Scientific Edge Computing: Linking Instruments to Supercomputers

Pete Beckman: Co-Director Northwestern Univ / Argonne Inst. for Science and Engineering
Collaborators: Ilkay Altintas, Charlie Catlett, Scott Collis, Nicola Ferrier, Kate Keahey, Eugene Kelly, Jim Olds, Mike Papka, Dan Reed, Raj Sankaran, Sean Shahkarami, Joe Swantek, Valerie Taylor, Doug Toomey, Frank Vernon, Rommel Zulueta, and many more….
Instrument  Data  Analysis
The Digital Continuum

IoT Facilities

Instrument

HPC/Cloud

Analysis

Analyse full resolution data, find highest value data for the science.
AI@Edge: Digital Continuum

Sensors
- LIDAR
- Software Defined Radios
- Hyperspectral Imaging

Facilities

Actuators
- Servos
- Dynamic adaptation

Edge Computing
- Scientific Data Analysis & Control
- Artificial Intelligence
  - Deep Learning Inference
  - Lightweight Training
  - Autonomous Action

Advanced Networking
- New inference (model)
- Adaptive controls / steering

Computation
- Cloud
- Data Center
- HPC
- Predictive Sim
- Digital Twins
- Data Analysis
- Machine Learning
Why Live on the Edge?

- More data than bandwidth
  - Imaging, LIDAR, SW defined radios, radar, audio, hyperspectral, large facilities, …
- Latency is important
  - Quick local decision, experimental control & actuation; adaptive sensing
- Privacy/Security requires short-lived data: process and discard
  - Compromised devices have no sensitive data to be revealed
- Resilience requires distributed processing, analysis, and control
  - Predictable service degradation, autonomy requires local (resilient) decision-making
- Quiet observation and energy efficiency
  - Vigilant low-power sensors, transmit only essential observations
Leadership Team

- **Pete Beckman** (NU: Director)
- **Nicola Ferrier** (UC: Deputy Dir.)
- **Scott Collis** (NU: Instruments, Atmos)
- **Valerie Taylor** (UC: Edu, Broader Impacts)
- **Eugene Kelly** (CSU: Ecosystems, NExT)
- **Ilkay Altintas** (SDSC: Data)
- **Scott Collis** (NU: Instruments, Atmos)
- **Charlie Catlett** (Illinois: Urban)
- **Jim Olds** (GMU: Life Sci, Risk)
- **Dan Reed** (Utah: Architecture)
- **Kathy Bailey** (Proj Mgmt)
- **Mike Papka** (NIU: Edu, Broader Impacts)
- **Raj Sankaran** (NU: Node Arch)

Cyberinfrastructure for AI at the Edge
sagecontinuum.org

**SAGE**

**Al@Edge Summer 2022**
(Student Outing: June 2022)

MSRI-1: 1935984
Start: October 1, 2019
Exciting, Hard, Challenging, CS Problems: From Instrument to the HPC/Cloud

- Reusable cyberinfrastructure for AI@Edge is new territory
- Digital Continuum programming models largely unexplored
- How can we build triggered simulations and autonomous reactions?
- Edge computing needs multi-tenancy for computation and actuation
- Remote, distributed instruments have unique cyberinfrastructure needs
- New AI methods and algorithms for the Edge
- New resource management for Science Goals
- Experimental cyberinfrastructure must first “do no harm” to operational facilities
Building on concepts from NSF Array of Things

Environmental and Air Quality Sensors

Edge Computing

PI: Charlie Catlett, Uchicago ~2016-2018
What is a “Software Defined Sensor”?

![Software Defined Sensor diagram]

Your software container running here

Analysis produces live results

AI-Based Measurement & Anomaly Detection

- Plant Species
- Pedestrian Flow
- Traffic Flow
- Wildlife
- Cloud Motion Vectors
- Wildfires: detecting smoke
- Flooding / surface water

AI-Based Measurement & Anomaly Detection

- Wildfires: detecting smoke
- Birdsong
- Flooding / surface water
Sage Cyberinfrastructure: Exciting Goals!

- Build new reusable cyberinfrastructure
  - High-quality, resilient, well-documented software
  - Leverage best Open Source frameworks
    - PyTorch, OpenCV, TensorFlow, Kubernetes, Docker, etc.
- Build community of AI@Edge scientists
  - New AI-based measurements
  - New AI algorithms for edge
- Deploy experimental testbed into production facilities
- Provide new capabilities for live data and triggered responses
- Teach and train students, explore new ideas

(Sensors sample at 40hz, aggregate to 30min)

Analyse full resolution data, find highest value data for the science
Key Sage Partners & Collaborators

AoT: Neighborhood scale urban environment and activity.

HPWREN/WIFIRE: Regional Environmental Conditions and Events.

