghasemian, H. Hosseinmardi, A. Galstyan, E. Airoldi, A. Clauset
Proceedings of the National Academy of Sciences

Structural Node Embedding in Signed Social Networks: Finding Online Misbehavior at Multiple Scales
M. Heimann, G. Muric, E. Ferrara
International Conference on Complex Networks and Their Applications

Tackling Misinformation: What Researchers Could Do with Social Media Data
I. Pasquetto et al.
The Harvard Kennedy School Misinformation Review

The Transsortative Structure of Networks
S. Ngo, A. Percus, K. Burghardt, K. Lerman
Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences

TILES-2018, a Longitudinal Physiologic and Behavioral Data Set of Hospital Workers
K. Mundnich, B. Booth, M. L'Hommedieu, T. Feng, B. Girault, J. L'Hommedieu, M. Wildman, S. Skaaden, A. Nadarajan, J. Villatte, T. Falk, K. Lerman, E. Ferrara, S. Narayanan
Scientific Data

Towards Automating Time Series Analysis for the Paleogeosciences
D. Khider, P. Athreya, V. Ratnakar, Y. Gil, F. Zhu, M. Kwan, J. Emile-Geay
MileTS '20: 6th KDD Workshop on Mining and Learning from Time Series

Towards Learning Representations of Binary Executable Files for Security Tasks
S. Arakelyan, C. Hauser, E. Kline, A. Galstyan
AAAI'20 Workshop on Artificial Intelligence for Cyber Security (AICS)

Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set
E. Chen, K. Lerman, E. Ferrara
JMIR Public Health and Surveillance

User-based Collaborative Filtering Mobile Health System
H. Kao, S. Yan, H. Hosseinmardi, S. Narayanan, K. Lerman, E. Ferrara
Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies

Using Dynamic Time Warping Self-organizing Maps to Characterize Diurnal Patterns in Environmental Exposures
K. Li, K. Sward, H. Deng, J. Morrison, R. Habre, M. Franklin, Y. Chiang, J. L. Ambite, J. Wilson, S. Eckel
Research Square

Using Task Descriptions in Lifelong Machine Learning for Improved Performance and Zero-shot Transfer
M. Rostami, D. Isele, E. Eaton
Journal of Artificial Intelligence Research

What Types of COVID-19 Conspiracies Are Populated by Twitter Bots?
E. Ferrara
First Monday

Multidisciplinary publications appear in all relevant areas.

Wisdom of the Expert Crowd Prediction of Response for 3 Neurology Randomized Trials
P. Atanasov, A. Diamantaras, A. MacPherson, E. Vinarov, D. Benjamin, I. Shrier, F. Paul, U. Dirnagl, J. Kimmelman
Neurology

Characterizing Social Media Manipulation in the 2020 U.S. Presidential Election
E. Ferrara, H. Chang, E. Chen, G. Muric, J. Patel
First Monday

Discovering Patterns of Online Popularity from Time Series
M. Ozer, A. Sapienza, A. Abeliuk, G. Muric, E. Ferrara
Expert Systems with Applications

Massive Cross-platform Simulations of Online Social Networks
G. Muric, A. Tregubov, J. Blythe, A. Abeliuk, D. Choudhary, K. Lerman, E. Ferrara
Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020)

MICROELECTRONICS AND ELECTRONICS SYSTEMS

3D Integrated Laser Attach Technology on 300-mm Monolithic Silicon Photonics Platform
Y. Bian et al.
2020 IEEE Photonics Conference (IPC)

45nm CMOS - Silicon Photonics Monolithic Technology (45CLO) for Next-generation, Low Power and High Speed Optical Interconnects
M. Rakowski, C. Meagher, K. Nummy, A. Aboketaf, J. Ayala, Y. Bian, B. Harris, K. Mclean, K. McStay, A. Sahin, L. Medina, B. Peng, Z. Sowinski, A. Stricker, T. Houghton, C. Hedges, K. Giewont, A. Jacob, T. Letavic, D. Riggs, A. Yu, J. Pellerin
Optical Fiber Communication Conference (OFC) 2020

