APPROACHING
ISI's
50th
ANNIVERSARY

LOOKING BEYOND AND INTO THE FUTURE

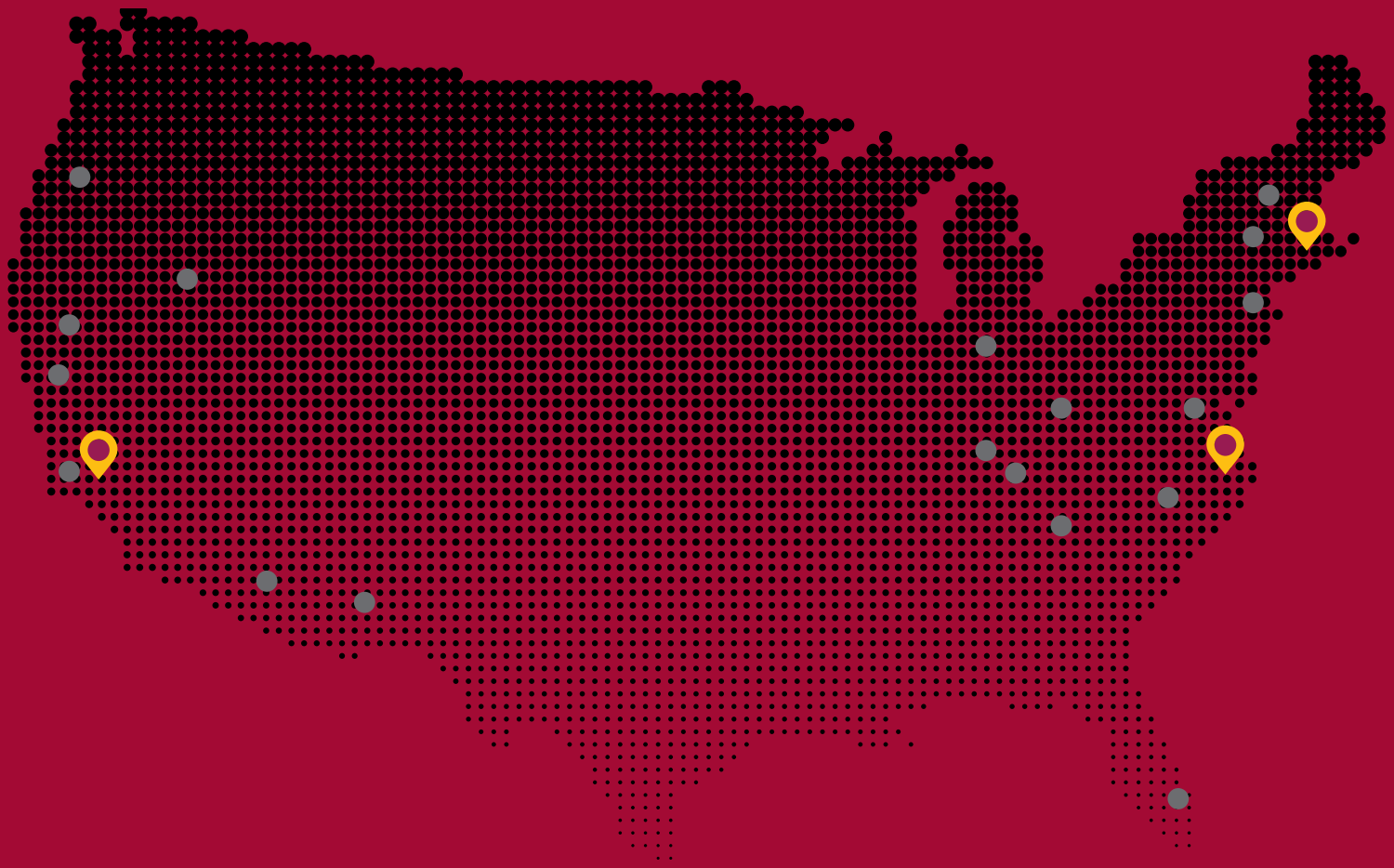
Information Sciences Institute

2021 ANNUAL REPORT

USC Viterbi
School of Engineering

INFORMATION SCIENCES INSTITUTE

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



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

MASSACHUSETTS
890 Winter Street #115
Waltham, MA 02451
781.622.9790

VIRGINIA
3811 Fairfax Drive #200
Arlington, VA 22203
703.812.3700

REMOTE LOCATIONS

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

Information Sciences Institute

USC Viterbi
School of Engineering

WELCOME TO ISI

This last year, ISI researchers tackled deepfake video detection, cybersecurity issues within the Bitcoin protocol, the spread of COVID-19, the spread of COVID-19 misinformation, nanosatellites, quantum annealing, and more.

2021 was a record year for ISI:

- ISI faculty advised 90+ PhD students.
- ISI researchers submitted 177 proposals in fiscal year 2021. Of those, five proposals were for projects over $10M; 10 were for projects from $5M to $10M.
- A record number of exploratory research proposals were submitted for funding from our endowment from philanthropist Michael Keston and ISI funds. Eight projects were selected; you can read about them in this report.
- I don't tally the number of honors, prizes, patents, and publications every year, but I suspect that 2021 was a record year in these areas, too.

ISI's Diversity, Equity, and Inclusion (DEI) Committee's work continued throughout the year, with three new subcommittees that focus on minority and women outreach and PhD recruiting, new hiring practices, and ensuring a supportive environment.

We are thoughtfully and analytically transforming ISI into a truly accommodating hybrid work environment. The inside cover of this report shows where we worked in 2021. We're now returning to ISI offices to work, but in a hybrid model shaped by a detailed survey of our staff. Each group is determining how they work best: remotely, at the office, or a combination of both. Not only are we redesigning the three ISI locations for collaboration and office sharing, but we're also providing furnishings and equipment for work-at-home offices.

In the coming year, ISI will celebrate its 50th anniversary. This institute began in 1972 when Keith Uncapher, who led the computer research effort at RAND, left with several researchers who "said they would be interested if [he] could set up a non-profit, or university-based center, R&D center." Later in the same 1989 interview, Uncapher said, "I guess the assumption was on my part, and I would guess on theirs, we would get together and do something useful."

ISI has, indeed, done something useful. You'll get an idea of the work that ISI has done, and continues to do, as you go through this report.

I hope you will help celebrate our 50-year anniversary by attending the Symposium on the Future of Computing Research on September 12 and 13. Expert speakers will examine the future of computing research, as a field and as a profession. You are welcome to attend—and participate!

It's been a challenging but rewarding year for ISI. Despite COVID-19 and our varied workplaces, ISI retained its vitality and sense of community. We thank the agencies, grantors, and donors who make our work—delivering the future—possible.

Craig A. Knoblock
Keston Executive Director
USC Information Sciences Institute

MARINA DEL REY, CALIFORNIA

ISI occupies nine floors of the Marina Towers South building, west of Los Angeles. The Santa Monica Mountains and Catalina Island are within sight, plus sunsets over the Pacific Ocean. Hotels, restaurants, shops, and apartments are within walking distance, in addition to the marina and a park. About 325 researchers, staff, graduate research assistants, faculty, and students are based here, working in all areas of computer science and engineering. The institute's administrative departments are also located here. The Space Engineering Research Center and the USC-Lockheed Martin Quantum Computing Center are both next door. Free shuttle service runs between ISI, the USC Institute for Creative Technologies, and the USC University Park campus. ISI is a 20-minute ride from Los Angeles International Airport; it is accessible by public transit and the 20-mile coastal bike trail.

ARLINGTON, VIRGINIA

ISI Arlington is located in Northern Virginia, near Washington, D.C. Three Arlington researchers hold appointments as research professors in the USC Ming Hsieh Department of Electrical and Computing Engineering. ISI's Secure and Robust Electronics Center and Application Specific Intelligent Computing Lab are both located here. 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 universities around the world to work at ISI Arlington, and from local and regional universities.

BOSTON, MASSACHUSETTS

Located in the Greater Boston area, the Waltham location is ISI's newest campus. Major research areas at ISI Boston include artificial intelligence, deep learning, multimedia and natural language understanding, and quantum information science. Researchers lead many projects in collaboration with universities and industry. ISI Boston also partners with local universities and universities around the country to host interns and co-op students. ISI's Laboratory for Quantum-Limited Information, which is dedicated to understanding and demonstrating the fundamental physical limits for extracting information from physical signals, is located here.

ISI FIRSTS

2022 — ISI 50th Anniversary

2020

- MOSIS Alliance with Intel Corporation | 2019
- ISI Boston | 2017
- Social Media Bot Detection | 2015
- Human Trafficking Search Tool | 2015
- Sports Data Analytics | 2014
- USC-Lockheed Martin Quantum Computing Center | 2012
- Nano-Satellite | 2010

2010

- 2009 | SoCal Seismic Hazards Map
- 2008 | Reconfigurable Robots
- 2003 | DETERLab Cybersecurity Testbed
- 2000 | Practical Rhetorical Theory
- 2000 | Internet2 Collaboration

2000

- USC Institute for Creative Technologies Spinoff | 1999
- ICANN Spinoff | 1998
- ISI Arlington | 1997
- FASTXchange Spinoff | 1997
- Online Payments | 1993
- Internet Society | 1992

1990

- 1988 | Rhetorical Structure Theory
- 1985 | Packet Video
- 1983 | Digital Phone Communications
- 1983 | Internet Domain Name System
- 1981 | MOSIS Semiconductor Implementation Service

1980

- Packet Radio Terminal Concept | 1979
- ARPANET Addressing and Request for Comments | 1977
- ARPANET Voice over IP | 1974
- ARPANET Interface | 1972
- ISI Founded by Keith Uncapher | 1972

1970

SYMPOSIUM ON THE

FUTURE OF COMPUTING RESEARCH

SEPTEMBER 12-13 2022 | IN-PERSON AND ONLINE

SPONSORED BY

USC Viterbi School of Engineering

TO REGISTER SIGN UP AT https://fcr-2022.net/register

USC/ISI INVITES YOU TO A SYMPOSIUM ON THE FUTURE OF COMPUTING RESEARCH

COMPUTING RESEARCH TODAY is increasingly shaped by forces and trends that separate it from the simpler circumstances of the field's founding era. Intellectual and technical maturing of the discipline, the ever-increasing centrality of computing to modern society, and changing expectations about career paths, collaborations, research impact, and work environments all contribute.

Our goal is to examine critical questions that will drive and define the field and profession of computing research in the future, and to contribute meaningfully to the ongoing conversations shaping this rapidly developing evolution.

WHO SHOULD ATTEND

- EXPERIENCED COMPUTER RESEARCH PROFESSIONALS
- EARLY CAREER COMPUTING RESEARCHERS
- PEER RESEARCHERS
- INDUSTRY TECHNOLOGISTS
- POLICYMAKERS
- RESEARCH POLICY PROFESSIONALS
- SCIENCE AND TECHNOLOGY JOURNALISTS

ISI IN NUMBERS

400+
STAFF, FACULTY, AND STUDENTS

2
POSTDOCS

108
PHD STUDENTS

FACULTY

33

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

STAFF AND FACULTY

7
MBA degrees

28
master's degrees

63
bachelor's degrees

92
PhD degrees

2021 NEW RESEARCH GRANTS

56

2021 RESEARCH EXPENDITURES

$71.43
million

QUANTUM COMPUTERS

1

SEMINARS HOSTED

76

TOTAL OFFICE SPACE

154,712
square feet

SUMMER INTERNS

24

ISI LOCATIONS


Marina del Rey, CA


Arlington, VA


Waltham, MA


United States

SCIENTIFIC AND TECHNICAL PORTFOLIO

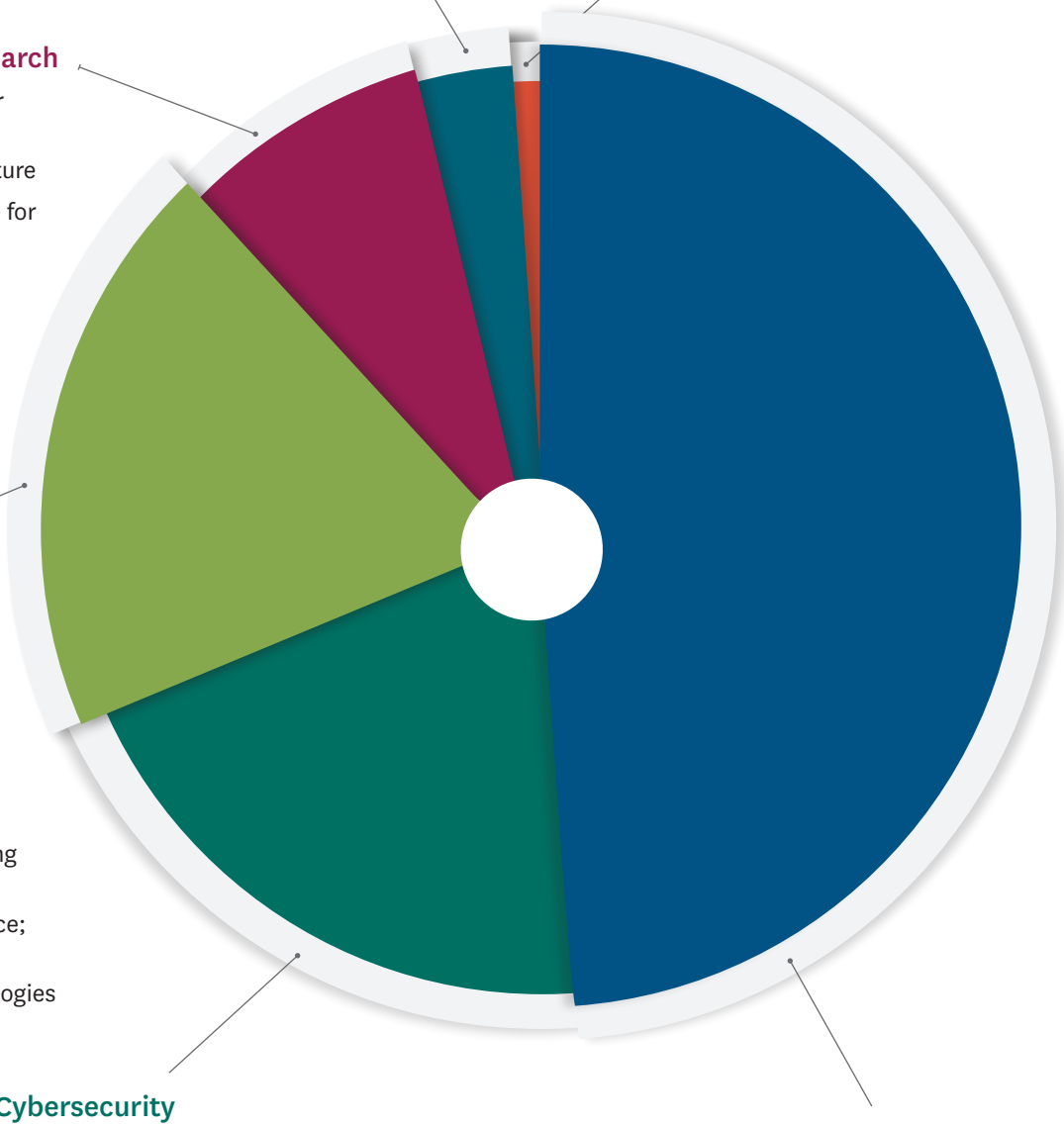
MOSIS Service

- Integrated circuit fabrication brokerage
- Multi-project integrated circuit wafers
- State-of-the-art semiconductor process technologies
- Integrated circuit design automation

Space Engineering Research

Informatics Systems Research

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



High Performance, Edge, and Quantum Computing

- Secure, robust, and trusted electronics
- Heterogeneous and cloud computing
- High performance computing architectures and software
- Quantum information science; quantum computing
- Science automation technologies

Computer Networks and Cybersecurity

- Network and security measurement, analysis, and defense
- Network infrastructure supporting science and applications
- Research, methods, and infrastructure for cyber experimentation
- Social engineering attacks
- Binary program analysis, vulnerability discovery, and reverse engineering
- Modeling human behavior for cybersecurity and social simulation
- Theory and practice of distributed computing

Artificial Intelligence

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

Information Sciences Institute

EXECUTIVE LEADERSHIP

Craig Knoblock, PhD
Keston Executive Director

Stephen Crago, PhD
Associate Director

James Whalen, MPhil
Associate Director
Chief Financial Officer

Eileen Lu, BS
Chief Information Officer
Director, Computing and Information Services

Yigal Arens, PhD
Senior Director for Administrative Affairs

John Wroclawski, MSEE
Senior Director for Strategic Initiatives

Elizabeth Boschee, BS
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

Elizabeth Boschee, BS
Associate Director, Artificial Intelligence Division

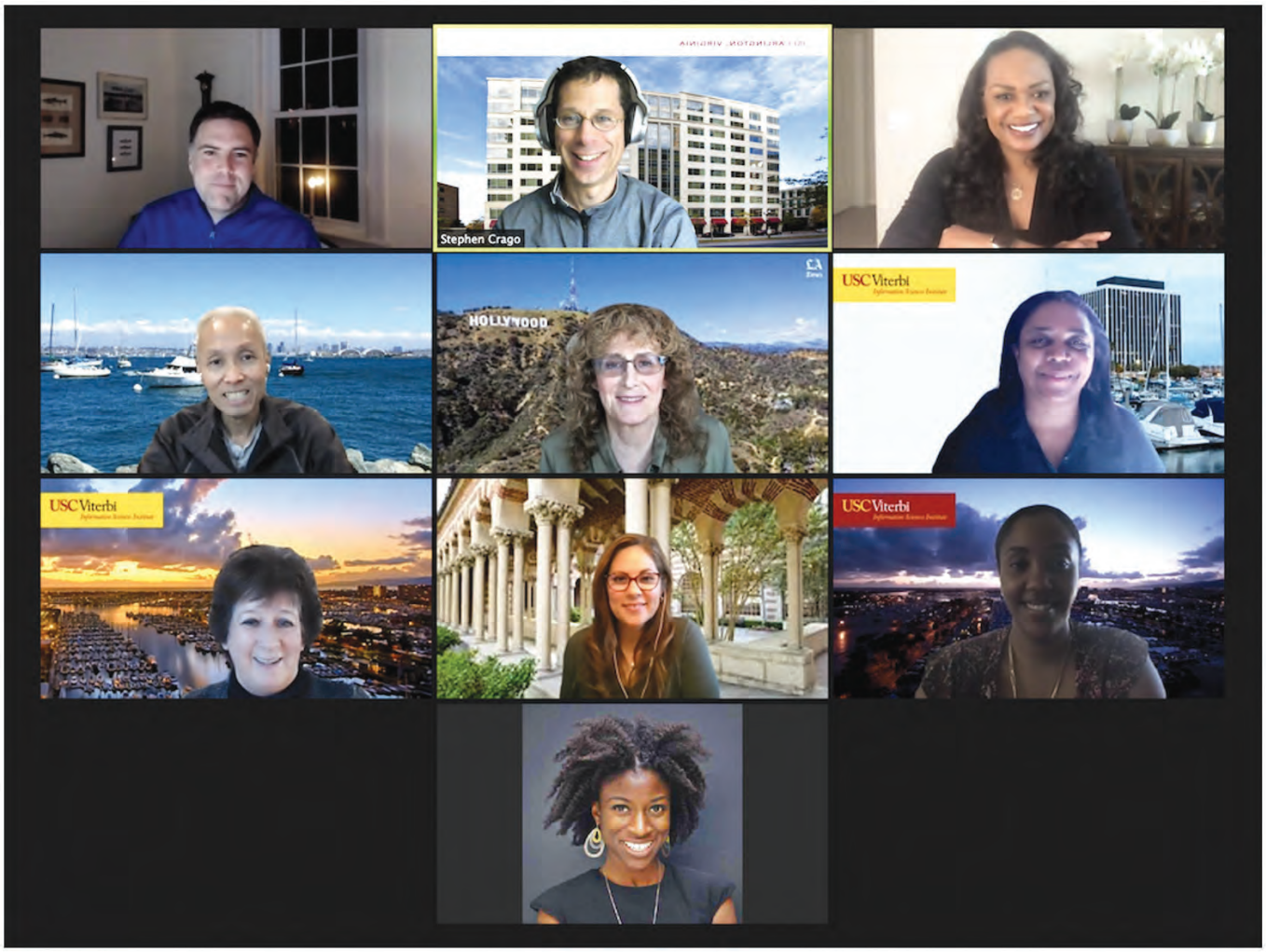
Yigal Arens, PhD
Director, Emerging Activities Group

The values of diversity, equity, and inclusion are fundamental to an organization's success and wellbeing. Diversity broadens perspective; equity fosters opportunity; inclusion builds community. Whether as a collective or as individuals, we can effect meaningful change by our actions, attitudes, and behaviors toward one another.

— Terry Benzel, director of the Networking and Cybersecurity Division, chairs the institute's DEI Committee

DEI Committee

Diverse, Equitable, and Inclusive



ARTIFICIAL INTELLIGENCE DIVISION

PEDRO SZEKLEY DIVISION DIRECTOR | **ELIZABETH BOSCHEE** ASSOCIATE DIVISION DIRECTOR

The Artificial Intelligence Division is one of the world's largest AI research groups, with over 150 researchers, research staff, and graduate students. Most researchers hold graduate degrees in computer science or related disciplines; many serve as research faculty in the Viterbi School of Engineering. The division conducts fundamental and applied research in these areas:

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

Natural language processing - low resource machine translation, multilingual representation learning, transfer learning, dialogue, decision-making, question answering, summarization, ontologies, information retrieval, extracting critical information about world events from foreign language sources

Knowledge graphs - using AI and machine learning techniques to construct and exploit large-scale knowledge bases and to induce taxonomies from data

Scientific data analysis and discovery - interactive knowledge capture, intelligent user interfaces, semantic workflows, provenance, large-scale data integration, analysis of biomedical data

Multi-modal understanding - image and video understanding for deepfake detection, visual misinformation identification and multimedia analysis, face recognition, biometric anti-spoofing, learning "common sense" knowledge from text, video, images, and audio, learning to automatically discover the structure of complex events from multilingual multimedia data

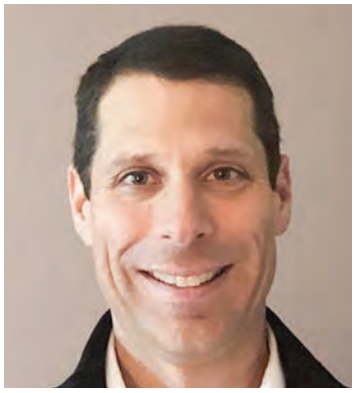
Common sense representation and reasoning - building cognitively inspired computational paradigms for evaluating common sense AI, human-centric dialog agents, game-theoretic simulators

Computational social science - structure detection and pattern matching in unusual complex systems with hidden information, large-scale, contextualized social media analysis, detecting and responding to social engineering attacks

AI fairness - detecting and mitigating bias, robustness against adversarial attacks, forecasting, crowdsourcing, identifying cultural values, polarization, and misinformation

COMPUTATIONAL SYSTEMS & 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.

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, circuit design, and fabrication.

INFORMATICS SYSTEMS RESEARCH DIVISION

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.

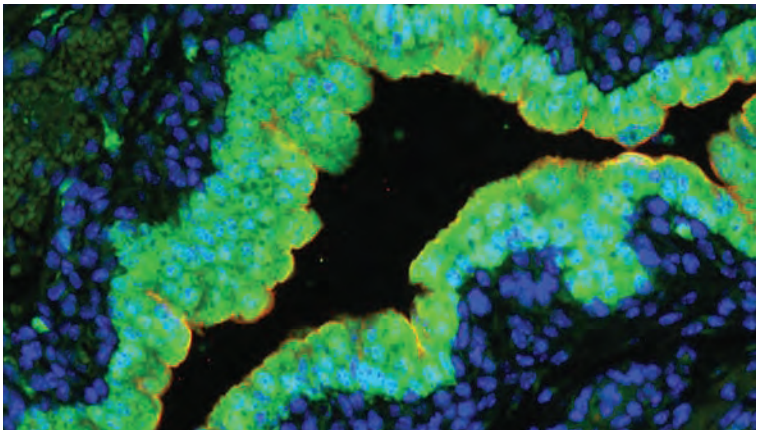


Photo credit: Immunohistochemistry image of mouse male bladder. Courtesy of Chad Vezina, GUDMAP Consortium.

NETWORKING AND CYBERSECURITY DIVISION

TERRY BENZEL DIVISION DIRECTOR

Networking is the backbone of the interconnected world, and cybersecurity is its guardian. The 50+ 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.

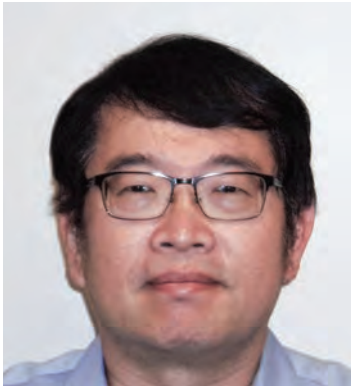
Theory and practice of distributed computing - today all computation and communication are fundamentally distributed, involving multiple participants and their interactions; division researchers address the challenges of enforcing safety, security, and robustness in these systems.

Social engineering attacks- 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.

THE MOSIS SERVICE

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 Foundry Services, and Samsung Foundry—to offer a range of semiconductor processes from 12nm FinFET to 350nm.

MOSIS supports the design enablement, design support, and fabrication sign-off for Intel Foundry Services' 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 Foundry Services.

The MOSIS Service ramped up to support Samsung 28nm projects in 2021. MOSIS is working with Samsung Foundry to promote projects with MOSIS in the Samsung Advanced Foundry Ecosystem (SAFE) Forum and other activities.

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 Foundry Services

MOSIS is also working with government and other organizations to establish silicon fabrication project platforms using MOSIS' design enablement and tapeout sign-off capabilities and services for academia and commercial design companies.

THE MOSIS SERVICE

The MOSIS Service carries out research in silicon fabrication and Design for Manufacturability (DFM) related areas, producing academic papers and patents. In addition to collaborating with other semiconductor design and manufacturing organizations, MOSIS staff pursue independent research. In 2021, MOSIS published a technical paper about MPW reticle yield technology in the SPIE Advanced Lithography Conference. MOSIS also published technical papers in industry forums, including a secure cloud-based design environment paper in the Synopsys Users Group (SNUG) World 2021 Conference and a reliability and manufacturability technology paper in the Samsung Advanced Foundry Ecosystem (SAFE) Forum.

CENTER ON KNOWLEDGE GRAPHS

JAY PUJARA DIRECTOR

The Center on Knowledge Graphs research group creates new approaches for amplifying artificial intelligence using structured knowledge. The group combines expertise in artificial intelligence, machine learning, the Semantic Web, natural language processing, databases, information retrieval, geospatial analysis, business, social sciences, and data science. The center comprises 16 senior ISI researchers, guiding the work of 17 PhD students, 12 MS students, and six researcher programmers.

The center has built tools and knowledge graphs to address challenging, real-world problems such as enabling common sense reasoning, reducing global food insecurity, fighting human trafficking, assessing medical and clinical data, fostering pharmacological discovery, ensuring scientific reproducibility, understanding supply chains, analyzing competition in business, integrating cultural heritage data, creating more engaging dialogue agents, and identifying social and moral norms in cultures. Supporting this array of projects are state-of-the-art tools for entity resolution and entity linking, automated semantic modeling, probabilistic reasoning, text generation, knowledge retrieval, table understanding, knowledge curation, error detection, event extraction, and visualization and exploration of knowledge.

The center's work includes the development of these systems:

- The Knowledge Graph Toolkit (KGTK) for creating and manipulating knowledge graphs
- KGTK-based browsers and dashboards for multi-relational graphs and time series data
- Table-to-Wikidata Mapping Language (T2WML) for efficient knowledge curation
- Table understanding tools for automated learning of table structure and content
- Record Linkage Toolkit and Table Linker for entity resolution and linking across datasets
- Graph-based probabilistic automatic semantic modeling
- Semi-supervised error detection in structured datasets
- The Karma system for semantic integration of diverse sources of data

For more information, software downloads, and tutorials, visit the Center on Knowledge Graphs website at https://usc-isi-i2.github.io/home/.

CYBERDEFENSE TECHNOLOGY EXPERIMENTAL RESEARCH LABORATORY

TERRY BENZEL DIRECTOR

ISI's Cyberdefense Technology Experimental Research Laboratory (DETERLab) is a scientific computing facility for cybersecurity researchers and academics to conduct cybersecurity experimentation and educational exercises. The DETERLab shared testbeds provide a platform for research in cybersecurity—serving academia, industry, and government. The lab is publicly available and Internet accessible, with integrated tools and methods for hands-on cybersecurity exploration.

In its almost two decades of operation, DETERLab has benefited more than 9,000 researchers and users from its targeted research community. DETERLab's educational materials and services have been used by over 17,000 learners and instructors from institutions worldwide. Of note, a third of the institutions using DETERLab are smaller-sized universities or community colleges that lack dedicated research infrastructure. DETERLab has enabled research and investigation that has resulted in hundreds of publications and numerous PhD and MS theses.

DETERLab facilities include:

- Local area network infrastructure elements and options
- Wide area networking between locations
- The ability to integrate one or more third-party network and computing facilities into experiments on-demand
- Nodes that can be configured with several operating system and application software stacks, virtual machine (VM) monitors, virtual VMs, network emulation elements, and network simulators
- A gatekeeper that protects the Internet-facing side of the testbeds and controls access to the user's private network
- Access to the node console by testbed users

Since 2004, DETERLab has operated continuously. Researchers, educators, and students access DETERLab remotely from hundreds of institutions worldwide:

- 9,000+ researchers conducting cybersecurity research on DETERLab
- 17,000+ students and teachers participating in hundreds of hands-on networking and cybersecurity classes using DETERLab educational exercises
- 48+ countries using DETERLab from a variety of institutions spanning the globe
- 300+ scientific publications based on DETERLab experimentation, including PhD and MS theses

Visit deterlab.net.