NEON: Continental scale ecology and environment.

Oregon Hazards Lab: AlertWildfire, ShakeAlert, and flooding hazards.
Delivering AI@Edge Platforms: Two Forms

**Wild Sage Node**

- Optical Rain Sensor
- Sensor Stevenson Shield
- Relative humidity
- Barometric pressure
- Ambient temperature
- Microphone
- 110/230V AC Power
- Ethernet
- Sky facing camera
- Additional Sensors
- Mounting Point
- NVIDIA Xavier NX GPU
- Wireless Network
- 4G, 5G, WiFi
- Sensor expansion
- ports: PoE & USB
- Ground Facing Camera

Ready for mounting outside, any PoE sensor can be easily added.

**Sage Blade**

Rugged server for instrument huts, new sensors easily added.

Research Credit: Edge Architecture led by Rajesh Sankaran, Northwestern University

Leverages Open Source, Open Hardware, extensible platform called *Waggle* developed at Argonne National Laboratory.
Sage Software Architecture

User “Plugins” run in “Waggle Edge Stack” (WES)
- Built on best Open Source AI packages
- Access to sensor and camera streams
- Libraries for efficient GPU usage
- Extreme cybersecurity
- Publish data to Beehive

Beehive manages
- Sage Edge Scheduler (SES)
- Sage Data Repository (log entries)
- Sage Data Repository (binary files)
- User Interface components

Cloud Infrastructure

AI@Edge “Plugin” from Edge Code Repository (ECR) (the “App Store”)

Research Credit: Software Team led by Joe Swantek, Northwestern University
Building Community for AI@Edge Sage Applications

The Edge Code Repository

AI Portfolio led by Sage Deputy Director: Nicola Ferrier, UChicago
Deploying Wild Sage Nodes in Chicago

Sean Shahkarami, Uchicago, overseeing installation (in brown jacket, by corner of building)
Undergraduate Research: Pedestrian Detection and Paths

NIU experimental node with wired network connection

- Experiment with sampling rate and resolution
- Work is now being ported to Sage node

YOLO based model for identifying people and to check for use of crosswalk

Pedestrian data processed to understand patterns and transformed for top-down view then bundled to highlight patterns

Justin Derus, Wesley Kwiecinski, Pratool Bharti, Michael Papka: Northern Illinois University
Avian diversity monitoring

- Bird diversity changes as a metric to track the current environmental conditions
- We automate Avian Diversity Monitoring by using a DNN, called BirdNET [1], capable of identifying 984 North American and European bird species by sound. Weekly cumulative detections of non-migratory species occurrence was highly correlated with human point count observations
- It will be possible to get exposure to many organisms occupying diverse areas without needing to detect them during demanding and expensive human fieldwork

Paintbrush Prairie Bird Detection

Total calls, for top 12 species, as a function of the hour of the day (UTC-06).

Top 25 bird calls recorded at the Paintbrush Prairie Natural Preserve (Nature Conservancy Site) from Sep 2020 to Dec 2021

Research Credit: Dario Dematties, Bhupendra Raut, Nicola Ferrier
Cloud Motion Vectors with Hemispheric Sky Camera

- Real-time cloud motion vectors from sky images have applications in meteorological analysis, nowcasting, and short-term prediction of solar irradiance.

- Camera contamination by rain and snow is identified by the ML algorithm and reported.

- The use of AI/ML with the other sensors produces valuable products (e.g., Solar irradiance, nowcasting, locale weather reporting).

Future: Instrument steering and Nowcasting
Cloud Motion Vectors with Hemispheric Sky Camera

- Real-time cloud motion vectors from sky images have applications in meteorological analysis, nowcasting, and short-term prediction of solar irradiance.
- Camera contamination by rain and snow is identified by the ML algorithm and reported.
- The use of AI/ML with the other sensors produces valuable products (e.g., solar irradiance, nowcasting, locale weather reporting).

Scott Collis, SAGE Co-PI, Northwestern University

Surprise Training Data
Measuring Water and Snow Depth

We are evaluating multiple approaches to estimate the water (or snow) level from images of rulers (in of a stream at a NEON site)
- Computer vision (CV) based
- Machine Learning algorithms
  - U-Net, ResNet
  - Self-supervised Learning

Some large errors with CV approach

Seongha Park, Northwestern University
Measuring Water and Snow Depth

Self-supervised Segmentation
• Exploring visual transformer ML
• ML model was trained using only images from IMAGENET (no labels and no NEON data)
• An Intersection over Union score > 0.5 is normally considered a “good” prediction.

Intersection/Union (IoU) = 0.729
Surface Water Detection

Linked with HPC, can be used to build hydrology models and predictive capabilities

Nicola Ferrier, UChicago
Sage and POWDER: Next Generation Wireless

Powder (the Platform for Open Wireless Data-driven Experimental Research) is flexible infrastructure enabling a wide range of software-defined experiments on the future of wireless networks.

