



# C3.ai Digital Transformation Institute

ANNUAL REPORT 2022-2023

**The C3.ai Digital Transformation Institute** is a research consortium dedicated to accelerating the benefits of artificial intelligence for business, government, and society. The Institute engages the world's leading scientists to conduct research and train practitioners in the new science of digital transformation, which operates at the intersection of artificial intelligence, machine learning, cloud computing, internet of things, big data analytics, organizational behavior, public policy, and ethics.

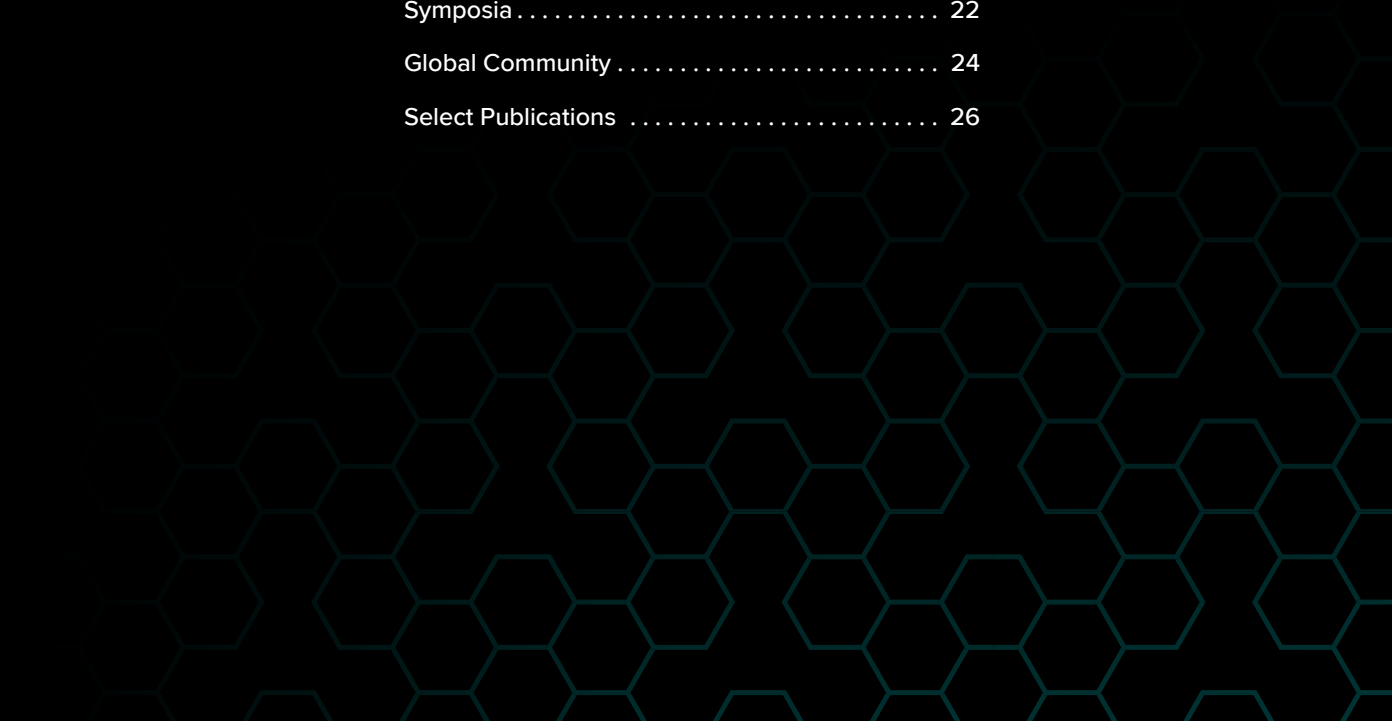




## TABLE OF CONTENTS

---

Introduction .....	2
Leadership .....	4
Executive Committee .....	6
Partner Institutions .....	7
Research Program .....	8
Awards & Honors .....	16
Colloquia .....	18
Workshops .....	20
Symposia .....	22
Global Community .....	24
Select Publications .....	26



# Advancing the Science of Digital Tr

THE FOUNDATION OF THE FUTURE

---

We began the C3.ai Digital Transformation Institute three years ago with the goal of developing the scientific foundation of the Digital Transformation of Societal Systems. Technologies at the nexus of AI/ Machine Learning, Internet of Things, and Data Analytics were creating opportunities for new business models and services in societal-scale systems, including transportation, energy, health care, and finance.

## RESEARCH IMPACT

Through our research grant program, we funded three rounds of proposals in areas ranging from health care and COVID-19 to distributed energy utilization and mitigating climate change to new defenses against cyber threats on financial and cyber-physical infrastructures. The 72 funded projects have already borne fruit. The pandemic accelerated the pace of digital transformation, additionally stimulating our research programs and impact.

DTI-funded research has resulted in the discovery of novel vaccines for COVID-19 with implications for cancer research, advanced epidemiological models for the spread of contagion, and new algorithms to prevent the kinds of blackouts that struck Texas in 2021. From DTI research, new techniques for monitoring emissions and leaks of hydrocarbons have emerged, along with early wildfire-warning systems, and AI-powered forecasts predicting hurricanes and El Niños soon enough to prepare, evacuate, and save lives. Our projects have developed new defenses against advanced persistent threats and data integrity attacks targeting financial infrastructures.

Our researchers have made new uses of the C3 AI Suite and cloud computing platforms for novel drug development, optimizing both crop yield and carbon capture in agriculture, and hardening electric power systems. Innovative DTI energy projects include work on an advanced electrode to capture “forever” chemicals and machine learning algorithms to navigate floating solar-powered kelp farms over tides to deposit sequestered carbon dioxide in deep ocean channels.

# ansformation

## SHAPING A NEW SCIENTIFIC DISCIPLINE

It is fair to say that from the immense volume of creative output that has been published, and presented at our colloquia and deep-dive workshops, a new discipline of the Science and Technology of Digital Transformation is taking shape. This new discipline combines machine learning and AI, data science, mathematical economics, and mechanism and incentive design into new normative models for introducing, assessing, and implementing new products and services. We expect these programs to grow rapidly in number and interest, at both business and engineering schools.