A CMOS Compatible Monolithic Fiber Attach Solution with Reliable Performance and Self-alignment
B. Peng, T. Barwicz, A. Sahin, T. Houghton, B. Hedrick, Y. Bian, M. Rakowski, S. Hu, J. Ayala, C. Meagher, Z. Sowinski, K. Nummy, A. Stricker, J. Lubguban, H. Chen, B. Fasano, I. Melville, Z. Wu, J. Cho, A. Jacob, D. Riggs, D. Berger, T. Letavic, A. Yu, J. Pellerin, K. Giewont
Optical Fiber Communication Conference (OFC) 2020

A Neuromorphic SLAM Architecture Using Gated-memristive Synapses
A. Jones, A. Rush, C. Merkel, E. Herrmann, A. Jacob, C. Thiem, R. Jha
Neurocomputing

Functional Read Enabling In-memory Computations in 1Transistor—1Resistor Memory Arrays
A. Jaiswal, R. Andrawis, A. Agrawal, K. Roy
IEEE Transactions on Circuits and Systems II: Express Briefs

Hybrid Vanadate Silicon Nanophotonic Platform for Extreme Light Management at Telecom Bands
Y. Bian, A. Jacob, W. S. Lee, B. Peng, M. Rakowski, A. Aboketaf, R. Augur
IEEE 70th Electronic Components and Technology Conference (ECTC 2020)

i-SRAM: Interleaved Wordlines for Vector Boolean Operations Using SRAMs
A. Jaiswal, A. Agrawal, M. Ali, S. Sharmin, K. Roy
IEEE Transactions on Circuits and Systems I: Regular Papers

IMAC: In-memory Multi-bit Multiplication and ACcumulation in 6T SRAM Array
M. Ali, A. Jaiswal, S. Kodge, A. Agrawal, I. Chakraborty, K. Roy
IEEE Transactions on Circuits and Systems I: Regular Papers

In Situ Unsupervised Learning Using Stochastic Switching in Magneto-electric Magnetic Tunnel Junctions
I. Chakraborty, A. Agrawal, A. Jaiswal, G. Srinivasan, K. Roy
Philosophical Transactions of the Royal Society A

Light Manipulation in a Monolithic Silicon Photonics Platform Leveraging 3D Coupling and Decoupling
Y. Bian, A. Jacob, A. Thomas, B. Peng, M. Rakowski, W. S. Lee, R. Augur
Frontiers in Optics / Laser Science

Pathways to Efficient Neuromorphic Computing with Non-volatile Memory Technologies
I. Chakraborty, A. Jaiswal, A. Saha, S. Gupta, K. Roy
Applied Physics Reviews

Towards Low-loss Monolithic Silicon and Nitride Photonic Building Blocks in State-of-the-art 300mm CMOS Foundry
Y. Bian, J. Ayala, C. Meagher, B. Peng, M. Rakowski, A. Jacob, K. Nummy, A. Stricker, Z. Sowinski, A. Sahin, A. Aboketaf, S. Hu, I. Stobert, K. Mclean, L. Medina, K. Dezfulian, B. Harris, S. Krishnamurthy, T. Houghton, W. S. Lee, M. Sorbara, D. Riggs, T. Letavic, A. Yu, K. Giewont, J. Pellerin
Frontiers in Optics / Laser Science

Ultra-compact Waveguide Crossings Using Cosine Tapers on a Monolithic Silicon Photonics Platform
S. Chandran, Y. Bian, P. Moreira, A. Jacob, M. S. Dahlem, Y. Gebregiorgis, M. Papadovasilakis, S. Alkaabi, J. Viegas, M. Rakowski, A. Stricker, K. Nummy, C. Meagher, B. Peng, A. Thomas, S. Hu, J. Petyk
OSA Advanced Photonics Congress (AP) 2020 (IPR, NP, NOMA, Networks, PVLED, PSC, SPPCom, SOF)

NATURAL LANGUAGE PROCESSING

"Don't quote me on that": Finding Mixtures of Sources in News Articles
A. Spangher, N. Peng, J. May, E. Ferrara
Proceedings of the Computation + Journalism Symposium

Combining Conceptual and Referential Annotation to Study Variation in Framing
M. Postma, L. Remijnse, F. Ilievski, A. Fokkens, S. Titarsolej, P. Vossen
Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet

Connecting the Dots: A Knowledgeable Path Generator for Commonsense Question Answering
P. Wang, N. Peng, F. Ilievski, P. Szekely, X. Ren
arXiv preprint arXiv:2005.00691

