Goals for the Meetings

Goal 1

Network with global thinkers: identify common challenges and opportunities for Pervasive Personalized Intelligence (PPI) for IoT systems.

Goal 2

Influence the direction of the Center: push back and challenge our research and make it better.

Goal 3

Validate the feasibility of the Center: analyze data and best practices from the National Science Foundation (NSF) on Industry/University Collaborative Research Centers (IUCRCs); analyze how PPI Center capabilities meet industry needs.

Goal 4: **List here your goals for the meeting**

Welcome from the Directors

Dear Participant,

We are glad you are here. We are grateful that you are bringing your best thinking into this meeting. This is what truly makes us better. The scale and complexity required to build tomorrow's Pervasive Personalized Intelligence is beyond what any organization can successfully build on their own. We need to unite our forces and our minds. Join the movers and shakers who make things happen in industry, academia, and government funding agencies.

We were looking forward to welcoming you in person on the Engineering Campus at University of Colorado Boulder for the launching of the new industry-university center. However, the pandemic changed our plans and moved our event virtually. The pandemic is accelerating trends like remote participation and the digitalization of the world through Internet of Things. It's pushing us to connect everything to the internet for remote monitoring and remote service and better analytics.

What's an R&D project that your company knew was already hard, but it is even harder now because of the pandemic? This is exactly what our new Center for Intelligent IoT systems can do to help you advance your R&D during difficult times.

How can you turn adversity into opportunity? You are already doing the first step. Innovation starts at connection. We created a program for multiple days with generous time to connect 1-on-1 or in small groups with other participants and access the diverse thinking and expertise in our group.

We invite you to fully validate information and ideas. This is truly two-way learning. You can read our academic research and analyze information on your own, but it doesn't allow you to push back and challenge it to make it better. Engage so that you influence the direction on which we are going as a center. It is this feedback that makes us better so that we can serve you better.

We also encourage you to analyze the data and best practices from the NSF and interact with the NSF representatives (Behrooz Shirazi and Dee Hoffman). Ask them how the Industry/University Collaborative Research Centers (IUCRCs) across the nation are providing tremendous value for industry. In countless conversations with them, we found Behrooz and Dee go the extra mile to serve the people in front of them. You will find this too.

One is too small of a number for significance. Putting together this event is truly a team effort. There are many people who worked behind the scenes to make this event possible, some of them you will not even meet, though their contribution was essential.

First and foremost, we are grateful to Colin Mahoney from the Two Tigers Productions and Michelle Ambruz from Garnet Peak Studio for being our dedicated production team. Everything that you are going to experience virtually, you will see and hear their professional touch.

We are very grateful to the teams in our home campuses who worked tirelessly to connect us with industry: Chris Yankee, Julie Chiron, Josh Rhoten, Emily Adams, Kellen Short, Rachel

Robertson, Tina Batten, Deborah Kurnik, Gale Sumida, Robert Miller, Abby Benson, Chris Muldrow, Amy Hill, Sara Buhr– you propel us to go higher.

We are grateful for the people who organize the high-tech professional associations in our states, Skip Newberry, Rylee O'Brien, Rob Schulberg (from the Technology Association of Oregon) and Frannie Mathews from the Colorado Technology Associations.

We are grateful for the support from our university administrators: Terri Fiez, Keith Molenaar, Bobby Schnabel, Ken Anderson (at CU Boulder), and Irem Tumer, Scott Ashford, Julie Brandis, Tom Weller (at OSU).

Last but not least, we each thank our spouse and children who had to accommodate us putting in long hours working on the Center.