The work is both normative and behavioral, in that the robust algorithms that we have developed adapt to changes in human patterns and practices. It is well-known that societal systems change and adapt to new digital transformation services and solutions. What is novel is how to incorporate these changes into the designs of algorithms integrating cognitive models of human decision-making.

The emergence of Generative AI and its methods for going beyond analytics to generate intelligent insights and predictions about the future will further advance this science. The ability for AI systems to work with humans, learn their reasoning styles, and adapt to provide meaningful advice is the starting point for well-integrated Human-AI teams. Generative AI has already shown what it can do in terms of generating text, images, and even software code. For use in societal systems, we will need to provide trust and high confidence guarantees about its conclusions.

This is an exciting intellectual agenda, and looking to the future we feel that C3.ai DTI is ideally positioned to take on the challenges of integrating tools such as Generative AI into Digital Transformation solutions.

**S. Shankar Sastry**, Co-Director,  
University of California, Berkeley  
C3.ai Digital Transformation Institute

**R. Srikant**, Co-Director,  
University of Illinois at Urbana-Champaign  
C3.ai Digital Transformation Institute

*The pandemic accelerated the pace of digital transformation, additionally stimulating our research programs and impact.*



# Leadership

---

“*Generative AI will fundamentally change the nature of the human-computer interaction model.*”

– Thomas M. Siebel,  
Chairman and CEO, C3 AI

The C3.ai Digital Transformation Institute was established in March 2020 by C3 AI, Microsoft Corporation, the University of California, Berkeley, and the University of Illinois at Urbana-Champaign. Institute partners include Carnegie Mellon University, KTH Royal Institute of Technology, Massachusetts Institute of Technology, Princeton University, Stanford University, University of Chicago, Lawrence Berkeley National Laboratory, and the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign. The Institute is jointly managed and hosted by the University of California, Berkeley and the University of Illinois at Urbana-Champaign.

#### ADVISORY BOARD



**Thomas M. Siebel**  
Chairman and CEO, C3 AI



**Eric Horvitz**  
Chief Scientific Officer, Microsoft



**S. Shankar Sastry**  
Co-Director, C3.ai DTI  
Thomas M. Siebel Professor of  
Computer Science



**R. Srikant**  
Co-Director, C3.ai DTI  
Fredric G. and Elizabeth H. Nearing  
Endowed Professor of Electrical  
and Computer Engineering



## EXECUTIVE COMMITTEE



### **Jonathan Carter**

Campus Lead, C3.ai DTI  
Associate Lab Director for  
Computing Sciences  
Lawrence Berkeley National Laboratory



### **William Sanders**

Campus Lead, C3.ai DTI  
Dr. William D. and Nancy W. Strecker  
Dean, College of Engineering  
Carnegie Mellon University



### **Michael Franklin**

Campus Lead, C3.ai DTI  
Liew Family Chairman of  
Computer Science  
University of Chicago



### **S. Shankar Sastry**

Co-Director, C3.ai DTI  
Thomas M. Siebel Professor of  
Computer Science  
University of California, Berkeley



### **Karl H. Johansson**

Campus Lead, C3.ai DTI  
VR Distinguished Professor of Electrical  
Engineering and Computer Science  
KTH Royal Institute of Technology



### **Costas Spanos**

Co-Chief Scientist, C3.ai DTI  
Andrew S. Grove Distinguished  
Professor of Electrical Engineering  
and Computer Sciences  
University of California, Berkeley



### **Christopher Manning**

Campus Lead, C3.ai DTI  
Thomas M. Siebel Professor of  
Machine Learning  
Stanford University



### **R. Srikant**

Co-Director, C3.ai DTI  
Fredric G. and Elizabeth H. Nearing  
Endowed Professor of Electrical  
and Computer Engineering  
University of Illinois at Urbana-  
Champaign



### **Asuman Ozdaglar**

Campus Lead, C3.ai DTI  
MathWorks Professor of Electrical  
Engineering and Computer Science  
Massachusetts Institute of Technology



### **Tandy Warnow**

Co-Chief Scientist, C3.ai DTI  
Grainger Distinguished Chair  
of Engineering  
University of Illinois at Urbana-  
Champaign



### **H. Vincent Poor**

Campus Lead, C3.ai DTI  
Michael Henry Strater University  
Professor  
Princeton University





## PARTNER INSTITUTIONS

The C3.ai Digital Transformation Institute consortium includes partners from academia, national laboratories, and industry. The Industry Partners program enables leading companies from around the world to engage with C3.ai DTI researchers and activities. Industry partners are also encouraged to collaborate on research projects, participate in the annual C3.ai DTI Research Symposium, and attend workshops, colloquia, and other special events.

### ACADEMIC PARTNERS



Carnegie Mellon University



Stanford University



### RESEARCH LABORATORIES



### INDUSTRY PARTNERS





# Research Award Program

---

The C3.ai Digital Transformation Institute supports teams of the best researchers in the world to advance AI techniques for industrial, commercial, and public sector applications. This rich ecosystem will help address some of the most complex issues inherent in a massive societal digital transformation and build the foundation for a new Science of Digital Transformation. C3.ai DTI awards cash grants and access to computing resources for research projects based at consortium universities.

In June 2020, the Institute awarded 26 projects to accelerate artificial intelligence research to mitigate the COVID-19 pandemic and future pandemics across multiple disciplines, including medicine, urban planning, public policy, and computer science, and the pandemic's impact on racial, economic, and healthcare disparities.

In June 2021, the Institute awarded 22 projects to advance AI for energy and climate security, including sustainability initiatives, carbon sequestration, AI for advanced energy and carbon markets, cybersecurity of power and energy infrastructure, smart-grid analytics, distributed energy resource management, climate change modeling, and improved natural catastrophe risk assessment.

In March 2022, the Institute awarded 24 projects for AI to transform cybersecurity and secure critical infrastructure in the areas of AI resilience; anomaly detection; advanced persistent threats; securing cyber-physical infrastructure; forensics; emerging financial infrastructure; vulnerability identification; and insider threats.