Connecting the Dots: Event Graph Schema Induction with Path Language Modeling
M. Li, Q. Zeng, Y. Lin, K. Cho, H. Ji, J. May, N. Chambers, C. Voss
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP2020)

Cross-lingual Structure Transfer for Zero-resource Event Extraction
D. Lu, A. Subburathinam, H. Ji, J. May, S. Chang, A. Sil, C. Voss
Proceedings of the 12th Language Resources and Evaluation Conference

Enabling Low-resource Transfer Learning across COVID-19 Corpora by Combining Event-Extraction and Co-training
A. Spangher, N. Peng, J. May, E. Ferrara
Proceedings of the 1st Workshop on NLP for COVID-19 at ACL2020

Experience Grounds Language
Y. Bisk, A. Holtzman, J. Thomason, J. Andreas, Y. Bengio, J. Chai, M. Lapata, A. Lazaridou, J. May, A. Nisnevich, N. Pinto, J. Turian
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP2020)

Finding the Optimal Vocabulary Size for Neural Machine Translation
T. Gowda and J. May
Findings of the Association for Computational Linguistics (EMNLP2020)

Grounding Conversations with Improvised Dialogues
H. Cho and J. May
Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics

KGTK: A Toolkit for Large Knowledge Graph Manipulation and Analysis
F. Ilievski, D. Garijo, H. Chalupsky, N. Divvala, Y. Yao, C. Rogers, R. Li, J. Liu, A. Singh, D. Schwabe, P. Szekely
International Semantic Web Conference

Large-scale Cross-lingual Language Resources for Referencing and Framing
P. Vossen, F. Ilievski, M. Postma, A. Fokkens, G. Minnema, L. Remijnse
Proceedings of the 12th Language Resources and Evaluation Conference

Multidisciplinary publications appear in all relevant areas.

Learning to Generalize for Sequential Decision Making
X. Yin, R. Weischedel, J. May
Findings of the Association for Computational Linguistics (EMNLP2020)

Man is to Person as Woman is to Location: Measuring Gender Bias in Named Entity Recognition
N. Mehrabi, T. Gowda, F. Morstatter, N. Peng, A. Galstyan
Proceedings of the 31st ACM Conference on Hypertext and Social Media

Predictive Engagement: An Efficient Metric for Automatic Evaluation of Open-Domain Dialogue Systems
S. Ghazarian, R. Weischedel, A. Galstyan, N. Peng
Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI-20)

Teaching Machine Comprehension with Compositional Explanations
Q. Ye, X. Huang, E. Boschee, X. Ren
Findings of the Association for Computational Linguistics (EMNLP2020)

The Role of Knowledge in Determining Identity of Long-tail Entities
F. Ilievski, E. Hovy, P. Vossen, S. Schlobach, Q. Xie
Journal of Web Semantics

GAIA: A Fine-grained Multimedia Knowledge Extraction System
M. Li, A. Zareian, Y. Lin, X. Pan, S. Whitehead, B. Chen, B. Wu, H. Ji, S. Chang, C. Voss, D. Napierski, M. Freedman
Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics System Demonstrations

LEAN-LIFE: A Label-efficient Annotation Framework Towards Learning from Explanation
D. Lee, R. Khanna, B. Lin, S. Lee, Q. Ye, E. Boschee, L. Neves, X. Ren
Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics System Demonstrations

SEARCHER: Shared Embedding Architecture for Effective Retrieval
J. Barry, E. Boschee, M. Freedman, S. Miller
Proceedings of the Workshop on Cross Language Search and Summarization of Text and Speech (CLSSTS2020)

NETWORKING

Accelerating FIU's Science Research and Education Towards Discovery and Innovation by Leveraging FIU's Science DMZ
V. Chergarova, J. Ibarra, C. D'Cunha, M. Kirgan, R. Cabrera, D. Rotella, J. Bezerra, R. Rotella, H. Morgan
Proceedings of AMCIS 2020

Advances in Collaboration in R&E Networking between Africa, Latin America, and the U.S.
H.Morgan, L. Lotz, L. Cadenas, A. Hazin
UbuntuNet Connect 2020

Clouding Up the Internet: How Centralized is DNS Traffic Becoming?
G. Moura, S. Castro, W. Hardaker, M. Wullink, C. Hesselman
Proceedings of the ACM Internet Measurement Conference

