

COMPUTING RESEARCH

ADDRESSING NATIONAL PRIORITIES AND SOCIETAL NEEDS



CCC

Computing Community Consortium
Catalyst

POSTERS

Poster Session List

| Name | Affiliation | Poster Title |
|---------------------|---|---|
| Sean Andrist | University of Wisconsin | Situated Gaze Mechanisms for Embodied Agents |
| Matthew Bietz | University of California, Irvine | Personal Data for Public Health: Opportunities and Challenges |
| Yonatan Bisk | University of Southern California | Natural Language Communication with Computers |
| Gedare Bloom | Howard University | Security and Provenance in the Internet of Things |
| Emma Brunskill | Carnegie Mellon University | Advancing Human Potential with Interactive Machine Learning |
| Tamara Denning | University of Utah | Human-Centered Computer Security and Privacy |
| Bistra Dilkina | Georgia Tech | Computational Sustainability @ GT |
| Lu Feng | University of Pennsylvania | Assuring the safety of cyber-physical systems interacting with human operators |
| Yong Ge | The University of North Carolina at Charlotte | Mining Patterns of Substance Use with Social Media Data |
| Vivian Genaro Motti | George Mason University | Wearable Health: Exploring Human-Centered Solutions of On-Body Technologies to Improve Healthcare |
| Samuel Gordon | George Mason University | Computing on Encrypted Data |
| Christopher Harris | Brown University | Processor Design Exploration for Vision Based Mobile Robots |
| Lane Harrison | Worcester Polytechnic Institute | Human-Centered Data Visualization |
| Michael Hay | Colgate University | Principled Evaluation of Differentially Private Algorithms using DPBench |
| Katie Henry | Johns Hopkins University | Automated in-patient monitoring in the ICU with application to septic shock prediction |
| Bert Huang | Virginia Tech | Weakly Supervised Cyberbullying Detection in Social Media |
| Patrick Gage Kelley | University of New Mexico | Privacy as Iconography: A Pictographic Collection and Comparison |
| Daniel Kifer | Penn State University | Measuring Information Leakage in Images |
| Edward Kim | Villanova University | Utilizing Deep Neural Networks for Classification in Cytopathology |
| Bart P. Knijnenburg | Clemson University | Adaptive Privacy Decision Support |
| Parisa Kordjamshidi | University of Illinois Urbana-Champaign | Declarative Learning Based Programming |
| Su-In Lee | University of Washington | Big data approach to identify novel biomarkers in cancer |

| Name | Affiliation | Poster Title |
|--------------------------|---|---|
| Xiaohui Liang | University of Massachusetts Boston | Towards Privacy-preserving Data Release in Mobile Healthcare |
| Jonathan May | University of Southern California | Advances in Rapid-Response Low-Resource Machine Translation |
| Mahsa Mirzargar | University of Utah | Empowering Uncertainty Characterization Using Scientific Visualization |
| Jamie Morgenstern | University of Pennsylvania | Informational Frictions leading to Rational Discrimination in Learning |
| Fay Cobb Payton | NC State University | Big Data - Opportunities & Challenges of Social Impacts |
| Pavithra Prabhakar | Kansas State University | Robust Verification of Cyber-Physical Systems |
| John Sampson | Penn State University | Non-Volatile Processors for Energy-Harvesting Platforms |
| Mohamed Sarwat | Arizona State University | GeoExpo - Interactive and Scalable Exploration of Big GeoSpatial Data |
| Ben Shapiro | University of Colorado Boulder | BlockyTalky: IoT, Networking, and Distributed Systems CS Education for All |
| Anshumali Shrivastava | Rice University | Scalable, Sustainable and Secure Machine Learning via Probabilistic Hashing |
| Jonathan Ullman | Northeastern University | Robust Traceability from Trace Amounts |
| Ting Wang | Lehigh University | Serendipity or Preparedness? Quantifying Creativity in Scientific Enterprise |
| Dong Wang | University of Notre Dame | Data Reliability in Cyber-Physical Systems for Smart Cities |
| Hongning Wang | University Of Virginia | Online Collaborative Learning from Interaction with Humans |
| Gloria Washington | Howard University | Affective Technologies for Improving the Lives of Persons With Chronic Disease |
| Chuan Yue | Colorado School of Mines | SafeSky: A Secure Cloud Storage Middleware for End-user Applications |
| Ting Zhu | University of Maryland, Baltimore County | Gait-Based Privacy Preserving Technique for Smart Cities |