“AI promises us new understandings of the foundations of human intelligence and the creation of new tools of reason and imagination. These new understandings will contribute profoundly to our wondrous arc of intelligence moving forward.”

– Eric Horvitz, Chief Scientific Officer, Microsoft



## BY THE NUMBERS



### GLOBAL CHALLENGES INVESTIGATED

- 2020: Mitigation of COVID-19 and Future Pandemics, 26 projects
- 2021: Digital Transformation and AI for Energy and Climate Security, 22 projects
- 2022: Cybersecurity and Secure Critical Infrastructure, 24 projects

### RESOURCES FOR RESEARCHERS

- \$100,000 to \$700,000 award range per project
- Up to \$2 million in Azure Cloud computing resources per project
- Up to 110,000 NCSA high-performance computing hours per project
- Up to 25 million NERSC high-performance computing hours per project
- Unlimited access to C3 AI Suite

## RESEARCH SPOTLIGHTS: AI TO MITIGATE THE COVID-19 PANDEMIC AND FUTURE PANDEMICS



In March of 2020, as the world reacted to a devastating new pandemic the likes of which had not been seen for generations, the C3.ai Digital Transformation Institute first launched with an urgent call for proposals: how to use the most advanced technology available to help mitigate this new COVID-19 pandemic. Three years later, results are still pouring in from that research – not just solutions for COVID-19 and future pandemics, and not just for healthcare and healthcare access more broadly – but beyond, to advances that could affect many other fields.

### CASCADING INNOVATION

**Maryellen Giger**, radiology professor at the **University of Chicago**, successfully achieved the aim of her DTI project: to build a clinical tool to detect SARS-CoV-2 infections in lung scans. Her research led to a patent-pending cascade-based deep transfer learning method for COVID-19 detection – which she later applied to accelerating head injury detection after cardiac arrest.

Early on, she recognized the need for a vast repository of thoracic images for more COVID research, and with NIH support launched the Medical Imaging and Data Resource Center, designed to be the largest open database of anonymized COVID-19 medical images, now including more than 300,000. Beyond that, the Center is developing ethical standards for data collection and curation that could serve as a model.

### ULTIMATE COVID VACCINE

**David Gifford**, professor of electrical engineering and computer science and of biological engineering at the **Massachusetts Institute of Technology**, designed a vaccine to work not only on the current strain of SARS-CoV-2 going around, but also on future variants.

Current vaccines deploy antibodies targeting the familiar spike protein, which can drift and change, rendering antibodies less effective. Gifford developed a novel technique to use the immune system's cellular response, with T-cells attacking infected cells, which targets core pieces of the virus that remain across variants. In mouse studies, the vaccine eliminated mortality; 80 percent of unvaccinated mice died.

Gifford is currently collaborating on a new vaccine for tuberculosis. "Another natural

thing to think about is combining the kind of COVID vaccine we have produced with a spike-based vaccine primarily designed to produce an antibody response,” says Gifford, “to get both going.”

---

### PROOF OF CONCEPT

The first news from the multi-university DTI COVID project headed up by **Ziv Bar-Joseph**, professor of Computational Biology and Machine Learning at **Carnegie Mellon University**, came in late 2020. After infecting tens of thousands of pluripotent stem cell-derived lung cells with SARS-CoV-2, the **Boston University** team tracked the resulting molecular reactions for possible therapeutic avenues. They discovered at least 18 existing, FDA-approved drugs to target the virus. Experimentally, five of those reduced coronavirus spread in human lung cells by more than 90 percent.

By 2023, the **Massachusetts Institute of Technology** team had been deploying their AI model to combat drug-resistant infections. Their algorithm searched nearly 7,000 potential compounds to identify one drug to kill *Acinetobacter baumannii*, a bacterium found in many hospital settings that can lead to pneumonia, meningitis, and other serious infections.

“The machine learning model can explore, in silico, large chemical spaces that can be prohibitively expensive for traditional experimental approaches,” says Co-PI **Regina Barzilay**, MIT professor of electrical engineering and computer science.

---

### TARGET, TEST, TELL

The **University of Illinois at Urbana-Champaign** has earned much acclaim for its highly successful S.H.I.E.L.D. program, a multi-faceted, highly coordinated all-units response to the pandemic that incorporated saliva testing, modeling, and a cellphone app to re-open the campus while suppressing the spread of the virus. Several DTI PIs worked on the campaign, including UIUC epidemiologist **Rebecca Smith**, who was among the authors of a June 2022 paper evaluating the university’s “SHIELD: Target, Test, Tell” response.

Comparing COVID-19 outcomes from July to December 2020 in all 251 counties in the U.S. that host a large university, the number of cases in Champaign County, where the university is located, was roughly a third less than expected. Moreover, Champaign County had the greatest reduction in deaths – more than 4X less than would be predicted based on other counties’ numbers.

“We don’t know what would have happened in Champaign-Urbana if we didn’t do this, but if we look at similar communities that have large universities, large student populations, we overperformed by a lot,” says Smith.



## RESEARCH SPOTLIGHTS: AI TO ADVANCE ENERGY AND CLIMATE SECURITY



By June 2021, the C3.ai Digital Transformation Institute released a call for proposals on another urgent worldwide crisis – the need to ensure energy and climate security and lead the way to lower-carbon, higher-efficiency economies. Two years later, results have been demonstrated in groundbreaking new fields from digital agriculture and AI-driven weather forecasting to distributed energy resources and electromobility.

### MITIGATING ‘FOREVER’ CHEMICALS

As the Environmental Protection Agency considers legislation to further protect the public from the health hazards associated with PFAs – called “Forever Chemicals” because they degrade so slowly – **Xiao Su**, assistant professor of chemical and biomolecular engineering at the **University of Illinois at Urbana-Champaign**, has designed an electrode to attract and capture short-chain PFAs from water. The work is a critical early step to improve water quality by removing short-chain PFAs, which have replaced long-chain PFAs in many industries.