Consistent and Secure Network Updates Made Practical
J. Lembke, S. Ravi, P-L. Roman, P. Eugster
Middleware '20: 21st International Middleware Conference

Improving Coverage of Internet Outage Detection in Sparse Blocks
G. Baltra, J. Heidemann
Proceedings of the Passive and Active Measurement Workshop

RoSCo: Robust Updates for Software-defined Networks
J. Lembke, S. Ravi, P. Eugster, S. Schmid
IEEE Journal on Selected Areas in Communication (JSAC)

Scalable, Serializable Networked Multi-actor Programming
B. Sang, P. Eugster, G. Petri, S. Ravi, P-L. Roman
Proceedings of the ACM on Programming Languages

Extended DNS Errors
W. Kumari, E. Hunt, R. Arends, W. Hardaker, D. Lawrence
Internet RFC 8914

PLASMA: Programmable Elasticity for Stateful Cloud Computing Applications
B. Sang, P-L. Roman, P. Eugster, H. Lu, S. Ravi, G. Petri
EuroSys '20: Fifteenth EuroSys Conference 2020

QUANTUM INFORMATION PROCESSING

Calculating the Divided Differences of the Exponential Function by Addition and Removal of Inputs
L. Gupta, L. Barash, I. Hen
Computer Physics Communications

De-Signing Hamiltonians for Quantum Adiabatic Optimization
E. Crosson, T. Albash, I. Hen, A. Young
Quantum

Discriminating Nonisomorphic Graphs with an Experimental Quantum Annealer
Z. Gonzalez Izquierdo, R. Zhou, K. Markstrom, I. Hen
Physical Review A

Dynamic Topology Reconfiguration of Boltzmann Machines on Quantum Annealers
J. Liu, K-T. Yao, F. Spedalieri
Entropy

Elucidating the Interplay between Non-stoquasticity and the Sign Problem
L. Gupta and I. Hen
Advanced Quantum Technologies

Hardness and Ease of Curing the Sign Problem for Two-local Qubit Hamiltonians
J. Klassen, M. Marvian, S. Piddock, M. Ioannou, I. Hen, B. Terhal
SIAM Journal on Computing

Noncommuting Conserved Charges in Quantum Many-body Thermalization
N. Yunger Halpern, M. Beverland, A. Kalev
Physical Review E

Optimal Fermion-to-qubit Mapping via Ternary Trees with Applications to Reduced Quantum States Learning
Z. Jiang, A. Kalev, W. Mruczkiewicz, H. Neven
Quantum

Permutation Matrix Representation Quantum Monte Carlo
L. Gupta, T. Albash, I. Hen
Journal of Statistical Mechanics: Theory and Experiment

SPACE SYSTEMS AND TECHNOLOGIES

Initial Safety Posture Investigations for Earth Regime Rendezvous and Proximity Operations
D. Barnhart, R. Rughani, J. Allam, K. Clarke
Journal of Space Safety Engineering

MAGNETO: Mapping the Earth's Magnetic Field at 300 km Using COTS Sensors
L. Villafana, J. Yuan, L. Broadus, W. Su, C. Carlton, A. Norrell, R. Rughani, E. Park, D. Barnhart
34th Annual Conference on Small Satellites

On-orbit Servicing Ontology Applied to Recommended Standards for Satellites in Earth Orbit
D. Barnhart and R. Rughani
Journal of Space Safety Engineering

Multidisciplinary publications appear in all relevant areas.

REACCH-reactive Electro-adhesive Capture Cloth Mechanism to Enable Safe Grapple of Cooperative/Non-cooperative Space Debris

S. Narayanan, D. Barnhart, R. Rogers, D. Ruffatto, E. Schaler, N. Van Crey, G. Dean, A. Bhanji, S. Bernstein, A. Singh et al.
AIAA Scitech 2020 Forum

Safe Construction in Space: Using Swarms of Small Satellites for In-space Manufacturing

R. Rughani and D. Barnhart
34th Annual Conference on Small Satellites

Using Genetic Algorithms for Safe Swarm Trajectory Optimization

R. Rughani and D. Barnhart
AIAA Scitech 2020 Forum

SCIENTIFIC COMPUTING

A Lightweight Method for Evaluating In Situ Workflow Efficiency

T. Do, L. Pottier, S. Caíno-Lores, R. Ferreira da Silva, M. Cuendet, H. Weinstein, T. Estrada, M. Taufer, E. Deelman
Journal of Computational Science