LABORATORY FOR QUANTUM-LIMITED INFORMATION

JONATHAN HABIF DIRECTOR

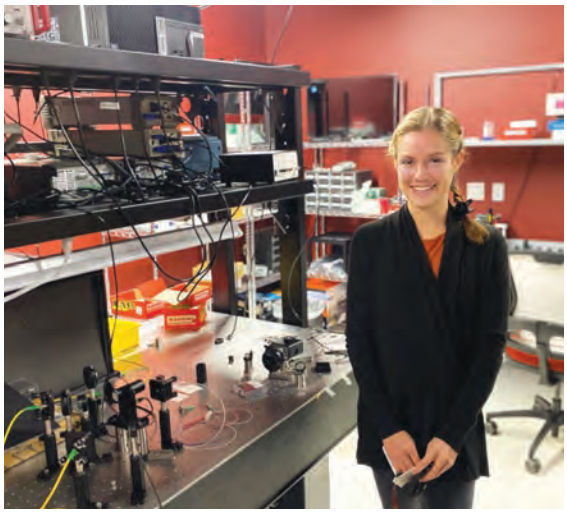
ISI's Laboratory for Quantum-Limited Information (QLIlab), located at ISI Boston, is dedicated to understanding and demonstrating the fundamental physical limits for measuring the weakest signals in the universe. The research carried out in the QLIlab is at the intersection of quantum physics, optics, and information theory, composing the field of quantum information science.

A complete understanding of a physical signal (such as electromagnetic energy like light or RF) is possible only when quantum mechanics is employed to mathematically describe the signal. This quantum mechanical description offers insight into the maximum amount of information that could be extracted from the physical signal. Research in the QLIlab examines sensing and communications tasks that are vital to current technology, such as laser ranging and secure communications. The research performs a completely quantum theoretical analysis of the preparation, transmission, and measurement of light. A key goal is to design, build, and demonstrate measurement tools that can detect the delicate quantum properties of these light fields. QLIlab researchers employ ultra-sensitive measurement tools to detect single photons (light particles) from wavelengths in the ultra-violet spectrum into the infrared spectrum.

Lab experiments also investigate the light emitted by warm bodies in the long-wave infrared spectrum—visible only by using thermal imaging technology. Researchers investigate how to harness this underutilized part of the spectrum for new types of communications protocols.

The QLIlab serves as a multi-disciplinary training ground for students interested in topics at the intersection of physics, mathematics, and engineering. A vibrant internship program gives undergrad and graduate students the opportunity to explore the quantum properties of light in a laboratory setting and learn how these unique properties could radically enhance current technologies and enable new capabilities that are impossible with a classical interpretation of the world.

QLIlab intern Phoebe Amory (Wellesley College) explaining her experiment to an aspiring scientist



Olivia Brasher (USC '23) in the QLIlab

SECURE AND ROBUST ELECTRONICS CENTER

MATTHEW FRENCH DIRECTOR

The Secure and Robust Electronics Center (SURE) strives to be a US center of excellence for secure, robust, hardware R&D to help protect national security, advance industry capabilities and competitiveness, and ensure that citizens can rely on the integrity of their own devices. SURE performs applied research to make chips trustworthy, secure, resilient, and reliable.

SURE investigates state-of-the-art manufacturing complexities that have significantly compromised the critical aspects of chip production. As experienced industry collaborators, SURE researchers also aim to accelerate the scale, pace, and impact of hardware robustness and security technology development. In addition to these technology developments, SURE seeks to meet an urgent cybersecurity need: electronics hardware security. This hardware security is the foundation on which software security rests. The more powerful integrated circuits become, the more opportunities materialize for their integrity to be compromised.

Government, industry, and higher education partners and supporters that SURE researchers have closely worked with (past and present) include DARPA, IARPA, NASA, the Army Research Office, the Air Force Research Laboratory, Defense Threat Reduction Agency, Xilinx Inc., Altera Corporation, Synopsys, Cadence, Stanford, Virginia Tech, Brigham Young University, Arizona State, University of North Carolina, and Georgia Tech Research Institute.

SPACE ENGINEERING RESEARCH CENTER

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 act as "doctors" working on real satellites and schedule-driven space systems for sponsors; students act as "residents" who work alongside. SERC teaches space systems research and offers hands-on training to build, test, and fly actual spacecraft, space systems, and satellites.

SERC hosts high school, community college, undergrad, graduate, and PhD students from the US and worldwide. Four cadets from the French Air Force Academy again worked and studied at SERC in 2021.

These activities in 2021 illustrate the breadth of research the center undertakes in space science and technology.

The STARFISH project, a unique soft crawling robotic device for spacecraft and satellite surfaces for on-orbit anomaly detection, achieved successful results and has been selected for additional funding to investigate more advanced capability. It uses technology from a previous biologically inspired SERC project, REACCH, a multi-tentacle gripper—modeled after an octopus—that uses electro-adhesion.

Through the CA Space Grant at UC San Diego and with UC Berkeley, SERC organized a NASA-sponsored university competition in lunar lander software simulation. As a result of the competition, the SERC team was awarded funding to use the SERC'S LEAPFROG lunar lander prototype flight vehicle and run a national competition with multiple LEAPFROG vehicles flying at radio-controlled flying aircraft fields across the US in 2022.

The SERC CLING-ERS project was selected to receive a free flight onboard the International Space Station (ISS) in conjunction with JPL and University of Nevada, Las Vegas collaborators. This is the first USC experiment to be tested aboard the ISS, and it will utilize the Astrobee flight platforms interacting with astronauts.

2021 continued to see growth in projects and activities at SERC. Additional advanced research includes:

- SERC will complete and launch USC's third satellite program in partnership with Lockheed Martin (La Jument). The satellite is scheduled to launch in 2022, and SERC will be the lead organization. It is the first optical payload in USC history.
- SERC was selected for Phase II Small Business Technology Transfer (STTR) funding with an industry partner in advanced solid rocket research in support of the Missile Defense Agency.
- SERC teamed with the Viterbi K-12 STEM Center to support a large multi-year proposal for K-12, college, and teacher training in space and aerospace disciplines.
- SERC will upgrade the USC ground station control center with command-and-control capability for the planned USC satellite launch in 2022. The modifications to the radio frequency (RF) front end and control systems will enable the USC Satellite Ground Station to be fully operational for USC and other satellite missions in the future.

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

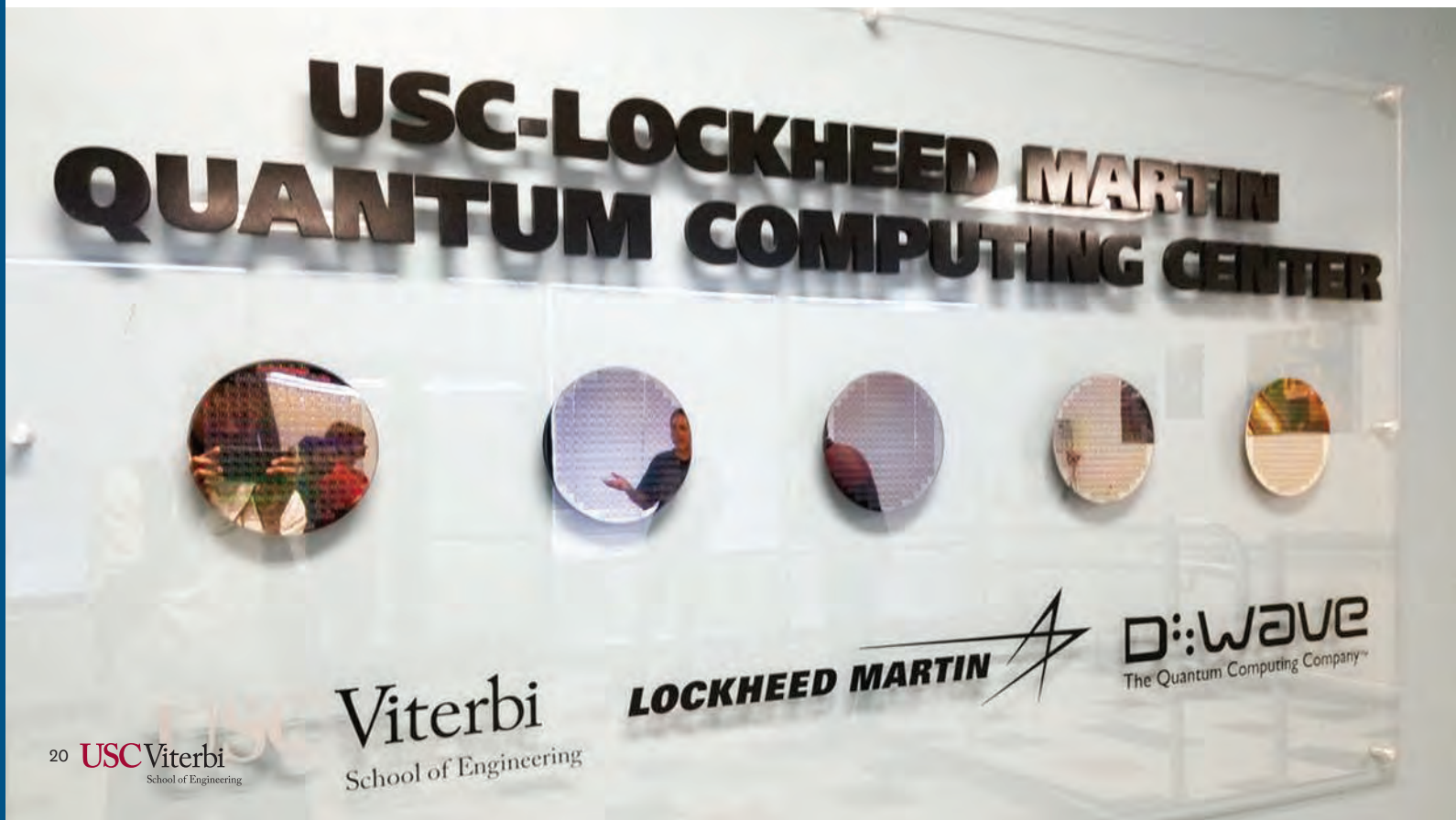
DANIEL LIDAR SCIENTIFIC AND TECHNICAL DIRECTOR
FEDERICO SPEDALIERI OPERATIONAL DIRECTOR
ROBERT LUCAS DIRECTOR OF RESEARCH PARTNERSHIPS

The USC-Lockheed Martin Quantum Computing Center (QCC) has housed a D-Wave quantum annealing system since it was established in 2011. QCC was the first organization outside the manufacturer D-Wave Systems 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 and how to use them for real-world applications.

USC and ISI researchers, faculty, and students perform basic and applied research to NISQ computing devices and collaborate with researchers worldwide. Five USC doctoral students and three postdoctoral associates are currently engaged in research at the QCC.

The QCC renewed its long-term relationship with D-Wave and Lockheed Martin in March 2020, leading to important upgrades at the facility. QCC is in the process of upgrading to D-Wave's Advantage™ system with more than 5000 qubits and increased connectivity. The upgrades will enable QCC to host a new Advantage generation of quantum annealers from D-Wave and provide USC researchers with access to D-Wave's Leap™ quantum cloud services.



VISUAL INTELLIGENCE AND MULTIMEDIA ANALYTICS LABORATORY

WAEL ABDALMAGEED FOUNDING DIRECTOR
MOHAMED HUSSEIN CO-DIRECTOR
LEONIDAS SPINOULAS CO-DIRECTOR
LIDA DIMITROPOULOU CO-DIRECTOR

In the Visual Intelligence and Multimedia Analytics Laboratory (VIMAL), researchers in all three ISI locations work on visual misinformation identification, face recognition, biometrics, optical character recognition, and health care applications of artificial intelligence. VIMAL's research addresses challenging problems in security, surveillance, and multimedia content analysis, including improving the security of biometric systems, making deep learning algorithms more robust against adversarial attacks and identifying deepfake videos and manipulated images. Key research areas include:

Multimedia forensics and visual misinformation identification: Forensic image analysis involves automatically assessing the truthfulness and integrity of visual information, including images and videos. VIMAL researchers work on characterizing signs of manipulated images and video (such as deepfakes) by detecting important indicators, including pixel-level attributes, the physics of the scene, and the semantics and genealogy of the image or video asset. Challenges include the variety of image-capturing devices, the increasing sophistication of manipulation tools and techniques (including rapid advances in computer graphics technology such as Photoshop), the emergence of generative adversarial networks (GANs), and the tremendous volume of analyzable data.

VIMAL
VISUAL INTELLIGENCE AND MULTIMEDIA ANALYTICS LABORATORY

Biometrics: In computer security, biometrics relate to automatic authentication of identities using measurable human characteristics, such as fingerprint recognition, face recognition, and iris recognition. Biometric authentication is used as a form of identification and access control at many facilities, commercial and government. VIMAL research aims to create systems and algorithms resilient to presentation attacks (spoofing) of biometric authentication systems, including face, iris, and fingerprint features.

Multimedia document analysis: VIMAL researchers are developing systems to extract and index information found in large collections of scientific and technical literature. Challenges include extracting information from figures and captions, detecting key aspects of an article such as assumptions and main findings, and relating the statements in separate but related articles. The goal is to create tools that facilitate access to knowledge in diverse fields of science.

VITERBI DATA SCIENCE PROGRAM

Several years ago, ISI researchers, including ISI director Craig Knoblock, worked with the Computer Science Department leadership to create and shape the Data Science Program in the USC Viterbi School of Engineering. Today this program is thriving with undergraduate and graduate offerings and is one of the largest programs in the school.

The Data Science Program now offers these degrees and programs:

- Bachelor of Arts in Data Science
- Master of Science in Environmental Data Science
- Master of Science in Applied Data Science
- Master of Science in Healthcare Data Science
- Master of Science in Cyber Security Engineering
- Master of Science in Public Policy Data Science
- Master of Science in Communication Data Science
- Master of Science in Spatial Data Science
- Undergraduate minor in Foundations of Data Science
- Graduate Certificate in Data Science Foundations
- Graduate Certificate in Applied Data Science
- Progressive degree programs

Some of these degrees are joint with other USC schools and are designed to introduce non-computer science majors to careers in data science.

LEADERSHIP

ISI's **Yolanda Gil** serves as Director of the Data Science Program; **Emilio Ferrara** is Associate Director. **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 teach in the Viterbi Data Science Program. They have designed introductory courses as well as advanced courses in scalable data systems, machine learning, knowledge graphs, and AI and data ethics.

Jeremy Abramson, *Lecturer*

Jose Luis Ambite, *Research Associate Professor*

Yigal Arens, *Research Professor*

Jim Blythe, *Lecturer*

Keith Burghardt, *Lecturer*

Emilio Ferrara, *Associate Professor*

Yolanda Gil, *Research Professor*

Ulf Hermjakob, *Lecturer*

Filip Ilevski, *Lecturer*

Carl Kesselman, *Professor*

Deborah Khider, *Lecturer*

Kristina Lerman, *Research Professor*

Jelena Mirkovic, *Research Associate Professor*

Fred Morstatter, *Research Assistant Professor*

Goran Muric, *Lecturer*

Cliff Neuman, *Associate Professor*

Jay Pujara, *Research Assistant Professor*

Mohammad Rostami, *Research Assistant Professor*

Pedro Szekely, *Research Associate Professor*

Satish Kumar Thittamaranahalli, *Research Assistant Professor*

USC + AMAZON CENTER ON SECURE AND TRUSTED MACHINE LEARNING

ISI researchers are significant participants in the new USC + Amazon Center on Secure and Trusted Machine Learning (Trusted AI), announced on January 28, 2021. The joint research center focuses on the development of new approaches to machine learning (ML) privacy, security, and trustworthiness. Trusted AI is housed at the USC Viterbi School of Engineering; it supports USC and Amazon researchers in the development of novel approaches to privacy-preserving ML solutions.

RESEARCH PROJECTS SELECTED FOR FUNDING

Each year, the center supports several research projects by USC faculty that focus on new methodologies for secure and privacy-preserving machine learning solutions. The research projects are selected in a competitive process. Of the five winning research projects selected in July 2021, four were proposed by ISI researchers jointly with USC colleagues:

Xiang Ren, Madhi Soltanolkotabi (Electrical and Computer Engineering)
Federated Learning for Natural Language Processing

Shri Narayanan
Federated Learning for Human-centered Experience and Perception Modeling with Biobehavioral Data

Keith Burghardt
Fast Fair Decentralized Learning

Jose Luis Ambite, Muhammad Naveed (Computer Science), **Paul Thompson** (Keck School of Medicine)
Efficient Federated Learning in Heterogeneous and Corrupted Environments by Secure Performance Weighting

AMAZON ML FELLOWSHIPS

The center also provides annual fellowships to doctoral students working in this research area in recognition of their promise and achievements. One of the two fellowship recipients named an Amazon ML Fellow is ISI PhD student **Ninareh Mehrabi.**

ADVISORY BOARD

The 2021-2022 Advisory Board consists of three Amazon scientists and three USC professors: **Kristina Lerman**, **Shri Narayanan**, and **Cyrus Shahabi** (Computer Science). Together, they play a crucial role in overseeing research efforts, participating in center activities, and supporting the center director and liaison.

For more information about the center, visit trustedai.usc.edu.

NEW DIRECTORS, RESEARCH TEAM LEADERS, AND RESEARCH LEADS

DIRECTORS

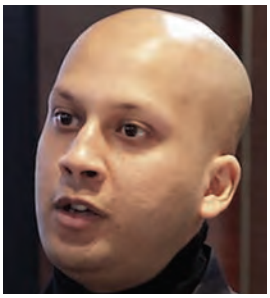
Wael AbdAlmageed
Research Director
Artificial Intelligence Division

Yolanda Gil
Viterbi Director of New Initiatives in AI and Data Science
Artificial Intelligence Division

Michael Haney
Director Advanced Photonics
Computational Systems and Technology Division

Jay Pujara
Director Center on Knowledge Graphs
Artificial Intelligence Division

RESEARCH TEAM LEADERS



Mayank Kejriwal
Research Team Leader
Artificial Intelligence Division

Erik Kline
Research Team Leader
Networking and Cybersecurity Division

Jonathan May
Research Team Leader
Artificial Intelligence Division

Fred Morstatter
Research Team Leader
Artificial Intelligence Division

RESEARCH LEADS

Christophe Hauser
Research Lead
Networking and Cybersecurity Division

Alefiya Hussain
Research Lead
Networking and Cybersecurity Division

Deborah Khider
Research Lead
Artificial Intelligence Division

Srivatsan Ravi
Research Lead
Networking and Cybersecurity Division

Rob Schuler
Research Lead
Informatics Systems Research Division

NEW FACULTY APPOINTMENT AND PROMOTIONS

Aram Galstyan
Research Professor
of Computer Science

Muhao Chen
Research Assistant Professor
of Computer Science

Xuezhe (Max) Ma
Research Assistant Professor
of Computer Science

Mohammad Rostami
Research Assistant Professor
of Computer Science

NEW ISI TEAM MEMBERS

In 2021, 22 computer scientists, programmers, project managers, and other staff members joined ISI. The institute sponsors mentoring talks, a proposal coaching seminar, and a "buddy" system to help integrate the ISIers who joined this year when most ISI colleagues are working remotely.

EXECUTIVE OFFICE

Magali Gruet, *Public Communications Specialist*

Jana-Lynn Louis, *Senior Project Manager*

Michael Pazzani, *Senior Supervising Computer Scientist*

NETWORKING AND CYBERSECURITY DIVISION

Jacob Lichtefeld, *Research Programmer*

Ganesh Sankaran, *Computer Scientist*

Paul Schmitt, *Research Computer Scientist*

Christopher Tran, *Research Programmer*

Quoc Tran, *Computer Scientist*

ARTIFICIAL INTELLIGENCE DIVISION

Min-Hsueh Chiu, *Programmer Analyst*

Emmanuel Johnson, *Postdoctoral Scholar*

COMPUTATIONAL SYSTEMS AND TECHNOLOGY DIVISION

Emre Akaturk, *Postdoctoral Scholar*

Kellie Canida, *Research Programmer*

Sujith Chandran, *Research Engineer*

Michael Haney, *Senior Supervising Computer Scientist*

Clynn Mathew, *VLSI Research Engineer*

Evgeny (Jenia) Mozgunov, *Research Scientist*

Benedict Reynwar, *Research Programmer*

Nicole Welch, *Project Manager*

ADMINISTRATIVE SERVICES

Jessica Coronel, *Financial Analyst*

COMPUTING AND INFORMATION SERVICES

Syed Hussain, *Network Engineer*

Dominik Staros, *Lead Web Developer*

EMERGING ACTIVITIES GROUP

Stephen Hunt, *Programmer Analyst*

STAFF PROMOTIONS

Congratulations to these ISI people who earned promotions in 2021.

ADMINISTRATIVE SERVICES

Winsome Chee, *Post Award Manager*

Eloisa Romo, *Accounting/Financial Director*

ARTIFICIAL INTELLIGENCE DIVISION

Manuel Ciosici, *Computer Scientist*

Goran Muric, *Computer Scientist*

COMPUTATIONAL SYSTEMS AND TECHNOLOGY DIVISION

Travis Haroldsen, *Research Computer Scientist*

Ryan Tanaka, *Research Programmer*

INFORMATICS SYSTEMS RESEARCH DIVISION

Alejandro Bugacov, *Computer Scientist*

Karl Czajkowski, *Senior Computer Scientist*

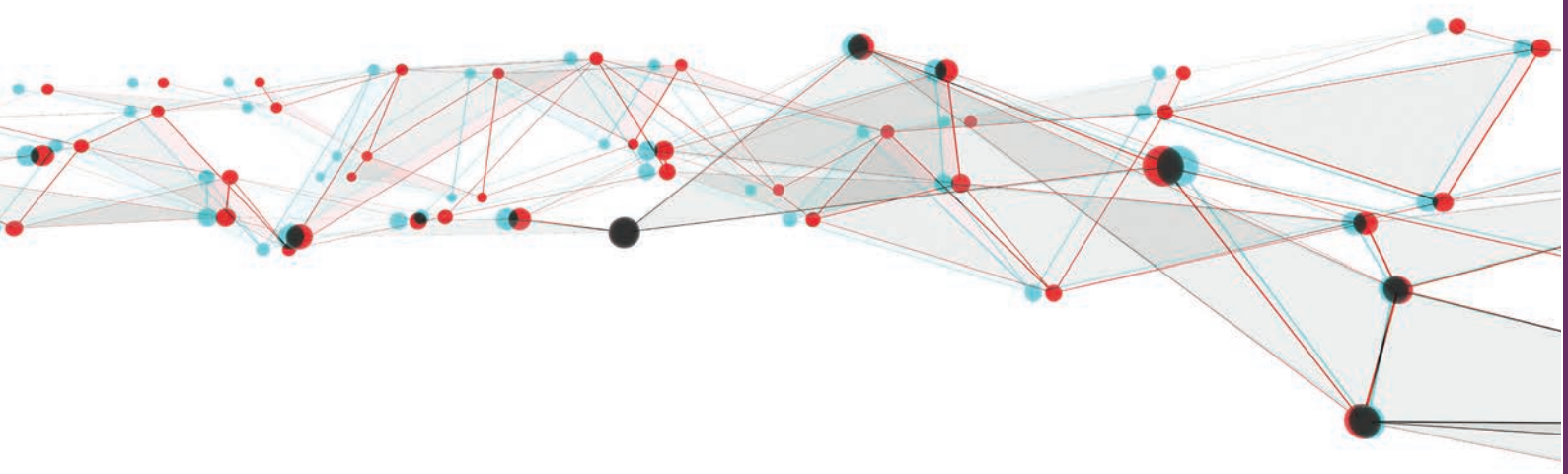
Hongsuda Tangmunarunkit, *Senior Computer Scientist*

NETWORKING AND CYBERSECURITY DIVISION

Joseph Cecil, *Research Programmer*

Joseph Cummings, *Research Programmer*

Christophe Hauser, *Research Computer Scientist*



KESTON 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 the 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. The Keston and ISI Exploratory Research Awards support research projects that are intellectually intriguing and significant to society, with the potential to produce a tangible result within about a year.

A committee of ISI researchers judged the proposals. Craig Knoblock presented eight exploratory research awards in January 2021. Overviews of the research follow.

Effective Interventions of Misinformation in Online Social Networks

Daniel Benjamin, Fred Morstatter, Abel Salinas, Priya Mane

Misinformation has led to many recent harms, including mistrust of institutions, disregard for public health guidelines, decreased vaccination rates, and civil unrest, among others. The fractured media landscape allows individuals to choose confirming information over credible information. Social sampling theory describes that our misperceptions of others are explained by the sample of people we encounter, and we are more likely to link to similar people online. ISI's NVISION team is striving to develop interventions to mitigate the spread of misinformation by visualizing social networks of popular content to help make these biases explicit. The team developed and tested three network visualizations, finding that a force-directed graph magnifies users' sense of polarization. The next experiment will test how individuals respond to news content when it is accompanied by this visualization.

SARFire: Rapid Wildfire Detection through Synthetic Aperture Radar

Andrew Rittenbach, John Paul Walters

The damage resulting from wildfires continues to increase. Ground-based detection methods and current satellite-based approaches both have limitations. This project is developing a deep learning-based onboard Synthetic Aperture Radar (SAR) imaging technique to provide near-constant overhead fire surveillance eventually. SAR is a remote sensing imaging technique used to form images from radar signals; this means that high-resolution imagery can be collected even when an area is covered by smoke, clouds, or at night. As part of the work, the research team is developing a prototype lightweight, end-to-end, onboard SAR processor that can perform both SAR image formation and wildfire detection onboard the satellite. The team is employing a novel deep learning-based approach to SAR image formation to take full advantage of hardware-based tensor accelerators that are expected to be available on next-generation SAR satellites. Looking ahead, the team is planning to apply this image-formation approach to other SAR applications.

ISI EXPLORATORY RESEARCH AWARDS

FairPRS: Fairly Predicting Genetic Risks for Personalized Medicine

Jose Luis Ambite, Greg ver Steeg, Keith Burghardt, Kristina Lerman, David Conti (USC Keck School of Medicine)

The increasing availability of genetic data promises better understanding and early diagnosis for many diseases. Polygenic risk scores (PRS) aggregate genetic information from genome-wide association studies to provide an estimate of the risk of a disease, but they perform poorly across populations. The FairPRS project is developing polygenic risk scores that apply across populations, including admixed populations, which have ancestry from multiple groups. To this end, the team is investigating methods from invariant representation learning and fairness, including information-theoretic regularization and orthogonal projection methods. The first application to be studied is prostate cancer.

Identifying Populations Susceptible to Anti-Science

Keith Burghardt, Goran Muric

Vaccine hesitancy has a long history and is driven recently by the anti-vaccine narratives shared online, degrading the efficacy of vaccination strategies such as those for COVID-19. Despite broad agreement in the medical community about the safety and efficacy of available vaccines, many social media users continue to be inundated with false information about vaccines. This study aimed to understand anti-vaccine sentiment better and reduce its impact by developing a system capable of identifying the users responsible for spreading anti-vaccine narratives. The team introduced a publicly available Python software package to analyze user tweets on Twitter and assess the likelihood of a user spreading anti-vaccine tweets. The software package was built using text embedding methods, neural networks, and automated dataset generation. It was trained on 100,000+ accounts and several million tweets. The model helps researchers and policymakers understand anti-vaccine discussion and misinformation to tailor targeted campaigns to debunk harmful anti-vaccination myths. Additionally, the researchers leveraged the data about such users to understand the moral and emotional characteristics of anti-vaccine spreaders.

AI2AI: Discovering and Assessing Vulnerability to AI-Generated Twin Identities

Mohamed Hussein, Wael AbdAlmageed

Modern AI methods can create pictures that match natural images, including photo-realistic face images. What if some of these AI-generated faces are virtually "identical twins" for real individuals? Such fake twins can, intentionally or unintentionally, cause harm. Orthogonally, face recognition models can be fooled into misidentification when specific malicious and imperceptible perturbations are applied to a person's image. This introduces another type of AI-generated twins, adversarial twins, which are easier to generate than fake twins and are harmful by construction. Although much research is dedicated to generating better-looking fake and adversarial twins, there has been little focus on discovering and assessing the vulnerability of individuals and demographic groups to the threats that fake and adversarial twins pose. Using multiple machine learning models, the AI2AI project (AI Investigating AI) is developing a service that assesses the vulnerability of individuals and demographic groups to malicious AI-generated twins to help communities and law enforcement agencies discover and assess that vulnerability. The front-end and back-end components of the system are nearly complete. The team will next integrate the whole system and collect data from the public.

Bio-PICS: Bio-Optical Point-of-Care Intelligent COVID-19 Sensor

Ajey Jacob, Sujith Chandran, Akhilesh Jaiswal, Neha Nanda (USC Keck School of Medicine)

This project is developing a bio-photonic breath analyzer using silicon photonics technology to detect COVID-19 directly. The sensing platform offers self-calibrated device response by eliminating the error due to changes in ambient temperature or fabrication uncertainties. This enables selective detection of ultra-low concentrations of pathogens (detection down to a single virus). The proposed sensing method can be successfully configured for environmental monitoring, detection of various viruses/pathogens, and cancer biomarkers. The CMOS compatibility of the sensor platform facilitates mass-manufacturability and hence a cost-effective point-of-care photonic diagnostic solution.