“Future studies will focus on coupling the electrodes developed in this study with electrochemical degradation methods, to ensure removal of these persistent contaminants from the environment,” says Su, who collaborated on the project with fellow UIUC chemical and biomolecular engineering professor **Diwakar Shukla**.

### FLOATING SEAWEED FARMS

In the Gulf of Mexico, a research team led by **Claire Tomlin**, electrical engineering and computer science chair at the **University of California, Berkeley**, is testing its floating seaweed farms, built in collaboration with the startup Phykos. The idea is for fast-growing seaweed to shed its carbon-absorbing biomass in the deep ocean. Researchers are using the C3 AI Suite to design novel AI-driven navigational models to enable the beds to “hitchhike” on ocean currents, a power savings a thousand times lower than conventional approaches.

“If we could grow seaweed on just 0.1 percent of the ocean and sink that biomass, we could remove all of Germany’s emissions,” team researcher **Marius Wiggert** told the audience at a Berlin conference this spring.



---

## CARBON COUNTS

Accurate and cost-effective carbon-credit accounting is key to establishing climate-smart agricultural practices and market-based agricultural carbon credit exchanges, according to **Kaiyu Guan**, associate professor of agroecosystem sensing and modeling at the **University of Illinois at Urbana-Champaign**. As carbon is a very accurate predictor of crop yields, computational models tracking carbon as it cycles through an agroecosystem also hold untapped potential to advance precision agriculture.

By example, identifying the prevalence of no-till and conservation tillage currently relies largely on self-reported stats compiled every five years. In Guan's study, University of Illinois scientists designed novel technology by integrating ground, airborne, and satellite imagery for cost-effective, real-time mapping of tilled land. Results reveal near 82 percent accuracy.

"Nobody has done this in the agricultural world," says Guan of his new cross-scale technology that significantly advances the capability "to create ground-truth information."



---

## FUTURE FORECASTING

Conventional wisdom says hot air rises, yet, in the tropical atmosphere, where humid air is lighter than dry air, the reverse is true. This phenomenon, known as vapor buoyancy, regulates low cloud cover over the subtropical ocean. Some commonly cited climate models fail to take this into account, according to **Da Yang**, faculty scientist at **Lawrence Berkeley National Laboratory**, and Yang's research demonstrates that the vapor buoyancy effect is more significant than previously understood. In climate models without vapor buoyancy, low cloud cover can be off by half as much in certain regions.

"In a warmer climate, the buoyancy effect of water vapor would be increasingly important due to more atmospheric water vapor," says Yang. "The biggest challenge in accurately predicting future climate change is clouds, so we have to get vapor buoyancy right."

*Seaweed can capture in its biomass the atmospheric carbon dioxide dissolved in the ocean. As the biomass decomposes, it makes its way to the deep ocean, where the carbon is buried for millennia. The idea behind our project is to use technology to amplify this natural process."*



– Claire Tomlin, Charles A. Desoer Professor of Engineering, University of California, Berkeley

## RESEARCH SPOTLIGHTS: AI TO HARDEN INFORMATION SECURITY AND SECURE CRITICAL INFRASTRUCTURE



By the spring of 2022, it was time to examine the vital structural integrity undergirding all this technology: Cybersecurity. Even the strongest ideas need the resilience to hold up to increasingly sophisticated and organized adversarial attacks. The C3.ai Digital Transformation Institute put out the call and the nation's top scientists responded, with proposals on how to identify and secure networks, systems, and infrastructure against vulnerabilities, threats, and adversaries.

---

### CRUISE CONTROL

In a five-day field trial on Interstate 24 outside of Nashville last fall, researchers led by **Alexandre Bayen**, professor of electrical engineering and computer science at the **University of California, Berkeley**, deployed a fleet of 100 semi-autonomous vehicles to test whether a new AI-powered cruise control system can help smooth the flow of traffic and improve fuel economy. "Preliminary results suggest that, even with a small proportion of these vehicles on the road, we can effectively change the overall behavior of traffic," says Bayen.

With more automated cars, however, comes the potential for more malicious adversaries that could compromise the system and degrade operations. Such attacks are difficult to detect with current techniques, according to **Carl A. Gunter**, engineering professor at the **University**

**of Illinois at Urbana-Champaign**, who is analyzing cybersecurity data from the road test.

"We are looking forward to seeing how well live data from I-24 Mobility can help advance our analytic techniques," says Gunter.

---

### SAFETY CAGES FOR I.T.

Industrial robots often operate within "safety cages" to protect workers from robots in operation. IT systems could use the same security, simple and explainable, according to **Cyrille Artho**, associate professor of theoretical computer science at the **KTH Royal Institute of Technology** in Stockholm, Sweden. Artho is designing a solution using machine learning and test generation, focusing on explainable AI so that the cage itself, along with its effects on network traffic, can be inspected and validated.

Lightweight approaches guarantee that such safety cages can be embedded in programmable networks or operating system kernels. Machine learning picks up behavioral models with roots in formal modeling (access policies, protocol states, Petri Nets); as a result, they are inherently readable by humans. Test-case generation validates diverse traces against the model and showcases malicious behavior, validating both positive and negative outcomes.

---

### **IDENTIFYING ATTACK SOURCES**

Critical infrastructures from power grids to transportation systems have increasingly been the target of unprecedented cyberattacks. At the consumer end, Internet of Things (IoT) enabled smart-home appliances, such as Wi-Fi-enabled air conditioners and plug-in-electric vehicles, present a new threat to power grid operations, with more than 30 billion devices globally. Yet their poor security features can invite malicious parties to gain access to the system. Unlike utility-side assets, compromised IoT-enabled end-user devices pose unique challenges to power grid operators.

**H. Vincent Poor**, Michael Henry Strater University Professor of Electrical and Computer Engineering at **Princeton University**, leads a research team aiming to locate attack sources and design new theoretical and algorithmic AI tools incorporating high accuracy, low computational complexity, and low sample complexity. Their 2022 study proposes data-driven algorithms to detect and identify IoT-enabled load-altering attacks against power grids. Numerical

results confirm that proposed data-driven algorithms outperform other benchmark techniques.