A Novel Metric to Evaluate In Situ Workflows

T. Do, L. Pottier, S. Thomas, R. Ferreira da Silva, M. Cuendet, H. Weinstein, T. Estrada, M. Taufer, E. Deelman
International Conference on Computational Science (ICCS)

An On-demand Weather Avoidance System for Small Aircraft Flight Path Routing

E. Lyons, D. Westbrook, A. Grote, G. Papadimitriou, K. Thareja, C. Wang, M. Zink, E. Deelman, A. Mandal, P. Ruth
Dynamic Data Driven Application Systems

An Open Ecosystem for Pervasive Use of Persistent Identifiers

R. Wagner, M. D'Arcy, K. Chard, R. Ananthakrishnan, I. Foster, J. Pruyne, B. McCollam, R. Schuler, C. Kesselman, P. Rocca-Serra
Practice and Experience in Advanced Research Computing (PEARC20)

Application Aware Software Defined Flows of Workflow Ensembles

G. Papadimitriou, E. Lyons, C. Wang, K. Thareja, R. Tanaka, P. Ruth, J. J. Villalobos, I. Rodero, E. Deelman, M. Zink, A. Mandal
2020 IEEE/ACM Innovating the Network for Data-Intensive Science (INDIS)

Characterizing, Modeling, and Accurately Simulating Power and Energy Consumption of I/O-intensive Scientific Workflows

R. Ferreira da Silva, H. Casanova, A. Orgerie, R. Tanaka, E. Deelman, F. Suter
Journal of Computational Science

Detecting Anomalous Packets in Network Transfers: Investigations Using PCA, Autoencoder and Isolation Forest in TCP

M. Kiran, C. Wang, G. Papadimitriou, A. Mandal, E. Deelman
Machine Learning

Developing Accurate and Scalable Simulators of Production Workflow Management Systems with WRENCH

H. Casanova, R. Ferreira da Silva, R. Tanaka, S. Pandey, G. Jethwani, W. Koch, S. Albrecht, J. Oeth, F. Suter
Future Generation Computer Systems

DyNamo: Scalable Weather Workflow Processing in the Academic MultiCloud

E. Lyons, M. Zink, A. Mandal, C. Wang, P. Ruth, C. Radhakrishnan, G. Papadimitriou, E. Deelman, K. Thareja, I. Rodero
100th American Meteorological Society Annual Meeting

End-to-end Online Performance Data Capture and Analysis for Scientific Workflows

G. Papadimitriou, C. Wang, K. Vahi, R. Ferreira da Silva, A. Mandal, Z. Liu, R. Mayani, M. Rynge, M. Kiran, V. Lynch, R. Kettimuthu, E. Deelman, J. Vetter, I. Foster
Future Generation Computer Systems

FaceBase 3 Analytical Tools and FAIR Resources for Craniofacial and Dental Research
B. Samuels, R. Aho, J. Brinkley, A. Bugacov, E. Feingold, S. Fisher, A. Gonzalez-Reiche, J. Hacia, B. Hallgrimsson, K. Hansen, M. Harris, T. Ho, G. Holmes, J. Hooper, E. Jabs, K. Jones, C. Kesselman, O. Klein, E. Leslie, H. Li, E. Liao, H. Long, N. Lu, R. Maas, M. Marazita, J. Mohammed, W. Prescott, R. Schuler, L. Selleri, R. Spritz, T. Swigut, H. van Bakel, A. Visel, I. Welsh, C. Williams, T. Williams, J. Wysocka, Y. Yuan, Y. Chai
Development

Gearing the DECam Analysis Pipeline for Multi-messenger Astronomy Using Pegasus Workflows
K. Vahi, D. Goldstein, G. Papadimitriou, P. Nugent, E. Deelman
Astronomical Data Analysis Software and Systems (ADASS) XXIX

Identifying Execution Anomalies for Data Intensive Workflows Using Lightweight ML Techniques
C. Wang, G. Papadimitriou, M. Kiran, A. Mandal, E. Deelman
2020 IEEE High Performance Extreme Computing Conference (HPEC)