Danny Dig, Weng-Keen Wong, Evan Chang, Shiv Mishra, Eugene Zhang -- Center Directors





Agenda

September 14, 2020

All times MDT

3:50 pm to 4:00 pm	**Speed Networking: Meet and Greet**
4:00 pm to 4:08 pm	**Opening & Welcome from CU Administration**  Terri Fiez Vice-Chancellor for Research and Innovation, CU
4:08 pm to 4:15 pm	**Welcome from OSU Administration**  Irem Tumer VP for Research, OSU
4:15 pm to 4:30 pm	**Vision and Goals for the PPI Center**  Danny Dig Founder and Executive Director of the PPI Center
4:30 pm to 5:00 pm	**Overview of the NSF Industry-University Cooperative Research Program**  Behrooz Shirazi NSF I/UCRC Program Director
5:00 pm to 5:30 pm	**Executive Industry Panel on Intelligent IoT Systems** Skip Newberry (CEO Technology Association of Oregon), **Moderator** Jeroen Leverman — IT Technology Director at NextEra Energy Resources (Renewable Energy) Roberto Aiello — Managing Director, Itron Idea Labs (Utilities) Dave Archer — Principal Investigator at Galois, Inc (High-Tech) Cynthia Wallace — Engineering AI Strategic and Functional Lead at Ball Aerospace Hector Dominguez — Open Data Coordinator at the City of Portland (SmartCity) Allison Saltzer -- Business Development Director - IOT and Artificial Intelligence, Microsoft (High-Tech) Eron Bucciarelli - Senior Product Manager In-Store Experience at Kroger Co (Food & Retail)
5:30 pm to 5:40 pm	**Mystery Speed Networking / Help Desk – The Yes Desk**

September 17, 2020

3:50 pm to 4:00 pm	**Speed Networking: Meet and Greet**
4:00 pm to 4:05 pm	**Open Set Detection**  Weng-Keen Wong OSU Faculty
4:05 pm to 4:10 pm	**Visual Analytics for Scalable AI Debugging**  Minsuk Kahng OSU Faculty
4:10 pm to 4:15 pm	**Structured Explanation Graphs for Image Classification**  Prasad Tadepalli OSU Faculty
4:15 pm to 4:20 pm	**Safe Programmatic Reinforcement Learning**  Ashutosh Trivedi CU Faculty
4:20 pm to 4:25 pm	**Improved Accessibility and Mobility Using a Topological Approach**  Eugene Zhang OSU Faculty
4:25 pm to 4:30 pm	**Device-Based Exploratory Visual Analytics**  Danielle Szafir CU Faculty
4:30 pm to 4:45 pm	**Speakers Ask-Me-Anything (AMA)** AMA session with speakers
4:45 pm to 4:50 pm	**Networking Break**
4:50 pm to 5:10 pm	**Small Groups 1: Industry Applications** Discuss project proposals' match to industry needs and challenges
5:10 pm to 5:30 pm	**Small Groups 2: Industry Applications** Discuss project proposals' match to industry needs and challenges
5:30 pm to 5:40 pm	**Mystery Speed Networking**

September 28, 2020

3:50 pm to 4:00 pm	**Speed Networking: Meet and Greet**
4:00 pm to 4:05 pm	**AI for Wireless Systems**   Dirk Grunwald and Sangtae Ha CU Faculty
4:05 pm to 4:10 pm	**System Support to Enable AI at the Edge**  Shiv Mishra CU Faculty
4:10 pm to 4:15 pm	**Enabling Secure and Federated Learning**  Yeongjin Jang OSU Faculty
4:15 pm to 4:20 pm	**Privacy-Preserving Data Reconciliation**  Mike Rosulek OSU Faculty
4:20 pm to 4:25 pm	**Conformance Verification for Machine Learning Models**  Sriram Sankaranarayanan CU Faculty
4:25 pm to 4:30 pm	**Intelligent Assistants for Creating PPI Applications**  Bor-Yuh Evan Chang CU Faculty
4:30 pm to 4:35 pm	**Code Migration for ML-based Software**  Danny Dig CU Faculty
4:35 pm to 4:45 pm	**Speakers Ask-Me-Anything (AMA)** AMA sessions with speakers
4:45 pm to 4:50 pm	**Networking Break**
4:50 pm to 5:10 pm	**Small Groups 1: Industry Applications** Discuss project proposals' match to industry needs and challenges
5:10 pm to 5:30 pm	**Small Groups 2: Industry Applications** Discuss project proposals' match to industry needs and challenges
5:30 pm to 5:40 pm	**Mystery Speed Networking**

Project Abstracts

September 17, 2020

Open Set Detection

Weng-Keen Wong (Center Faculty, OSU)

Classifiers deployed in the real-world, such as image recognition systems on autonomous vehicles, face the challenge of encountering data from classes that were not seen in the training set. For instance, an image recognition system can be trained to recognize stop signs and speed limit signs from a training set, but it can also encounter yield signs in the real world. Traditional machine learning systems are often unpredictable in how they handle these unseen classes, leading to concerns over their real-world behavior.