ISI EXPLORATORY RESEARCH AWARDS (continued)

3D Facial Muscle Screening Tool for Disease Diagnosis

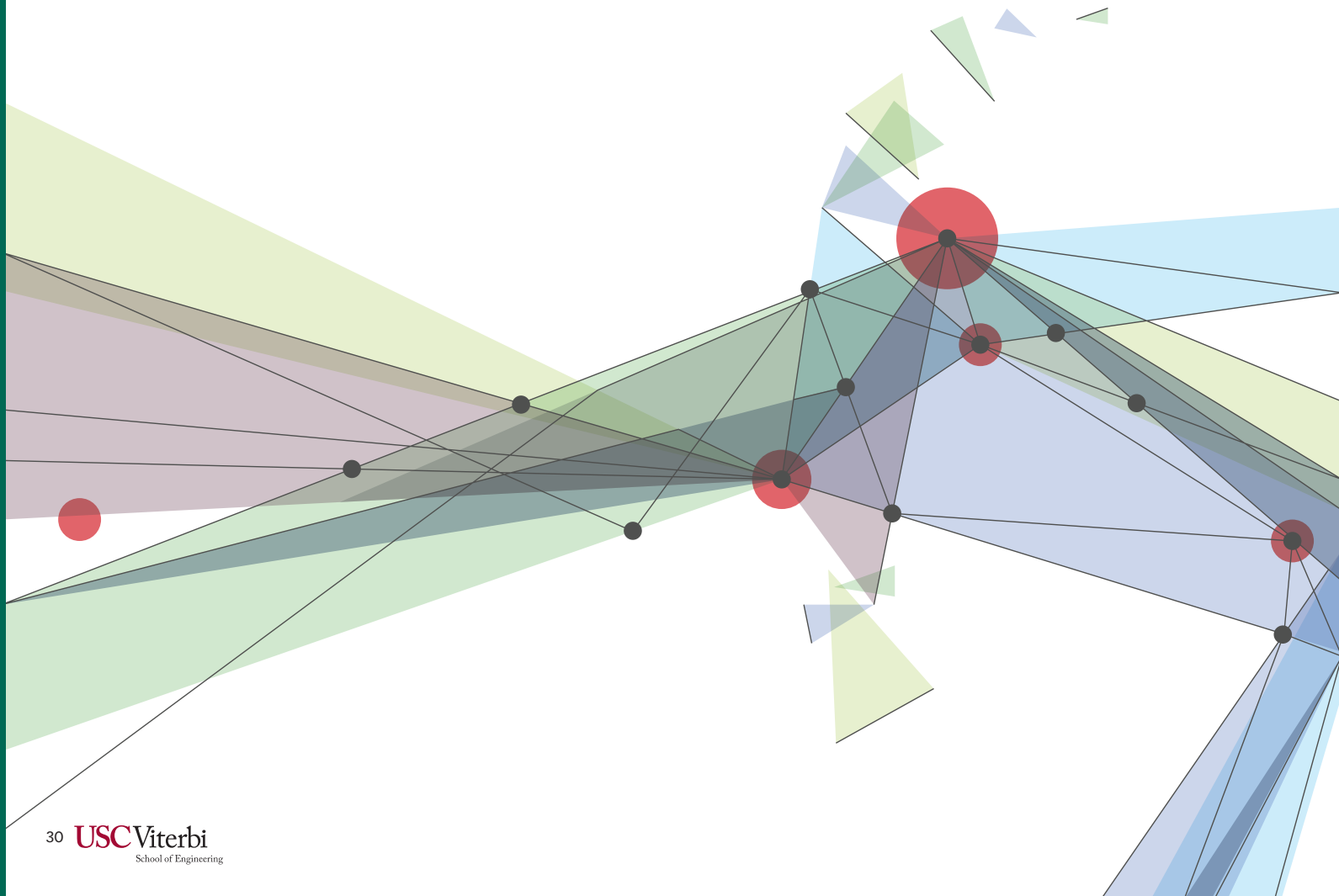
Wael AbdAlmageed

This project initially proposed to develop a facial expression screening tool for the diagnosis of Parkinson's disease. The work is now being applied to congenital central hypoventilation syndrome, a disorder that affects breathing. The project is developing a 3D facial expression screening tool over a fine-grained 3D mesh of the face to derive more information than 2D tools. The project leverages the mobile 3D camera unit developed with support from an earlier Keston Exploratory Research Award. A 3D mesh per frame of the captured video is computed from videos of subjects by fusing RGB images and depth data. The reconstructed 3D meshes are analyzed and quantified to study and evaluate spasticity and rigidity of facial muscles.

Learning Fair AI Models Across Distinct Domains

Mohammad Rostami, Aram Galstyan, Serban Stan

As societies increasingly rely on AI for automatic decision-making, concerns about bias and fairness in AI are growing. Fairness in AI is not merely an ethical issue; bias undermines efficiency and productivity in the labor market. A common approach for studying fairness is investigating whether model decisions are related to sensitive attributes such as gender or race. Most current works simply focus on debiasing AI models for a single domain. However, a fair model trained for one domain may be used in other domains, but there is no guarantee that it will generalize fairly to other domains. The team addressed this challenge within a domain adaptation formulation. They adapted a pre-trained fair model to generalize well in a target domain using solely unlabeled target domain datasets. Instead of starting training again with a new unbiased dataset, the team used the knowledge gained during the original debiasing to preserve the fairness of the model in the new domain using unannotated target domain data.



NEW SPONSORED RESEARCH AWARDS

In 2021, ISI received 56 new funded research awards. Sponsored by the federal government, industry, and philanthropy, the awards range from tens of thousands of dollars to millions. The basic and applied research supported by these awards spans many areas: quantum information science, machine learning, microelectronics security, quantifying and analyzing synapse structure in zebrafish brains, and more.

AIR FORCE RESEARCH LABORATORY

Predicting Space Systems Operations
David Barnhart, *Principal Investigator*
(via Kayhan Space Corporation)

Access To Intel's 22FFL Technology through the MOSIS-Intel Alliance (ATMI)
Lifu Chang, *Principal Investigator*

OpTrust Capability for Field Programmable Gate Array (FPGA) Multiprocessor System on Chip (MPSoC)
Matthew French, *Principal Investigator*
(via Graf Research)

ARMY RESEARCH OFFICE

A Flexible Testbed for Quantum-Limited Sensing
Arunkumar Jagannathan, *Principal Investigator*

DEFENSE ADVANCED RESEARCH AGENCY DEFENSE SCIENCES OFFICE (DARPA DSO)

VENICE: Verifying Implicit Cultural Models
Fred Morstatter, *Principal Investigator*

INDIRECT: Initiating Data-Informed Interventions to Reduce Escalation while Confronting Threats
Greg Ver Steeg, *Principal Investigator*

Towards Enhanced Quantum Annealing In Learning and Simulation
Itay Hen, *Principal Investigator*

QUASAR: Quantum Assisted Sampling for Machine Learning
Federico Spedalieri, *Principal Investigator*

Food & Agricultural Assurance & Supply Chains Testbed (FAAST)
Craig Knoblock, *Principal Investigator*
(via University of Maryland)

DARPA INFORMATION INNOVATION OFFICE (I2O)

Universal Population Segmentation and Characterization Algorithms for Online Environments (UPSCALE)
Emilio Ferrara, *Principal Investigator*

DARMA: Dialogue Agent for Reducing Malicious Acts
Jonathan May, *Principal Investigator*

Social Analysis of Code to Counter Human Engineering Threats
James Blythe, *Principal Investigator*

Self-Stabilizing Synchronization over Secure Federated Knowledge Graphs for Multi-lateral Market Structures
Srivatsan Ravi, *Principal Investigator*

EDIFICE: Early Detection of Influence Indicators with Machine Intelligence
Kristina Lerman, *Principal Investigator*

NEW SPONSORED RESEARCH AWARDS (continued)

DARPA MICROSYSTEMS TECHNOLOGY OFFICE (MTO)

TREBUCHET
Matthew French, *Principal Investigator*
(via Duality Technologies)

Recurrent Neural Network Processing In-Pixel for Efficient Low Energy Heterogeneous Systems (RPIXELS)
Ajey Jacob, *Principal Investigator*

DEPARTMENT OF DEFENSE

WRT-1041: SMC Production Corps Mission Engineering and Integration of Emerging Technologies
Michael Orosz, *Principal Investigator*
(via Stevens Institute of Technology)

SERC OSD Agile Report Project
Michael Orosz, *Principal Investigator*
(via Stevens Institute of Technology)

Galactic Gopher
Matthew French, *Principal Investigator*
(via Georgia Institute of Technology)

DEPARTMENT OF ENERGY

Using Machine Learning to Understand Material Science
Keith Burghardt, *Principal Investigator*
(via Sandia National Laboratories)

PosEiDon: Platform for Explainable Distributed Infrastructure
Ewa Deelman, *Principal Investigator*

Q4Q: Quantum Computation for Quantum Prediction of Materials and Molecular Properties
Itay Hen, *Principal Investigator*

INTELLIGENCE ADVANCED RESEARCH PROJECTS ACTIVITY (IARPA)

Generative Range and Altitude Identity Learning (GRAIL)
Wael AbdAlmageed, *Principal Investigator*

LOCKHEED MARTIN

Pharos: Short Duration Cubesat Missions for Communication Dynamics Research
David Barnhart, *Principal Investigator*

Directed Energy Detection
Jonathan Habif, *Principal Investigator*

MISSILE DEFENSE AGENCY

KBox: A KVM-Based Virtual Platform Simulation Support for Discrete-Event Simulation System
John Paul Walters, *Principal Investigator*
(via Intelligent Automation, Inc.)

Inception: An Extendable Virtual Execution Framework to Support Dynamic Code Analysis of Embedded Systems
John Paul Walters, *Principal Investigator*
(via Intelligent Automation, Inc.)

NASA

MA - HPSC Design Study Phase 1
John Paul Walters, *Principal Investigator*
(via Microchip Technology, Inc.)

Deep Learning Processing Subsystem (DLPS): An HPSC-Compatible Deep Learning Coprocessor
John Paul Walters, *Principal Investigator*
(via Intelligent Automation, Inc.)

NATIONAL SCIENCE FOUNDATION

EAGER: QIA: A Quantum Algorithm for Detecting Quantum Information Leakage in Qubit Systems
Amir Kalev, *Principal Investigator*

REU Site: SURF-I: Safe, Usable, Resilient, and Fair Internet
Jelena Mirkovic, *Principal Investigator*

Collaborative Research: EAGER: VisDict - Visual Dictionaries for Enhancing the Communication between Domain Scientists and Scientific Workflow Providers
Ewa Deelman, *Principal Investigator*

CRII: OAC: Scalability of Deep-Learning Methods on HPC Systems: An I/O-Centric Approach
Loic Pottier, *Principal Investigator*

CI CoE: An NSF Cyberinfrastructure Center of Excellence for Navigating the Major Facilities Data Lifecycle
Ewa Deelman, *Principal Investigator*

Quantum Optics with Nonlinear Organic Small Molecule Enhanced Integrated Photonics Devices
Jonathan Habif, *Principal Investigator*

Collaborative Research: Elements: Simulation-Driven Evaluation of Cyberinfrastructure Systems
Loic Pottier, *Principal Investigator*

MA - Computing Innovation Fellows 2021 Project
Yolanda Gil, *Principal Investigator*
(via Computing Research Association)

Collaborative Research: NCS: FO: Enhancing Episodic Memory through Real-World Integration of Brain Recording and Stimulation with Semantic Alignment of Human and IoT Perception
Luis Garcia, *Principal Investigator*

EarthCube Capabilities: PaleoCube: Enabling Cloud-Based Paleoclimatology
Deborah Khider, *Principal Investigator*

Collaborative Research: Creating Mechanisms to Make Integrative Structures of Large Macromolecular Assemblies Available from the Protein Data Bank
Carl Kesselman, *Principal Investigator*

FMitF: Track I: Synthesis of Quantitative Network Analytics: From Left-of-Launch to Right-of-Boom
Michael Collins, *Principal Investigator*

CC* Regional: A Purpose-Built SoCal Science DMZ for Catalyzing Scientific Research Collaborations
Carl Kesselman, *Principal Investigator*

CCRI: NEW: CLASSNET: Community Labeling and Sharing of Security and Networking Test Datasets
Jelena Mirkovic, *Principal Investigator*

CRII: III: Knowledge Graph Completion with Transferable Representation Learning
Muhao Chen, *Principal Investigator*

NEW SPONSORED RESEARCH AWARDS (continued)

NSF Convergence Accelerator - Track D: Artificial Intelligence and Community Driven Wildland Fire Innovation via a WIFIRE Commons Infrastructure for Data and Model Sharing
Yolanda Gil, *Principal Investigator*
(via UC San Diego)

NSF-BSF: RI: Small: Efficient Bi- and Multi-Objective Search Algorithms
Satish Thittamaranahalli Ka, *Principal Investigator*
(via University of Southern California)

Collaborative Research: OAC Core: Simulation-Driven Runtime Resource Management for Distributed Workflow Applications
Loic Pottier, *Principal Investigator*

NATIONAL INSTITUTES OF HEALTH

Synapse Circuits Underlying Threat and Safety
Carl Kesselman, *Principal Investigator*

OFFICE OF NAVAL RESEARCH

Towards Reflection Competencies for AI Scientists: Developing a Conceptual Framework and Open Research Platform
Yolanda Gil, *Principal Investigator*

Persuasion, Identity, and Morality in Social-Cyber Environments
Fred Morstatter, *Principal Investigator*
(via Carnegie Mellon University)

ROSE FOUNDATION FOR COMMUNITIES AND THE ENVIRONMENT

Auditing for Discrimination in On-Line Advertisements
John Heidemann, *Principal Investigator*

SANDIA NATIONAL LABORATORIES

Consultation on Sandias Strategy for Emulytics
Terry Benzel, *Principal Investigator*

SUNGKYUNKWAN UNIVERSITY

Studying Human Aspects of Cybersecurity
Jelena Mirkovic, *Principal Investigator*

US AIR FORCE

Horizon Drive - A Radical Propellentless Propulsion Concept for Commercial and Defense Applications
David Barnhart, *Principal Investigator*
(via Positron Dynamics)

US GOVERNMENT – OTHER

BIGFOOT Phase 2
John Paul Walters, *Principal Investigator*
(via Peraton, Inc.)

Cross-Lingual Information Extraction for JHU SCALE 2021
Scott Miller, *Principal Investigator*
(via Johns Hopkins University)

MAJOR RESEARCH AWARDS

These significant, collaborative, multi-year research projects were awarded to ISI principal investigators in 2021.

ACCESS TO INTEL'S 22FFL TECHNOLOGY THROUGH THE MOSIS-INTEL ALLIANCE (ATMI)

Lifu Chang, *Principal Investigator*
Co-PI: Michael Haney, ISI

The ATMI project, supported by the Air Force Research Laboratory (AFRL), enables universities, defense industry companies, and government labs to utilize Intel's state-of-the-art Intel 16 FinFET silicon fabrication technology to produce chips from research- and prototype-oriented semiconductor designs. ISI's MOSIS Service provides project management, technical support, and foundry project milestones services to these customers.
Award: $18M
Duration: 2 years

CYBERINFRASTRUCTURE CENTER OF EXCELLENCE FOR NAVIGATING THE MAJOR FACILITIES DATA LIFECYCLE (CI COE)

Ewa Deelman, *Principal Investigator*
Co-PIs:
- Anirban Mandal, Renaissance Computing Institute
- Angela Murillo, Indiana University
- Jaroslaw Nabrzyski, Notre Dame University
- Valerio Pascucci, University of Utah

Collaborators:
- Kerk Kee, Texas Tech University
- Rob Ricci, University of Utah

Sponsored by National Science Foundation (NSF), this Cyberinfrastructure Center of Excellence enhances the cyberinfrastructure underlying the data lifecycle of NSF major facilities. It enables the transformation of raw data captured by scientific instruments into interoperable and integration-ready data products that can be visualized, disseminated, and converted into insights and knowledge. The project engages with NSF major facilities and contributes knowledge and expertise by offering its services, including evaluating cyberinfrastructure plans, helping design new architectures and solutions, developing proofs of concept, and assessing the applicability and performance of existing cyberinfrastructure solutions.
Award: $8M
Duration: 5 years

GENERATIVE RANGE AND ALTITUDE IDENTITY LEARNING (GRAIL)

Wael AbdAlmageed, *Principal Investigator*
Co-PIs:
- Mohamed Hussein, ISI
- Leonidas Spinoulas, ISI

Collaborators:
- Aggelos Katsaggelos, Northwestern University
- Sébastien Marcel, Idiap Research Institute
- Ram Nevatia, USC Computer Science
- Amit Roy-Chowdhury, University of California, Riverside
- J.P. Walters, ISI

Sponsored by the Intelligence Advanced Research Projects Activity (IARPA), this research project addresses the yet unsolved problem of multi-biometric identification in unconstrained environments. The issues addressed include acute viewing angles, large distances, and hazy, noisy imagery. The project is developing generative deep learning methods for human identification from large distances and altitudes, from the face (a hard biometric) and gait and body shape (soft biometrics).
Award: $12.5M
Duration: 2.5 years

HONORS AND AWARDS

PROFESSIONAL AND ACADEMIC RECOGNITION

Jose Luis Ambite

Keynote invited talk, 26th ACM SIGKDD Workshop on Knowledge-Infused Mining and Learning, *Semantics in Biomedical Data Science*

Invited talk, USC Biostatistics Seminar, *Scalable Identity-By-Descent Mapping in Massive Genetic Datasets*

Eva Deelman

HPC Wire Readers' Choice Award: Best HPC Collaboration (Academia/Government/Industry) O*rganizations Partner to Rescue Petabytes of Data from the Arecibo Observatory*

Matthew French

Invited panelist, IEEE Council on Electronic Design Automation (CEDA), *Security Assessment and Verification for Microelectronics: A New Future or More of the Same?*

Judge, CSAW'21 Cybersecurity and Games Conference, Logic Locking Contest

Yolanda Gil

Fellow, Institute of Electrical and Electronics Engineers (IEEE)

Invited panelist, Cambridge University, *AI for Interdisciplinary Scientific Discovery*

Invited talk, National Academies of Sciences, Engineering, and Medicine workshop series on critical technologies for national security, *AI Advances: Past and Future*

Invited panelist, IEEE Computer Society's EvoPro Conference 75th Anniversary, *A Twenty-Year Research Roadmap for the US*

Invited keynote, Automated Knowledge Base Construction (AKBC) Conference, *Extracting Knowledge from Text about Models and Workflows: Transparency, Reproducibility, and Automation in Science*

Invited keynote, China National Computer Congress (CNCC) of the China Computer Federation (CCF), *Artificial Intelligence for Scientific Discoveries through World-Wide Collaborations*

Itay Hen

Invited talk, "Gleb Wataghin" Institute of Physics (IFGW) Quantum Technologies School, *Analog Quantum Computing*

Invited talk, Quantum Computing on the Cloud – Early Adopters Meeting (QEAM21), *Quantum Simulations in the NISQ Era: Simulating Hamiltonian Dynamics on a Quantum Computer Using the Off-Diagonal Series Expansion*

Invited talk, XXXII International Union of Pure and Applied Physics (IUPAP) Conference on Computational Physics, *Permutation Matrix Expansion Quantum Monte Carlo*

Invited talk, IBM Quantum, *Quantum Algorithm for Simulating Hamiltonian Dynamics with the Off-Diagonal Series Expansion*

I**nvited talk,** Adiabatic Quantum Computing Conference 2021, *Non-Stoquasticity and Non-Simulability in Adiabatic Quantum Computing*

Invited talk, University of California, Santa Barbara Winter School on Quantum Science, *Analog Quantum Computing: Theory and Practice*

Filip Ilievski

Invited keynote, The Knowledge Graph Conference Knowledge-infused Learning Workshop, *Building Agents with Common Sense*

Research gift, Bosch Research North America

Connor Imes

Outstanding Paper Award, IEEE High Performance Extreme Computing Conference (HPEC 2021)

Carl Kesselman

Invited keynote, Collaborating for the Advancement of Interdisciplinary Research in Benign Urology Annual Meeting (CAIRIBU 2021), *Translating the Grid: How a Translational Approach Shaped the Development of Grid Computing*

William M. Keck Chair in Engineering, USC Viterbi School of Engineering

Mayank Kejriwal

Award for Excellence in Mentoring (Graduate Student Category), USC

Finalist, Early Career Award for Public Engagement with Science, American Association for the Advancement of Science (AAAS)

Craig Knoblock

Invited talk, Open Data Science Conference (ODSC) West, *Acquiring and Exploiting the Semantics of Data*

AAIA Fellow, Asia-Pacific Artificial Intelligence Association

Dong-Ho Lee

Best Paper Award, North American Chapter of the Association for Computational Linguistics (NAACL) Workshop on Trustworthy Natural Language Processing (TrustNLP), *AutoTriggER: Named Entity Recognition with Auxiliary Trigger Extraction*

Kristina Lerman

Invited talk, Association for Computing Machinery (ACM) Special Interest Group on Knowledge Discovery and Data Mining Conference (KDD2021) epiDAMIK4.0 Workshop, *Understanding the Infodemic: Political Polarization and Mistrust of Science in the Age of COVID-19*

Invited talk, International Association for the Advancement of Artificial Intelligence (AAAI) Conference on Web and Social Media (ICWSM) Mediate 2021 Workshop, *The Geography of Science Skepticism in the US*

Invited talk, University of Chicago Computational Social Science Summer School, *The Challenges of Learning from Heterogeneous Data: Bias in Data and Algorithmic Fairness*

Invited talk, Society for Industrial and Applied Mathematics (SIAM) Minisymposium, *Friendship Paradox and Emergence of Bias in Directed Networks*

Xuezhe Ma

Invited talk, TechBeat, *Optimization for Deep Neural Networks*

Jonathan May

Invited talk, Amazon Machine Learning Conference (AMLC) Workshop on Multilingual NLP, *Rethinking Assumptions in Multilingual NLP*

George Papadimitriou et al.

Best Paper Award, Future Generation Computer Systems 2021, *End-to-End Online Performance Data Capture and Analysis for Scientific Workflows*

HONORS AND AWARDS (continued)

Abdul Qadeer and John Heidemann

Special Paper Award, IEEE Conference on Cloud Computing, *Efficient Processing of Streaming Data using Multiple Abstractions*

Xiang Ren

2021 National Science Foundation CAREER Award

Ke Shen

Best Paper Award, Specialist Group on Artificial Intelligence, AI-2021: The Forty-first SGAI International Conference, *On the Generalization Abilities of Fine-Tuned Commonsense Language Representation Models*

Haoda Wang

Goldwater Scholar, Barry Goldwater Scholarship and Excellence in Education Foundation

2021 US-ISSUED PATENTS

Young Cho, Siddharth Bhargav, Andrew Goodney

Digital Signal Transition Counters for Digital Integrated Circuits (patent number 10,963,026)

Akhilesh Jaiswal, Ajey Jacob, Steven Soss

Circuit structure and method for memory storage with memory cell and MRAM stack (patent number 11,145,348)

MRAM device comprising random access memory (RAM) and embedded read only memory (ROM) (patent number 10,964,367)

Akhilesh Jaiswal and **B. C. Paul**

Low variability reference parameter generation for magnetic random access memory (patent number 11,120,857)

Akhilesh Jaiswal and **Ajey Jacob**

Integrated pixel and three-terminal non-volatile memory cell and an array of cells for deep in-sensor, in-memory computing (patent number 11,069,402)

Akhilesh Jaiswal and **Mudit Bhargava**

Method, system and device for magnetic memory (patent number 10,991,406)

Method, system and device for integration of volatile and non-volatile memory bitcells (patent number 10,971,229)

D-MRAM devices and methods for replicating data and read and write operations (patent number 10,896,730)

Michael Sika

Systems and Methods for Mitigating Faults in Combinatory Logic (patent number 10,901,836)

ISI ACHIEVEMENT AWARDS

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

Karen Rawlins for her unceasing work ethic, willingness to serve above and beyond, leadership by example, and—especially—her role in making ISI a great place to work

Eloisa Romo for her commitment and the many hours of work that made it possible for ISI to use the university's enterprise software system to support ISI's research mission

2021 PHD GRADUATES

ISI congratulates these terrific USC graduates for completing their PhD degrees in 2021.

Ruhollah (Roohy) Shemirani
Dissertation: *Ultra Rapid Identity-By-Descent Mapping in Massive Genetic Datasets*
Current position: Postdoctoral Fellow, Institute for Genomic Health at the Icahn School of Medicine at Mount Sinai

Jose Luis Ambite, *Advisor*

Rahul Rughani
Dissertation: *Relative-Motion Trajectory Generation and Maintenance for Multi-Spacecraft Swarms*
Current position: Lead Systems Engineer, Arkisys, Inc.

David Barnhart, *Advisor*

Adbul Qadeer
Dissertation: *Efficient Processing of Streaming Data in Multi-User and Multi-Abstraction Workflows*
Current position: Member of the Technical Staff, Educative.io

John Heidemann, *Advisor*

Nazanin Alipourfard
Dissertation: *Emergence and Mitigation of Bias in Heterogeneous Data*
Current position: Software Engineer, Twitter

Yuzi He
Dissertation: *Learning Fair Models with Biased Heterogeneous Data*
Current position: Data Scientist, Facebook

Nazgol Tavabi
Dissertation: *Modeling Dynamic Behaviors in the Wild*
Current position: Postdoctoral Researcher, Harvard University

Kristina Lerman, *Advisor*

Xusen Yin
Dissertation: *Generalized Sequential Decision-Making via Language*
Current position: Applied Scientist, Amazon

Jonathan May, *Advisor*

Xiyue Deng
Dissertation: *Studying Malware Behavior Safely and Efficiently*
Current position: Software Engineer, Google

Sivaramakrishnan Satyamangalam Ramanathan
Dissertation: *Leveraging Programmability and Machine Learning for Distributed Network Management to Improve Security and Performance*
Current position: Software Engineer, Facebook

Jelena Mirkovic, *Advisor*

Victor Martinez Palacios
Dissertation: *Computational Narrative Models of Character Representations to Estimate Audience Perception*
Current position: Machine Engineer, Apple

Karel Mundnich
Dissertation: *Learning Multi-Annotator Subjective Label Embeddings*
Current position: Applied Scientist, Amazon

Karan Singla
Dissertation: *Improving Language Understanding and Summarization by Leveraging Auxiliary Information Through Self-Supervised or Unsupervised Learning*
Current position: Senior Inventive Scientist, Interactions, LLC

Shri Narayanan, *Advisor*

PHD STUDENTS

ISI is proud to recognize this impressive cohort of doctoral students and their advisors. The PhD students here include those who are supported at ISI and students whose advisors are associated with ISI.

Jiazhi Li | *Debiasing face recognition*

Hayley Song | *Generative modeling and representation learning*

Hanchen Xie | *Learning with less-labeled data*

Jiageng Zhu | *Invariant representation learning*

Wael AbdAlmageed, *Advisor*

Dimitris Stripelis | *Federated learning*

Jose Luis Ambite, *Advisor*

James Yipeng Huang | *Natural language processing, knowledge-incorporated natural language inference*

Wenxuan Zhou | *Natural language processing, robust learning, inference for knowledge extraction from text*

Muhao Chen, *Advisor*

Yang Hu | *Edge systems for machine learning and inference*

Haonan Wang | *Edge systems for machine learning and inference*

Steve Crago, *Advisor*

Tu Mai Anh Do | *In situ workflows, high performance computing*

Patrycja Krawczuk | *AI for optimizing distributed workflows and systems*

Georgios Papadimitriou | *Distributed resource management*

Ewa Deelman, *Advisor*

Alexander Bisberg | *Machine learning*

Herbert Chang | *Machine learning*

Emily Chen | *Machine learning*

Julie Jiang | *Machine learning*

Hsien-Te Kao | *Machine learning*

Akira Matsui | *Machine learning*

Karishma Sharma | *Machine learning*

Alexander Spangher | *Machine learning (co-advisor: Jonathan May)*

Shen Yan | *Machine learning*

Yilei Zeng | *Machine learning*

Meiqing Zhang | *Computational social science (co-advisor)*

Weiwei Zheng | *Computational social science*

Emilio Ferrara, *Advisor*

Tainã Coleman | *Scientific workflows and distributed computing*

Rafael Ferreira Da Silva, *Advisor*

Sarik Ghazarian | *Dialogue systems*

Rujun Han | *Natural language understanding*

Mehrnoosh Mirtaheri | *Knowledge graphs*

Aram Galstyan, *Advisor*

Guillermo Baltra Elorriaga | *Measurement of network reliability, improving and understanding active sources*

Asma Enayet | *Measurement of network reliability, understanding sensitivity of passive sources*

Basileal Yoseph Imana | *Algorithm privacy in online advertising*

ASM Rizvi | *DDoS defense for critical services*

Xiao Song | *Network measurement to understand reactions to COVID-19*

John Heidemann, *Advisor*

Lalit Gupta | *Quantum Monte Carlo simulations of many-body systems*

Nicholas Ezzell | *Zero-temperature Quantum Monte Carlo methods*

Jing Wu | *Quantum many-body simulations*

Itay Hen, *Advisor*

Zihan Yin | *Novel computing paradigms for smart sensors and memory structures for data-intensive applications and edge intelligence*

Akhilesh Jaiswal, *Advisor*

Ke Shen | *Machine commonsense reasoning, open-world AI*

Mayank Kejriwal, *Advisor*

Robert Schuler | *Data management, informatics research*

Carl Kesselman, *Advisor*

Weiwei Duan | *Deep learning for geographic feature recognition from historical maps*

Minh Tran Xuan Pham | *Data cleaning for knowledge graph construction*

Basel Shbita | *Building spatio-temporal knowledge graphs from historical maps*

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

Craig Knoblock, *Advisor*

Nathan Bartley | *Computational social science*

Rebecca Dorn | *Fairness, science of science*

Fiona Guo | *Social media analysis, computational social science*

Yuzi He | *Fairness*

Zihao He | *Social media analysis, natural language processing*

Negar Mokhberian | *Computational social science*

Ashwin Shreyas Mohan Rao | *Network science*

Kristina Lerman, *Advisor*

Ming-Chang Chiu | *Robustness in deep representation learning*

Jiao Sun | *Trustworthy natural language generation*

Xuezhe Ma, *Advisor*

Nada Aldarrab | *Natural language processing, decipherment*

Jacob Bremerman | *Natural language processing*

Hyundong Cho | *Natural language processing, dialogue (co-advisor: Xiang Ren)*

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*

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

Jonathan May, *Advisor*

Sima Arasteh | *Binary analysis*

Pithayuth Charnsethikul | *Privacy, anti-phishing*

Rajat Tandon | *DDoS and privacy*

Nicolaas Weideman | *Binary analysis*

Wei-Cheng Wu | *Binary analysis*

Jelena Mirkovic, *Advisor*

Bahareh Harandizadeh | *Topic modeling, cultural modeling, machine learning, AI*

Yuzhong Huang | *Reproducibility, scientific networks, embedding, machine learning, AI*

Ninareh Mehrabi | *Algorithmic fairness in machine learning and natural language processing*

Fred Morstatter, *Advisor*

Sabyasachee Baruah | *Natural language processing, affective computing, computational media intelligence*

Victor Ardulov | *Human-centered machine intelligence and applications, especially focused on children*