---

### **SOK: DEFI**

With the development of blockchain technology, decentralized finance (DeFi) has become an important player in today's economy, accumulating a peak total value of more than \$253 billion USD. Such rapid growth has drawn proportional attacks, totaling an estimated \$3.24 billion in losses in the past four years. Yet DeFi security remains in a nascent stage, according to **Dawn Song**, professor of electrical engineering and computer sciences at the **University of California, Berkeley**.

In an April 2023 paper, the researchers provide the first framework for reasoning about the DeFi system and threat models. After investigating 77 academic papers, 30 audit reports, and 181 real-world incidents, the team outlines a wide spectrum of adversarial goals, assumptions, prior knowledge, and capabilities, as well as common causes for potentially harmful DeFi incidents to create a standard model for related works.





# Awards & Honors

---

The C3.ai Digital Transformation Institute works with top researchers from the world's leading research universities to advance the Science of Digital Transformation. In recent years, many DTI Principal Investigators have been recognized for outstanding contributions in fields from engineering and computer science to environmental science and public health, among others.

---

## RECOGNITION CONFERRED ON DTI P.I.s & LEADERSHIP

### American Association for the Advancement of Science Fellow

**Michael Franklin**, University of Chicago, for “outstanding research contributions, with a focus on the creation and direction of the AMPLab” at the University of California, Berkeley

### National Academy of Engineering

**Klara Nahrstedt**, University of Illinois at Urbana-Champaign, and her mother, **Ruzena Bajcsy**, University of California, Berkeley, the first mother-daughter pair ever elected to NAE

### American Academy of Arts and Sciences

**Gerbrand Ceder**, University of California, Berkeley, and Lawrence Berkeley National Laboratory

### National Academy of Inventors Fellow

**Lawrence Thomas Pileggi**, Carnegie Mellon University

### Association for Computing Machinery Fellow

**Ben Y. Zhao**, University of Chicago

### Electrochemical Society Fellow

**Gerbrand Ceder**, University of California, Berkeley

### Honorary Doctorates

**Christopher Manning**, Stanford University, from the University of Amsterdam, Netherlands

**Alberto Sangiovanni-Vincentelli**, University of California, Berkeley, from AGH University of Science and Technology, Krakow, Poland

**Bin Yu**, University of California, Berkeley, from Université de Lausanne, Switzerland

### American Society of Mechanical Engineers Rufus Oldenburger Medal

**S. Shankar Sastry**, University of California, Berkeley

### Association for Computing Machinery SIGMETRICS Achievement Award

**R. Srikant**, University of Illinois Urbana-Champaign

### IEEE Fellows

Six DTI PIs have recently been elevated to IEEE Fellows, a distinction of achievement reserved for less than one-tenth of one percent of the organization's voting members worldwide.

**Alexandre Bayen**, University of California, Berkeley

**Carolyn L. Beck**, University of Illinois at Urbana-Champaign

**Alejandro Dominguez-Garcia**, University of Illinois at Urbana-Champaign

**Javad Lavaei**, University of California, Berkeley

**Asuman Ozdaglar**, Massachusetts Institute of Technology

**Henrik Sandberg**, KTH Royal Institute of Technology

## National Science Foundation CAREER Awards

**Chuchu Fan**, Massachusetts Institute of Technology, “DeepCertify: Data-driven Formal Approach to Safe Autonomy”

**Sanmi Koyejo**, Stanford University, “Probabilistic Models for Spatiotemporal Data with Applications to Dynamic Brain Connectivity”

**Da Yang**, Lawrence Berkeley National Laboratory, “Exploring the Role of Vapor Buoyancy in Clouds, Circulation and Climate”

## Association of American Publishers Prose Award

**Stefano Bertozzi**, University of California, Berkeley, for “Innovations in Journal Publishing” as editor-in-chief of “Rapid Reviews: COVID-19,” open-access overlay journal from MIT Press

## World Laureates Association Prize

**Michael I. Jordan**, University of California, Berkeley, for “fundamental contributions to the foundations of machine learning and its application”

## Skip Ellis Early Career Award

**Sanmi Koyejo**, Stanford University

## IEEE AI’s 10 to Watch

**Bo Li**, University of Illinois at Urbana-Champaign

## Elsevier Prize for Green Electrochemistry

**Xiao Su**, University of Illinois at Urbana-Champaign

## Highly Cited Researchers

Three DTI PIs were named to the 2022 global list of Highly Cited Researchers for demonstrating “significant and broad influence reflected in their publication of multiple highly cited papers over the last decade,” with papers ranking in the top one percent by numbers of citations in their field in a publication year.

**Gerbrand Ceder**, University of California, Berkeley, in Materials Science

**Kaiyu Guan**, University of Illinois at Urbana-Champaign, in Cross-field Research

**H. Vincent Poor**, Princeton University, cited in two fields, Computer Science and Engineering

## Test-of-Time Awards

Three DTI researchers were recognized with Test-of-Time awards for the highest-impact papers over 10 years or more, as determined by how often papers have been cited.

**Vern Paxton**, University of California, Berkeley, for 1998 paper cited 3,852 times, “Bro: A system for Detecting Network Intruders in Real Time,” USENIX

**Dawn Song** and **David Wagner**, University of California, Berkeley, for 2011 paper, “Android Permissions Demystified,” ACM SIGSAC

**Hanghang Tong**, University of Illinois at Urbana-Champaign, for 2011 paper, “BIG-ALIGN: Fast Bipartite Graph Alignment,” IEEE International Conference on Data Mining

## Best Paper Awards

**Saurabh Amin**, Massachusetts Institute of Technology, “Interpretable Machine Learning Models for Modal Split Prediction in Transportation Systems,” IEEE Transactions on Intelligent Transportation Systems

**Marta Gonzalez**, University of California, Berkeley, “Safe Driving Using Mobile Phones,” IEEE Transactions on Intelligent Transportation Systems

**Michael I. Jordan**, University of California, Berkeley, “On-Demand Sampling: Learning Optimally from Multiple Distributions,” NeurIPS 2022

**H. Vincent Poor**, Princeton University, “Learning Mixtures of Linear Dynamical Systems,” 39th International Conference on Machine Learning



# Colloquia

## ON DIGITAL TRANSFORMATION SCIENCE

The Colloquium on Digital Transformation is a series of weekly online talks on how artificial intelligence, machine learning, and big data can lead to scientific breakthroughs with large-scale societal benefit. From rising stars to venerable authorities, the series hosts the nation's top experts in artificial intelligence each term.