Poster Descriptions

Sean Andrist • University of Wisconsin

Situated Gaze Mechanisms for Embodied Agents

This poster summarizes my PhD research on how embodied agents--both virtual agents and physical robots--can achieve positive social and communicative outcomes through the use of situated gaze mechanisms. It will present several projects I have carried out to design and test models of gaze behavior for agents in various contexts. These projects include (1) how agents can produce gaze shifts that target specific high-level interaction outcomes, (2) how agents can effectively utilize gaze aversions in conversation, (3) how agents can coordinate their gaze with the user's gaze while collaborating on a physical task, and (4) how agents can adapt their gaze behaviors to the personality of their users for rehabilitation.

Matthew Bietz • University of California, Irvine

Personal Data for Public Health: Opportunities and Challenges

New forms of data can help us understand human health. Streams of data and traces of human behavior are being generated by mobile apps, wearable devices, online activity, social media and the Internet of Things. My research investigates the opportunities and challenges for using these new forms of personal data for research in public health, including issues of privacy, research ethics, and the potential biases in these datasets.

Yonatan Bisk • University of Southern California

Natural Language Communication with Computers

We propose a framework for devising testable algorithms for bridging the communication gap between humans and robots. We begin with a setting in which humans give instructions to robots using unrestricted natural language commands, with instruction sequences aimed at building complex goal configurations in a blocks world. We then flip the paradigm to also address the problem of language generation, where a human performs commands produced by a machine to demonstrate two-way communication.

Gedare Bloom • Howard University

Security and Provenance in the Internet of Things

Embedded and cyber-physical systems are increasingly vulnerable to attack especially because of ubiquitous network access: the Internet of Things (IoT) refers to the proliferation of such systems. The biggest challenge that IoT brings is the multifaceted attack surfaces spanning embedded systems, mobile applications, clouds services, and the gateways and networks connecting them. This project aims to improve device security using low-cost solutions in open source embedded system software for IoT end points.

Emma Brunskill • Carnegie Mellon University

Advancing Human Potential with Interactive Machine Learning

There is an increasing need to assist people to reach their full potential-- for example, in 2007 over one-fifth of entering college students required remedial education. Relying on human experts cannot provide the scale and temporal persistence required. Since it is now easy to store and process huge amounts of information, interactive machine learning systems could be a key part of the solution.

Tamara Denning • University of Utah
Human-Centered Computer Security and Privacy

Technology is becoming increasingly capable (e.g., sensors, actuators, connectivity, mobility), and we are incorporating it into a growing portion of our lives. As a result, we are exposed to unprecedented security and privacy threats. Unfortunately (or fortunately, depending on how you look at it), security and privacy are not solely technical problems; frequently the weakest link in the chain is the interface between the technical and the human, logistical, or legal.

Bistra Dilkina • Georgia Tech
Computational Sustainability @ GT

We present some of the work in Computational Sustainability done at Georgia Tech. In particular, we focus on large-scale combinatorial optimization models and methods to facilitate cost-effective biodiversity and landscape connectivity conservation. With increasing habitat fragmentation and shifts due to urban growth, climate change and sea level rise, it is key to develop effective data-driven and scalable decision support tools that can help policy makers take informed choices.

Lu Feng • University of Pennsylvania
Assuring the safety of cyber-physical systems interacting with human operators

Cyber-physical systems (CPS) are increasingly everywhere, providing new capabilities to improve quality of life and transform many critical areas. However, significant new challenges are also posed for assuring the safety of CPS. This poster presents an overview of my work on assuring the safety of CPS that interact with human operators, focusing on addressing CPS safety challenge due to the uncertainty of human interaction, with applications to medical devices and unmanned aerial vehicles.

Yong Ge • The University of North Carolina at Charlotte
Mining Patterns of Substance Use with Social Media Data

This project collects and analyzes Twitter data to extract public dialogues related to substance use and then uses the results to uncover patterns of substance use. We will develop advanced data mining methods to extract tweets related to substance use by leveraging substance slang collected from online dictionaries, and determine patterns of substance use, such as emerging types of substances, new combinations of substance use, and dynamic geographic and temporal distributions of substance.

Vivian Genaro Motti • George Mason University
Wearable Health: Exploring Human-Centered Solutions of On-Body Technologies to Improve Healthcare

Wearable technologies have a large potential to support healthcare. Their solutions not only span across different stages and domains of medical care but also benefit care givers and patients. By understanding the potential opportunities that emerge from wearables in healthcare from a human centered perspective, we aim at bringing technologies seamlessly in to users' lives, supporting their daily activities and bridging the gap between what users need and what wearable technologies provide them.