Modeling the Performance of Scientific Workflow Executions on HPC Platforms with Burst Buffers
L. Pottier, R. Ferreira da Silva, H. Casanova, E. Deelman
2020 IEEE International Conference on Cluster Computing (CLUSTER)

Peachy Parallel Assignments
H. Casanova, R. Ferreira da Silva, A. Gonzalez-Escribano, W. Koch, Y. Torres, D. Bunde
2020 IEEE/ACM Workshop on Education for High-Performance Computing (EduHPC)

The Pegasus Workflow Management System: Translational Computer Science in Practice
E. Deelman, R. Ferreira da Silva, K. Vahi, M. Rynge, R. Mayani, R. Tanaka, W. Whitcup, M. Livny
Journal of Computational Science

The Workflow Trace Archive: Open-access Data from Public and Private Computing Infrastructures
L. Versluis, R. Mathá, S. Talluri, T. Hegeman, R. Prodan, E. Deelman, A. Iosup
IEEE Transactions on Parallel and Distributed Systems

Towards Co-evolution of Data-centric Ecosystems
R. Schuler, K. Czajkowski, M. D'Arcy, H. Tangmunarunkit, C. Kesselman
32nd International Conference on Scientific and Statistical Database Management (SSDBM '20)

Workflow Submit Nodes as a Service on Leadership Class Systems
G. Papadimitriou, K. Vahi, J. Kincl, V. Anantharaj, E. Deelman, J. Wells
Proceedings of the Practice and Experience in Advanced Research Computing

WorkflowHub: Community Framework for Enabling Scientific Workflow Research and Development
R. Ferreira da Silva, L. Pottier, T. Coleman, E. Deelman, H. Casanova
2020 IEEE/ACM Workflows in Support of Large-Scale Science (WORKS)

CHiSEL: A User-oriented Framework for Simplifying Database Evolution
R. Schuler and C. Kesselman
Distributed and Parallel Databases

FaceBase 3.0: A Comprehensive Resource on Craniofacial Development
R. Schuler, C. Kesselman, Y. Chai
Craniofacial Morphogenesis and Tissue Regeneration Gordon Research Conference

VISION, IMAGE, SPEECH, AND TEXT ANALYTICS

Affect Estimation with Wearable Sensors
S. Yan, H. Hosseinmardi, H. Kao, S. Narayanan, K. Lerman, E. Ferrara
Journal of Healthcare Informatics Research

Affective Conditioning on Hierarchical Attention Networks Applied to Depression Detection from Transcribed Clinical Interviews
D. Xezonaki, G. Paraskevopoulos, A. Potamianos, S. Narayanan
Proceedings of Interspeech

Multidisciplinary publications appear in all relevant areas.

An Empirical Analysis of Information Encoded in Disentangled Neural Speaker Representations
R. Peri, H. Li, K. Somandepalli, A. Jati, S. Narayanan
Proceedings of Odyssey

Automated Empathy Detection for Oncology Encounters
Z. Chen, J. Gibson, M. Chiu, Q. Hu, T. Knight, D. Meeker, J. Tulsky, K. Pollak, S. Narayanan
Proceedings of International Conference on Healthcare Informatics (ICHI 2020)

Automatic Prediction of Suicidal Risk in Military Couples Using Multimodal Interaction Cues from Couples Conversations
S. Chakravarthula, M. Nasir, S. Tseng, H. Li, T. Park, B. Baucom, C. Bryan, S. Narayanan, P. Georgiou
Proceedings of IEEE International Conference on Audio, Speech and Signal Processing (ICASSP)

Bringing in the Outliers: A Sparse Subspace Clustering Approach to Learn a Dictionary of Mouse Vocalizations
J. Wang, K. Mundnich, A. Knoll, P. Levitt, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Clinical State Tracking in Serious Mental Illness through Computational Analysis of Speech
A. Arevian, D. Bone, N. Malandrakis, V. Martinez, K. Wells, D. Miklowitz, S. Narayanan
PLoS ONE

Conic Affinity Measures to Design Speech Enhancement Systems Operating in Unseen Noise Conditions
P. Papadopoulos and S. Narayanan
Proceedings of Interspeech

Conversational Correlates of Rapid Social Judgments of Children and Adolescents With and Without ASD
A. Shield, X. Wang, D. Bone, S. Narayanan, R. Grossman
Clinical Linguistics & Phonetics