With a focus on Federated Learning, the fall 2022 lineup included several top tech industry scientists and entrepreneurs as well as academic researchers from leading universities. Google research scientist Brendan McMahan, whose team pioneered the concept of Federated Learning, was among industry leaders represented.

DTI talks are open to the public at no charge; all recordings are posted at [YouTube.com/C3DigitalTransformationInstitute](https://www.youtube.com/C3DigitalTransformationInstitute).



**6**  
Series on  
AI topics



**80**  
Countries  
tuning in



*What we're going to see moving forward is the transformation of almost every industry according to the principles of data-driven learning and machine intelligence."*

*– Zico Kolter, Associate Professor, Carnegie Mellon University*





$$f(\mathbf{x}_1, \dots, \mathbf{x}_K) = \frac{1}{K} \sum_{i=1}^K f_i(\mathbf{x}_i)$$

## FALL SERIES 2022

August 25

**The Many Facets of Robust Machine Learning: from Mathematical Guarantees to Real-world Shifts**

Aditi Raghunathan, Carnegie Mellon University

---

September 1

**Two Surprises When Optimization Meets Machine Learning**

Suvrit Sra, Massachusetts Institute of Technology

---

September 8

**Trustworthy Machine Learning: Robustness, Privacy, Generalization, and their Interconnections**

Bo Li, University of Illinois at Urbana-Champaign

---

September 15

**Federated Learning with Formal User-Level Differential Privacy Guarantees**

Brendan McMahan, Google

---

September 22

**New Approaches to Detecting and Adapting to Domain Shifts in Machine Learning**

Zico Kolter, Carnegie Mellon University

---

October 6

**Machine Learning at All Levels: A Pathway to “Autonomous” AI**

Eric Xing, Carnegie Mellon University

---

October 13

**Adversarial Machine Learning from a Privacy Perspective**

Tom Goldstein, University of Maryland

---

October 20

**Improving Communication for Differential Privacy: Insight from Human Behavior**

Rachel Cummings, Columbia University

---

October 27

**AI Model Inspector: Towards Holistic Adversarial Robustness for Deep Learning**

Pin-Yu Chen, IBM

---

November 10

**Underspecified Foundation Models Considered Harmful**

Nicholas Carlini, Google Brain

---

November 17

**Tackling Computational and Data Heterogeneity in Federated Learning**

Gauri Joshi, Carnegie Mellon University

---

December 1

**Improved Adversarial Attacks and Certified Defenses Via Nonconvex Relaxations**

Richard Y. Zhang, University of Illinois at Urbana-Champaign

---

December 8

**The Power of Adaptivity in Representation Learning: From Meta-learning to Federated Learning**

Sewoong Oh, University of Washington

# Workshops

## ON DIGITAL TRANSFORMATION SCIENCE

---

The Workshops on Digital Transformation Science are deep dives into foundational topics in digital transformation. These multi-day events are intended for researchers, practitioners, policymakers, and others interested in gaining insights and understanding of topics and trends from leading experts and scientists. Find all workshop videos on the C3.ai DTI YouTube channel at [YouTube.com/C3DigitalTransformationInstitute](https://www.youtube.com/C3DigitalTransformationInstitute).

October 10-12, 2022

### DATA, LEARNING, AND MARKETS

#### Abstract:

Online markets that match buyers and sellers are ubiquitous. Well-known examples include Airbnb, Uber, Lyft, and eBay. All of these online markets rely on machine learning algorithms to learn from past data to maximize profit and/or utility of the platform to users. Additionally, new markets are emerging for data itself: an example is the emerging research on the design of privacy-preserving mechanisms for the possible release of private data by individuals for the greater good of society. This workshop explores theoretical aspects of the interplay between data science, machine learning, game theory, and economics for the design of markets using data and for the design of data markets.

#### Tutorials:

Tutorials on dynamic spatial matching, ranking problems for platforms with a mix of real and fake users, new signaling schemes, transforming offline algorithms to their online counterparts, and designing exchanges to allow arbitrary pairs of assets to be traded against each other without the presence of a reserve currency.

#### Organizers:

Rasoul Etesami (University of Illinois at Urbana-Champaign), Eric Mazumdar (California Institute of Technology), R. Srikant (University of Illinois at Urbana-Champaign)

#### Speakers:

Subhonmesh Bose (University of Illinois at Urbana-Champaign), Ozan Candogan (University of Chicago), Rasoul Etesami (University of Illinois at Urbana-Champaign), Raul Castro Fernandez (University of Chicago), Nikhil Garg (Cornell University), Ashish Goel (Stanford University), Negin Golrezaei (Massachusetts Institute of Technology), Thibaut Horel (Massachusetts Institute of Technology), Nicole Immorlica (Microsoft Research), Yash Kanoria (Columbia University), Eric Mazumdar (California Institute of Technology), Kamesh Munagala (Duke University), Rad Niazadeh (University of Chicago), Amin Saberi (Stanford University), Vijay Vazirani (University of California, Irvine)



**9** Multi-day online seminars



**60** Hours of talks by research scientists and industry leaders

#### DTI WORKSHOPS 2020-2022

- Epidemics and (Mis)Information
- Trusted Autonomous Systems
- Machine Learning for Carbon-free Electricity
- Machine Learning Networks
- Analytical Foundations of Deep Learning
- Data-driven Decision-making in Socio-technical Systems
- Data Analytics in Security and Privacy
- Smart Buildings
- Data, Learning, and Markets

See all presentations at: [YouTube.com/C3DigitalTransformationInstitute](https://www.youtube.com/C3DigitalTransformationInstitute).