Powder supports software-programmable experimentation on 5G and beyond, massive MIMO, ORAN, spectrum sharing and CBRS, RF monitoring, and anything else that can be supported on software-defined radios.
Wildfire Detection and Prediction

Exploring wildfire detection at the edge linked to HPC simulations

ALERTWildfire: A unique wildfire detection and monitoring system

Collaboration: Doug Toomey, UOregon

Frank Vernon, UCSD, HPWREN
Mt Wilson Fire
Exploring New Methods for Wildfire Smoke Detection

Two approaches to improve the predictions are
1. Use of thermal IR camera, and
2. Incorporating motion of the smoke in the DL models.

- Prescribed burns and real wildfire data is needed to train the AI models.
- Cloud temperatures can be used to estimate cloud-base heights and cloudiness.

Bhupendra Raut, Northwestern University
NEON Mobile Deployment Platform (MPD) with Sage Konza Prairie for controlled burn: April 2022.

Special Thanks: NEON Team! Rommel Zulueta @ Battelle

Sage Co-PI: Eugene Kelly, Colorado State
eugene.kelly@colostate.edu

Optical Rain Sensor
Sensor Stevenson Shield
Relative humidity barometric pressure ambient temperature
Microphone
110/230V AC Power Ethernet
Sky facing camera
Additional Sensors Mounting Point
NVIDIA Xavier NX GPU
Wireless Network 4G, 5G, WiFi
Sensor expansion ports: PoE & USB
Ground Facing Camera
Sage NEON NSF Controlled Burn
Konza Prairie, Kansas
Data from the experiment already available to the community!

Ismael Perez and Ilkay Altintas, UCSD
Collaboration with CSIRO in Australia

City Environment Sensor Network

Key Science objectives

- Model-data fusion to increase spatial and temporal resolution of modeled local weather (and climate)
  - integration with physically based models like CCAM and Spark
  - for applications with machine learning techniques
- Calibration and validation to ensure accuracy of data
- Application of sensor data in disaster management and for city resilience
- Provide data to inform policy and strategy decisions on environmental and urban growth

Gold Coast has rapid population growth and population densification, an existing city owned network of IoT sensors (LoRaWAN) and is an existing member of Open and Agile Smart Cities group (OASC).

Sensor network will be deployed in the SA3 zones in the City of Goldcoast

Collaboration with Mahesh Prakash and Nikhil Garg, CSIRO
Partnership with Ojibwe Nation
Study Climate Change Impact on Manoomin (Wild Rice)

- Sage node at Bad River Fish Hatchery
- Deployment approved by Tribal Council

Collaboration with Kim Marion Suiseeyaa, Northwestern University
Partnership with Native Hawai‘ian Community
Study Climate Change Impact

Collaboration with Josiah Hester, Northwestern University
Sage Cyberinfrastructure: Key Point

Your science, your sensors. Sage is the cyberinfrastructure

Sage can interface with any instrument or sensor*

* We don’t do Windows
Data can be downloaded live via API interface, downloaded as CSV TGZ Bundles, or browsed.
Sage Resources

Getting started with Sage! - https://docs.sagecontinuum.org
Sage AI@Edge Apps - https://portal.sagecontinuum.org/apps/explore
Sage Data - https://portal.sagecontinuum.org/data
Sage Konza MDP Campaign - https://mdp.sagecontinuum.org
Overall Sage system status - https://admin.sagecontinuum.org/status

Portal showing the current set of science jobs executing on the various nodes will be available to the public soon…
Questions?

Leadership Team

Professors Aaron Packman and William Miller, Northwestern University
Gensburg-Markham Prairie, The Nature Conservancy
Photo Credits: Liliana Hernandez-Gonzalez, Northwestern University
Dec 2015
Special Thanks

Students!

Ilkay Altintas
Kathy Bailey
Daniel Balouek-Thomert
Pete Beckman
John Blair
Eric Bruning
Adam Brust
Charlie Catlett
Scott Collis
Neal Conrad

Geoff Davis
Dario Dematties
Nicola Ferrier
Jannick Fischer
Larry Hartman
Robert Jackson
Euguene Kelly
Yongho Kim
Nick Maggio
Seth Magle

Bill Miller
Patrick O'Neal
Jim Olds
Aaron Packman
Mike Papka
Seongha Park
Ismael Perez
Bhupendra Raut
Dan Reed
Mike SanClements

Raj Sankaran
Sean Shahkarami
Sergey Shemyakin
Joe Swantek
Helen Taaffe
Valerie Taylor
Doug Toomey
Frank Vernon
Rommel Zulueta