Samuel Gordon • George Mason University
Computing on Encrypted Data

With data breaches becoming more prevalent and their impact more severe, encrypting data while it is at rest no longer suffices. More advanced techniques in cryptography, such as secure computation, homomorphic encryption and functional encryption provide the possibility of performing computations on data while it remains encrypted, allowing us to guarantee its security while preserving its utility.

Christopher Harris • Brown University

Processor Design Exploration for Vision Based Mobile Robots

Vision based mobile robots such as UAVs require increasingly complex onboard processors. These systems require high-performance, low power computing be integrated with actuators and data-intensive sensors in a system providing real-time performance. This research describes a new modeling framework for the design of specialized system-on-chip (SoC) modules for sensor based robots. This work represents a new and untapped frontier at the intersection of robotics and computer architecture.

Lane Harrison • Worcester Polytechnic Institute

Human-Centered Data Visualization

The world is becoming data-driven, and people are relying on visualizations to make high-impact, life-critical decisions. While visualization research has delivered new tools and techniques, we lack an understanding of exactly how much information people perceive and process. Human-centered visualization leverages cognitive and perceptual principles to advance the science of visualization, while creating opportunities for next-generation systems.

Michael Hay • Colgate University

Principled Evaluation of Differentially Private Algorithms using DPBench

DPBench is a framework for standardized evaluation of differentially private algorithms based on a set of novel principles that ensure sound empirical evaluation. Using DPBench we study 15 algorithms for answering multi-dimensional range queries, observing new insights into algorithm behavior and clarifying the state of the art. With NSF funding, DPBench is being developed into an online site designed to support a broad community, including data owners interested in using differential privacy.

Katie Henry • Johns Hopkins University

Automated in-patient monitoring in the ICU with application to septic shock prediction

Clinicians are continually monitoring which patients in the ICU are at high of risk of developing an adverse event. We can leverage machine learning to develop models that use higher dimensional input to identify developing events earlier than traditional approaches, which typically rely on paper-based scores. In particular we demonstrate the ability to automatically identify septic shock, a leading cause of death in the United States, many hours earlier than current clinical practices.

Bert Huang • Virginia Tech

Weakly Supervised Cyberbullying Detection in Social Media

A growing majority of human communication occurs online. Advances in mobile and connected technology amplify individuals' abilities to interact. Unfortunately, the amplification of social connectivity includes a disproportionate amplification of detrimental behavior such as cyberbullying. We are developing automated, social-aware, and data-driven methods for cyberbullying detection. These methods could enable technologies that mitigate the harm and toxicity created by these harmful behaviors.

Patrick Gage Kelley • University of New Mexico

Privacy as Iconography: A Pictographic Collection and Comparison

The promise of icons as a silver bullet for simplifying the complexity of privacy policies, permissions, data sharing, and other similarly tangled privacy applications remains unmet. This poster explores the visual language of these icons, across platforms and contexts, how they have evolved over the last decade, and attempts to intuit what the future of privacy icons holds.

Daniel Kifer • Penn State

Measuring Information Leakage in Images

In this poster we present recent developments in ways of measuring the effective number of bits leaked by various image transformations. Their goal is to enable rigorous privacy definitions for smart cameras. The measures satisfy fundamental criteria such as the data processing inequality and linear composition.

Edward Kim • Villanova University

Utilizing Deep Neural Networks for Classification in Cytopathology

Cytopathology is the study of disease at the cellular level and often used as a screening tool for cancer. With recent advances in the digitization of pathology and machine learning techniques, there are transformative opportunities for computers to assist pathologists in their analysis and diagnosis of disease. Toward this goal, we describe the utilization of a particular methodology, deep convolutional neural networks, to the application of cytopathology classification.

Bart P. Knijnenburg • Clemson University

Adaptive Privacy Decision Support

Privacy concerns are an important barrier to the growth of user data-driven applications. To help users balance the benefits and risks of information disclosure in a user-friendly manner, I developed a “privacy adaptation procedure” that offers tailored privacy decision support. This procedure predicts users’ privacy preferences and behaviors based on their past behavior and known characteristics. It then provides automatic adaptive default settings in line with users’ disclosure profile.