Open set detection is an area of machine learning which attempts to recognize unseen classes of data during deployment while still accurately classifying classes seen during training. The majority of open set detection methods use a softmax layer followed by a cross entropy loss function. This approach has known problems and we propose to investigate these issues more deeply. In addition, we propose to develop an improved approach based on density ratio estimation.

Visual Analytics for Scalable AI Debugging

Minsuk Kahng (Center Faculty, OSU)

Even state-of-the-art AI models fail to make correct predictions for many data instances, which can be a big issue especially for safety-critical applications (e.g., medical diagnosis, self-driving cars). How can we help people effectively analyze ***when* AI fails** and infer ***why* it fails**, so that they get actionable insights for improving the model accuracy? We propose a new *human-in-the-loop* approach to this challenging problem, by creating a **novel visual analytics tool** for *debugging* AI models with a focus on *large* datasets used for training them. We will build upon our successful experience of developing visual analytics systems for industry-scale deep learning models (e.g., ActiVis deployed at Facebook). Specifically, the proposed tool will visually summarize a large number of failed cases using scalable data visualization techniques and provide a list of potential reasons for these mispredictions caused by common dataset-related issues such as insufficient data, distribution shift, or labeling errors. With this information, users can get actionable insights for improving the model accuracy (e.g., adding more outdoor images to training set if such images are insufficient), which reduces their time for debugging and fosters their trust in AI.

Structured Explanation Graphs for Image Classification

Prasad Tadepalli (Center Faculty, OSU)

Explanations for image classification with deep neural networks typically rest on different kinds of `activation maps' that show which parts of the image are responsible for the classification. However, due to the non-linearity of the neural network functions, this kind of maps do not explain the behavior of the network under different occlusion conditions. Indeed, different parts of the image contribute to the classification at varying levels depending on which parts of the image are visible. Structured Explanation Graphs (SAGs) explicate this behavior of the network under different occlusion conditions in the form of a graph over activation maps. Our extensive user studies reveal that SAGs with an interactive user interface gives significantly more insight to the users in predicting the behavior of the network under unseen conditions. In the current proposal, we seek to extend this work in three directions: 1) build a fully interactive interface for SAGs where the users can query the network's behavior under any occlusion conditions, 2) develop and evaluate algorithms for generating complete SAGs for any desired confidence level and 3) extend SAGs to work with different images of the same class and generate class-level explanations.

Based on joint work with Vivswan Shitole [grad student], Fuxin Li, Alan Fern, and Minsuk Kahng

Safe Reinforcement Learning

Ashutosh Trivedi (Center Faculty, CU)

Reinforcement learning (RL)-based programming is a rapidly emerging programming paradigm that seamlessly integrates human creativity with AI-guided explorations. The research objective of this project is to improve the scalability, usability, and trustworthiness of programmatic RL by integrating RL with automatic program abstractions guided by learning requirements. Software systems pose scalability challenges for RL due to, among other things, numerical data types, and potentially unbounded procedural call stacks. Current RL approaches overcome this challenge by relying on programmers' expertise and insights in providing suitable approximation schemes. As the complexity of software grows, manual approximations get prohibitively expensive and lead to low productivity and low assurance. Moreover, the resulting lack of automation results in a significant barrier to entry in this promising field. The PI aims to democratize the RL-based programming paradigm by tying automatic software abstractions with programmatic RL to speedup learning and to provide better assurance. The long-term vision of this project is to democratize programmatic RL by developing principled methodologies and powerful tools to improve the scalability, usability, and trustworthiness of programmatic RL.

Improved Accessibility and Mobility Using a Topological Approach

Eugene Zhang (Center Faculty, OSU)

Accessibility and Mobility is one of the most fundamental needs in our lives. It is dependent on the available routes between the source and the destination as well as the current traffic condition of these routes at the time of travel.