Deblurring for Spiral Real-time MRI Using Convolutional Neural Networks
Y. Lim, S. Narayanan, K. Nayak
Proceedings of Medical Imaging with Deep Learning (MIDL 2020)

Deblurring for Spiral Real-time MRI Using Convolutional Neural Networks
Y. Lim, Y. Bliesener, S. Narayanan, K. Nayak
Magnetic Resonance in Medicine

Dynamical Systems Modeling of Day-to-day Signal-based Patterns of Emotional Self-regulation and Stress Spillover in Highly-demanding Health Professions
G. Hadjiantonis, P. Paromita, K. Mundnich, A. Nadarajan, B. Booth, S. Narayanan, T. Chaspari
Proceedings of 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2020)

Extending the Beta Divergence to Complex Values
C. Vaz and S. Narayanan
Pattern Recognition Letters

Fifty Shades of Green: Towards a Robust Measure of Inter-annotator Agreement for Continuous Signals
B. Booth and S. Narayanan
Proceedings of ICMI

Graph Vertex Sampling with Arbitrary Graph Signal Hilbert Spaces
B. Girault, A. Ortega, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Group-specific Models of Healthcare Workers' Well-being Using Iterative Participant Clustering
V. Ravuri, P. Paromita, K. Mundnich, A. Nadarajan, B. Booth, S. Narayanan, S. Shrikanth, T. Chaspari
2020 Second International Conference on Transdisciplinary AI (TransAI)

How an Aglossic Speaker Produces an Alveolar-like Percept Without a Functional Tongue Tip
A. Toutios, D. Byrd, L. Goldstein, S. Narayanan
Journal of the Acoustic Society of America

Identifying Truthful Language in Child Interviews
V. Ardulov, Z. Durante, S. Williams, T. Lyon, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Joint Estimation and Analysis of Risk Behavior Ratings in Movie Scripts
V. Martinez, K. Somandepalli, Y. Tehranian-Uhls, S. Narayanan
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)

Joint Multi-dimensional Model for Global and Time-series Annotations
A. Ramakrishna, R. Gupta, S. Narayanan
IEEE Transactions on Affective Computing

Learning Behavioral Representations from Wearable Sensors
N. Tavabi, H. Hosseinmardi, J. Villatte, A. Abeliuk, S. Narayanan, E. Ferrara, K. Lerman
Proceedings of International Conference on Social Computing, Behavioral Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS 2020)

Learning Domain Invariant Representations for Child-adult Classification from Speech
R. Lahiri, M. Kumar, S. Bishop, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Leveraging Linguistic Context in Dyadic Interactions to Improve Automatic Speech Recognition for Children
M. Kumar, S. Kim, C. Lord, T. Lyon, S. Narayanan
Computer, Speech and Language

Linguistically Aided Speaker Diarization Using Speaker Role Information
N. Flemotomos, P. Georgiou, S. Narayanan
Proceedings of Odyssey

Meta-learning for Robust Child-adult Classification from Speech
N. Rao, M. Kumar, S. Kim, C. Lord, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Modeling Behavior as Mutual Dependency Between Physiological Signals and Indoor Location in Large-scale Wearable Sensor Study
T. Feng, B. Booth, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Modeling Behavioral Consistency in Large-scale Wearable Recordings of Human Bio-behavioral Signals
T. Feng and S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Modeling Vocal Entrainment in Conversational Speech Using Deep Unsupervised Learning
M. Nasir, B. Baucom, C. Bryan, S. Narayanan, P. Georgiou
IEEE Transactions on Affective Computing

Modeling Movement Behavior Among Nursing Profession
T. Feng and S. Narayanan
Proceedings of 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)

Multitask Learning for DARPA Lorelei's Situation Frame Extraction Task
K. Singla and S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Neural Speech Decoding During Audition, Imagination and Production
R. Sharon, S. Narayanan, M. Sur, H. Murthy
IEEE Access

Revealing True Identity: Detecting Makeup Attacks in Face-based Biometric Systems
M. Arab, P. Moghadam, M. Hussein, W. AbdAlmageed, M. Hefeeda
ACM Multimedia

Multidisciplinary publications appear in all relevant areas.