Zhuohao Chen | *Behavioral signal processing and machine intelligence*

Tiantian Feng | *Trustworthy machine intelligence, behavioral modeling, wearables*

Nikolaos Flemotomos | *Behavioral signal processing and machine intelligence*

Amrutha Nadarajan | *Trustworthy audio processing, human-centered machine intelligence*

Raghuveer Peri | *Robust audio processing and machine learning*

Shri Narayanan, *Advisor*

Jiaxin Cheng | *Computer vision*

I-Hung Hsu | *Information extraction, named entity recognition, relation extraction, event extraction*

Soumyaroop Nandi | *Computer vision*

Ekraam Sabir | *Digital image forensics (detection of manipulations in images)*

Prem Natarajan, *Advisor*

PHD STUDENTS (continued)

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

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

Dong-Ho Lee | *Few-shot natural language processing, contrastive representation learning, temporal knowledge graphs*

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

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

Jay Pujara, *Advisor*

Charles Bickham | *Formal methods for provenance in distributed networks*

Weizhao Jin | *Multi-party computation for the dispersed edge*

Srivatsan Ravi, *Advisor*

Shushan Arakelyan | *Machine learning for program analysis*

Aaron Chan | *Model explainability, explanation-based learning*

Woojeong Jin | *Improving language understanding using visual information*

Brihi Joshi | *Interpretability, explanation-based learning*

Yuchen (Bill) Lin | *Commonsense reasoning, federated learning, cross-task generalization*

Soumya Sanyal | *Logical reasoning in language models, explanation-based learning*

Jun Yan | *Model robustness, commonsense reasoning*

Qinyuan Ye | *Few-shot learning, meta-learning, explanation-based learning*

Pei Zhou | *Commonsense reasoning, dialogue systems, interpretability (co-advisor: Jay Pujara)*

Xiang Ren, *Advisor*

Serban Stan | *Unsupervised domain adaptation*

Mohammad Rostami, *Advisor*

Hannes Leipold | *Quantum computing*

Federico Spedalieri, *Advisor*

Ehsan Qasemi | *Commonsense reasoning*

Peifeng Wang | *Commonsense reasoning*

Pedro Szekley, *Advisor*

Ang Li | *Fast graph algorithms with applications in data science*

Satish Kumar Thittamaranahalli, *Advisor*

Sami Abu-El-Haija | *Graph neural networks, machine learning (co-advisor: Aram Galstyan)*

Robert Brekelmans | *Information geometry, latent factor models, machine learning*

Umang Gupta | *Privacy and fairness, machine learning*

Hrayr Harutyunyan | *Deep learning theory, machine learning*

Neal Lawton | *Optimization, machine learning*

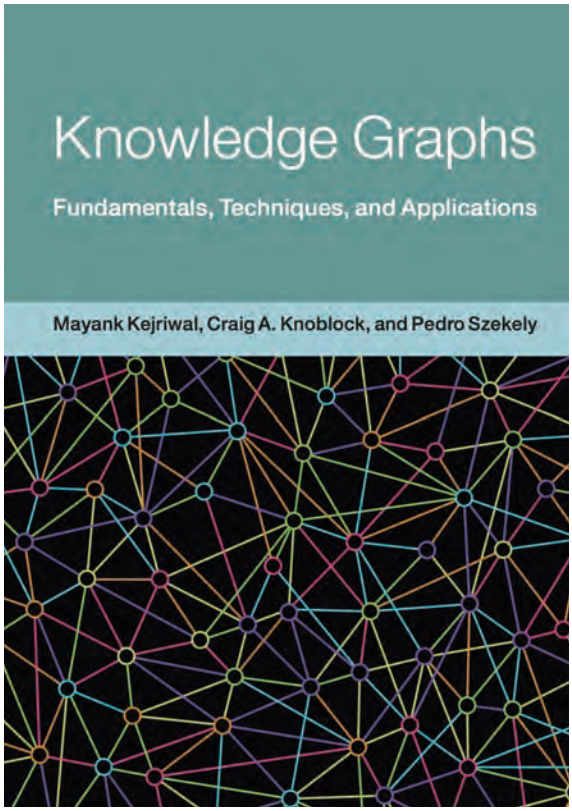
Elan Markowitz | *Graph neural networks, machine learning*

Myrl Marmarelis | *Point processes, diffusion, machine learning*

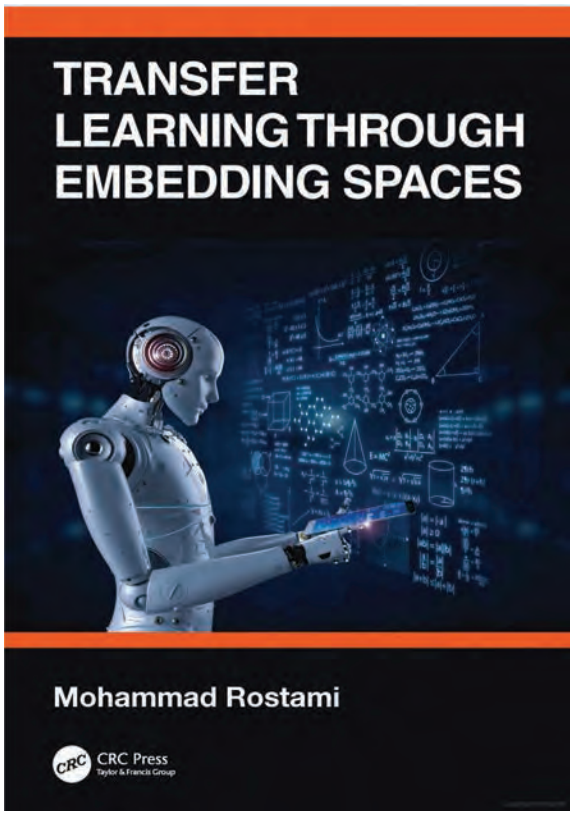
Kyle Reing | *Information decomposition, interpretability, machine learning*

Greg Ver Steeg, *Advisor*

NEW BOOKS BY ISI AUTHORS



The field of knowledge graphs, which allows computer scientists to model, process, and derive insights from complex real-world data, has emerged as an active and interdisciplinary area of artificial intelligence over the last decade, drawing on such fields as natural language processing, data mining, and the semantic web. Current projects involve predicting cyberattacks, recommending products, and even gleaning insights from thousands of papers on COVID-19. This textbook offers rigorous and comprehensive coverage of the field. It focuses systematically on the major approaches, both those that have stood the test of time and the latest deep learning methods.



Recent progress in artificial intelligence has revolutionized everyday life. This success, however, is conditioned on the availability of huge, annotated datasets to train AI models. Data annotation is a time-consuming and expensive task that still is being performed by human workers. Learning efficiently from less data is a step toward making AI more like natural intelligence. Transfer learning can reduce the need for data annotation by transferring knowledge across similar tasks and using similarities and previously learned knowledge to learn more efficiently. This book provides a brief background on transfer learning and then focuses on the idea of transferring knowledge through intermediate embedding spaces. The idea is to couple and relate different learning through embedding spaces that encode task-level relations and similarities. The book covers various machine learning scenarios and demonstrates that transfer learning through embedding spaces can overcome the challenges of zero-shot learning, few-shot learning, domain adaptation, continual learning, lifelong learning, and collaborative learning.

RISING STAR MS INTERNSHIP 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 in an ISI team on an ongoing research project—and work directly with an individual mentor. The internships last throughout the course of the master's degree.

The program offers multiple "tracks" to provide each intern with experience tuned to their specific educational and career goals:

- 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

Wes Hardaker, in the Networking and Cybersecurity Division, is mentoring **Spencer Stingley.** Spencer was selected for the program in 2021 and is concentrating on analyzing malware samples being collected for the Global Analysis of Weak Signals for Enterprise Event Detection (GAWSEED) project. Wes continues to mentor **Erin Szeto**, who was selected for the program in 2020. Erin, too, is working on the GAWSEED program.

Filip Ilievski, in the Artificial Intelligence Division, is mentoring **Nicholas Klein**, who was selected for the program in 2020. They are working on a project about knowledge graph profiling.

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.

Instead of coming to an ISI location, most of the interns worked remotely with their ISI mentors and research teams this year. In addition to their research, they participated in seminars, reading groups, and social activities. The 2021 ISI interns included undergrads, master's students, PhD students (over half women), and one high school student. Four of the students had interned at ISI previously and returned to continue their work.

Ten senior ISI researchers served as mentors to interns:

Wael AbdAlmageed, *Advisor*

Ismat Jarin, University of Michigan – PhD student
Research area: *Developing image manipulation detection neural network architectures*

Marjorie Freedman, *Advisor*

Ashwin Kumar, University of Louisville – undergraduate student
Research area: Natural language processing

Justin Martin, University of Louisville – undergraduate student
Research area: Natural language processing

Yolanda Gil, *Advisor*

Haoyang Chen, USC – master's student
Research area: *Machine learning for genomics*

Christopher Heidelberg, USC - undergraduate student
Research area: *Automated generation of explanations for data analysis workflows*

Megan Josep, USC - master's student
Research area: *Data science reports*

Lu Niu, USC - master's student
Research area: *Healthcare data science*

Ming Yan, USC - master's student
Research area: *Machine learning for neuroimaging*

Tingyu Zhang, USC - undergraduate student
Research area: *Ontology-enhanced search for clinical trials*

Jonathan Habif, *Advisor*

Olivia Brasher, USC – undergraduate student
Research area: *Experimental measurement of achieving the quantum Chernoff bound for quantum sensing*

Zi Chua, Wellesley College – undergraduate student
Research area: *Quantum-limited estimation of coherence in quantum mixed states*

Rishabh Jain, University of Massachusetts Amherst – undergraduate student
Research area: *Experimental testbed for demonstration of quantum steganography*

Christine Kinzfogl, Lincoln-Sudbury Regional High School
Research area: *Measurement of atmospheric scattering of laser light for quantum estimation of laser pointing angle*

Travis Haroldsen, *Advisor*

Shlok Agarwal, Virginia Tech – undergraduate student
Research area: *FPGA hardware testing*

Hunain Ali, Virginia Tech – undergraduate student
Research area: *FPGA hardware testing*

Aldin Shiri, University of Maryland – PhD student
Research area: *FPGA characterization and test*

Deborah Khider, *Advisor*

Shravya Manety, USC – master's student
Research area: *Intelligent guidance for climate data annotation*

Kristina Lerman, *Advisor*

Yulin Yu, University of Michigan – PhD student
Research area: *Social networks, the science of science*

Jonathan May, *Advisor*

Shanxiu He, UCLA – undergraduate student
Research area: *Natural language processing*

Leo Zeyu Liu, University of Washington – undergraduate student
Research area: Natural language processing

Shira Wein, Georgetown University – PhD student
Research area: *Natural language processing*

Vivek Menon, *Advisor*

Richard Becker, University of Arkansas - PhD student
Research area: *Asynchronous IC design, hardware security, polymorphic circuits*

Rakibul Hassan, George Mason University - PhD student
Research area: *Machine learning, hardware security*

Andrew Schmidt, *Advisor*

Sathwika Bavikadi, George Mason University - PhD student
Research area: *Reconfigurable computing, processing in memory and machine learning*

RESEARCH EXEPERIENCES FOR UNDERGRADUATES PROGRAM

ISI sponsored ten undergrads from colleges and universities across the country to spend nine weeks working with ISI research faculty and their teams. This year's theme was *SURF-I: Safe, Usable, Resilient, and Fair Internet*.

The NSF Research Experiences for Undergraduates (REU) program offers an intellectually and socially stimulating research experience to undergrads while providing:

- Paid training in research and working in research teams
- Information about research careers
- Guidance on writing for publication

Each student concentrates on a specific research project, and they work closely with ISI research faculty and other ISI researchers.

ISI was selected to host an REU site during the summer beginning in 2018, with Jelena Mirkovic, an ISI cybersecurity researcher and project leader, as the site supervisor.

The undergrads worked with ISI researchers remotely in 2021 instead of coming to ISI to work. While each student worked with their own mentor, they attended twice-weekly seminars that explored a variety of research areas and tools. The 2021 cohort included four male students and six female students. On August 6, as the program concluded, the students presented their research as poster sessions at the ISI REU symposium, where they presented to their ISI mentors, researchers, and students.

12 ISI researchers served as mentors to the undergraduate researchers:

Jeremy Abramson, *Advisor*
Chloe Eggleston, UMass Amherst
Research topic*: Woolery: Extending Frame Semantics to Structured Documents**

Genevieve Bartlett, *Advisor*
Katherine Au, New Mexico State University
Research topic*: Automatically Classifying Phishing Emails into Common Types*

Keith Burghardt, Goran Muric, *Advisors*
Kaitlyn Ko, University of Chicago
Research topic*: Enter at Your Own Risk: The Impacts of Joining a Hateful Subreddit**

Emilio Ferrara, Emily Chen, *Advisors*
Oliver Allen, Indiana University Bloomington
Research topic*: Pictures as a Form of Protest: A Survey and Analysis of Images Posted During the Stop Asian Hate Movement on Twitter**

Luis Garcia, *Advisor*
Jasmine Unterburger, Santa Monica College
Research topic*: Towards a Cyber-physical Fuzzing Simulation Framework*

Christophe Hauser, *Advisor*
Rene Reyes, Massachusetts Institute of Technology
Research topic*: Privacy Preserving Email Search*

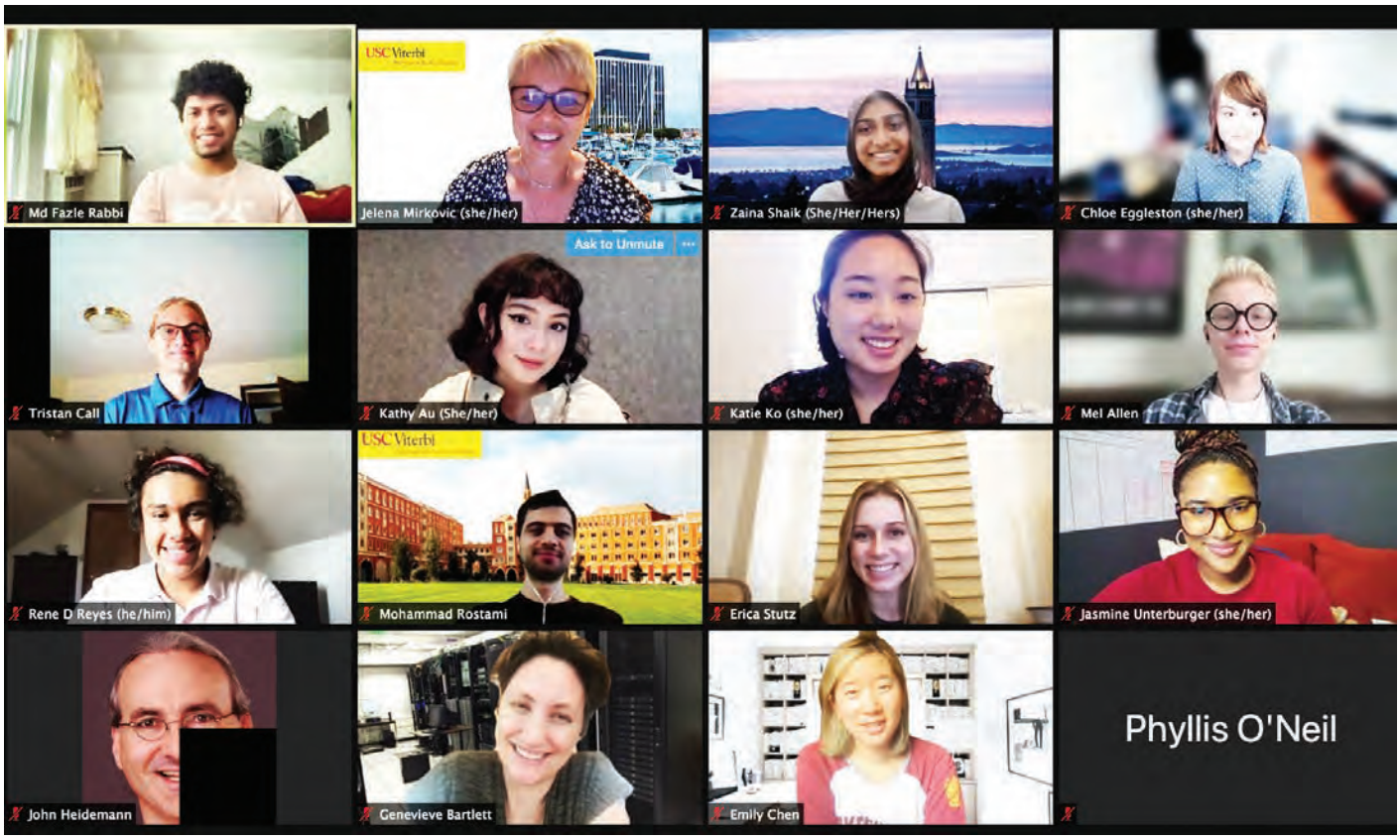
John Heidemann, *Advisor*
Erica Stutz, Swarthmore College
Research topic*: Mapping COVID-19 Work-from-Home*

Filip Ilievski, *Advisor*
Zaina Shaik, University of California, Berkeley
Research topic*: Race and Country Bias in Wikidata**

Fred Morstatter, *Advisor*
Tristan Call, Gonzaga University
Research topic*: Analyzing Gameability in Citation Networks*

Mohammed Rostami, *Advisor*
Md Fazle Rabbi, Fordham University
Research topic*: Hyperparameter Tuning for the Derivative Compressive Sampling*

**Posters were selected for presentation at the Seventh National Workshop for REU Research in Networking and Systems (REUNS 2021).*



from the students . . .

"I learned 5 new coding languages, had my first actual paycheck, worked on adding to a cool interface, strengthened my presentation skills, and got to meet so many amazing and influential people in this field. I have also learned many life skills that I will take with me everywhere. I am in the process of writing a paper, which is really exciting and will be extremely useful when applying to graduate school!"

"I loved that this REU felt so collaborative—not just amongst the cohort, but between all the students and the supervisors as well. I feel as if I got to know all of the supervisors pretty well, which felt reassuring throughout the program to know that I had 9 other people I could go to for help."

RESEARCH HIGHLIGHTS

Saving the Arecibo Telescope's Astronomical Data

The Arecibo Telescope, in the Arecibo Observatory in Puerto Rico, was completed in 1963. It was the world's largest single-aperture telescope for 53 years. But in 2020, breaks in cables supporting the receiver platform ultimately caused the instrument platform to crash through the dish, shearing off the tops of the support towers.

An international community is working to archive the astronomical data from the Arecibo radio telescope, a massive volume—over 3 petabytes—of valuable astronomy data. The community of researchers is also expanding access to this scientific data: more than 50 years of sky scans.

ISI researchers, along with Indiana University, Notre Dame, the University of North Carolina at Chapel Hill, and other research teams, participated in the Cyberinfrastructure Center of Excellence Pilot project (CI CoE Pilot) supported by the NSF.

The CI CoE Pilot collaborators helped ensure that the data was managed and made accessible for astronomers and scientists. They promoted the exchange of cyberinfrastructure knowledge among Arecibo, other major NSF facilities, and the broader community.

The $8 million NSF-sponsored Cyberinfrastructure Center of Excellence for Navigating the Major Facilities Data Lifecycle project (CI Compass) now continues the work that the three-year CI CoE Pilot began.

CI Compass provides expertise and active support to cyberinfrastructure practitioners. It has developed a data lifecycle model for NSF scientific major facilities, including workflow management, data archiving, identity management, semantic technologies, and visualization.

CI Compass enhances the overall NSF cyberinfrastructure ecosystem by providing expertise to enhance and evolve the major facilities' cyberinfrastructure, capturing and disseminating cyberinfrastructure knowledge and best practices that power the major facilities' scientific breakthroughs, and brokering connections to enable knowledge sharing among major facilities cyberinfrastructure professionals and the broader cyberinfrastructure community.

Contact: Ewa Deelman, deelman@isi.edu

Credit: Arecibo Observatory

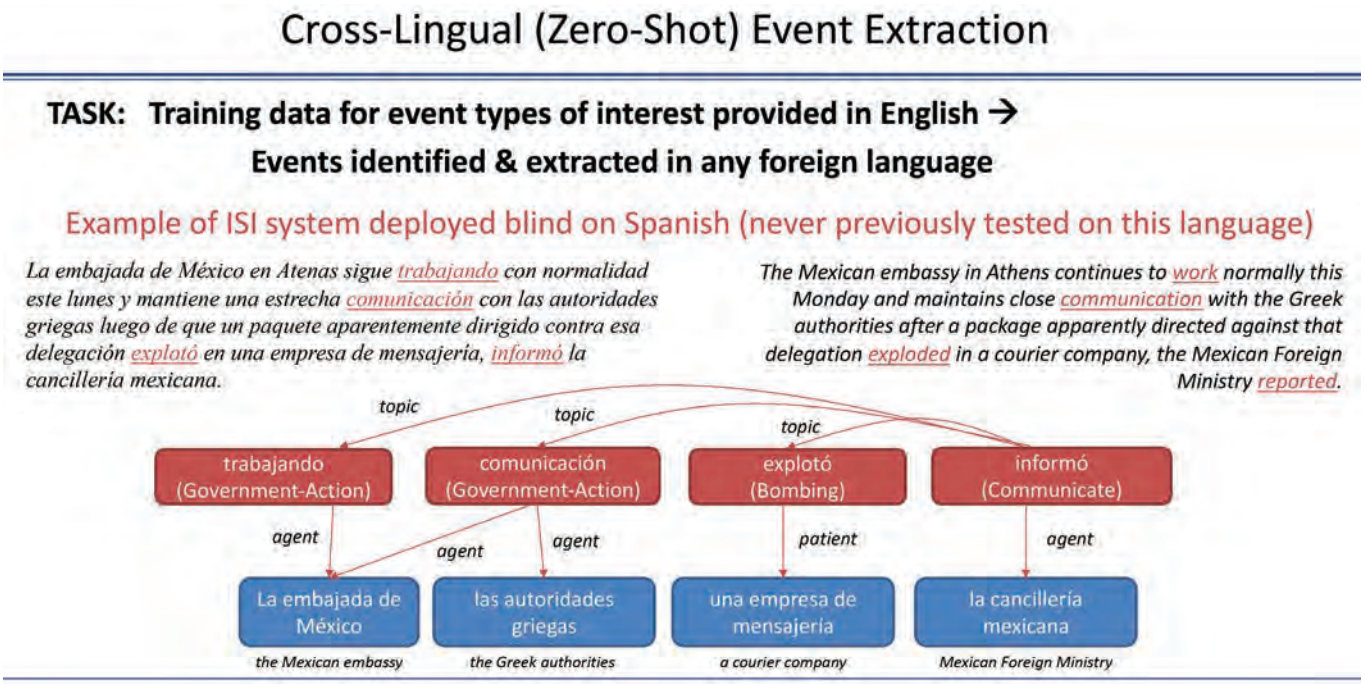
Tracking and Understanding Events Around the Globe

Protests, epidemics, terrorist attacks, refugee crises, natural disasters—these all occur around the globe but are important to people far beyond the local communities in which they occur. Unfortunately, local reporting is often in a language not spoken by analysts or decision-makers assessing global conditions. ISI's CLEAR effort (Cross-Lingual Event and Argument Extraction) endeavors to break that language barrier for world events, using English-only resources to build models that can effectively identify and track events in 100+ languages around the globe.

To accomplish this, the research team exploits the power of language models that learn to understand and categorize words from a massive amount of raw multi-lingual data. By reading billions of words of text in languages from around the globe, the models can learn not just that *protested* and *demonstrated* occur in similar contexts and mean something similar, and not only that *protested* (English) and *protestato* (Italian) occur in similar contexts (which we might guess since they look the same), but also that 抗議した (Japanese) ορορδιαμαρτυρήθηκε (Greek) fall into the same category.

The CLEAR system can then learn a model of events using English examples and apply that model to any foreign language in the language model, identifying not only that φοιτητές διαμαρτυρήθηκαν για τους Ολυμπιακούς Αγώνες refers to a protest event, but also the people participating in the protest (students) and the target of the protest (the Olympics). This gives CLEAR system users a much more nuanced way to analyze or aggregate the data, for instance, enabling them to easily quantify how many protests were reported on each of several distinct topics.

The figure below shows a graph of events and their participants extracted from a paragraph written in Spanish, a language that the event model has never seen before:



Contact: Elizabeth Boschee, boschee@isi.edu

RESEARCH HIGHLIGHTS

Employing Data Science to Clarify COVID-19's Spread

COVID-19 continued to be big news in 2021—and a big story to a team of ISI researchers who studied the disease's spread in the US and the world. To understand the disease dynamics and develop mitigation strategies, policymakers and epidemiologists need to know how the disease is spreading in their communities.

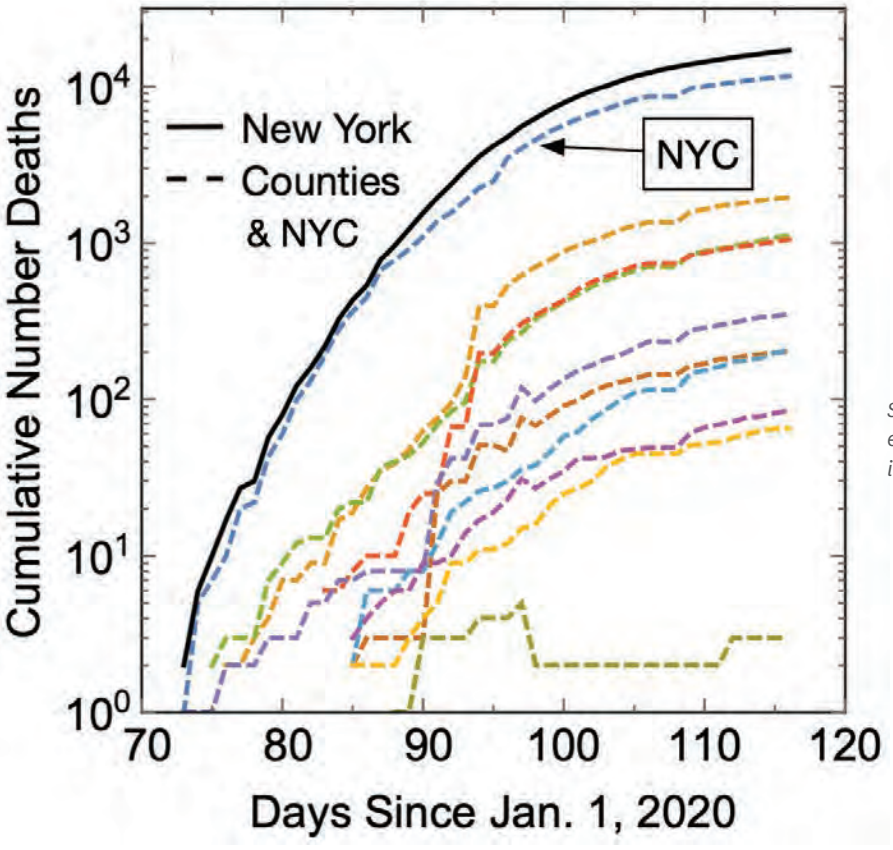
ISI researchers analyzed confirmed infections and deaths over multiple geographic scales (e.g., county vs. state vs. nation) to show that COVID-19's spread is highly unequal: many regions had almost no infections or deaths in the first few months, while others were hot spots with over-burdened hospitals. The researchers attributed the unequal distribution of cases and deaths to a model in which the disease arrives in regions and grows exponentially at different rates. For example, statistics for a state are overwhelmed by infections within fast-growing hot spot counties. As a result, COVID-19 appears to spread faster within that state than in most counties in the state.

Interestingly, the team also found that the growth rate of COVID-19 has consistently slowed with each surge.

These results demonstrate a tradeoff when modeling how COVID-19 spreads: measuring cases across a wide area makes modeling easier (more data means less error) but introduces biases in which analysts consistently overestimate the growth of the disease and therefore overestimate its initial impact. Public policy and epidemic modeling should take note of this bias and address it by carefully aggregating data.

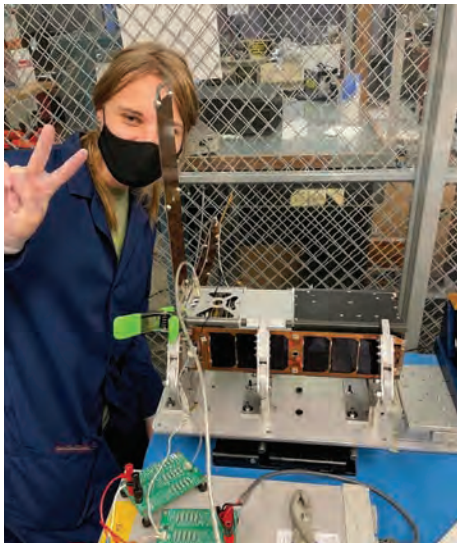
This is one of many efforts related to COVID-19 that the ISI team pursued in 2021. The research team analyzed online conversations about COVID-19 and developed methods to quantify political polarization and susceptibility to misinformation. In addition, researchers studied the pandemic's impact on gender disparities in science, uncovering new disadvantages for women who work in science due to the pandemic.

Contact: Keith Burkhart, keithab@isi.edu



Statistics are dominated by New York City, even though the growth rate can be slower in non-NYC counties.

SERC Builds Dodona, Third Satellite in USC History



Graduate student Julia Schatz in the cleanroom at SERC, working on Dodona

At SERC, the joint ISI and Department of Astronautical Engineering Space Engineering Research Center, USC students and faculty built Dodona, a miniaturized satellite (CubeSat) the size of breadbox, and tested Lockheed Martin's newest payload technologies, including new software that allows for rapid mission changes while in orbit. This technology is part of Lockheed Martin's La Jument program and Dodona is the first in a series of demonstration flights.

The process of developing and launching a satellite requires vigorous design work and testing, including "shake and bake" testing, which literally checks how the satellite fares when shaken and heated to simulate launch.