In this research, we plan to apply advanced topological analysis, which is primarily concerned with the connectivity of spaces and has found many applications in many areas in science, engineering, and medicine, to improve accessibility and mobility in urban areas. Working closely with our industrial partners, we plan to develop tools that allow efficient topological analysis and visualization of their data. In addition, the users can use our tools, for example, to modify existing road networks to simulate potential traffic disruption due to natural disasters, add or remove entities such as hospitals and gas stations to simulate their effects on the accessibility of the community, and change the geographical and demographical information based on projected future growth. While these functionalities are available in many software, applying advanced topological analysis and visualization to them is at the core of this industry-academic collaboration.

We plan to leverage on our past research that links the topological singularities such as wedges (dead ends) and trisectors (junction points) in the network to path planning and road network design.

Device-Based Exploratory Visual Analytics

Danielle Szafir (Center Faculty, CU)

As technology becomes increasingly pervasive, we create new opportunities to leverage data to inform action anywhere. However, traditional visual analytics tools are bound to the mouse and monitor: data exists in diverse and dynamic environments, but our means for understanding it are bound to static desktop metaphors. To enable people to make decisions with and act on data in dynamic environments, the proposed work will develop a toolkit for creating responsive visualizations that accommodate the range of devices seen in IoT applications. We will generate a suite of empirically-driven visualization designs optimized to communicate key patterns in data across displays ranging from mobile phones to embedded displays on public hubs to heads-up displays used in fieldwork. Our toolkit will automatically convert relevant queried data across display types, allowing people to interact with data using either conventional methods (e.g., touchscreens) or natural language queries. We will augment the toolkit with adaptive collaborative machine learning algorithms that accelerate data processing and discovery to allow for efficient and accurate insight into data at the times and places where data can best inform human action.

September 28, 2020

AI for Wireless Systems


Dirk Grunwald (Center Faculty, CU), and Sangtae Ha (Center Faculty, CU)

Traditionally, mobile networks' control and management has been predominantly done in the "cellular core" instead of at the edge or the client, despite the fact that network quality is highly variable and dependent on local conditions. We approach this research thrust by exploiting the client inference on a global network state using theoretical tools and machine learning. We show that the practical systems built from this client-side inference outperform the current state-of-the-art systems while requiring no coordination or change in the infrastructure.

Privacy-Preserving Data Reconciliation

Mike Rosulek (Center Faculty, OSU)

Certain analytics often require data from different organizations. When that data is privacy-sensitive, it cannot be collected in a centralized location to perform these analytics. In this presentation, I will discuss state-of-the-art cryptographic tools that allow analytics to be run on data in a privacy-preserving and decentralized way.

The first step in distributed analytics is to reconcile the data sets to a common "language" -- for example, two organizations will use different identifiers to refer to the same customer. I will propose new methods to obtain common identifiers in a privacy-preserving way, so that the two organizations obtain the same identifier for matching records, without learning which records are actually in common. I will give examples of how such common identifiers can be leveraged to efficiently perform nontrivial analytics on the corresponding data.

System Support to Enable AI at the Edge

Shiv Mishra (Center Faculty, CU)

Edge computing introduces middle-tier compute servers much closer to the sensors and end users to build context aware IoT and mobile applications, thereby minimizing latency and preserving privacy. Applications include recognizing and monitoring people in work and public spaces, smart management of city operations including water and energy distribution, smart agriculture, managing traffic and transportation systems, law enforcement, and augmented reality applications. These applications require distributed execution of complex, compute-intensive algorithms such as deep learning, gesture recognition, acoustic detection and classification, and anomaly detection. The main goal of our research project is to develop core system-level services to enable a distributed, microservice-based edge system architecture that facilitates building such complex applications. The key features of this system include incorporating humans in the loop, optimized placement of compute and data elements in a dynamically changing environment, and

computing over a diverse set of processing elements including traditional CPUs and ASICs such as GPUs and FPGAs. The proposed system aims to integrate and augment elements of current edge solutions such as EdgeX, Kura, OpenEdge and Azure IoT Edge.

This research is partly funded by multiple grants from the National Science Foundation.