Parisa Kordjamshidi • University of Illinois Urbana-Champaign

Declarative Learning Based Programming

Developing intelligent problem solving systems for real world applications is very challenging. I present Saul, a declarative machine-learning based programming language that is designed to help experts in various domains who are not expert in machine learning, to design complex data-driven intelligent systems. Such a language helps reusability and replicability of research results and facilitates the use of structured learning algorithms, new data resources and background knowledge.

Su-In Lee • University of Washington

Big data approach to identify novel biomarkers in cancer

Cancer is full of mysteries. Two individuals with seemingly similar tumors sometimes have very different responses to chemotherapy and other treatments, as well as drastically different survival outcomes. In order to better understand this phenomenon, researchers have developed ways to obtain a molecular snapshot of an individual’s tumor, which amounts to gigabytes of data. I will present our recent development of novel machine learning approaches to identify biomarkers from these big data.

Xiaohui Liang • University of Massachusetts Boston

Towards Privacy-preserving Data Release in Mobile Healthcare

Mobile healthcare (mHealth) integrates advanced sensing and communication technologies to continuously monitor the data that affects and reflects the private health status of patients. The privacy-preserving data release in mHealth is challenging because the release process over a huge volume of data must be efficient, fast, and accurate. In this poster, we will present new query-based data release and differentially private data release approaches to address the unique challenges of mHealth.

Jonathan May • University of Southern California

Advances in Rapid-Response Low-Resource Machine Translation

In 2010, response to the devastating earthquake in Haiti was hampered by an inability of international aid workers to understand Haitian Creole. Machine Translation (MT) has helped bridge the language barrier, however, state-of-the-art quality is generally limited to cases where hundreds of millions of words of human translation have already been collected. I will show how recent advances in neural network learning can be leveraged to yield state-of-the-art MT results with limited resources.

Mahsa Mirzargar • University of Utah

Empowering Uncertainty Characterization Using Scientific Visualization

Data-intensive methods are transforming scientific discovery. Visualization, as an integral component of the data analysis, can facilitate the exploration and communication of the data and the uncertainty. In this poster, I will present novel uncertainty visualization paradigms that use nonparametric statistical analysis to derive robust summaries from uncertain data. I will demonstrate their utility for analyzing and visualizing simulation outputs used to predict the path hurricanes will take.

Jamie Morgenstern • University of Pennsylvania

Informational Frictions leading to Rational Discrimination in Learning

A large part of digital infrastructure is designed to automate decision making as a function of historical data. If this infrastructure is to be truly trustworthy, it must have several important features, of which fairness is of primary importance. We study a well-known learning algorithm which is harsher when classifying smaller (equal quality, equally classifiable) group over large groups. We investigate the cost of possible interventions for this inherent bias against smaller samples.

Fay Cobb Payton • NC State University

Big Data - Opportunities & Challenges of Social Impacts

This poster will cover the current findings in the big data research. The foci will include trends in the literature, questions around computing and will outline opportunities/challenges where the field falls short of addressing social impacts of big data. The content will also include findings from my research publications to highlight where voids are in the literature.

Pavithra Prabhaka • Kansas State University

Robust Verification of Cyber-Physical Systems

The poster describes the recent results and software tools for automated analysis of robustness of cyber-physical systems, with particular focus on the interaction of discrete and continuous dynamics that arise due to the interaction of digital controllers with physical systems. The techniques borrow ideas from formal methods, dynamical systems theory and control theory.

John Sampson • Penn State University

Non-Volatile Processors for Energy-Harvesting Platforms

Nonvolatile processors (NVPs) are a promising solution for energy-harvesting scenarios in which the available power supply is unstable, intermittent, and unpredictable. We explore the space of NVP architectures and policies to maximize forward progress and show how heterogeneity in microarchitecture and prediction mechanisms in power management can greatly enhance conversion of harvested energy into useful computation.

Mohamed Sarwat • Arizona State University

GeoExpo - Interactive and Scalable Exploration of Big GeoSpatial Data

Recently, the volume of spatial data increased tremendously. Such data includes but not limited to: weather maps, socioeconomic data, and vegetation indices. Making sense of spatial data is beneficial for several applications that may transform science and society, e.g., socio-economic, climate change analysis, urban planning, road network design, and fast disaster response. This poster presents GeoExpo a data management system that enables interactive and scalable exploration of spatial data.

Ben Shapiro • University of Colorado Boulder

BlockyTalky: IoT, Networking, and Distributed Systems CS Education for All

Most technologies that kids use daily, e.g. texting, online games, and, increasingly, Internet of Things devices, apply ideas from CS that are missing from current K12 CS Education. Simultaneously, many researchers are looking for new ways to motivate members of under-represented populations to learn computer science. We are investigating how enabling kids to build distributed and physical computing systems could broaden participation in computer science and train students in modern CS.