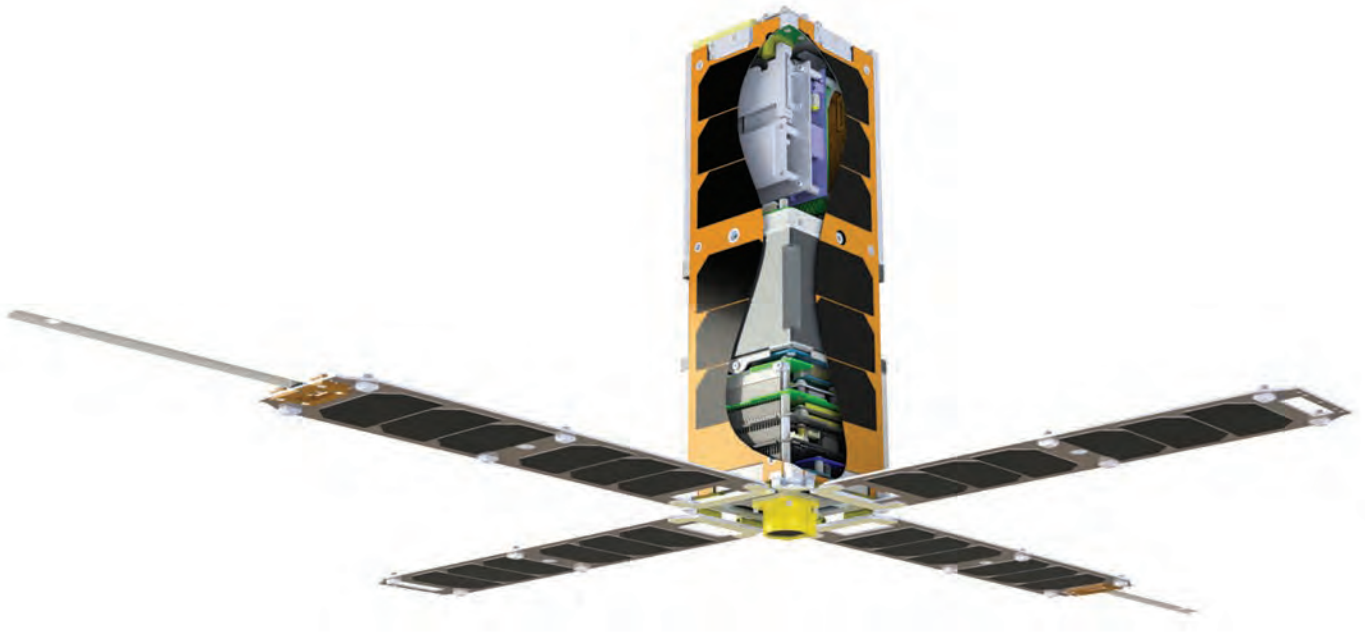
In addition to Lockheed's new technology suite, the SERC team tested a new "B dot" controller—a new guidance control algorithm that interacts with the Earth's magnetic field to stabilize the satellite in orbit.

Satellites must adhere to global regulations, including making sure the team can bring the satellite down within an agreed-upon period of time and ensuring the satellite's orbit will not interfere with other satellites. The team also needed to receive FCC approval for operating Dodona's small radio, an off-the-shelf radio made for CubeSats, so the satellite can receive commands sent up from the ground station and downlink data on the health and status of the satellite while in orbit.

Dodona is the first satellite that requires advanced guidance and control for an optical payload. In another first, SERC commands the satellite from a USC ground station.

SERC launched its first satellite, Careus—developed in partnership between USC and Northrop Grumman—in 2010, and its second, Aeneas, in 2012.

Contact: David Barnhart, barnhart@isi.edu



RESEARCH HIGHLIGHTS

Advancing Data-Intensive Research in Craniofacial Development and Dysmorphology

Craniofacial abnormalities such as cleft lip and palate are among the most common birth defects. To better understand the underlying causes of craniofacial syndromes and uncover potential therapies, dental and craniofacial researchers employ a range of data-intensive techniques. These techniques include imaging, microscopy, genomics, and genetics, involving human subjects and animal models of craniofacial development and dysmorphology.

All types of sciences increasingly depend on large volumes of data. However, the limited capacity to "curate" data—collect, organize, describe, and then disseminate results—impedes scientific progress.

The FaceBase Consortium, funded by the NIH's National Institute of Dental and Craniofacial Research, was established in 2009 in recognition that dental and craniofacial research is increasingly data-intensive. Initially, FaceBase operated as a "hub and spoke" consortium with approximately ten "spoke" research projects generating data on critical aspects of craniofacial development and contributing datasets and online resources.

The consortium moved to a new approach, "self-curation," where the consortium hub focuses on developing tools and infrastructure to empower scientists to curate and publish their data with minimal training and oversight from hub curators. In 2019, FaceBase began an ambitious experiment in data curation—opening the hub to the community and employing self-curation methodology with the research community at large.

Training a broad base of researchers to curate data to a level of quality that is reusable and reproducible had been generally considered unattainable. But researchers in the Informatics Systems Research Division have advanced the cyberinfrastructure and socialized best practices for research data stewardship to ensure robust and reproducible research methods among the international community of craniofacial biologists served by the FaceBase Consortium. The research team has developed automated pipelines for dataset quality control, online image preparation, and single-cell transcriptomics visualization.

In the two years of the open community model, FaceBase has doubled the number of contributing research projects to nearly 40, compared to the previous ten years.

In 2021, FaceBase passed the milestone of 1,000 datasets. The datasets offer the research community over 40,000 files of imaging, genomic, and other data generated from 2,000 experiments involving more than 10,000 biological samples. It is a diverse collection of rich information for designing experiments and analyses into the causes and treatments for craniofacial abnormalities. More than 7,000 scientists accessed the FaceBase website in the last six months and downloaded 41,000+ data files for offline analysis. These data are a vital resource for the international craniofacial research community to drive scientific discovery and translate research results into patient outcomes.

Contact: Rob Schuler, schuler@isi.edu

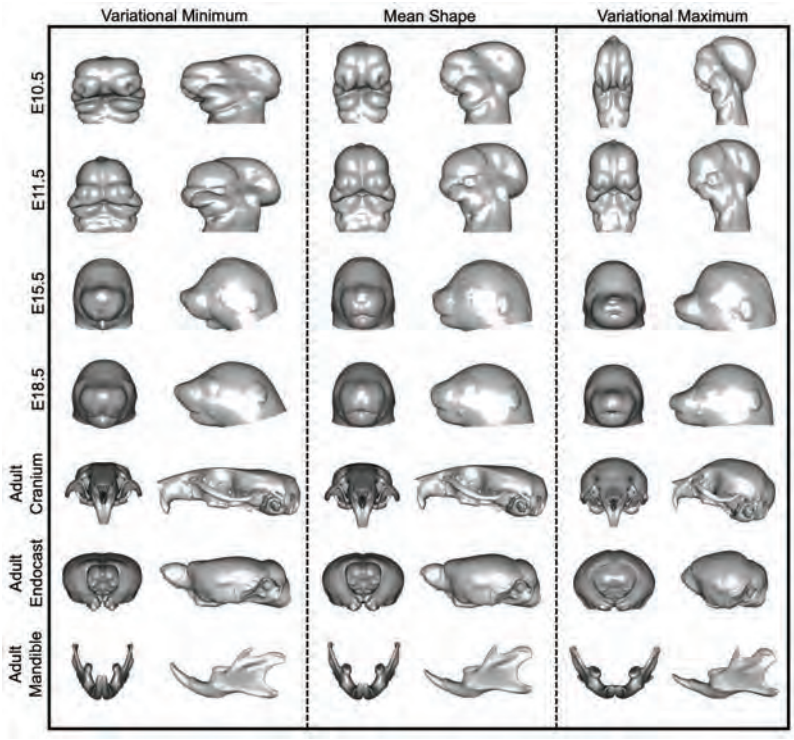


Figure from FaceBase's new "MusMorph" (mouse morphology) dataset: a large collection of high fidelity imaging data, labeled with biological and phenotypic characteristics available for driving new discovery using artificial intelligence and machine learning methods

Designing Next-Generation Quantum Annealers to Speed Up Discrete Optimization

The FluQS (Flux-Based Quantum Speedup) project was funded by IARPA and DARPA to design and fabricate the next generation of quantum annealers. Researchers from government, industry, and academia, based in the US, Europe, and Japan, collaborated on the five-year project.

Quantum annealers are special-purpose quantum computing devices for exploiting quantum mechanical effects such as superposition, tunneling, and entanglement, with the goal of speeding up the search for high-quality solutions in discrete optimization problems. Discrete optimization problems have many applications, including logistics, machine learning, drug discovery, network design, and manufacturing. These problems are notoriously hard to solve. Even a small improvement in the quality of their solutions can have an economic impact worth millions of dollars.

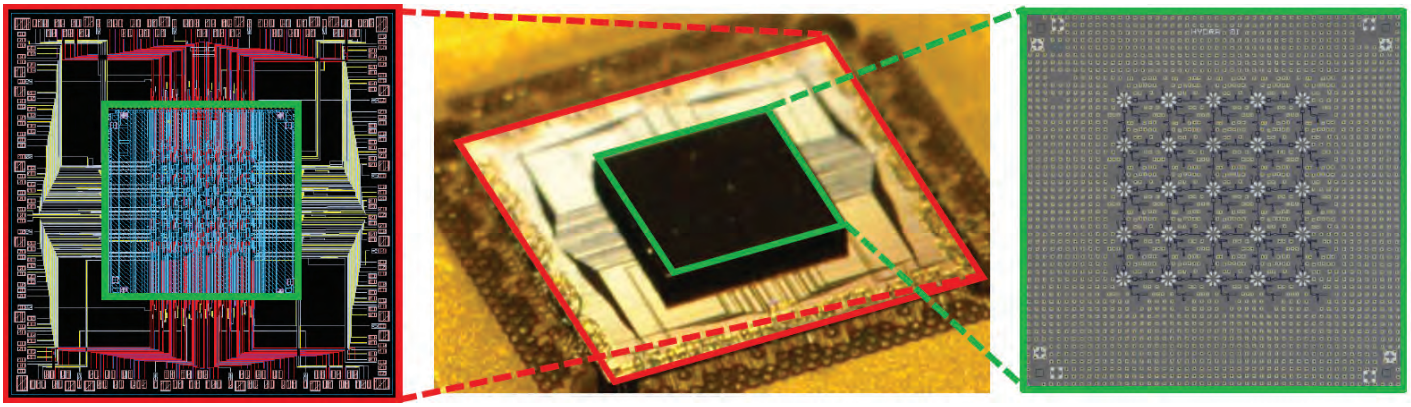


Image courtesy of MIT Lincoln Laboratory and Northrup Grumman Corporation

The project team designed and fabricated quantum annealers that use the magnetic fluxes generated by a circuit of superconducting devices to encode and process information in order to find good solutions for discrete optimization problems. One key feature of quantum information is its fragility: any unwanted interaction with an uncontrolled environment degrades quantum superpositions and denies any potential computational advantage. The FluQS testbed achieved a higher degree of coherence than similar devices, which means quantum effects are sustained for longer times, enabling more precise computations.

The testbed was designed and fabricated in collaboration with Northrop Grumman and MIT Lincoln Laboratory. It consists of 25 qubits arranged on a square lattice chip (plus packaging and control equipment). The FluQS testbed will enable researchers to design new algorithms and develop insights into the ultimate power of quantum computing. The testbed is currently being used in multiple DARPA projects that are investigating the potential of quantum annealing for applications in machine learning and simulating quantum systems.

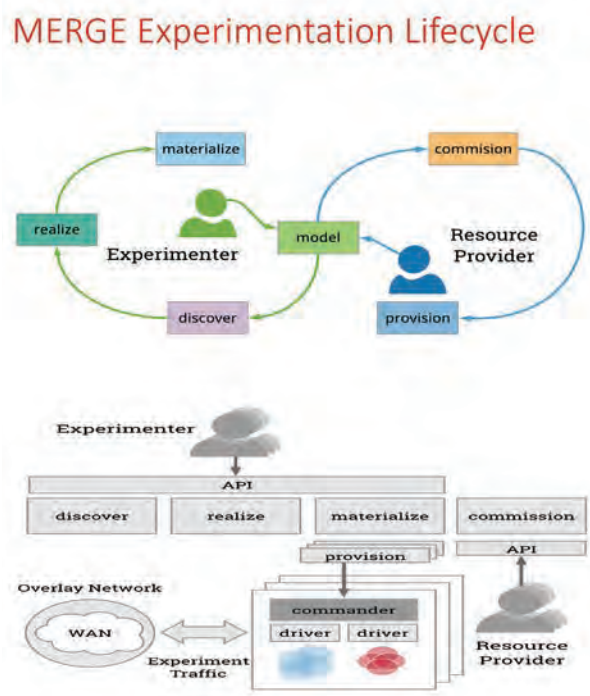
Contact: Federico Spedalieri, fspedali@isi.edu

RESEARCH HIGHLIGHTS

Mod DETER: Modernizing Cyber Research Experimentation in DETERLab

The DETERLab cybersecurity testbed was started by ISI in 2003 and continues to be operated and enhanced by ISI. DETERLab came from the DETER Project, the research program focused on advancing the science of cybersecurity and funded by NSF, DHS, DoD, and DARPA. The lab was built with the broader research community in mind—academia, industry, and government—to bring about greater synergy, including education across intersecting areas of disciplines. (DETERLab is described earlier in this report's Centers section.)

While DETERLab has a long history and established user base, it is critical to ongoing cybersecurity research and education that DETERLab continues beyond its initial project lifecycle and that its software and hardware align with the needs of researchers today and in the future. With funding from the National Science Foundation (CCRI program) and support for new equipment from the Army Research Office (DURIP program), the Mod DETER project is modernizing DETERLab hardware and software to bring advances in several directions. Hardware modernization will increase the lab's capacity and capabilities. Modernizing its management software with the new ISI-developed Merge software will provide support for large-scale, virtualized experiments with reconfigurable topologies. These substantial enhancements will enable Mod DETER to accommodate more researchers and support larger-scale experiments. The Mod DETER effort will also streamline the experimentation process with integrated traffic generators, orchestration software, and malware experimentation environments.



The impetus for modernizing DETERLab software came from the NSF-funded Cybersecurity Experimentation for the Future (CEF) project, exploring current and expected experimentation infrastructure needs. CEF discovered that multi-organizational cybersecurity experimentation capabilities are key to supporting tomorrow's research. To address this need, ISI developed the Merge testbed management software. Merge is a modular platform that provides the ability to conduct experimental research on networked systems using a code-driven approach to experiment development. It enables experimentation ecosystems to catalyze cybersecurity research across a broader spectrum of distributed, networked facilities.

Merge software has powered DARPA's DCOMP testbed for its Dispersed Computing program since April 2019. DCOMP running Merge supports up to 1,400-node experiments. Additionally, DARPA's LightHouse testbed, developed for the DARPA Searchlight program for enterprise QoS running Merge software, commenced operation in 2021. This notable progress, along with user acceptance, has paved the way for DETERLab to replace its legacy testbed management software with Merge.

Mod DETER will launch a beta version in early 2022, focusing on transitioning Classic DETERLab users to the new user interfaces and experiment paradigm. The transition will introduce automated support along with webinars and training sessions. The DETER team is actively soliciting volunteers to participate in the beta release. The team anticipates that the official Mod DETER testbed launch will occur in mid-2022. ISI will continue to operate Classic DETERLab in parallel with the Mod DETER transition to ensure its user base has minimal disruption to their research and time to learn and engage with the new platform. The DETER Education Cluster will continue to operate through 2022.

Mod DETER will bring experimentation templates and build capabilities that will offset the asymmetric paradigm of cyberspace protectors, enabling greater preparedness, higher assuredness, and better-fielded solutions—all essential to further broadening and diversifying DETERLab's user base.

To learn more about Merge, visit https://mergetb.org.

Contact: Terry Benzel, tbenzel@isi.edu

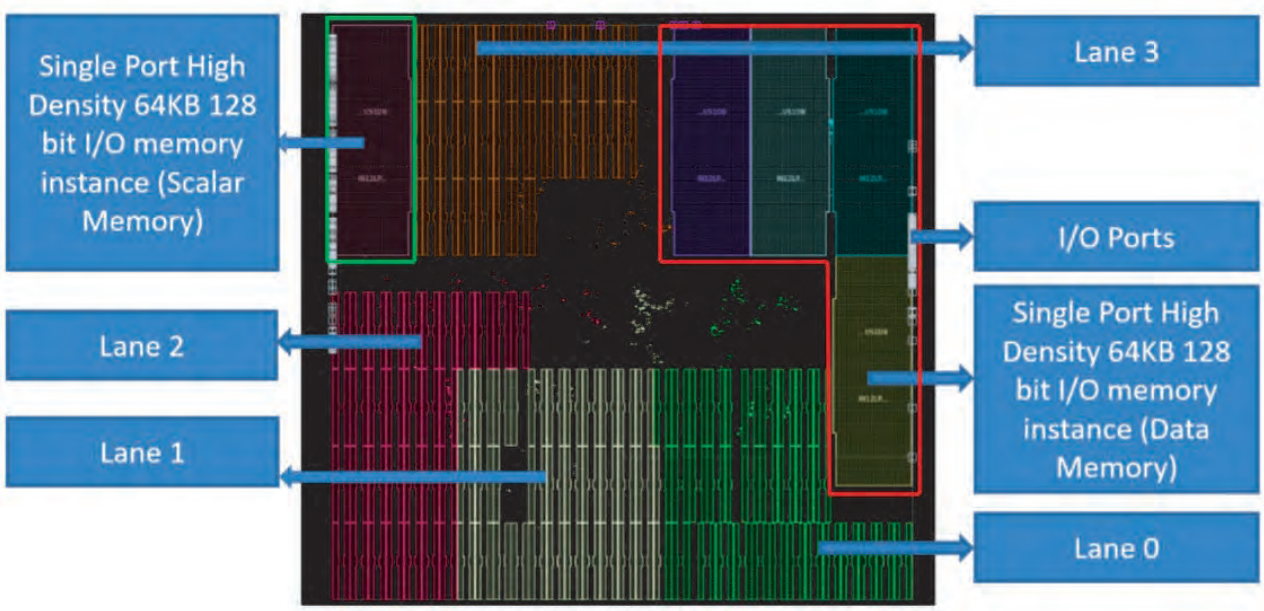
1,120 legacy cables pulled; 736 new cables installed

Data Protection in Virtual Environments

ISI researchers in the Computational Systems and Technology Division are engaged with a team of researchers and engineers to develop the world's first custom Fully Homomorphic Encryption (FHE) hardware accelerator. FHE enables computation directly on encrypted data without the need to decrypt and re-encrypt, thus providing a substantially higher degree of security and privacy. Although FHE offers better security because data remains encrypted, it requires a prohibitive amount of compute power and time.

This work is supported by the DARPA Data Protection in Virtual Environments (DPRIVE) program, which seeks to reduce the compute runtime of current software-based FHE approaches by 10,000x. In addition to the ISI researchers, the Trebuchet team includes researchers from USC's Ming Hsieh Department of Computer and Electrical Engineering, Duality Technologies, New York University, Carnegie Mellon University, and Two Six Technologies.

FHE requires native data sizes of 128 to 4096 bits. When these algorithms are implemented on traditional 32- or 64-bit computer architectures, they incur substantial overhead in basic computations and data movement because each operation must be parsed into multiple, smaller operations with additional control instructions to aggregate the results. The Trebuchet team's approach is to develop a computer architecture that natively operates on, stores, and moves data at FHE native word sizes in order to eliminate these bottlenecks. The approach centers around the co-design of every ASIC development stage—from algorithms to transistors—to achieve optimal efficiency.



Floorplan of tile with four HPLE lanes and five instances of 64KB 128-bit I/O single port high-density SRAM

The ISI Trebuchet team is implementing a design space exploration electronic design automation (EDA) tool suite, called Fulcrum, that enables co-design of FHE algorithms, software, computer architecture, and FHE optimized hardware components. In addition they are investigating the use of novel non-volatile memories (MRAM) and novel SRAM designs to reduce area footprints by up to 2x and inventing new verification methodologies that scale to support large bit-widths.

The USC Trubechet team is researching Large Algorithmic Word Size modulo multipliers that can support 128- to 4096-bit-width operations in efficient, high performance, and scalable manners.

The Trebuchet architecture is coarse-grained, reconfigurable arrays consisting of a 2D array of tiles. The tile design enables flexibility to rapidly pivot to different implementations ranging from large data-center devices to smaller area- and power-efficient edge computing devices. The chief computation is performed in high performance LAWS engines (HPLE), or lanes, which can be scaled according to the size of the problem and available footprint. The DPRIVE program is targeting a 150mm2 device in GF 12nm, and initial benchmarking is demonstrating over 2,000 128-bit giga operations per second.

Contact: Matthew French, mfrench@isi.edu

RESEARCH HIGHLIGHTS

Developing and Measuring Machine Intelligence from Common Sense

Although expected in human intelligence, common sense is lacking in machine intelligence. If a recommender system, decision aid, or autonomous car suggests something that defies common sense, we are not likely to trust it—it may even seem laughable. But human common sense can help us when learning a challenging task; we need only learn what is specific to the new task.

How can an AI system acquire common sense? Encoding common sense has been attempted since the mid-80s, but despite vast manual effort, encoding common sense does not approach what a human learns naturally. Deep neural nets have recently offered a new approach to capture knowledge automatically in an *embedding space*, in which meaning is represented in a high-dimensional vector space where the "distance" between two vectors measures semantic similarity.

An ISI research team, working with researchers from Columbia University, University of California Los Angeles, and University of Massachusetts Amherst, is devising algorithms that mine the Web for common sense from text, video, images, and knowledge bases, capturing knowledge in an embedding space.

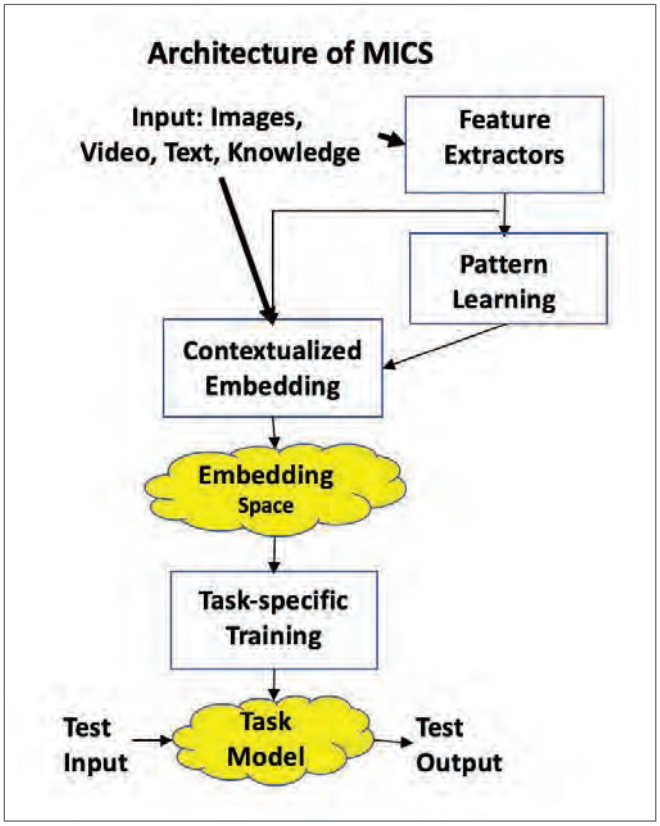
One focus of ISI researchers is developing a method of measuring and displaying the value of machine common sense. Essentially, they present an entry-level university textbook on social science (*American Government*) or the humanities (*U.S. History*) to the Machine Intelligence from Common Sense (MICS) system to "read." The MICS system then answers true-or-false questions, for example: "Is the following true or false: The goal of the Tea Act of 1773 was to help revive the struggling East India Company."

Test materials are available via a leaderboard on the Web so that any researcher can download a textbook, provide it to their AI system, have their system take the test, and receive a score. The leaderboard displays these measurements:

- Content in the textbook that is prior knowledge of the researcher's system
- How much knowledge is assimilated by the researcher's system after reading the textbook once
- How performance is improved if the researcher's system can access the textbook during the test

The researcher can then compare their system's results against other systems to see how theirs performs, absolutely and relative to others. Since the test is not downloadable and cannot be viewed by software or researcher, a system's performance can be tracked across system changes. Thus far, when the textbook is not accessible during the test, no system performs significantly better than chance. However, when systems can access the textbook during the test, scores are in the low 70s.

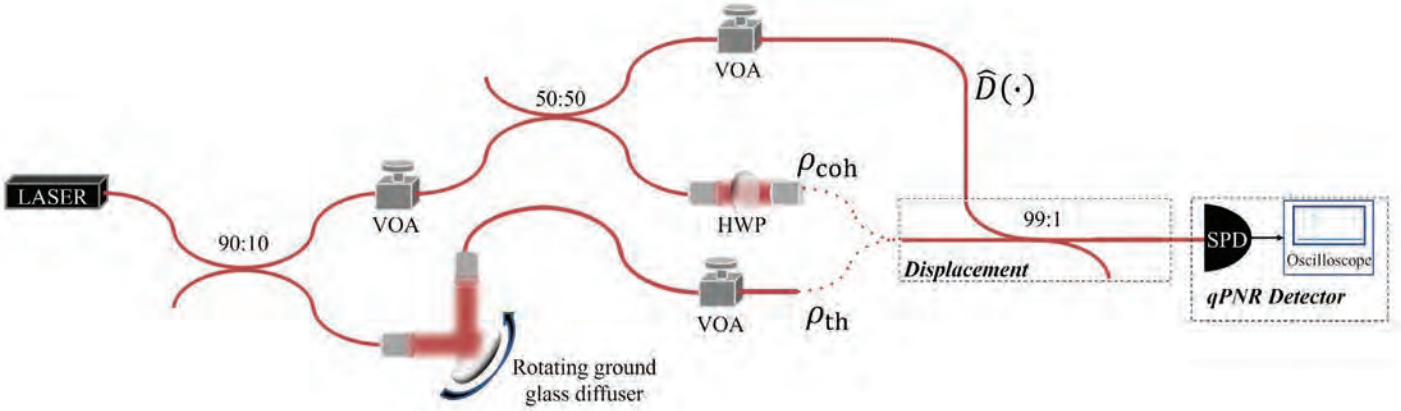
The newest challenge is testing whether a system with common sense can say: "I'm not sure" and provide a confidence estimate on its answer. To assess the quality of confidence estimates, the ISI researchers adopted the Brier score, which is used to measure the accuracy of forecasts. Each possible outcome is assigned a confidence estimate. After the correct answer is known, the Brier score measures the gap between confidence-weighted and correct outcomes. The ISI researchers created a leaderboard for confidence estimates on one type of task, filling a number in a blank such as "An octopus typically has ____ legs." The researchers are also exploring adding the evaluation of confidence estimates to additional benchmark tests of machine common sense.



Contact: Ralph Weischedel, weisched@isi.edu

Reaching the Quantum Limits for Sensing with Optical Sensing

Sponsored by the Army Research Office, researchers at ISI's Laboratory for Quantum-Limited Information (QLIlab) carried out the first demonstration of a quantum measurement to achieve the quantum Chernoff bound for state discrimination. The quantum Chernoff bound represents a fundamental quantum limit to the sensitivity of an optical sensor.



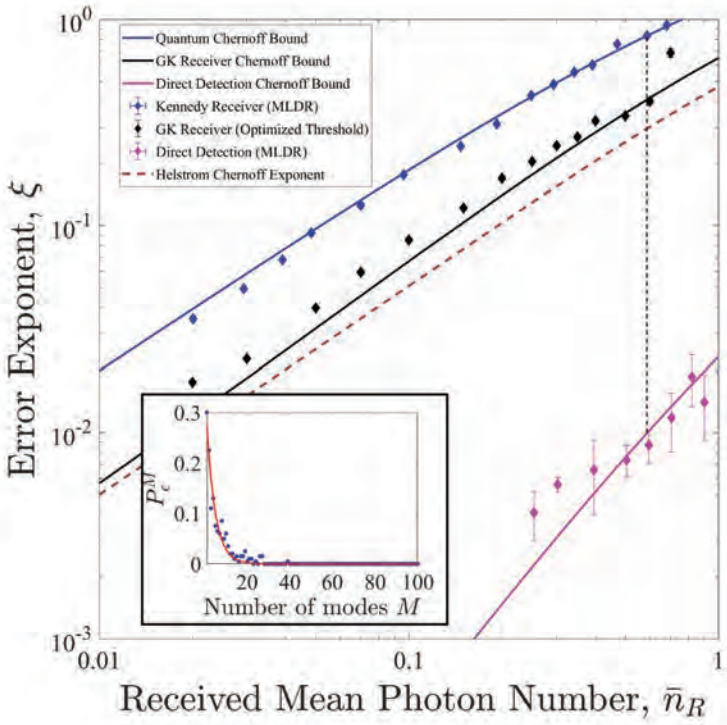
The experimental setup in the QLIlab demonstrates the most sensitive quantum measurement reported for discriminating between quantum laser states and quantum states from natural sources

The measurement of quantum states has been notoriously challenging, as it requires exquisite sensitivity in the measurement apparatus. Although researchers have worked for decades on precision measurements of quantum states, there has been little work investigating the measurement of quantum states found in the real world, such as sunlight. Understanding the full complexity of quantum states requires using quantum theory because the equations from classical physics are insufficient.

Working with physicists from the University of Arizona and MIT, the QLIlab team built a quantum measurement experiment that discriminates quantum pure states (states of light emitted from a laser) from mixed states (light found in the environment) at the fundamental limit of sensitivity allowed by quantum theory.

In concert with this experimental work, the team presented a foundational theory for quantum measurement bounding the performance of discrimination when a single state or many states are available for measurement.

Contact: Jonathan Habif, habif@usc.edu



Data shows experimental validation of the measurement to achieve the quantum Chernoff bound

RESEARCH HIGHLIGHTS

Scalable All-Optical Processing

Data flows through optical fiber networks at blazing speeds approaching terabit-per-second (Tbps) speeds. To perform any type of data processing on the data flowing through the optical fibers, however, the data must be funneled into a computer, a huge bottleneck.