Conformance Verification for Machine Learning Models
Sriram Sankaranarayanan (Center Faculty, CU)

Machine learning is making a huge impact in areas such as autonomous systems and medicine, wherein data driven models are used to drive decisions such as how to steer a UAV to avoid an obstacle or how much insulin to provide a patient with type-1 diabetes to control their blood glucose levels. However, despite learning highly accurate models on high quality data, these models often hide dangerous errors that we term "conformance" violations. We will briefly demonstrate such violations in neural network models and discuss how verification tools can be used to expose these errors. We will then discuss some future research directions in this area.

Intelligent Assistants for Creating PPI Applications
Bor-Yuh Evan Chang (Center Faculty, CU)

The demands on software developers to create rich PPI applications that are safe, secure, and privacy-preserving will be immense. PPI applications will be built on rich software frameworks that abstract the commonalities amongst such applications. While rich software frameworks enable software engineers to build complex applications on sophisticated platforms, developing against them safely and securely relies on following often complex and incompletely documented protocols. To make PPI application-development feasible, we consider techniques that itself leverage algorithmic and data-driven techniques to assist PPI application developers in finding, fixing, and understanding defects.

Code Migration for ML-based Software
Danny Dig (Center Faculty, CU)

It is widely known that at least two-thirds of software costs are due to evolution, with some industrial surveys claiming 90%. ML software and models need to evolve to respond to internal & external changes. An important example of evolution is to migrate code to use the newer version of ML libraries or to optimize performance. Due to many non-backwards-compatible API changes, this often requires engineers to rewrite their code & models from scratch. Our goal is to mechanize such tasks that are expensive, time-consuming, and error-prone. In this proposal we first automatically mine evolutionary code changes from a wide repository of open-source ML codebases. Grounded on these formative studies, we will design, implement, and evaluate refactoring tools to help engineers evolve their codebases safely and effectively.

October 5 & 6, 2020

Faculty Response and Student Posters

October 14, 2020

Industry Project Voting, Response from the PPI Leadership, and Next Steps

Guidelines for Effectively Participating in Level of Interest and Feedback Evaluation (LIFE) Process

Introduction: *The LIFE feedback process is not a project selection methodology but is meant to inform whatever project selection approach your center uses. There are a number of purposes served by asking industry representatives and PIs to complete LIFE feedback on project proposals: 1) Q&A time is usually limited and having member organizations provide written feedback allows everyone the chance to have input; 2) Written feedback gives PIs a chance to consider industry concerns and provide a thoughtful reply; 3) Feedback and replies can be debriefed as a group and help surface areas of agreement and disagreement and reach a consensus on the need for and feasibility of project changes.* 4) *Reviewing the interest rating distribution allows members to understand whether a few or many members are interested in a project and use this information to decide how to vote during project selection.*

Steps for Completing Feedback and Responses:

1. Website: www.iucrc.com
2. Select Center Meeting : PPI (Sept 17)
3. Enter PW (case sensitive, include spaces) = PPI Begins!
4. Select role: **Industry** for industrial participants or **Faculty** for faculty project leads (and students)
5. IAB
    a. Click [Evaluate Project] and select a level of interest rating based on your firm's needs and interests.
    b. Provide comments, questions, and/or suggestions you have about the project. **The most valuable feedback is"actionable" comments like suggestions and questions that help the PI / student improve the project**. If you rated the project "Needs change" make sure to add a comment or suggestions on what needs to be changed or what needs to be done to get the project on course.
    c. Enter your Name and your Organization.
    d. SELECT SUBMIT AFTER EACH PROJECT.
    e. Repeat for each project.
6. PI / Student
    a. Read instructions and click [Continue to Response Page].
    b. Find your project and select [Response to Comments]
    c. Read comments provided by industry members and respond as necessary (not every comment may require a response).
    d. SELECT SUBMIT AFTER RESPONDING TO EACH page of comments. (there are usually more than one page of comments so then continue to the next page).
    e. Once you have completed responding to comments and submitted, scroll up to the blue links under the ratings to respond to the questions, and then the suggestions.
7. Both
    a. You can review the feedback and responses to each project by selecting [Summary] next to each project.
    b. If you would like to review responses to all projects presented at the meeting, you may use the [Review Meeting] link at the top of the project list page (PDF and Word versions are also available).