Anshumali Shrivastava • Rice University

Scalable, Sustainable and Secure Machine Learning via Probabilistic Hashing

Modern big-data settings pose a new set of challenges which primarily include scalability, energy efficiency and privacy. I design probabilistic hashing algorithms which provide a practical and provable solutions to all the three major challenges. These algorithms trade a very small amount of certainty, with often exponential gains in the computations, memory and energy efficiency. They also provide secure probabilistic encoding of the data which do not reveal the attributes information.

Jonathan Ullman • Northeastern University

Robust Traceability from Trace Amounts

We demonstrate a simple attack that takes seemingly innocuous summary statistics on a dataset (e.g., allele frequencies in a genome-wide association study) and information about a targeted individual, and can determine whether or not that individual is in the dataset. The summary statistics can be extremely noisy and our attacker needs very little information about the population. Our attack demonstrates the surprising ways information can leak, and the necessity of formal privacy protections.

Ting Wang • Lehigh University

Serendipity or Preparedness? Quantifying Creativity in Scientific Enterprise

The building blocks of scientific innovations are often embodied in existing knowledge, yet it is creativity that blends disparate ideas. Here, by correlating researchers' information consumption with their publications, we find remarkable predictability in scientific creative processes. Further, we develop a mechanistic model that not only effectively predict disparate ideas likely to be linked by creativity, but also identifies critical references for such linkings to happen.

Dong Wang • University of Notre Dame

Data Reliability in Cyber-Physical Systems for Smart Cities

Consider a Cyber-Physical System (CPS) application that uses crowdsensing to collect data about the physical environment. The data reliability challenge in CPS refers to designing a state estimator that takes raw unreliable crowdsensing data as input and outputs reliable estimates of the underlying physical states and appropriate error bounds on estimations. This poster presents interesting perspectives of addressing this emerging challenge and its connection to future smart city applications.

Hongning Wang • University Of Virginia

Online Collaborative Learning from Interaction with Humans

It is vital for an online service system to learn from its interactions with humans. In our work, we have developed bandit-based algorithms to enable a system to 1) actively learn from its interaction with a crowd of users; 2) identify latent factors that facilitate online learning. We rigorously prove nice theoretical properties of the proposed online learning algorithms, and verified their utilities with extensive experiments in various real-world applications.

Gloria Washington • Howard University

Affective Technologies for Improving the Lives of Persons With Chronic Disease

Behavioral and physical characteristics can be used as input to technologies that monitor stress, anxiety, and/or depression in humans. These technologies are commonly found in academic laboratories and not where we need them: mobile devices. This work seeks to improve the lives of sickle cell disease sufferers through motivational, encouraging technologies that also provide dietary tips. Results from this work will help to alleviate the emotional toll for managing and handling chronic disease.

Chuan Yue • Colorado School of Mines

SafeSky: A Secure Cloud Storage Middleware for End-user Applications

It is desirable and even essential for end-user applications to have the cloud storage capability to improve their functionality, usability, and accessibility. However, incorporating the cloud storage capability into applications must be done in a secure manner to ensure the confidentiality, integrity, and availability of users' data in the cloud. We introduce our SafeSky, a middleware system that can immediately enable an application to use the cloud storage services securely and efficiently.

Ting Zhu • University of Maryland, Baltimore County

Gait-Based Privacy Preserving Technique for Smart Cities

With the advent of the Internet of Things (IoT) and big data, high fidelity localization and tracking systems that employ cameras, RFIDs, and attached sensors intrude on personal privacy. However, the benefit of localization information is important for smart city applications. To address this challenge, we introduce Wobly, an attribute based signature (ABS) which measures gait. Wobly uses the physical layer channel and the unique human gait as a means of encoding a person's identity.



CCC

Computing Community Consortium
Catalyst

The mission of Computing Research Association's Computing Community Consortium (CCC) is to catalyze the computing research community and enable the pursuit of innovative, high-impact research. CCC conducts activities that strengthen the research community, articulate compelling research visions, and align those visions with pressing national and global challenges. CCC communicates the importance of those visions to policymakers, government and industry stakeholders, the public, and the research community itself.

This material is based upon work supported by the National Science Foundation under Grant No. (1136993).

cra.org/ccc



CRA

Computing Research
Association