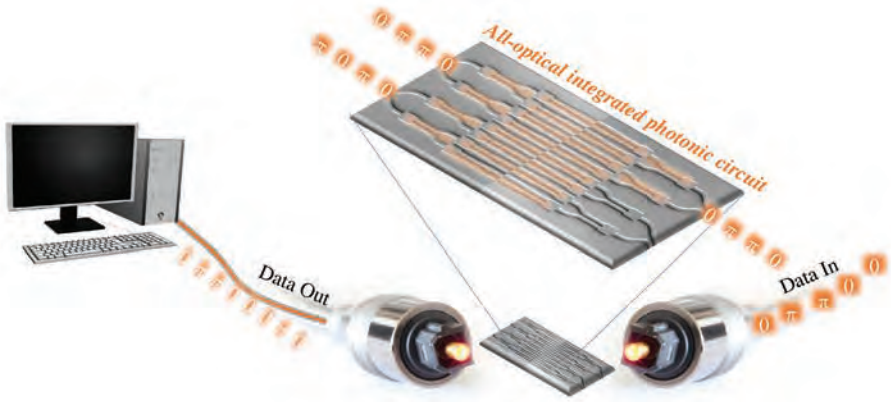
Researchers in the Computational Systems and Technology Division are designing, building, and demonstrating circuits that can perform data processing functions directly on optical data as it flows through fiber optical links.

If successful, this capability will alleviate the bottleneck associated with moving data into and out of computers. The ISI physicists are collaborating with Professors Alan Willner and Hossein Hashemi from USC'S Ming Hsieh Department of Electrical and Computer Engineering and a group of graduate students.

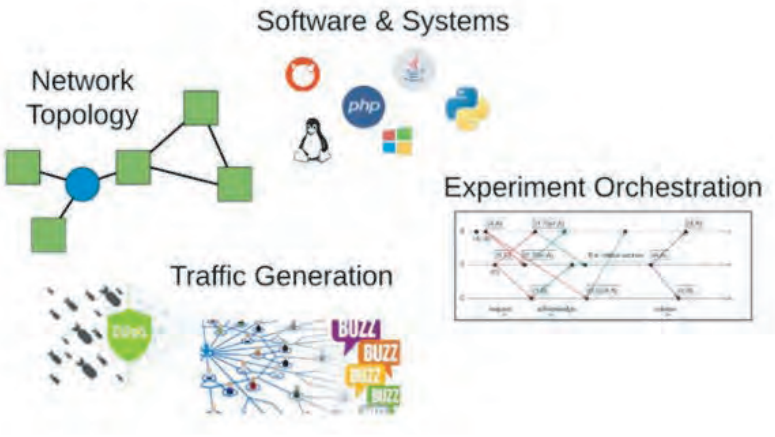
The concept for the optical systems was developed at ISI. Professor Willner's team designed the photonic systems and validated them on the benchtop on campus.

Professor Hashemi's group transitioned the optical designs into silicon photonic integrated circuits fabricated at Tower Semiconductor. These circuits are currently in test.

Contact: Jonathan Habif, jhabif@isi.edu



Can Optical Interconnect Technologies Reduce Data Center Power Consumption?



Data centers are consuming increasing amounts of energy, in the US and the world. ISI's STEAM project (System Testbed, Evaluation, and Architecture Metrics) aims to reduce data energy consumption by a factor of two, focusing on optical interconnects and their potential to increase data center efficiency.

With support from the Advanced Research Projects Agency-Energy (Department of Energy), the STEAM research team is building a framework and testbed to model the effect of optical interconnects at the data center level. These optical interconnects are enabled by photonic technologies developed by Columbia University, IBM, the University of California, Santa Barbara, and the University of California San Diego.

In the last two years, the STEAM project showed that this energy reduction goal is feasible by increasing throughput through improved interconnect performance.

Next, the team will model new workloads. The researchers will perform large-scale simulations of thousands of nodes based on component characteristics derived from hardware emulations and more detailed smaller-scale simulations to demonstrate that data center energy consumption can be reduced using new photonic interconnect technologies.

Contact: Stephen Crago, crago@isi.edu

A Knowledge-Centric Approach to Reproducible Research

The scientific method, positing hypotheses and evaluating them through experimentation, has been the foundation of scientific progress. However, this progress hinges on the ability to trust the results of prior scientific research and build upon those results.

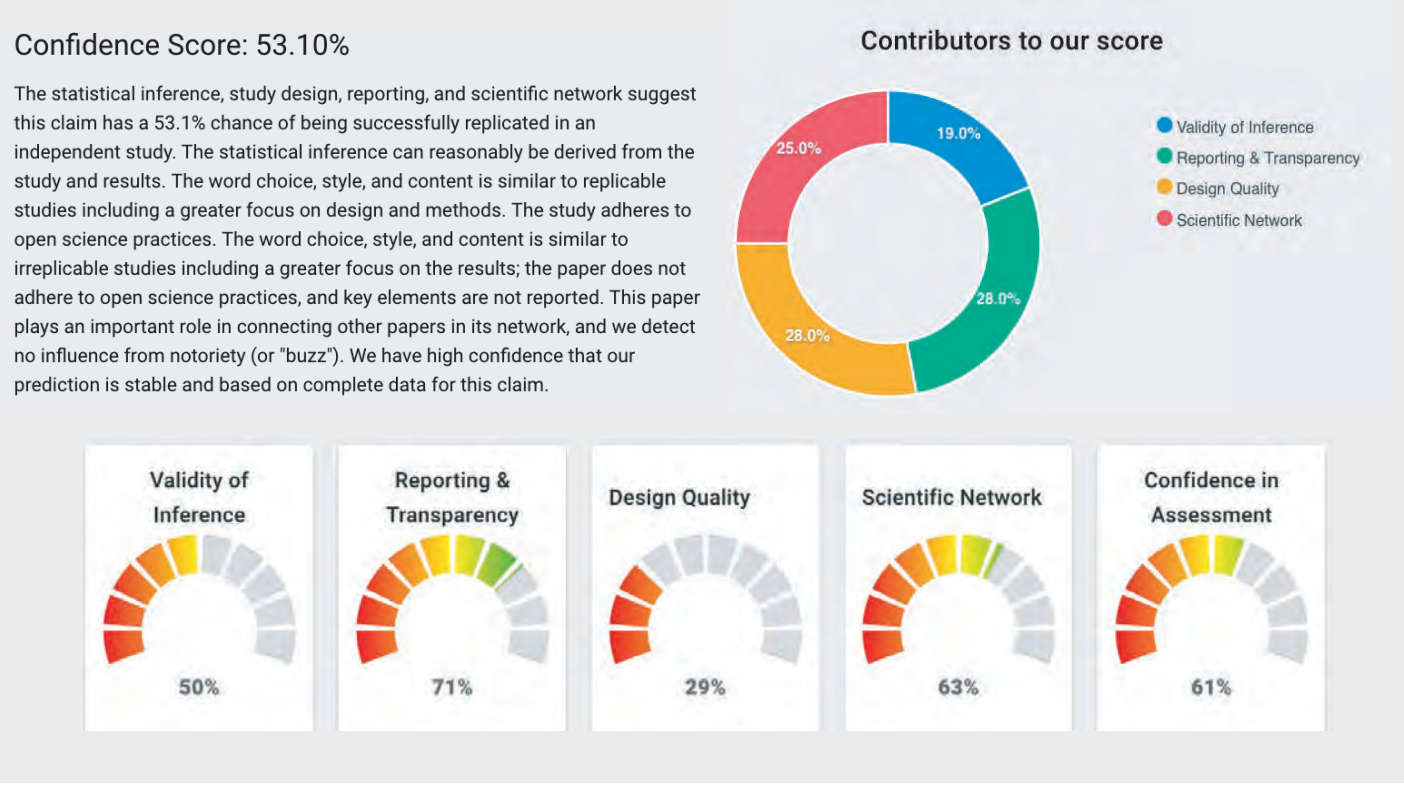
The difficulty in reproducing past results and the volume of new findings have led to a crisis in scientific research. As part of DARPA's Systematizing Confidence in Open Research and Evidence (SCORE) project, ISI artificial intelligence researchers developed a knowledge graph to help decision-makers (for example, scientists, analysts, or journalists) assess scientific research. The system automatically scores scientific articles in journals and conferences and gives a confidence score that predicts whether the experimental outcomes reported in these articles can be reproduced.

The system, named MACROSCORE, combines features that include "micro" level experimental details and "macro" level relationships with the scientific community, such as citations and references. Micro features comprise detailed experimental parameters and outcomes such as sample size, statistical models, statistical significance, and mentions of open data, open materials, and funding disclosures. Macro features use hundreds of millions of citations and references to characterize the breadth and depth of the scientific communities that the paper impacts.

The micro and macro features are fed to advanced multimodal representation learning techniques and probabilistic models, which then provide a set of explainable prediction scores.

An easy-to-understand dashboard breaks down prediction scores based on reference validity, open science practices, methodological design, and the related scientific community. The dashboard also displays a meta-score: the MACROSCORE model's confidence in its prediction score.

Contact: Jay Pujara, jpujara@isi.edu



RESEARCH HIGHLIGHTS

Artificial Intelligence to Create Complex Models of the Effects of Climate Change

Predicting the effects of extreme climate events requires integrating models of natural processes (e.g., climate, hydrology, agriculture) and human processes (e.g., planting, commodity markets) that are interrelated and complex. It typically takes many months, and even years, to create the integrated models needed to understand such complex systems.

A team of ISI researchers is building a novel framework that uses AI to make modeling more efficient and useful for decision making.

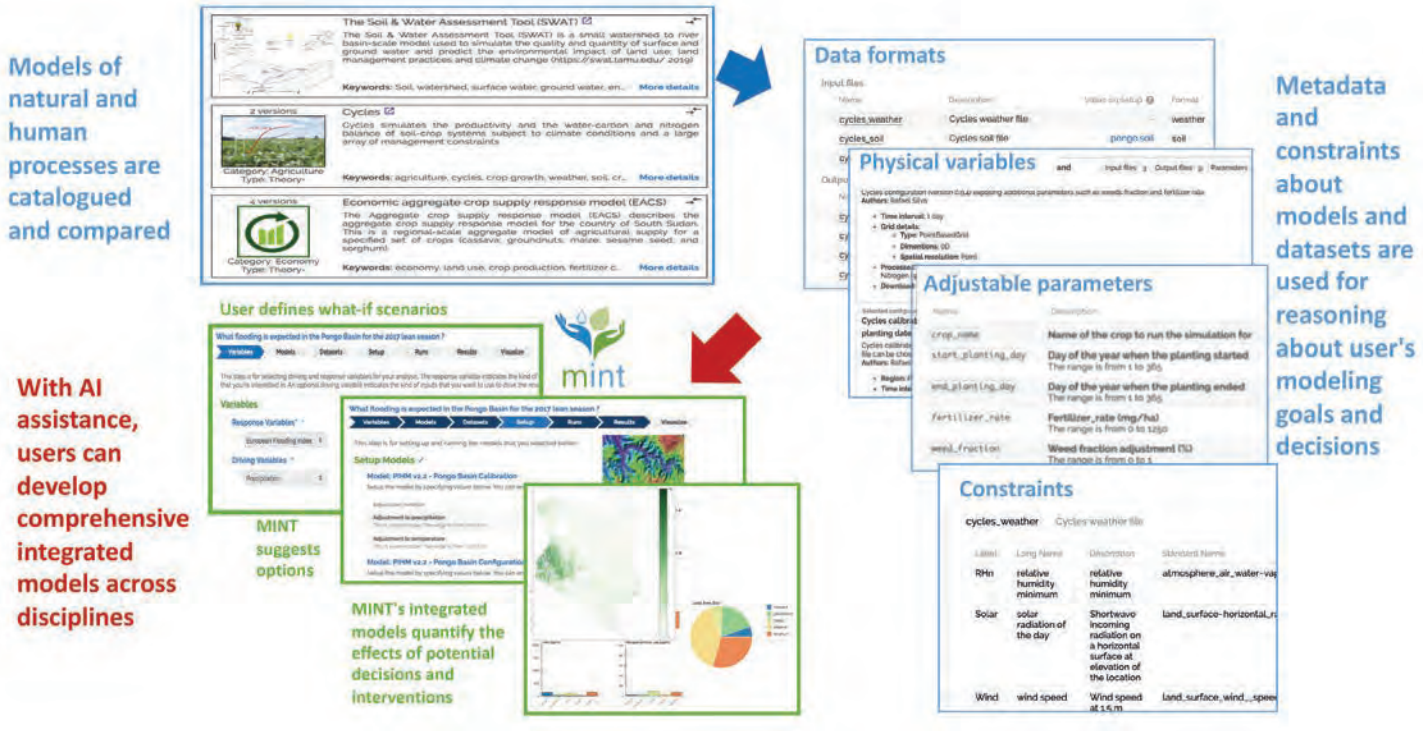
The MINT framework uses semantic technologies to represent knowledge about model versions and settings, execution requirements, and key model variables that enable the use of modeling goals to constrain modeling choices. By reasoning about that knowledge, this framework guides users to pose well-formed modeling questions, specify possible scenarios, find appropriate models and data, set up model parameters, run simulations, and understand new findings.

By integrating meteorological, hydrological, and agricultural models, ISI researchers are using the system to understand the effects of climate change on food insecurity in Sub-Saharan Africa, particularly in Ethiopia, Kenya, and South Sudan.

In addition, the researchers are using this system to study the impact of natural disasters such as flooding and droughts in different regions of the United States.

The AI techniques under development are also being applied to wildfires. As climate change leads to more uncontrolled wildfires, the damage could be reduced by pre-treating vegetation with controlled fires. These techniques allow non-experts to use science-grade fire models to develop simulations quickly. The goal is to increase the number of controlled fires that planners can manage by tenfold.

Contact: Yolanda Gil, gil@isi.edu



Exploiting Network Asynchrony to Disrupt the Bitcoin Protocol

As the use of cryptocurrencies and non-fungible tokens (NFTs) grows, it is essential to understand the security vulnerabilities that could be exploited. Researchers in the Networking and Cybersecurity Division are studying the inherent flaws that attackers might exploit.

Bitcoin blockchain safety relies on strong network synchrony (no delay in sending data from node to node). But the team's research shows that network synchrony does not hold in the real-world Bitcoin network. Currently, strong adversaries who control a mining pool with ~51% hash rate could violate blockchain safety. However, the asynchrony could be exploited to lower the hash rate required to violate blockchain safety and chain quality.

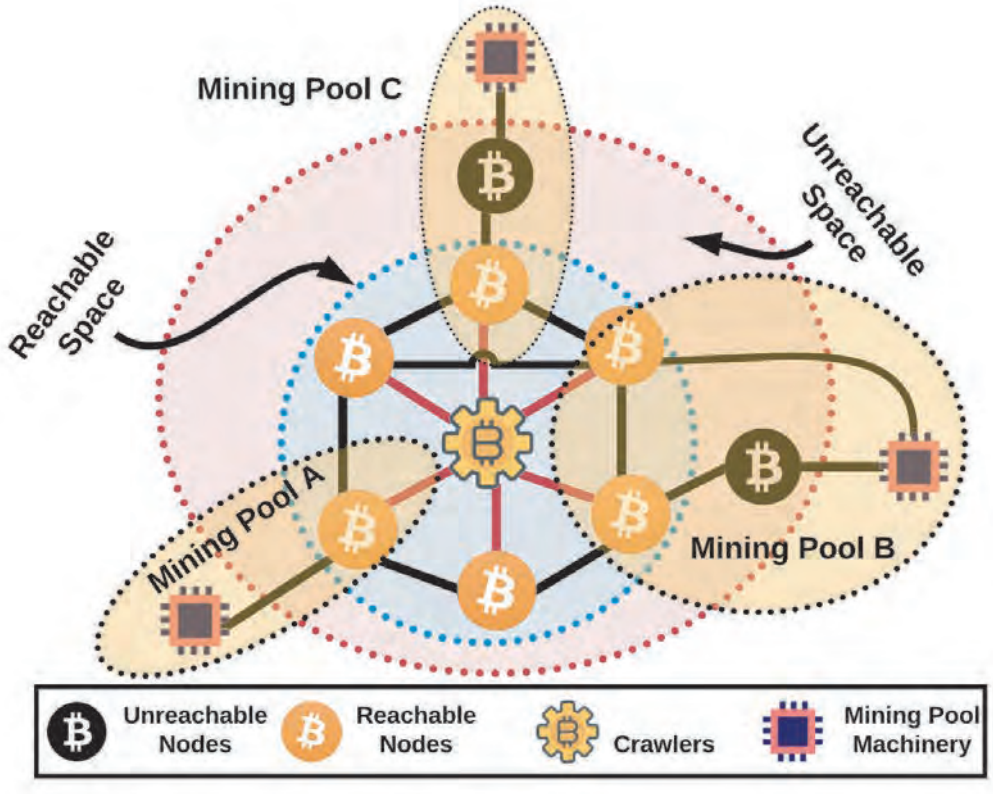
The researchers constructed the Bitcoin ideal-world functionality to formally specify its ideal execution model in a synchronous network. They then developed a large-scale data collection system through which they connected with more than 36,000 IP addresses of Bitcoin nodes and identified 359 mining nodes. They observed a non-uniform block propagation pattern among the mining nodes, showing that the Bitcoin network is asynchronous in practice, with anomalies that could be exploited.

To demonstrate the threat of an asynchronous network, the team developed the HashSplit attack, which allows an adversary to orchestrate concurrent mining on multiple branches of the blockchain—violating common address-prefix and chain-quality properties that ensure the integrity of the blockchain protocol.

The team also proposed countermeasures to such attacks by releasing a Bitcoin Core version that closely models the Bitcoin ideal-world functionality.

The research team's measurements, theoretical modeling, HashSplit attack, and countermeasures open new directions for evaluating the security of blockchain systems such as Bitcoin cryptocurrency

Contact: Srivatsan Ravi, srivatsr@usc.edu



PUBLICATIONS

COMPUTING ARCHITECTURES AND SYSTEMS

Integrated Constellation Analysis Tools to Support New Observing Strategy Mission Design
C. D. Ball, I. J. Tapia-Tamayo, M. Paolieri, A. J. O'Brien, M. French, J. T. Johnson, and P. Grogan
CubeSats and SmallSats for Remote Sensing V

Serverless Containers - Rising Viable Approach to Scientific Workflows
K. Burkat, M. Pawlik, B. Balis, M. Malawski, K. Vahi, M. Rynge, R. F. d. Silva, and E. Deelman
2021 IEEE 17th International Conference on eScience

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

Canary: An FPGA Assurance Plugin for Vendor EDA Tools
D. Glick, A. G. Schmidt, J. Nifong, T. Haroldsen, J. Monson, E. M. Ruiz, and M. French
Government Microcircuit Applications and Critical Technology Conference (GOMACTech)

Independent Testing of Untrusted FPGAs for Faulty Interconnnect
T. Haroldsen, M. French, T. Sung, D. Glick, J. Danner, and L. Lerner
Government Microcircuit Applications and Critical Technology Conference

Distributed and Heterogeneous SAR Backprojection with Halide
C. Imes, T. Li, M. Glines, R. Khan, and J. P. Walters
2021 IEEE High Performance Extreme Computing Conference (HPEC)

Design and Performance Evaluation of Multispectral Sensing Algorithms on CPU, GPU, and FPGA
V. V. Menon, S. A. Siddiqui, S. Rao, A. G. Schmidt, M. French, V. Chirayath, and A. Li
IEEE Aerospace Conference

Fight Club: Maturing Defense in Depth Obfuscation Techniques
V. V. Menon, U. Sharma, S. Roshanisefat, S. S. Shukla, A. G. Schmidt, M. French, P. A. Beerel, and P. Nuzzo
Government Microcircuit Applications and Critical Technology Conference (GOMACTech)

End-to-End Online Performance Data Capture and Analysis for Scientific Workflows
G. Papadimitriou, C. Wang, K. Vahi, R. F. d. Silva, A. Mandal, L. Zhengchun, R. Mayani, M. Rynge, M. Kiran, V. E. Lynch, R. Kettimuthu, E. Deelman, J. S. Vetter, and I. Foster
Future Generation Computer Systems

Demonstration of a Fully Neural Network Based Synthetic Aperture Radar Processing Pipeline for Image Formation and Analysis
A. Rittenbach and J. P. Walters
Sensors, Systems, and Next-Generation Satellites XXV

Detection and Localization of Near Infrared Lasers from Atmospheric Laser Scattering
A. Rittenbach, C. Finnerty, and J. L. Habif
Emerging Imaging and Sensing Technologies for Security and Defence VI

Root Cause Analysis of Data Integrity Errors in Networked Systems with Incomplete Information
Y. Xin, S. Fu, A. Mandal, I. Baldin, R. Tanaka, M. Rynge, K. Vahi, E. Deelman, I. Abhinit, and W. Von
2021 International Conference on Information and Communication Technology Convergence (ICTC)

CYBERSECURITY

Optimal Concurrency for List-Based Sets
V. Aksenov, V. Gramoli, P. Kuznetsov, D. Shang, and S. Ravi
16th International Conference on Parallel Computing Technologies (PaCT)

Building Reproducible Video Streaming Traffic Generators
C. Ardi, A. Hussain, and S. Schwab
Proceedings of 14th Cyber Security Experimentation and Test Workshop (CSET)

PyCT: A Python Concolic Tester
Y. T. Chen
Programming Languages and Systems

Agenda Pushing in Email to Thwart Phishing
H. Cho, G. Bartlett, and M. Freedman
Proceedings of the 1st Workshop on Document-Grounded Dialogue and Conversational Question Answering (DialDoc2021)

Mini-Me, You Complete Me! Data-Driven Drone Security via DNN-Based Approximate Computing
A. Ding, P. Murthy, L. Garcia, P. Sun, M. Chan, and S. Zonouz
24th International Symposium on Research in Attacks, Intrusions, and Defenses

Case Studies in Experiment Design on a Minimega Based Network Emulation Testbed
B. Kocoloski, A. Hussain, M. Troglia, C. Ardi, S. Cheng, D. DeAngelis, C. Symonds, M. Collins, R. Goodfellow, and S. Schwab
Proceedings of 14th Cyber Security Experimentation and Test Workshop (CSET)

Aerogel: Lightweight Access Control Framework for WebAssembly-Based Bare-Metal IoT Devices
R. Liu, L. Garcia, and M. Srivastava
The Sixth ACM/IEEE Symposium on Edge Computing (SEC '21)

Multidisciplinary publications appear in all relevant areas.

SecDeep: Secure and Performant On-Device Deep Learning Inference Framework for Mobile and IoT Devices
R. Liu, L. Garcia, Z. Liu, B. Ou, and M. Srivastava
Proceedings of the International Conference on Internet-of-Things Design and Implementation

User Experiences on Network Testbeds
J. Mirkovic and P. Pusey
Proceedings of 14th Cyber Security Experimentation and Test Workshop (CSET)

Revisiting Nakamoto Consensus in Asynchronous Networks: A Comprehensive Analysis of Bitcoin Safety and Chain Quality
M. Saad, A. Anwar, S. Ravi, and D. Mohaisen
ACM Conference on Computer and Communications Security (SIGSAC)

I Always Feel Like Somebody's Sensing Me! A Framework to Detect, Identify, and Localize Clandestine Wireless Sensors
A. D. Singh, L. Garcia, J. Noor, and M. Srivastava
30th USENIX Security Symposium

Secure Neuroimaging Analysis using Federated Learning with Homomorphic Encryption
D. Stripelis, H. Saleem, T. Ghai, N. Dhinagar, U. Gupta, C. Anastasiou, G. V. Steeg, S. Ravi, M. Naveed, P. M. Thompson, and J. L. Ambite
17th International Symposium on Medical Information Processing and Analysis

Defending Web Servers Against Flash Crowd Attacks
R. P. Tandon
Applied Cryptography and Network Security

PERFUME: Programmatic Extraction and Refinement for Usability of Mathematical Expression
N. Weideman, V. K. Felkner, W. Wu, J. May, C. Hauser, and L. Garcia
Proceedings of the 2021 Research Workshop on Offensive and Defensive Techniques in the Context of Man At The End (MATE) Attacks

DeepSQA: Understanding Sensor Data via Question Answering
T. Xing, L. Garcia, F. Cerutti, L. Kaplan, A. Preece, and M. Srivastava
Proceedings of the International Conference on Internet-of-Things Design and Implementation

AMPPERE: A Universal Abstract Machine for Privacy-Preserving Entity Resolution Evaluation
Y. Yao, T. Ghai, S. Ravi, and P. A. Szekely
The 30th ACM International Conference on Information and Knowledge Management

KNOWLEDGE TECHNOLOGIES

Predicting Youth at High Risk of Aging Out of Foster Care using Machine Learning Methods
E. Ahn, Y. Gil, and E. Putnam-Horstein
Child Abuse and Neglect

BD2K Training Coordinating Center's ERuDIte: The Educational Resource Discovery Index for Data Science
J. L. Ambite, L. Fierro, J. Gordon, G. A. P. C. Burns, F. Geigl, K. Lerman, and J. D. V. Horn
IEEE Transactions on Emerging Topics in Computing

Toward a Fine-Scale Population Health Monitoring System
G. M. Belbin, S. Cullina, S. Wenric, E. R. Soper, B. S. Glicksberg, D. Torre, A. Moscati, G. L. Wojcik, R. Shemirani, N. D. Beckmann, A. Cohain, E. P. Sorokin, D. S. Park, and J. Ambite
Cell

Creating and Querying Personalized Versions of Wikidata on a Laptop
H. Chalupsky, P. Szekely, F. Ilievski, D. Garijo, and K. Shenoy
Wikidata-21 workshop

Viola: A Topic Agnostic Generate-and-Rank Dialogue System
H. Cho, B. Shbita, K. Shenoy, S. Liu, N. Patel, H. Pindikanti, J. Lee, and J. May
Proceedings of the Alexa Prize Socialbot Grand Challenge 4

3D Convolutional Neural Networks for Classification of Alzheimer's and Parkinson's Disease with T1-Weighted Brain MRI
N. J. Dhinagar, S. I. Thomopoulos, C. Owens-Walton, D. Stripelis, J. L. Ambite, G. V. Steeg, D. Weintraub, P. Cook, C. McMillan, and P. M. Thompson
17th International Symposium on Medical Information Processing and Analysis (SIPAIM)

More Computing and Less Programming: A Proposal to Broaden Participation in Data Science
Y. Gil
Harvard Data Science Review

Will AI Write Scientific Papers in the Future?
Y. Gil
AI Magazine

Artificial Intelligence for Modeling Complex Systems: Taming the Complexity of Expert Models to Improve Decision Making
Y. Gil, D. Garijo, D. Khider, C. A. Knoblock, V. Ratnakar, M. Osorio, H. Vargas, M. Pham, J. Pujara, B. Shbita, B. Vu, Y. Chiang, D. Feldman, Y. Lin, H. Song, V. Kumar et al.
ACM Transactions on Interactive Intelligent Systems

Membership Inference Attacks on Deep Regression Models for Neuroimaging
U. Gupta, D. Stripelis, P. K. Lam, P. M. Thompson, J. L. Ambite, and G. V. Steeg
Medical Imaging with Deep Learning (MIDL)

PUBLICATIONS

Socioeconomic Correlates of Anti-Science Attitudes in the US
M. Hu, A. Rao, M. Kejriwal, and K. Lerman
Future Internet

Dimensions of Commonsense Knowledge
F. Ilievski, A. Oltramari, K. Ma, B. Zhang, D. L. McGuinness, and P. Szekely
Knowledge-Based Systems (KBS)

Story Generation with Commonsense Knowledge Graphs and Axioms
F. Ilievski, J. Pujara, and H. Zhang
Workshop on Commonsense Reasoning and Knowledge Bases

CSKG: The CommonSense Knowledge Graph
F. Ilievski, P. Szekely, and B. Zhang
Extended Semantic Web Conference (ESWC)

A Meta-Engine for Building Domain-Specific Search Engines
M. Kejriwal
Software Impacts

Link Prediction Between Structured Geopolitical Events: Models and Experiments
M. Kejriwal
Frontiers of Big Data

Unsupervised DNF Blocking for Efficient Linking of Knowledge Graphs and Tables
M. Kejriwal
Information

A Multi-Agent Simulator for Generating Novelty in Monopoly
M. Kejriwal and S. Thomas
Simulation Modelling Practice and Theory

Knowledge Graphs: Fundamentals, Techniques, and Applications (book)
M. Kejriwal, C. A. Knoblock, and P. Szekely

An Evaluation and Annotation Methodology for Product Category Matching in E-Commerce
M. Kejriwal, K. Shen, C. Ni, and N. Torzec
Computers in Industry

An Empirical Study of Emoji Usage on Twitter in Linguistic and National Contexts
M. Kejriwal, Q. Wang, H. Li, and L. Wang
Online Social Networks and Media

Empirical Best Practices on Using Product-Specific Schema.org
M. Kejriwal, R. K. Selvam, C. Ni, and N. Torzec
35th Conference on Artificial Intelligence (AAAI 2021)

Generating Explainable Abstractions for Wikidata Entities
N. Klein, F. Ilievski, and P. Szekely
Proceedings of the 11th Knowledge Capture Conference

W-TSS : A Wavelet-Based Algorithm for Discovering Time Series Shapelets
K. Li, H. Deng, J. Morrison, R. Habre, M. Franklin, Y. Chiang, K. Sward, F. D. Gilliland, J. L. Ambite, and S. P. Eckel
Sensors

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. Ambite, J. P. Wilson, and S. P. Eckel
Scientific Reports

Synthetic Map Generation to Provide Unlimited Training Data for Historical Map Text Detection
Z. Li, R. Guan, Q. Yu, Y. Chiang, and C. A. Knoblock
Proceedings of the 4th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery

Exploring Strategies for Generalizable Commonsense Reasoning with Pre-trained Models
K. Ma, F. Ilievski, J. Francis, S. Ozaki, E. Nyberg, and A. Oltramari
EMNLP 2021

Knowledge-Driven Data Construction for Zero-Shot Evaluation in Commonsense Question Answering
K. Ma, F. Ilievski, J. Francis, Y. Bisk, E. Nyberg, and A. Oltramari
35th Conference on Artificial Intelligence (AAAI 2021)

SPADE: A Semi-Supervised Probabilistic Approach for Detecting Errors in Tables
M. Pham, C. A. Knoblock, M. Chen, B. Vu, and J. Pujara
Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)

From Tables to Knowledge: Recent Advances in Table Understanding
J. Pujara, P. Szekely, H. Sun, and M. Chen
Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

TOMATE: A Heuristic-Based Approach to Extract Data from HTML Tables
J.C. Roldán, P. Jiménez, P. Szekely, and R. Corchuelo
Information Sciences

Analyzing Race and Citizenship Bias in Wikidata
Z. Shaik, F. Ilievski, and F. Morstatter
2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS)

iCURL: Refining Estimated Identical-by-Descent Segments Via Community Detection
R. Shemirani, C. Gignoux, N. Zaitlen, C. Avery, E. Kenny, G. Belbin, and J. L. Ambite
Annual Meeting of the American Society of Human Genetics

Rapid Detection of Identity-by-Descent Tracts for Mega-Scale Datasets
R. Shemirani, G. M. Belbin, C. L. Avery, E. E. Kenny, C. R. Gignoux, and J. L. Ambite
Nature Communications

Multidisciplinary publications appear in all relevant areas.