Robust Speaker Recognition Using Unsupervised Adversarial Invariance
R. Peri, M. Pal, A. Jati, K. Somandepalli, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Screenplay Quality Assessment: Can We Predict Who Gets Nominated?
M. Chiu, T. Feng, X. Ren, S. Narayanan
Proceedings of the First Joint Workshop on Narrative Understanding, Storylines, and Events

Sentence Level Estimation of Psycholinguistic Norms Using Joint Multidimensional Annotations
A. Ramakrishna and S. Narayanan
Proceedings of Interspeech

Speaker Diarization for Naturalistic Child-adult Conversational Interactions Using Contextual Information
M. Kumar, S. Kim, C. Lord, S. Narayanan
Journal of the Acoustical Society of America

Speaker Diarization Using Latent Space Clustering in Generative Adversarial Network
M. Pal, M. Kumar, R. Peri, T. Park, S. Kim, C. Lord, S. Bishop, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Speaker-invariant Affective Representation Learning via Adversarial Training
H. Li, M. Tu, J. Huang, S. Narayanan, P. Georgiou
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

The INTERSPEECH 2020 Far-field Speaker Verification Challenge
X. Qin, M. Li, H. Bu, W. Rao, R. Das, S. Narayanan, H. Li
Proceedings of Interspeech

The Role of Annotation Fusion Methods in the Study of Human-Reported Emotion Experience During Music Listening
T. Greer, K. Mundnich, M. Sachs, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

TILES-2018: A Longitudinal Physiologic and Behavioral Data Set of Hospital Workers
K. Mundnich, B. Booth, M. L'Hommedieu, T. Feng, B. Girault, J. L'Hommedieu, M. Wildman, S. Skaaden, A. Nadarajan, J. Villatte, T. Falk, K. Lerman, E. Ferrara, S. Narayanan
Scientific Data (Nature Research)

Towards End-2-end Learning for Predicting Behavior Codes from Spoken Utterances in Psychotherapy Conversations
K. Singla, Z. Chen, D. Atkins, S. Narayanan
Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics

Trapezoidal Segment Sequencing: A Novel Approach for Fusion of Human-produced Continuous Annotations
B. Booth and S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Variability in Individual Constriction Contributions to Third Formant Values in American English
S. Harper, L. Goldstein, S. Narayanan
Journal of the Acoustical Society of America

Vocal Tract Articulatory Contour Detection in Real Time Magnetic Resonance Images Using Spatio Temporal Context
A. Hebbar, R. Sharma, K. Somandepalli, A. Toutios, S. Narayanan
Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Vocal Tract Shaping of Emotional Speech
J. Kim, A. Toutios, S. Lee, S. Narayanan
Computer, Speech and Language

Bill Manning, an early, long-time operator of the B-Root Domain-Name-Server, passed away on January 25, 2020.

Manning was a researcher at ISI from 1994–2015 while also working with the WIDE Project in Japan, which played an important role in the formation of the Internet. At ISI, he worked on the Los Nettos project, a regional network project in the early days of the Internet, as well as the Routing Arbiter Project, the INT. domain, the RS.NET testbed, and others.

Manning was integral to B-Root, one of the 13 root servers in the world that are essential to supporting the infrastructure and traffic control of the Internet, which ISI has operated since 1987. As an operator and program manager, Manning helped steer the evolution of the root server, and was passionate about autonomy in thought and speech on the Internet.

Bill Manning, pictured here with a friend's daughter. Photo courtesy of Rod Van Meter

"Bill fought fiercely for the importance and independence of the root server system and we all owe him a great deal of thanks for his work on the root," said Wes Hardaker, senior computer scientist in the Networking and Cybersecurity Division, who worked alongside Manning on B-Root.

With his expertise in Internet infrastructure management, DNS development and operation, and Internet instrumentation, Manning served on various technical advisory boards and was a valuable member of numerous organizations, including IEEE, ACM, USENIX, the Asia Pacific Internet Association (APIA), and the Internet Service Providers' Consortium (ISP/C).

"Bill fought for the fundamental principles of everyone on the Internet in ways that few people will appreciate," Hardaker said. "We owe him a great deal of thanks."


USC
VITERBI
SCHOOL OF
ENGINEERING
INFORMATION
SCIENCES
INSTITUTE

Information Sciences Institute

4676 Admiralty Way #1001
Marina del Rey, CA 90292
isi.edu | media@isi.edu

USC Viterbi
School of Engineering