# Digital Transformation At Scale

C3.AI DIGITAL TRANSFORMATION INSTITUTE 2022 ANNUAL RESEARCH SYMPOSIUM

The annual C3.ai Digital Transformation Institute Research Symposium brings together leaders from academia, research laboratories, industry, and government. In 2022, the two-day symposium was held in Miami, Florida, and hosted two keynotes and talks and panels with researchers on project progress and results.

## KEYNOTE 1 - DEFENDING FORWARD: A PROACTIVE POSTURE FOR BUILDING U.S. CYBER RESILIENCY



**Lt. Gen. Charles L. Moore, Jr.**  
Deputy Commander  
United States Cyber Command

“The successful implementation of integrated deterrence will require a level of cybersecurity and defense far beyond what we can accomplish without the use of Artificial Intelligence. We must utilize it to realize the speed, scale, security, defense, and offense necessary to be successful.

“Touching on the role that AI must play going forward if we are to maintain our military security, deter conflict, and prevail should an actual war occur, we all understand that we are living in a time where data has become a resource and a commodity. From a military perspective, I would offer that data is the ammunition, information is the weapon, and influencing behavioral change is the ultimate effect. And based on the realities of digital convergence, I am convinced that a strategy focused on information superiority is required to ensure our long-term success.”

## KEYNOTE 2 - THE GREAT GAME: COMING OUT ON TOP IN CHESS AND CYBERSECURITY



**Garry Kasparov**  
World Chess Champion and Chairman  
Human Rights Foundation

“We need new and creative ways of using the advantage of free minds – people from the free world who can innovate – and most important, what’s the biggest advantage of people in the free world versus people who work in Russia and China? We have freedom to fail. That’s everything. Failure is unacceptable in authoritarian systems. Yes, they can fund a hundred startups, but they have to know exactly which one will be Google. Failure is the element that makes our system more flexible, more resilient, and more productive.

“We know what matters most; machines don’t. I believe we should be optimists about the future. These machines, they are not harbingers of utopia or dystopia – they are tools, made by us. We here in the free world have the huge advantage of entertaining free minds, freedom of competition, freedom of innovation, freedom to fail.”

M



Held in Miami, the symposium brought together leadership, researchers, and students from around the consortium with industry and government leaders and professionals.



Awarded projects of the C3.ai Digital Transformation Institute's Third Call for Proposals were announced at the event. Co-director Shankar Sastry of UC Berkeley hosted a panel discussion on AI for cybersecurity that included new awardees from around the consortium.



Students presented posters from their work on research projects related to AI for COVID-19 mitigation and for energy and climate security.





# Global Community

---



Since the launch of the C3.ai Digital Transformation Institute in March 2020, the Institute has built a worldwide community of subscribers, followers, fans, attendees, and audience members across online platforms from social media platforms to Zoom webinars.

*Sign up to join the conversation at the C3DTI.ai website and on our LinkedIn, Twitter, Facebook, and YouTube channels.*

---

## GLOBAL AUDIENCE



**10K**  
Followers



**15K**  
Interactions

### TOP 10 COUNTRIES

1. United States
2. Mexico
3. Germany
4. India
5. Armenia
6. Puerto Rico
7. Canada
8. Hong Kong
9. Nigeria
10. Thailand

### TOP 10 LANGUAGES SPOKEN BY AUDIENCES

1. English
2. Japanese
3. Spanish
4. Chinese
5. Korean
6. Swedish
7. Portuguese
8. Arabic
9. Farsi
10. French

## SELECT PUBLICATIONS

### COVID-19 MITIGATION RESEARCH

Detection of Covid-19 from Joint Time and Frequency Analysis of Speech, Breathing and Cough Audio, ICASSP 2022, Narendra Ahuja, et al.

Optimal Information Provision for Strategic Hybrid Workers, IEEE CDC 2022, Saurabh Amin, et al.

Non-communicable diseases and inequalities increase risk of death among COVID-19 patients in Mexico, PLoS One, Oct. 2020, Stefano Bertozzi, et al.

Detecting and mitigating simultaneous waves of COVID-19 infections. Nature Scientific Reports, Oct. 2022, Subhonmesh Bose, et al.

Data-driven control of COVID-19 in buildings: a reinforcement-learning approach, Dec. 2022, Munther Dahleh, et al.

U.S. eviction filing patterns in 2020, Socus 7, 2021, Co-PIs Matthew Desmond, Peter Hepburn, and Timothy A. Thomas, et al.

A pan-variant mRNA-LNP T cell vaccine protects HLA transgenic mice from mortality after infection with SARS-CoV-2 Beta, Frontiers in Immunology, Mar. 2023, David K. Gifford, et al.

Role of standard and soft tissue chest radiography images in deep-learning-based early diagnosis of COVID-19, Journal of Medical Imaging vol. 8, 2021, Maryellen Giger, et al.

Cascaded deep transfer learning on thoracic CT in COVID-19 patients treated with steroids, Journal of Medical Imaging vol. 8, 2021, Maryellen Giger, et al.

Robust and personalized federated learning with spurious features: an adversarial approach, NSF-PAR 2021, Sanmi Koyejo, Klara Nahrstedt, et al.

Host heterogeneity and epistasis explain punctuated evolution of SARS-CoV-2, PLOS, Feb. 2023, Simon A. Levin, et al.

Lockdown interventions in SIR models: Is the reproduction number the right control variable?

IEEE CDC, Dec. 2021, Asuman Ozdaglar, Daron Acemoglu, et al.

A Data-Informed Approach for Analysis, Validation, and Identification of COVID-19 Models, ACS 2021, Prashant Mehta, Tamer Başar, Rebecca L. Smith, Matthew West, et al.

A 680,000-person megastudy of nudges to encourage vaccination in pharmacies, NAS PNAS, Feb 2022, Sendhil Mullainathan, et al.

Fundamental Limitations on Efficiently Forecasting Certain Epidemic Measures in Network Models, NAS Proceedings vol. 119, Jan. 2022, H. Vincent Poor, et al.