EPS: Automated Feature Selection in Case-Control Studies using Extreme Pseudo-Sampling
R. Shemirani, S. Wenric, E. E.Kenny, and J. L. Ambite
Bioinformatics

On the Generalization Abilities of Fine-Tuned Commonsense Language Representation Models
K. Shen and M. Kejriwal
SGAI International Conference on Artificial Intelligence

A Study of the Quality of Wikidata
K. Shenoy, F. Ilievski, D. Garijo, D. Schwabe, and P. Szekely
Journal of Web Semantics

Building Survivable Software Systems by Automatically Adapting to Sensor Changes
Y. Shi, A. Li, T. K. S. Kumar, and C. A. Knoblock
Applied Sciences

Secure Neuroimaging Analysis using Federated Learning with Homomorphic Encryption
D. Stripelis, H. Saleem, T. Ghai, N. Dhinagar, U. Gupta, C. Anastasiou, G. V. Steeg, S. Ravi, M. Naveed, P. M. Thompson, and J. L. Ambite
17th International Symposium on Medical Information Processing and Analysis (SIPAIM)

Scaling Neuroscience Research using Federated Learning
D. Stripelis, J. L. Ambite, P. Lam, and P. Thompson
IEEE 18th International Symposium on Biomedical Imaging (ISBI)

Numeracy Enhances the Literacy of Language Models
A. Thawani, J. Pujara, and F. Ilievski
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

Representing Numbers in NLP: A Survey and a Vision
A. Thawani, J. Pujara, F. Ilievski, and P. Szekely
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics

Combining Remote-Sensing-Derived Data and Historical Maps for Long-Term Back-Casting of Urban Extents
J. H. Uhl, S. Leyk, Z. Li, W. Duan, B. Shbita, Y. Chiang, and C. A. Knoblock
Remote Sensing

A Graph-Based Approach for Inferring Semantic Descriptions of Wikipedia Tables
B. Vu, C. A. Knoblock, P. Szekely, M. Pham, and J. Pujara
The Semantic Web (ISWC2021)

Table-Based Fact Verification with Salience-Aware Learning
F. Wang, K. Sun, J. Pujara, P. Szekely, and M. Chen
Findings of the Association for Computational Linguistics (EMNLP2021)

Retrieving Complex Tables with Multi-Granular Graph Representation Learning
F. Wang, K. Sun, M. Chen, J. Pujara, and P. Szekely
Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval

NERO: A Biomedical Named-Entity (Recognition) Ontology with a Large, Annotated Corpus Reveals Meaningful Associations Through Text Embedding
K. Wang, R. Stevens, H. Alachram, Y. Li, L. Soldatova, R. King, S. Ananiadou, A. M. Schoene, M. Li, F. Christopoulou, J. L. Ambite, J. Matthew, S. Garg, U. Hermjakob, D. Marcu, and E. Sheng
Systems Biology and Applications

Do Language Models Perform Generalizable Commonsense Inference?
P. Wang, F. Ilievski, M. Chen, and X. Ren
ACL Findings

AMPPERE: A Universal Abstract Machine for Privacy-Preserving Entity Resolution Evaluation
Y. Yao, T. Ghai, S. Ravi, and P. Szekely
Proceedings of the 30th ACM International Conference on Information & Knowledge Management

MACHINE LEARNING AND DATA SCIENCE

Implicit SVD for Graph Representation Learning
S. Abu-El-Haija, H. Mostafa, M. Nassar, V. Crespi, G. V. Steeg, and A. Galstyan
Advances in Neural Information Processing Systems (NeurIPS)

Fast GRL with Unique Optimal Solutions
S. Abu-El-Haija, V. Crespi, G. V. Steeg, and A. Galstyan
ICLR 2021 Workshop on Geometrical and Topological Representation Learning

Fast Graph Learning with Unique Optimal Solutions
S. Abu-El-Haija, V. Crespi, G. V. Steeg, and A. Galstyan
ICLR 2021 Workshop on Geometrical and Topological Representation Learning

Disaggregation via Gaussian Regression for Robust Analysis of Heterogeneous Data
N. Alipourfard, K. Burghardt, and K. Lerman
Handbook of Computational Social Science, Volume 2: Data Science, Statistical Modelling, and Machine Learning Methods

Topics of Nicotine-Related Discussions on Twitter: Infoveillance Study
J. Allem, A. Dormanesh, A. Majmundar, J. B. Unger, M. G. Kirkpatrick, A. Choube, A. Aithal, E. Ferrara, T. B. Cruz et al.
Journal of Medical Internet Research

Pictures as a Form of Protest: A Survey and Analysis of Images Posted During the Stop Asian Hate Movement on Twitter
O. M. Allen, E. Chen, and E. Ferrara
2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS)

Bin2vec: Learning Representations of Binary Executable Programs for Security Tasks
S. Arakelyan, S. Arasteh, C. Hauser, E. Kline, and A. Galstyan
Cybersecurity

PUBLICATIONS

Auditing Algorithmic Bias on Twitter
N. Bartley, A. Abeliuk, E. Ferrara, and K. Lerman
13th ACM Web Science Conference 2021

The Emergence of Heterogeneous Scaling in Research Institutions
K. A. Burghardt, Z. He, A. G. Percus, and K. Lerman
Communications Physics

Having a Bad Day? Detecting the Impact of Atypical Events using Wearable Sensors
K. Burghardt, N. Tavabi, E. Ferrara, S. Narayanan, and K. Lerman
International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction, and Behavior Representation in Modeling and Simulation

Social Bots and Social Media Manipulation in 2020: The Year in Review
H. H. Chang, E. Chen, M. Zhang, G. Muric, and E. Ferrara
Handbook of Computational Social Science, Volume 1: Theory, Case Studies, and Ethics

Digital Civic Participation and Misinformation During the 2020 Taiwanese Presidential Election
H. H. Chang, S. Haider, and E. Ferrara
Media and Communication

Election 2020: The First Public Twitter Dataset on the 2020 US Presidential Election
E. Chen, A. Deb, and E. Ferrara
Journal of Computational Social Science

COVID-19 Misinformation and the 2020 US Presidential Election
E. Chen, H. Chang, A. Rao, K. Lerman, G. Cowan, and E. Ferrara
The Harvard Kennedy School Misinformation Review

COVID-19 Infodemiology at Planetary Scale: Charting the Information and Misinformation Landscape to Characterize Misinfodemics Spread on Social Media
E. Chen, J. Jiang, H. H. Chang, G. Muric, and E. Ferrara
JMIR Infodemiology

Identifying Botnet IP Address Clusters using Natural Language Processing Techniques on Honeypot Command Logs
V. Crespi, W. Hardaker, S. Abu-El-Haija, and A. Galstyan
SIAM International Conference on Data Mining (SDM21)

3D Convolutional Neural Networks for Classification of Alzheimer's and Parkinson's Disease with T1-Weighted Brain MRI
N. J. Dhinagar, S. I. Thomopoulos, C. Owens-Walton, D. Stripelis, J. L. Ambite, G. V. Steeg, D. Weintraub, P. Cook, C. McMillan, and P. M. Thompson
International Symposium on Medical Information Processing and Analysis (SIPAIM)

Explaining Classification Performance and Bias via Network Structure and Sampling Technique
L. Espin-Noboa, F. Karimi, B. Ribeiro, K. Lerman, and C. Wagner
Applied Network Science

Predicting Progression from Mild Cognitive Impairment to Alzheimer's Disease using MRI-Based Cortical Features and a Two-State Markov Model
E. Ficiara, V. Crespi, S. P. Gadewar, S. I. Thomopoulos, J. Boyd, P. M. Thompson, N. Jahanshad, and F. Pizzagalli
Proceedings of the International Symposium on Biomedical Imaging (ISBI)

Zero-Shot Synthesis with Group-Supervised Learning
Y. Ge, S. Abu-El-Haija, G. Xin, and L. Itti
International Conference on Learning Representations

Plot-Guided Adversarial Example Construction for Evaluating Open-Domain Story Generation
S. Ghazarian, Z. Liu, A. SM, R. Weischedel, A. Galstyan, and N. Peng
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

Artificial Intelligence for Modeling Complex Systems: Taming the Complexity of Expert Models to Improve Decision Making
Y. Gil, D. Garijo, D. Khider, C. A. Knoblock, V. Ratnakar, M. Osorio, H. Vargas, M. Pham, J. Pujara, B. Shbita, B. Vu, Y. Chiang, D. Feldman, Y. Lin, H. Song, V. Kumar, and A. Khandelwal
ACM Transactions on Interactive Intelligent Systems

Controllable Guarantees for Fair Outcomes via Contrastive Information Estimation
U. Gupta, A. Ferber, B. Dilkina, and G. V. Steeg
35th AAAI Conference on Artificial Intelligence (AAAI 21)

Membership Inference Attacks on Deep Regression Models for Neuroimaging
U. Gupta, D. Stripelis, P. K. Lam, P. Thompson, J. L. Ambite, and G. V. Steeg
Medical Imaging with Deep Learning

Improved Brain Age Estimation with Slice-Based Set Networks
U. Gupta, P. Lam, G. V. Steeg, and P. Thompson
IEEE International Symposium on Biomedical Imaging (ISBI)

Estimating Informativeness of Samples with Smooth Unique Information
H. Harutyunyan, A. Achille, G. Paolini, O. Majumder, A. Ravichandran, R. Bhotika, and S. Soatto
International Conference on Learning Representations

Information-Theoretic Generalization Bounds for Black Box Learning Algorithms
H. Harutyunyan, M. Raginsky, G. V. Steeg, and A. Galstyan
Advances in Neural Information Processing Systems (NeurIPS)

Identifying Shifts in Collective Attention to Topics on Social Media
Y. He, A. Rao, K. Burghardt, and K. Lerman
International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction, and Behavior Representation in Modeling and Simulation

Multidisciplinary publications appear in all relevant areas.

Heterogeneous Effects of Software Patches in a Multiplayer Online Battle Arena Game
Y. He, C. Tran, J. Jiang, K. Burghardt, E. Ferrara, E. Zheleva, and K. Lerman
16th International Conference on the Foundations of Digital Games (FDG) 2021

Speaker Turn Modeling for Dialogue Act Classification
Z. He, L. Tavabi, K. Lerman, and M. Soleymani
Findings of the Association for Computational Linguistics: EMNLP 2021

Detecting Polarized Topics using Partisanship-Aware Contextualized Topic Embeddings
Z. He, N. Mokhberian, A. Camara, A. Abeliuk, and K. Lerman
Findings of the Association for Computational Linguistics: EMNLP 2021

Socioeconomic Correlates of Anti-Science Attitudes in the US
M. Hu, A. Rao, M. Kejriwal, and K. Lerman
Future Internet

Story Generation with Commonsense Knowledge Graphs and Axioms
F. Ilievski, J. Pujara, and H. Zhang
AKBC Workshop on Commonsense Reasoning and Knowledge Bases

The Wide, the Deep, and the Maverick: Types of Players in Team-based Online Games
J. Jiang, D. Maldeniya, K. Lerman, and E. Ferrara
Proceedings of the ACM Conference on Human-Computer Interaction

Learning Graph Representations of Biochemical Networks and its Application to Enzymatic Link Prediction
J. Jiang, L. Liu, and S. Hassoun
Bioinformatics

Social Media Polarization and Echo Chambers in the Context of COVID-19: Case Study
J. Jiang, X. Ren, and E. Ferrara
JMIRx med

ForecastQA: A Question Answering Challenge for Event Forecasting with Temporal Text Data
W. Jin, R. Khanna, S. Kim, D. Lee, F. Morstatter, A. Galstyan, and X. Ren
ACL 2021

Learn Continually, Generalize Rapidly: Lifelong Knowledge Accumulation for Few-Shot Learning
X. Jin, B. Y. Lin, M. Rostami, and X. Ren
Findings of the Association for Computational Linguistics: EMNLP 2021

Finding Pragmatic Differences Between Disciplines
L. Kezar and J. Pujara
NAACL Workshop on Scholarly Document Processing

Social Bots and Their Coordination During Online Campaigns: A Survey
T. Khaund, B. Kirdemir, N. Agarwal, H. Liu, and F. Morstatter
IEEE Transactions on Computational Social Systems

Enter At Your Own Risk: The Impacts of Joining a Hateful Subreddit
K. Ko, K. Burghardt, and G. Muric
2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS)

Partner-Assisted Learning for Few-Shot Image Classification
J. Ma, H. Xie, G. Han, S. Chang, A. Galstyan, and W. AbdAlmageed
Proceedings of the IEEE/CVF International Conference on Computer Vision

Graph Traversal with Tensor Functionals: A Meta-Algorithm for Scalable Learning
E. S. Markowitz, K. Balasubramanian, M. Mirtaheri, S. Abu-El-Haija, B. Perozzi, G. V. Steeg, and A. Galstyan
International Conference on Learning Representations (ICLR)

q-Paths: Generalizing the Geometric Annealing Path using Power Means
V. Masrani, R. Brekelmans, T. Bui, F. Nielsen, A. Galstyan, G. V. Steeg, and F. Wood
Conference on Uncertainty in Artificial Intelligence (UAI)

The Impact of Peer Review on the Contribution Potential of Scientific Papers
A. Matsui, E. Chen, Y. Wang, and E. Ferrara
PeerJ

Using Word Embedding to Reveal Monetary Policy Explanation Changes
A. Matsui, X. Ren, and E. Ferrara
Proceedings of the Third Workshop on Economic and Natural Language Processing

geoChronR - An R Package to Model, Analyze, and Visualize Age-Uncertain Data
N. McKay, J. Emile-Geay, and D. Khider
Geochronology

A Survey on Bias and Fairness in Machine Learning
N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan
ACM Computing Surveys (CSUR)

Exacerbating Algorithmic Bias through Fairness Attacks
N. Mehrabi, M. Naveed, F. Morstatter, and A. Galstyan
35th Conference on Artificial Intelligence (AAAI 2021)

Lawyers are Dishonest? Quantifying Representational Harms in Commonsense Knowledge Resources
N. Mehrabi, P. Zhou, F. Morstatter, J. Pujara, X. Ren, and A. Galstyan
Conference on Empirical Methods in Natural Language Processing

One-Shot Learning for Temporal Knowledge Graphs
M. Mirtaheri, M. Rostami, X. Ren, F. Morstatter, and A. Galstyan
Automated Knowledge Base Construction (AKBC 2021)

PUBLICATIONS

Identifying and Analyzing Cryptocurrency Manipulations in Social Media
M. Mirtaheri, S. Abu-El-Haija, F. Morstatter, G. V. Steeg, and A. Galstyan
IEEE Transactions on Computational Social Systems

Explaining Face Presentation Attack Detection using Natural Language
H. Mirzaalian, M. E. Hussein, L. Spinoulas, J. May, and W. AbdAlmageed
2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)

Graph Signal Recovery using Restricted Boltzmann Machines
A. Mohan, A. Nakano, and E. Ferrara
Expert Systems with Applications

I'll Play on My Other Account: The Network and Behavioral Differences of Sybils
F. Morstatter, D. O. Kim, N. Jonckheere, C. Liu, M. Seth, and D. Williams
Proceedings of the ACM Conference on Human-Computer Interaction

Gender Disparity in the Authorship of Biomedical Research Publications During the COVID-19 Pandemic: Retrospective Observational Study
G. Muric, K. Lerman, and E. Ferrara
Journal of Medical Internet Research

COVID-19 Vaccine Hesitancy on Social Media: Building a Public Twitter Data Set of Antivaccine Content, Vaccine Misinformation, and Conspiracies
G. Muric, Y. Wu, and E. Ferrara
JMIR Public Health and Surveillance

Detecting Cryptocurrency Pump-and-Dump Frauds using Market and Social Signals
H. Nghiem, G. Muric, F. Morstatter, and E. Ferrara
Expert Systems with Applications

Layer-Wise Neural Network Compression via Layer Fusion
J. O'Neill, G. V. Steeg, and A. Galstyan
Asian Conference on Machine Learning (ACML)

SPADE: A Semi-Supervised Probabilistic Approach for Detecting Errors in Tables
M. Pham, C. Knoblock, M. Chen, B. Vu, and J. Pujara
International Joint Conference on Artificial Intelligence (IJCAI)

Political Partisanship and Antiscience Attitudes in Online Discussions About COVID-19: Twitter Content Analysis
A. Rao, F. Morstatter, M. Hu, E. Chen, K. Burghardt, E. Ferrara, K. Lerman et al.
Journal of Medical Internet Research

Discovering Higher-Order Interactions Through Neural Information Decomposition
K. Reing, G. V. Steeg, and A. Galstyan
Entropy

Influence Decompositions for Neural Network Attribution
K. Reing, G. V. Steeg, and A. Galstyan
The 24th International Conference on Artificial Intelligence and Statistics (AISTATS)

Lifelong Domain Adaptation via Consolidated Internal Distribution
M. Rostami
Advances in Neural Information Processing Systems

Transfer Learning Through Embedding Spaces (book)
M. Rostami

Detection and Continual Learning of Novel Face Presentation Attacks
M. Rostami, L. Spinoulas, M. Hussein, J. Mathai, and W. AbdAlmageed
Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)

Zero-Shot Image Classification using Coupled Dictionary Embedding
M. Rostami, S. Kolouri, Z. Murez, Y. Owekcho, E. Eaton, and K. Kim
Journal of Machine Learning with Applications

Limiting Tags Fosters Efficiency
T. Santos, K. Burghardt, K. Lerman, and D. Helic
13th ACM Web Science Conference 2021

Models, Markets, and the Forecasting of Elections
R. Sethi, J. Seager, E. Cai, D. M. Benjamin, and F. Morstatter
SSRN

Analyzing Race and Citizenship Bias in Wikidata
Z. Shaik, F. Ilievski, and F. Morstatter
2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS)

Identifying Coordinated Accounts on Social Media through Hidden Influence and Group Behaviours
K. Sharma, Y. Zhang, E. Ferrara, and Y. Liu
KDD '21

Follow the Leader: Documents on the Leading Edge of Semantic Change Get More Citations
S. Soni, K. Lerman, and J. Eisenstein
Journal of the Association for Information Science and Technology

Multi-Modal Fingerprint Presentation Attack Detection: Evaluation on a New Dataset
L. Spinoulas, H. Mirzaalian, M. E. Hussein, and W. AbdAlmageed
IEEE Transactions on Biometrics, Behavior, and Identity Science

Multispectral Biometrics System Framework: Application to Presentation Attack Detection
L. Spinoulas, M. E. Hussein, D. Geissbühler, J. Mathai, O. G. Almeida, G. Clivaz, S. Marcel, and W. AbdAlmageed
IEEE Sensors Journal

Multidisciplinary publications appear in all relevant areas.

Unsupervised Model Adaptation for Continual Semantic Segmentation
S. Stan and M. Rostami
Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI 2021)

Hamiltonian Dynamics with Non-Newtonian Momentum for Rapid Sampling
G. V. Steeg and A. Galstyan
Advances in Neural Information Processing Systems (NeurIPS)

Secure Neuroimaging Analysis using Federated Learning with Homomorphic Encryption
D. Stripelis, H. Saleem, T. Ghai, N. Dhinagar, U. Gupta, C. Anastasiou, G. V. Steeg, S. Ravi, M. Naveed, P. M. Thompson, and J. L. Ambite
International Symposium on Medical Information Processing and Analysis (SIPAIM)

Scaling Neuroscience Research using Federated Learning
D. Stripelis, J. L. Ambite, P. Lam, and P. Thompson
IEEE 18th International Symposium on Biomedical Imaging (ISBI)

Tabular Functional Block Detection with Embedding-Based Agglomerative Cell Clustering
K. Sun, F. Wang, M. Chen, and J. Pujara
Conference on Information and Knowledge Management

A Hybrid Probabilistic Approach for Table Understanding
K. Sun, H. Rayudu, and J. Pujara
35th Conference on Artificial Intelligence (AAAI 2021)

Numeracy Enhances the Literacy of Language Models
A. Thawani, J. Pujara, and F. Ilievski
Conference on Empirical Methods in Natural Language Processing

Representing Numbers in NLP: A Survey and a Vision
A. Thawani, J. Pujara, P. Szekely, and F. Ilievski
Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)

Identifying Distributional Perspectives from Colingual Groups
Y. Tian, T. Chakrabarty, F. Morstatter, and N. Peng
Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media

A Graph-Based Approach for Inferring Semantic Descriptions of Wikipedia Tables
B. Vu, C. A. Knoblock, P. Szekely, M. Pham, and J. Pujara
International Semantic Web Conference (ISWC 2021)

Table-Based Fact Verification with Salience-Aware Learning
F. Wang, K. Sun, J. Pujara, P. Szekely, and M. Chen
Findings of the Association for Computational Linguistics (EMNLP2021)

Retrieving Complex Tables with Multi-Granular Graph Representation Learning
F. Wang, K. Sun, M. Chen, J. Pujara, and P. Szekely
ACM Conference on Research and Development in Information Retrieval (SIGIR)

NERO: A Biomedical Named-Entity (Recognition) Ontology with a Large, Annotated Corpus Reveals Meaningful Associations through Text Embedding
K. Wang, R. Stevens, H. Alachram, Y. Li, L. Soldatova, R. King, S. Ananiadou, A. M. Schoene, M. Li, F. Christopoulou, J. L. Ambite, J. Matthew, S. Garg, U. Hermjakob, D. Marcu, and E. Sheng
NPJ Systems Biology and Applications

Muscle: Strengthening Semi-Supervised Learning via Concurrent Unsupervised Learning Using Mutual Information Maximization
H. Xie, M. E. Hussein, A. Galstyan, and W. AbdAlmageed
Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision

Mitigating the Bias of Heterogeneous Human Behavior in Affective Computing
S. Yan, H. Kao, K. Lerman, S. Narayanan, and E. Ferrara
2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)

Disrupting the COVID-19 Misinfodemic with Network Interventions: Network Solutions for Network Problems
L. E. Young, E. Sidnam-Mauch, M. Twyman, L. Wang, J. J. Xu, M. Sargent, T. W. Valente, E. Ferrara, J. Fulk, and P. Monge
American Journal of Public Health

A Survey of Human Judgement and Quantitative Forecasting Methods
M. Zellner, A. E. Abbas, D. V. Budescu, and A. Galstyan
Royal Society Open Science

Commonsense-Focused Dialogues for Response Generation: An Empirical Study
P. Zhou, K. Gopalakrishnan, B. Hedayatnia, S. Kim, J. Pujara, X. Ren, Y. Liu, and D. Hakkani-Tur
Proceedings of the Special Interest Group on Discourse and Dialogue

Probing Commonsense Explanation in Dialogue Response Generation
P. Zhou, P. Jandaghi, H. Cho, B. Y. Lin, J. Pujara, and X. Ren
Findings of the Association for Computational Linguistics (EMNLP2021)

RICA: Evaluating Robust Inference Capabilities Based on Commonsense Axioms
P. Zhou, R. Khanna, S. Lee, B. Y. Lin, D. Ho, J. Pujara, and X. Ren
Conference on Empirical Methods in Natural Language Processing

MATHEMATICS

Well-Poised Hypersurfaces
J. Cecil, N. Dutta, C. Manon, B. Riley, and A. Vichitbandha
Communications in Algebra

PUBLICATIONS

MICROELECTRONICS AND ELECTRONICS SYSTEMS

IMPULSE: A 65-nm Digital Compute-in-Memory Macro with Fused Weights and Membrane Potential for Spike-Based Sequential Learning Tasks
A. Agrawal, A. Mustafa, K. Minsuk, R. Nitin, J. Akhilesh, and K. Roy
IEEE Solid-State Circuits Letters, Vol. 4

Monolithically Integrated Silicon Nitride Platform
Y. Bian, C. Meagher, K. Nummy, S. Chandran, W. S. Lee, A. Aboketaf, D. Ma, R. Sporer, M. Rakowski, N. Yuksek et al.
2021 Optical Fiber Communications Conference and Exhibition (OFC)

Improving Reliability and Manufacturability by Maximizing Via Insertion Rates
L. Chang
Samsung SAFE Forum 2021

Reducing Stress Effects on Multi-Project-Wafer Reticles by Optimizing Metal Densities and Density Gradients in an MPW Placement Flow
L. Chang
Proceedings of SPIE

Apodized Distributed Bragg Reflector (DBR) Bends for Compact cWDM Filter
Y. Gebregiorgis, S. Chandran, M. Papadovasilakis, Y. Bian, M. Rakowski, F. O. Afzal, R. Augur, and J. Viegas
Integrated Photonics Research, Silicon, and Nanophotonics

Neural Computing with Magnetoelectric Domain-Wall-Based Neurosynaptic Devices
A. Jaiswal, A. Agrawal, P. Panda, and K. Roy
IEEE Transactions on Magnetics

Low-Loss 2 * 2 Wavelength-Independent Coupler using MZI Based on Bezier Curves
M. Papadovasilakis, S. Chandran, Y. Gebregiorgis, Y. Bian, M. Rakowski, R. Augur, and J. Viegas
Laser Science

Fabrication Tolerant MZI Design for Broadband WDM Filters on Scalable SiP Platform
M. Papadovasilakis, S. Chandran, Y. Gebregiorgis, Y. Bian, M. Rakowski, S. Krishnamurthy, R. Augur, and J. Viegas
Integrated Photonics Research, Silicon and Nanophotonics

Compact MMI-Based AWGs in a Scalable Monolithic Silicon Photonics Platform
A. M. Taha, S. Chandran, J. Viegas, Y. Bian, M. Rakowski, R. Augur, A. Jacob, and M. S. Dahlem
IEEE Photonics Journal

Secure Cloud-Based Lynx Virtual Design Environment Enabling Accurate and Efficient Physical Verification and Fabrication Sign-Off
F. K. G. Thorne
SNUG World 2021

Intrinsic Spike Timing Dependent Plasticity in Stochastic Magnetic Tunnel Junctions Mediated by Heat Dynamics
H. I. Velarde, J. Nagaria, Z. Yin, A. Jacob, and A. Jaiswal
IEEE Magnetics Letters

NATURAL LANGUAGE PROCESSING

Can Sequence-to-Sequence Models Crack Substitution Ciphers?
N. Aldarrab and J. May
Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing

A Grounded Approach to Modeling Generic Knowledge Acquisition
D. Beser, J. Cecil, M. Freedman, J. Lichtefeld, M. P. Marcus, S. Payne, and C. D. Yang
Proceedings of the Annual Meeting of the Cognitive Science Society

Keynote Abstract: Events on a Global Scale: Towards Language-Agnostic Event Extraction
E. Boschee
Proceedings of the 4th Workshop on Challenge and Applications of Automated Extraction of Sociopolitical Events from Text (CASE2021)

The Emergence of Heterogeneous Scaling in Research Institutions
K. A. Burghardt, Z. He, A. G. Percus, and K. Lerman
Communications Physics

CaSiNo: A Corpus of Campsite Negotiation Dialogues for Automatic Negotiation Systems
K. Chawla, J. Ramirez, R. Clever, G. Lucas, J. May, and J. Gratch
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

Event-Centric Natural Language Understanding
M. Chen, H. Zhang, Q. Ning, M. Li, H. Ji, and D. Roth
AAAI Tutorials

Event-Centric Natural Language Processing
M. Chen, H. Zhang, Q. Ning, M. Li, H. Ji, K. McKeown, and D. Roth
Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Tutorial Abstracts

Cross-Lingual Entity Alignment with Incidental Supervision
M. Chen, W. Shi, B. Zhou, and D. Roth
Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics

Multidisciplinary publications appear in all relevant areas.

Probabilistic Box Embeddings for Uncertain Knowledge Graph Reasoning
X. Chen, M. Boratko, M. Chen, S. S. Dasgupta, X. L. Li, and A. McCallum
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

Perhaps PTLMs Should Go to School—A Task to Assess Open Book and Closed Book QA
M. Ciosici, J. Cecil, D. Lee, A. Hedges, M. Freedman, and R. Weischedel
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

Machine-Assisted Script Curation
M. R. Ciosici, J. Cummings, M. DeHaven, A. Hedges, Y. Kankanampati, D. Lee, R. M. Weischedel, and M. Freedman
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations (NAACL-HLT)

Plot-Guided Adversarial Example Construction for Evaluating Open-Domain Story Generation
S. Ghazarian, Z. Liu, A. S. M, R. Weischedel, A. Galstyan, and N. Peng
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

Cross-Attention is All You Need: Adapting Pretrained Transformers for Machine Translation
M. Gheini, X. Ren, and J. May
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

Macro-Average: Rare Types Are Important Too
T. Gowda, W. You, C. Lignos, and J. May
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics Human Language Technologies

Many-to-English Machine Translation Tools, Data, and Pretrained Models
T. Gowda, Z. Zhang, C. Mattmann, and J. May
Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing

WARP: Word-Level Adversarial ReProgramming
K. Hambardzumyan, H. Khachatrian, and J. May
Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing

Speaker Turn Modeling for Dialogue Act Classification
Z. He, L. Tavabi, K. Lerman, and M. Soleymani
Findings of the Association for Computational Linguistics (EMNLP 2021)

Detecting Polarized Topics using Partisanship-Aware Contextualized Topic Embeddings
Z. He, N. Mokhberian, A. Camara, A. Abeliuk, and K. Lerman
Findings of the Association for Computational Linguistics (EMNLP 2021)

Learning Where to Cut from Edited Videos
Y. Huang, X. Bai, O. Wang, F. Caba, and A. Agarwala
2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)

JEDI: Circular RNA Prediction Based on Junction Encoders and Deep Interaction among Splice Sites
J. Jiang, C. J. Ju, J. Hao, M. Chen, and W. Wang
Bioinformatics

Unsupervised Real-Time Induction and Interactive Visualization of Taxonomies over Domain-Specific Concepts
M. Kejriwal and K. Shen
Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

An Evaluation and Annotation Methodology for Product Category Matching in E-Commerce
M. Kejriwal, K. Shen, C. Ni, and N. Torzec
Computers in Industry

ParsiNLU: A Suite of Language Understanding Challenges for Persian
D. Khashabi, A. Cohan, S. Shakeri, P. Hosseini, P. Pezeshkpour, M. Alikhani, M. Aminnaseri, M. Bitaab, F. Brahman, S. Ghazarian, M. Gheini et al.
Transactions of the Association for Computational Linguistics

AutoTriggER: Named Entity Recognition with Auxiliary Trigger Extraction
D. Lee, R. K. Selvam, S. M. Sarwar, B. Y. Lin, M. Agarwal, F. Morstatter, J. Pujara, E. Boschee, J. Allan, and X. Ren
Proceedings of the First Workshop on Trustworthy Natural Language Processing

Visual Pivoting for (Unsupervised) Entity Alignment
F. Liu, M. Chen, D. Roth, and N. Collier
Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI 2021)

X-METRA-ADA: Cross-Lingual Meta-Transfer Learning Adaptation to Natural Language Understanding and Question Answering
M. M'hamdi, D. S. Kim, F. Dernoncourt, T. Bui, X. Ren, and J. May
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics

HyperExpan: Taxonomy Expansion with Hyperbolic Representation Learning
M. D. Ma, M. Chen, T. Wu, and N. Peng
Findings of the Association for Computational Linguistics (EMNLP 2021)

PUBLICATIONS

Decoupling Global and Local Representations via Invertible Generative Flows
X. Ma, X. Kong, S. Zhang, and E. Hovy
Proceedings of the 9th International Conference on Learning Representations (ICLR-2021)

SPADE: A Semi-Supervised Probabilistic Approach for Detecting Errors in Tables
M. Pham, C. A. Knoblock, M. Chen, B. Vu, and J. Pujara
International Joint Conference on Artificial Intelligence

From Tables to Knowledge: Recent Advances in Table Understanding
J. Pujara, P. Szekely, H. Sun, and M. Chen
Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

On the Generalization Abilities of Fine-Tuned Commonsense Language Representation Models
K. Shen and M. Kejriwal
International Conference on Innovative Techniques and Applications of Artificial Intelligence

COM2SENSE: A Commonsense Reasoning Benchmark with Complementary Sentences
S. Singh, N. Wen, Y. Hou, P. Alipoormolabashi, T. Wu, X. Ma, and N. Peng
Findings of the Association for Computational Linguistics (ACLIJCNLP 2021)

Multitask Semi-Supervised Learning for Class-Imbalanced Discourse Classification
A. Spangher, J. May, S. Shiang, and L. Deng
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

The Danish Gigaword Corpus
L. Stromberg-Derczynski, M. Ciosici, R. Baglini, M. H. Christiansen, J. A. Dalsgaard, R. Fusaroli, P. J. Henrichsen, R. Hvingelby, A. Kirkedal, and A. Kjeldsen
Proceedings of the 23rd Nordic Conference on Computational Linguistics

AESOP: Paraphrase Generation with Adaptive Syntactic Control
J. Sun, X. Ma, and N. Peng
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

Tabular Functional Block Detection with Embedding-Based Agglomerative Cell Clustering
K. Sun, F. Wang, M. Chen, and J. Pujara
Proceedings of the 30th ACM International Conference on Information and Knowledge Management

Knowing the No-Match: Entity Alignment with Dangling Cases
Z. Sun, M. Chen, and W. Hu
Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics

Numeracy Enhances the Literacy of Language Models
A. Thawani, J. Pujara, and F. Ilievski
Conference on Empirical Methods in Natural Language Processing

Representing Numbers in NLP: a Survey and a Vision
A. Thawani, J. Pujara, P. Szekely, and F. Ilievski
Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)

Table-Based Fact Verification with Salience-Aware Learning
F. Wang, K. Sun, J. Pujara, P. Szekely, and M. Chen
Findings of the Association for Computational Linguistics (EMNLP 2021)

Retrieving Complex Tables with Multi-Granular Graph Representation Learning
F. Wang, K. Sun, M. Chen, J. Pujara, and P. Szekely
Proceedings of the 44th ACM SIGIR Conference on Research and Development in Information Retrieval

Learning Constraints and Descriptive Segmentation for Subevent Detection
H. Wang, H. Zhang, M. Chen, and D. Roth
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

NERO: A Biomedical Named-Entity (Recognition) Ontology with a Large, Annotated Corpus Reveals Meaningful Associations through Text Embedding
K. Wang, R. Stevens, H. Alachram, Y. Li, L. N. Soldatova, R. King, S. Ananiadou, A. M. Schoene, M. Li, F. Christopoulou, J. Ambite, J. Y. Matthew, S. Garg, U. Hermjakob, D. Marcu, E. Sheng, and T. Beissbart
NPJ Systems Biology and Applications

Do Language Models Perform Generalizable Commonsense Inference?
P. Wang, F. Ilievski, M. Chen, and X. Ren
Findings of the Association for Computational Linguistics (ACL 2021)

PERFUME: Programmatic Extraction and Refinement for Usability of Mathematical Expression
N. Weideman, V. K. Felkner, W. Wu, J. May, C. Hauser, and L. Garcia
Proceedings of the 2021 Research on Offensive and Defensive Techniques in the Context of Man At The End (MATE) Attacks

Personalized Response Generation via Generative Split Memory Network
Y. Wu, X. Ma, and D. Yang
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics

Summary-Oriented Question Generation for Informational Queries
X. Yin, L. Zhou, K. Small, and J. May
Proceedings of the 1st Workshop on Document-grounded Dialogue and Conversational Question Answering (DialDoc2021)

Multidisciplinary publications appear in all relevant areas.

Salience-Aware Event Chain Modeling for Narrative Understanding
X. Zhang, M. Chen, and J. May
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

Examining and Combating Spurious Features under Distribution Shift
C. Zhou, X. Ma, P. Michel, and G. Neubig
Proceedings of the 38th International Conference on Machine Learning

Learning from Noisy Labels for Entity-Centric Information Extraction
W. Zhou and M. Chen
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

Contrastive Out-of-Distribution Detection for Pretrained Transformers
W. Zhou, F. Liu, and M. Chen
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

Learning from History: Modeling Temporal Knowledge Graphs with Sequential Copy-Generation Networks
C. Zhu, M. Chen, C. Fan, G. Cheng, and Y. Zhang
Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI 2021)

NETWORKING

Building Reproducible Video Streaming Traffic Generators
C. Ardi, A. Hussain, and S. Schwab
Proceedings of 14th Cyber Security Experimentation and Test Workshop (CSET)

Deploying Per-Packet Telemetry in a Long-Haul Network: The AmLight Use Case
J. Bezerra, I. Brito, A. Quintana, J. Ibarra, V. Chergarova, R. Frez, H. Morgan, M. LeClerc, and A. Paneri
2021 IEEE Workshop on Innovating the Network for Data-Intensive Science (INDIS)

NRE-012: In-Band Network Telemetry @ AmLight
J. Bezerra, J. Ibarra, A. Q. Torres, I. V. D. S. Brito, H. Morgan, V. Chergarova, A. Paneri, N. B. Ker, and M. LeClerc
Supercomputing Conference (SC21)

Introducing New Technologies, Innovations, and Collaborations in R&E Networking Between Africa, Latin America, Europe, and the US through New International Projects
J. Bezerra, L. E. Cadenas, V. Chergarova, D. Cox, D. Greaves, E. Grizendi, A. Hazin, J. Ibarra, L. Lopez, L. Lotz, S. Mammen, H. Morgan, M. Stanton, M. Teixeira, and S. Wiener
WACREN 2021

Hysteresis in Combinatorial Optimization Problems
Y. Guan, A. Li, S. Koenig, S. Haas, and S. Thittamaranahalli
Proceedings of the Thirty-Fourth International FLAIRS Conference (FLAIRS-2021)

Observing the Global IPv4 Internet: What IP Addresses Show
J. Heidemann
SKC Science and Technology Webinar Series

Collecting, Labeling, and Using Networking Data: The Intersection of AI and Networking
J. Heidemann, J. Mirkovic, W. Hardaker, and M. Kallitsis
2021 NSF Workshop on AI for Networking

New Directions in Automated Traffic Analysis
J. Holland, P. Schmitt, N. Feamster, and P. Mittal
Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security

Designing for Tussle in Encrypted DNS
A. Hounsel, P. Schmitt, K. Borgolte, and N. Feamster
Proceedings of the 20th ACM Workshop on Hot Topics in Networks

Encryption without Centralization: Distributing DNS Queries Across Recursive Resolvers
A. Hounsel, P. Schmitt, K. Borgolte, and N. Feamster
Proceedings of the Applied Networking Research Workshop

Auditing for Discrimination in Algorithms Delivering Job Ads
B. Imana, A. Korolova, and J. Heidemann
Proceedings of the 2021 International Conference on World Wide Web

Institutional Privacy Risks in Sharing DNS Data
B. Imana, A. Korolova, and J. Heidemann
Proceedings of the Applied Networking Research Workshop

Anycast In Context: A Tale of Two Systems
T. Koch, K. Li, C. Ardi, E. Katz-Bassett, M. Calder, and J. Heidemann
Proceedings of the 2021 ACM SIGCOMM 2021 Conference

Case Studies in Experiment Design on a Minimega Based Network Emulation Testbed
B. Kocoloski, A. Hussain, M. Troglia, C. Ardi, S. Cheng, D. DeAngelis, C. Symonds, M. Collins, R. Goodfellow, and S. Schwab
Proceedings of 14th Cyber Security Experimentation and Test Workshop (CSET)

NRE-016: AutoGOLE/SENSE: End-to-End Network Services and Workflow Integration
T. Lehman, M. Schwarz, H. Trompert, H. Newman, J. Balcas, R. Sirvinskas, J. Chang, X. Yang, J. MacAuley, C. Guok, I. Monga, J. Hess, F. I. Yeh, J. H. Chen, J. M., J. Graham, T. Defanti, T. Hutton et al.
Supercomputing Conference (SC21)

Lifelong Multi-Agent Path Finding in Large-Scale Warehouses
J. Li, A. Tinka, S. Kiesel, J. Durham, S. Thittamaranahalli, and S. Koenig
Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI 2021)

PUBLICATIONS

tsuNAME: Exploiting Misconfiguration and Vulnerability to DDoS DNS
G. C. Moura, S. Castro, J. Heidemann, and W. Hardaker
Proceedings of the 21st ACM Internet Measurement Conference

Efficient Processing of Streaming Data using Multiple Abstractions
A. Qadeer and J. Heidemann
Proceedings of the IEEE International Conference on Cloud Computing

A Minimal Resource High-Speed Routing Lookup Mechanism for Servers with NetFPGAs
G. C. Sankaran and K. M. Sivalingam
Transactions on Emerging Telecommunications Technologies

Leveraging In-Network Computing and Programmable Switches for Streaming Analysis of Scientific Data
G. C. Sankaran, J. Chung, and R. Kettimuthu
2021 IEEE 7th International Conference on Network Softwarization (NetSoft)

P4 and NetFPGA Based Secure In-Network Computing Architecture for AI-Enabled Industrial Internet of Things
G. C. Sankaran, K. M. Sivalingam, and H. Gondaliya
IEEE Internet of Things Journal

Pretty Good Phone Privacy
P. Schmitt and B. Raghavan
30th USENIX Security Symposium (USENIX Security 21)

Building Survivable Software Systems by Automatically Adapting to Sensor Changes
Y. Shi, A. Li, S. Thittamaranahalli, and C. Knoblock
Special Issue Article on Advances in Aerial, Space, and Underwater Robotics in Journal of Applied Sciences (JAS-2021)

Visualizing Internet Measurements of Covid-19 Work-from-Home
E. Stutz, Y. Pradkin, X. Song, and J. Heidemann
Proceedings of the National Symposium for NSF REU Research in Data Science

Message Digest for DNS Zones
D. Wessels, P. Barber, M. Weinberg, W. Kumar, and W. Hardaker
IETF Request for Comments (RFC) 8976

A Hierarchical Approach to Multi-Agent Path Finding
H. Zhang, M. Yao, Z. Liu, J. Li, L. Terr, S. Chan, S. Thittamaranahalli, and S. Koenig
Proceedings of the 14th International Symposium on Combinatorial Search (SOCS-2021)

Temporal Reasoning with Kinodynamic Networks
H. Zhang, N. Tiruviluamala, S. Koenig, and S. Thittamaranahalli
Proceedings of the 31st International Conference on Automated Planning and Scheduling (ICAPS-2021)

Workshop on Overcoming Measurement Barriers to Internet Research (WOMBIR 2021) Final Report
K. Claffy, D. Clark, F. E. Bustamante, J. Heidemann, M. Jonker, A. Schulman, and E. Zegura
ACM Computer Communication Review

QUANTUM INFORMATION SENSING, COMMUNICATION, AND PROCESSING

Optical Signal Processing Performance Dependence on Non-Ideal MZI Operation in a Tapped-Delay-Line
F. Alishahi, A. Fallahpour, K. Zou, A. Minoofar, H. Zhou, C. Liu, M. Tur, J. Habif, and A. E. Willner
Conference on Lasers and Electro-Optics

Experimental Demonstration of Remotely Controlled and Powered Tunable Optical 2-4 Taps Correlator of a 20-100 Gbit/s QPSK Channel Based on Laser-Delivered Bias and Control Signals
F. Alishahi, A. Minoofar, A. Fallahpour, K. Zou, H. Zhou, J. Habif, M. Tur, and A. E. Willner
2021 Optical Fiber Communications Conference and Exhibition (OFC)

Demonstration of a Tunable Optical Correlation of a 10-15 Gbaud QPSK Data Signal using Nonlinear Wave Mixing at a Remotely Controlled Node
F. Alishahi, K. Zou, A. Minoofar, H. Zhou, M. Tur, J. L. Habif, and A. E. Willner
2021 IEEE Photonics Conference (IPC)

Quantum Algorithm for Time-Dependent Hamiltonian Simulation by Permutation Expansion
Y. Chen, A. Kalev, and I. Hen
PRX Quantum

Quantum-Limited Estimation of Coherence under Thermal Noise in Photon-Starved States
Z. Chua, J. Habif, and F. Spedalieri
Bulletin of the American Physical Society

Experimental Demonstration of Remotely Powered, Controlled, and Monitored Optical Switching Based on Laser-Delivered Signals
A. Fallahpour, A. Minoofar, F. Alishahi, K. Zou, S. Idres, H. Hashemi, J. Habif, M. Tur, and A. E. Willner
Optical Letters

Quantum-Limited Discrimination of Laser Light and Thermal Light
J. L. Habif, A. Jagannathan, S. Gartenstein, P. Amory, and S. Guha
Optical Express

Determining Quantum Monte Carlo Simulability with Geometric Phases
I. Hen
Physical Review Research

Testing a Quantum Annealer as a Quantum Thermal Sampler
Z. G. Izquierdo, I. Hen, and T. Albash
ACM Transactions on Quantum Computing

Optical Steganography using Phase Encoding of Coherent States in a Noisy Thermal Channel
R. Jain, A. Jagannathan, and J. Habif
Frontiers in Optics

Multidisciplinary publications appear in all relevant areas.

An Integral-Free Representation of the Dyson Series using Divided Differences
A. Kalev and I. Hen
New Journal of Physics

Quantum Algorithm for Simulating Hamiltonian Dynamics with an Off-diagonal Series Expansion
A. Kalev and I. Hen
Quantum

Constructing Driver Hamiltonians for Optimization Problems with Linear Constraints
H. Leipold and F. M. Spedalieri
Quantum Science and Technology

Estimating Expectation Values using Approximate Quantum States
M. Paini, A. Kalev, D. Padilha, and B. Ruck
Quantum

Detection and Localization of Near Infrared Lasers from Atmospheric Laser Scattering
A. Rittenbach, C. Finnerty, and J. L. Habif
Emerging Imaging and Sensing Technologies for Security and Defence VI

SCIENTIFIC COMPUTING

Reproducing GW150914: The First Observation of Gravitational Waves from a Binary Black Hole Merger
D. Brown, K. Vahi, M. Taufer, V. Welch, and E. Deelman
Computing in Science and Engineering

Serverless Containers—Rising Viable Approach to Scientific Workflows
K. Burkat, M. Pawlik, B. Balis, M. Malawski, K. Vahi, M. Rynge, R. F. d. Silva, and E. Deelman
2021 IEEE 17th International Conference on eScience (eScience)

Emerging Frameworks for Advancing Scientific Workflows Research, Development, and Education
H. Casanova, E. Deelman, S. Gesing, M. Hildreth, S. Hudson, W. Koch, J. Larson, M. A. McDowell, N. Meyers, J. Navarro, G. Papadimitriou, R. Tanaka, I. Taylor, D. Thain, S. M. Wild et al.
2021 IEEE Workshop on Workflows in Support of Large-Scale Science (WORKS)

Teaching Parallel and Distributed Computing Concepts in Simulation with WRENCH
H. Casanova, R. Tanaka, W. Koch, and R. F. d. Silva
Journal of Parallel and Distributed Computing

WfChef: Automated Generation of Accurate Scientific Workflow Generators
T. Coleman, H. Casanova, and R. F. d. Silva
2021 IEEE 17th International Conference on eScience (eScience)

Evaluating Energy-Aware Scheduling Algorithms for I/O-Intensive Scientific Workflows
T. Coleman, H. Casanova, T. Gwartney, and R. F. d. Silva
International Conference on Computational Science (ICCS)

Blueprint: Cyberinfrastructure Center of Excellence
E. Deelman, A. Mandal, A. P. Murillo, J. Nabrzyski, V. Pascucci, R. Ricci, I. Baldin, S. Sons, L. Christopherson, C. Vardeman, R. F. d. Silva, J. Wyngaard, S. Petruzza, and M. Rynge
Zenodo

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

Modeling the Linux Page Cache for Accurate Simulation of Data-Intensive Applications
H. Do, V. Hayot-Sasson, R. F. d. Silva, C. Steele, H. Casanova, and T. Glatard
IEEE Cluster

Assessing Resource Provisioning and Allocation of Ensembles of In Situ Workflows
T. M. A. Do, L. Pottier, R. F. d. Silva, S. Caino-Lores, M. Taufer, and E. Deelman
50th International Conference on Parallel Processing Workshop

A Lightweight Method for Evaluating In Situ Workflow Efficiency
T. M. A. Do, L. Pottier, S. Caino-Lores, R. F. d. Silva, M.A. Cuendet, H. Weinstein, T. Estrada, M. Taufer, and E. Deelman
Journal of Computational Science

VisDict: Enhancing the Communication between Workflow Providers and User Communities via a Visual Dictionary
S. Gesing, R. F. d. Silva, E. Deelman, M. Hildreth, M. A. McDowell, M. Natalie, T. Ian, and T. Douglas
2021 IEEE/ACM Workflows in Support of Large-Scale Science (WORKS)

Artificial Intelligence for Modeling Complex Systems: Taming the Complexity of Expert Models to Improve Decision Making
Y. Gil, D. Garijo, D. Khider, C. A. Knoblock, V. Ratnakar, M. Osorio, H. Vargas, M. Pham, J. Pujara, B. Shbita, B. Vu, Y. Chiang, D. Feldman, Y. Lin, H. Song, V. Kumar, et al.
ACM Transactions on Interactive Intelligent Systems

GLUME: A Strategy for Reducing Workflow Execution Times on Batch-Scheduled Platforms
E. Hataishi, P. Dutot, R. F. d. Silva, and H. Casanova
Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP)

2021 SciTech and Friends Research Symposium Proceedings
C. Hayes, C. Kulkarni, E. D. Milman, O. Okunloye, A. Olshansky, R. Oruche, K. Kee, P. C. S. Moreira, C. Vardeman, T. Coleman, T. M. A. Do, A. Jain, P. Krawczuk, and K. Lam
2021 SciTech and Friends Research Symposium

Distributed and Heterogeneous SAR Backprojection with Halide
C. Imes, T. Li, M. Glines, R. Khan, and J. P. Walters
2021 IEEE High Performance Extreme Computing Conference (HPEC)

PUBLICATIONS

A Case Study in Scientific Reproducibility from the Event Horizon Telescope (EHT)
R. Ketron, J. Leonard, B. Roachell, R. Patel, R. White, S. Caíno-Lores, N. Tan, P. Miles, K. Vahi, E. Deelman, D. Brown, and M. Taufer
2021 IEEE 17th International Conference on eScience (eScience)

A Performance Characterization of Scientific Machine Learning Workflows
P. Krawczuk, G. Papadimitriou, R. Tanaka, T. M. A. Do, S. Subramany, S. Nagarkar, A. Jain, K. Lam, A. Mandal, L. Pottier, and E. Deelman
2021 IEEE/ACM Workflows in Support of Large-Scale Science (WORKS)

Anomaly Detection in Scientific Workflows using End-to-End Execution Gantt Charts and Convolutional Neural Networks
P. Krawczuk, G. Papadimitriou, S. Nagarkar, M. Kiran, A. Mandal, and E. Deelman
Practice and Experience in Advanced Research Computing

CrisisFlow: Multimodal Representation Learning Workflow for Crisis Computing
P. Krawczuk, S. Nagarkar, and E. Deelman
2021 IEEE 17th International Conference on eScience (eScience)

Predicting Flash Floods in the Dallas-Fort Worth Metroplex using Workflows and Cloud Computing
E. Lyons, D. Seo, S. Kim, H. Habibi, G. Papadimitriou, R. Tanaka, E. Deelman, M. Zink, and A. Mandal
2021 IEEE 17th International Conference on eScience (eScience)

FlyNet: A Platform to Support Scientific Workflows from the Edge to the Core for UAV Applications
E. Lyons, H. Saplakoglu, M. Zink, K. Thareja, A. Mandal, C. Qu, S. Wang, P. Calyam, G. Papadimitriou, R. Tanaka, and E. Deelman
Proceedings of the 14th IEEE/ACM International Conference on Utility and Cloud Computing

A Lightweight GPU Monitoring Extension for Pegasus Kickstart
G. Papadimitriou and E. Deelman
2021 IEEE/ACM Workflows in Support of Large-Scale Science (WORKS)

End-to-End Online Performance Data Capture and Analysis for Scientific Workflows
G. Papadimitriou, C. Wang, K. Vahi, R. F. d. Silva, A. Mandal, L. Zhengchun, R. Mayani, M. Rynge, M. Kiran, V. E. Lynch, R. Kettimuthu, E. Deelman, J. S. Vetter, and I. Foster
Future Generation Computer Systems

Fair Sharing of Network Resources among Workflow Ensembles
G. Papadimitriou, E. Lyons, C. Wang, K. Thareja, R. Tanaka, P. Ruth, I. Rodero, E. Deelman, M. Zink, and A. Mandal
Cluster Computing

A Community Roadmap for Scientific Workflows Research and Development
R. F. d. Silva, H. Casanova, K. Chard, I. Altintas, R. M. Badia, B. Balis, T. Coleman, F. Coppens, F. D. Natale, B. Enders, T. Fahringer, R. Filgueira, G. Fursin, D. Garijo, and C. Goble
2021 IEEE Workshop on Workflows in Support of Large-Scale Science (WORKS)

A Roadmap to Robust Science for High-Throughput Applications: The Scientists' Perspective
M. Taufer, E. Deelman, R. F. d. Silva, T. Estrada, and M. Hall
2021 IEEE 17th International Conference on eScience (eScience)

A Roadmap to Robust Science for High-Throughput Applications: The Developers' Perspective
M. Taufer, E. Deelman, R. F. d. Silva, T. Estrada, M. Hall, and M. Livny
2021 IEEE International Conference on Cluster Computing (CLUSTER)

Mining Workflows for Anomalous Data Transfers
H. Tu, G. Papadimitriou, M. Kiran, C. Wang, A. Mandal, E. Deelman, and T. Menzies
2021 IEEE/ACM 18th International Conference on Mining Software Repositories (MSR)

New System for Archiving Integrative Structures
B. Vallat, B. Webb, M. Fayazi, S. Voinea, H. Tangmunarunkit, S. J. Ganesan, C. L. Lawson, J. D. Westbrook, C. Kesselman, A. Sali, and H. M. Berman
Acta Crystallographica Section D Structural Biology

Root Cause Analysis of Data Integrity Errors in Networked Systems with Incomplete Information
Y. Xin, S. Fu, A. Mandal, I. Baldin, R. Tanaka, M. Rynge, K. Vahi, E. Deelman, I. Abhinit, and W. Von
2021 International Conference on Information and Communication Technology Convergence (ICTC)

SPACE SYSTEMS AND TECHNOLOGIES

Sensor Fusion Kalman Filtering for Stability and Control of Satellite Swarms
R. Rughani and D. A. Barnhart
16th International Conference on Space Operations

Integrated Constellation Analysis Tools to Support New Observing Strategy Mission Design
C. D. Ball, I. J. Tapia-Tamayo, M. Paolieri, A. J. O'Brien, M. French, J. T. Johnson, and P. Grogan
CubeSats and SmallSats for Remote Sensing V

VISION, IMAGE, SPEECH, AND TEXT ANALYTICS

A Grounded Approach to Modeling Generic Knowledge Acquisition
D. Beser, J. Cecil, M. Freedman, J. A. Lichtefeld, M. Marcus, S. R. Payne, and C. Yang
Proceedings of the Annual Meeting of the Cognitive Science Society

Multidisciplinary publications appear in all relevant areas.

Leveraging Non-Target Language Resources to Improve ASR Performance in a Target Language
J. Billa
Proceedings of INTERSPEECH 2021

Customizable Camera Verification for Media Forensic
H. Cao and W. AbdAlmageed
16th International Conference on Document Analysis and Recognition (ICDAR 2021)

An Analysis of Observation Length Requirements for Machine Understanding of Human Behaviors from Spoken Language
S. N. Chakravarthula, B. R. Baucom, S. Narayanan, and P. Georgiou
Computer Speech and Language

Agenda Pushing in Email to Thwart Phishing
H. Cho, G. Bartlett, and M. Freedman
Proceedings of the 1st Workshop on Document-Grounded Dialogue and Conversational Question Answering (DialDoc 2021)

Perhaps PTLM's Should Go to School—A Task to Assess Open Book and Closed Book QA
M. Ciosici, J. Cecil, D. Lee, A. Hedges, M. Freedman, and R. Weischedel
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing

Machine-Assisted Script Curation
M. Ciosici, J. Cummings, M. DeHaven, A. Hedges, Y. Kankanampati, D. Lee, R. Weischedel, and M. Freedman
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics Human Language Technologies: Demonstrations

Arabic Scene Text Recognition in the Deep Learning Era: Analysis on a Novel Dataset
H. Hassan, A. El-Mahdy, and M. E. Hussein
IEEE Access

SCAN: Sequence-Character Aware Network for Text Recognition
H. Hassan, M. Torki, and M. E. Hussein
International Conference on Computer Vision Theory and Applications (VISAPP)

Temporal Dynamics of Workplace Acoustic Scenes: Egocentric Analysis and Prediction
A. Jati, A. Nadarajan, R. Peri, K. Mundnich, T. Feng, B. Girault, and S. Narayanan
IEEE/ACM Transactions on Audio, Speech, and Language Processing

Developing Neural Representations for Robust Child-Adult Diarization
S. Krishnamachari, M. Kumar, S. H. Kim, C. Lord, and S. Narayanan
Proceedings of IEEE Spoken Language Technology Workshop

Information-Theoretic Bias Assessment of Learned Representations of Pretrained Face Recognition
J. Li and W. AbdAlmageed
2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)

Explaining Face Presentation Attack Detection using Natural Language
H. Mirzaalian, M. E. Hussein, L. Spinoulas, J. May, and W. AbdAlmageed
IEEE International Conference on Automatic Face and Gesture Recognition

Intrapersonal and Interpersonal Vocal Affect Dynamics during Psychotherapy
A. Paz, E. Rafaeli, E. Bar-Kalifa, E. Gilboa-Schectman, S. Gannot, B. Laufer-Goldshtein, S. Narayanan, J. Keshet, and D. Atzil-Slonim
Journal of Consulting and Clinical Psychology

Detection and Continual Learning of Novel Face Presentation Attacks
M. Rostami, L. Spinoulas, M. Hussein, J. Mathai, and W. AbdAlmageed
Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)

RNN Based Incremental Online Spoken Language Understanding
G. P. Sivakumar, N. Kumar, P. Georgiou, and S. Narayanan
Proceedings of IEEE Spoken Language Technology Workshop

Computational Media Intelligence: Human-Centered Machine Analysis of Media
K. Somandepalli, T. Guha, V. Martinez, N. Kumar, H. Adam, and S. Narayanan
Proceedings of the IEEE

Multi-Modal Fingerprint Presentation Attack Detection: Evaluation on a New Dataset
L. Spinoulas, H. Mirzaalian, M. E. Hussein, and W. AbdAlmageed
IEEE Transactions on Biometrics, Behavior, and Identity Science

Multispectral Biometrics System Framework: Application to Presentation Attack Detection
L. Spinoulas, M. E. Hussein, D. Geissbühler, J. Mathai, O. G. Almeida, G. Clivaz, S. Marcel, and W. Almageed
IEEE Sensors Journal

MUSCLE: Strengthening Semi-Supervised Learning via Concurrent Unsupervised Learning using Mutual Information Maximization
H. Xie, M. E. Hussein, A. Galstyan, and W. AbdAlmageed
Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)

Improved 3D Real-Time MRI of Speech Production
Z. Zhao, Y. Lim, D. Byrd, S. Narayanan, and K. S. Nayak
Magnetic Resonance in Medicine

The best thing about the future is that it comes one day at a time.

— Abraham Lincoln

Marina Towers, lit in blue and yellow

Information Sciences Institute

4676 Admiralty Way #1001

Marina del Rey, CA 90292

310.822.1511 | isi.edu | media@isi.edu

USC Viterbi

School of Engineering