Backcasting COVID-19: A Physics-Informed Estimate for Early Case Incidence, Royal Society Open Science, Dec. 2022, Zoi Rapti, et al.

Enabling accurate and early detection of recently emerged SARS-CoV-2 variants of concern in wastewater, Nature Communications, May 2023, Todd Treangen, et al.

### ENERGY & CLIMATE SECURITY RESEARCH

Smoothed Adversarial Linear Contextual Bandits with Knapsacks, ICML 2022, Arindam Banerjee, et al.

Identifying climate model structural inconsistencies allows for tight constraint of aerosol radiative forcing, 2022, Hamish Gordon, et al.

Optimizing Nitrogen Management with Deep Reinforcement Learning and Crop Simulations, AAMAS 2023 Proceedings, Naira Hovakimyan, et al.

Winning the CityLearn Challenge: Adaptive Optimization with Evolutionary Search under Trajectory-based Guidance, AAAI 2023, Ming Jin, et al.

When Should We Prefer Offline Reinforcement Learning Over Behavioral Cloning? ICLR 2022, Sergey Levine, et al.

Dynamic graph-based anomaly detection in the electrical grid, IEEE Transactions on Power Systems 2021, Amritanshu Pandey, et al.

Mobile storage for demand charge reduction, IEEE Transactions on Intelligent Transportation Systems vol. 7, July 2022, Kameshwar Poolla, et al.

A Meta-Learning Approach to the Optimal Power Flow Problem Under Topology Reconfigurations, IEEE Open Access Journal of Power and Energy, vol. 9, Jan. 2022, H. Vincent Poor, et al.

Data-Driven Detection and Identification of IoT-Enabled Load-Altering Attacks in Power Grids, IET Smart Grid vol. 5, June 2022, H. Vincent Poor, et al.

Green Routing Game: Strategic Logistical Planning Using Mixed Fleets of ICEVs and EVs, IEEE CDC, 2022, Henrik Sandberg, et al.

Predicting the Activities of Drug Excipients on Biological Targets using One-Shot Learning, Journal of Physical Chemical B, Feb. 2022, Diwakar Shukla, et al.

Imparting Selective Fluorophilic Interactions in Redox Copolymers for the Electrochemically Mediated Capture of Short-Chain Perfluoroalkyl Substances, Journal of the American Chemical Society 2023, March 2023, Xiao Su, Diwakar Shukla, et al.

Hamilton-Jacobi multi-time reachability, 2022 IEEE CDC, Dec. 2022, Claire Tomlin, et al.

Navigating underactuated agents by hitchhiking forecast flows, IEEE CDC 2022, Claire Tomlin, et al.

Physics-informed Neural Network-based Online Impedance Identification of Voltage Source Converters, IEEE Transactions on Industrial Electronics 2023, Qianwen Xu, et al.

Substantial influence of vapour buoyancy on tropospheric air temperature and subtropical cloud, National Geoscience 15, 2022, Da Yang, et al.

## CYBERSECURITY RESEARCH

An Active Learning Approach to Dynamic Alert Prioritization for Real-time Situational Awareness, IEEE CNS, Oct. 2022, György Dán, et al.

Compositional Neural Certificates for Networked Dynamical Systems, L4DC 2023, Chuchu Fan, et al.

EDoG: Adversarial edge detection for graph neural networks, IEEE SaTML 2023, Carl Gunter, Bo Li, et al.

Contrastive Learning with Complex Heterogeneity. KDD 2022, Jungrui He, et al.

Byzantine-robust federated learning with optimal statistical rates and privacy guarantees, Proceedings of Machine Learning Research, 2023, Dawn Song, Michael. I. Jordan, et al.

An Efficient Homotopy Method for Solving the Post-contingency Optimal Power Flow to Global Optimality, IEEE Access vol. 10, Nov. 2022, Javad Lavaei, et al.

Certifiably Robust Collaborative Inference via Feature Purification, NeurIPS 2022, Bo Li, Sanmi Koyejo, et al.

Degradation Attacks on Certifiably Robust Neural Networks, Transactions on ML Research, Nov. 2022, Corina Pasareanu, et al.

Data-Driven Detection and Identification of IoT-Enabled Load-Altering Attacks in Power Grids, IET Smart Grid vol. 5, June 2022, H. Vincent Poor, et al.

Learning Monitorable Operational Design Domains for Assured Autonomy, ATVA, October 2022, Sanjit A. Seshia, et al.

SoK: Decentralized Finance (DeFi) Incidents, IEEE S&P 2023, Dawn Song, et al.

BFT Protocol Forensics, ACM CCS 2021, Pramod Viswanath, et al.

Post-breach Recovery: Protection against White-box Adversarial Examples for Leaked DNN Models, ACM CCS, Nov. 2022, Ben Y. Zhao, et al.







## **CONTACT**

### **C3.ai Digital Transformation Institute @ Berkeley**

University of California, Berkeley  
750 Sutardja Dai Hall, MC 1764  
Berkeley, California 94720-1764

### **C3.ai Digital Transformation Institute @ Illinois**

University of Illinois at Urbana-Champaign  
1205 W. Clark Street, MC-257, Room 1008  
Urbana, Illinois 61801

### **Inquiries and more information**

For general inquiries:  
[\*info@c3dti.ai\*](mailto:info@c3dti.ai)

Sign up for DTI newsletter:  
[\*C3DTI.ai/contact\*](https://C3DTI.ai/contact)

Copyright: C3.ai Digital Transformation Institute © 2023





# C3.ai Digital Transformation Institute



 [C3DTI.ai](https://www.c3dti.ai)

 [@C3DTI](https://twitter.com/C3DTI)

 [facebook.com/C3DTI](https://www.facebook.com/C3DTI)

 [linkedin.com/company/c3-ai-digital-transformation-institute](https://www.linkedin.com/company/c3-ai-digital-transformation-institute)

 [youtube.com/C3DigitalTransformationInstitute](https://www.youtube.com/C3DigitalTransformationInstitute)