SAGE “AI@Edge” Science Workshop

April 29-30, 2019
Argonne, Illinois

Some 30 scientists from organizations convened at Argonne National Laboratory to (a) review the state technology and scientific services for a set of urban-, regional-, and continental-scale instruments, (b) explore new science questions to be enabled by integrating edge computation into those instruments, and by integrating the instruments with common cyberinfrastructure, and (c) translate new, convergent scientific research opportunities into specific infrastructure technology, services, and operations plans.

The workshop brought together five specific instruments, each with unique measurement (and associated data) features and spatial, scale, and temporal measurement attributes:

- Array of Things (AoT) - an urban sciences instrument with over 100 (soon over 200) AI@edge-capable “nodes” in Chicago and smaller deployments in a growing number and diversity of cities in the U.S. and abroad.
- Atmospheric Radiation Measurement (ARM) Facility - measuring atmospheric phenomena through a combination of fixed and deployable assets stationed to support regional and national-scale research.
- National Ecological Observatory Network (NEON) - a new instrument providing environmental, biological, and ecological measurements through over 80 major ground stations and over 100 specialized stations, covering 17 major ecological regions across the U.S., including Alaska, Hawaii, and Puerto Rico.
- High Performance Wireless Research and Education Network (HPWREN) - a regional environmental network, measuring conditions and impacts related to natural disruptors, from wildfires to earthquakes, providing both longitudinal measurements and real-time imagery and sensor data for emergency responders.
- Urban Wildlife Information Network (UWIN) - an urban wildlife detection and monitoring instrument providing data for analyzing the ecology and behavior of urban animal, insect, bird, and plant species. UWIN includes partners in over 20 cities across North America.

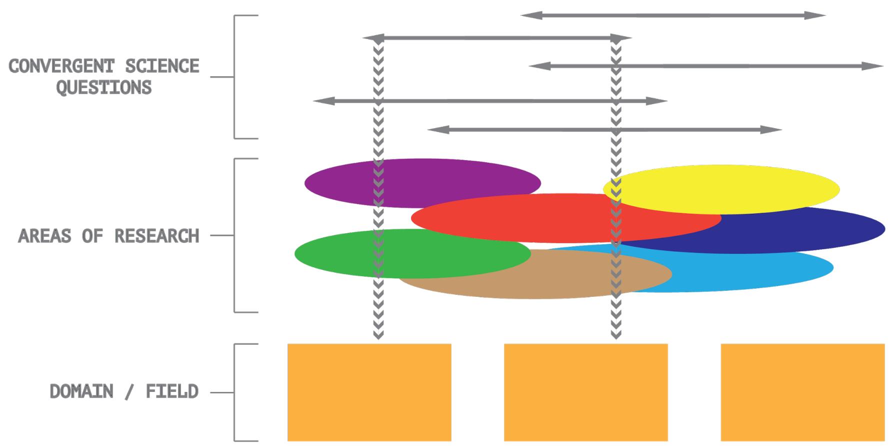
Following one day of plenaries, four working groups fleshed out initial plans during the day two, covering (a) convergent science questions and measurement requirements, concentrating in particular on new measurements enabled by edge computation, (b) cyberinfrastructure requirements for creating a research infrastructure to both integrate edge capabilities with existing instruments and to provide the research community with mechanisms for harnessing “AI at the edge” and the combined reach and features of multiple instruments, (c) opportunities to combine broadening participation, training, education, and dissemination programs across multiple instruments, catalyzing both the adoption and expansion of best-of-breed, and the development of new, scalable programs.

Table of Contents

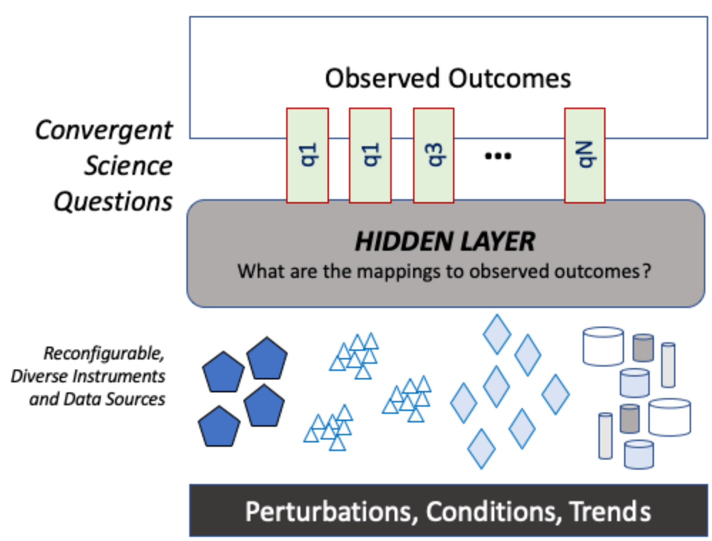
C. Catlett (cec@uchicago.edu) and P. Beckman (beckman@anl.gov)

Vision: Convergence Science

The architecture and functionality of the proposed research infrastructure has emerged from convergent science questions. That is, consequential societal questions that cannot be fully explored without scientific insight from multiple disciplines. The SAGE science agenda is focused in particular on convergent questions that require some combination of (a) measurements spanning existing instruments, (b) new, AI@Edge-enabled measurements, and (c) sufficient deployment scale to support repeatable, and comparable, measurements and experiments in diverse locations, for instance a transportation policy evaluation in multiple cities or the response to extreme weather in multiple rural, urban, and regional settings.



These broad societal questions drive specific research inquiries, which themselves span multiple disciplines (Figure 1). Mapping the convergent science questions to a set of instruments integrated as research infrastructure can be illustrated as a strategy to understand a convergence science question, such as, "why do some individuals, neighborhoods, communities, or cities recover more quickly from catastrophic events?." Many of the salient factors influencing this resilience are multi-dimensional, complex, and coupled, and they may also be difficult to measure directly. Consequently, an infrastructure to support such questions must be reconfigurable, capable of measuring many related factors to provide context, and must include mechanisms to combine measurements and data from both instruments and external sources, such as education, economic, or infrastructure performance data (figure 2), mobility data, or data from instruments outside of the SAGE infrastructure such as indoor air quality networks or social science surveys.



C. Catlett (cec@uchicago.edu) and P. Beckman (beckman@anl.gov)

SAGE will catalyze new disciplinary investigations through the combined measurement power of the five instruments, comprising unique but overlapping focus on ecology and environment, and biology (NEON at continental scales; UWIN at urban regional scales), climate and weather (ARM at regional scales; AoT at urban and neighborhood scales), mobility and human interactions (AoT), extreme events and associated environmental conditions (HPWREN), and the interplay between wildlife, ecology, and human systems (UWIN, AoT). For these investigations, SAGE will provide an expanded suite of measurements through the federation of instruments. Moreover, SAGE will enable one instrument to detect conditions or events, alerting another instrument in order to trigger alternative measurement settings.

SAGE will also bring new, edge computing-enabled measurement capabilities to each partner instrument. ARM will add a LIDAR instrument to its deployable stations in Houston, requiring edge analytics due to the high-bandwidth nature of these devices, which make it infeasible to stream data to cloud servers. Moreover, AoT nodes in Houston will provide localized data to develop and demonstrate the utility of multimessenger functions where AoT data triggers the ARM LIDAR focus on particular areas of the city for observation of hyper-local weather events, or incoming storm fronts. Similarly, HPWREN will be able to actuate pan/zoom cameras automatically based on event triggers such as particulate matter from wildfires or seismic activity.

As exciting as are these new capabilities, which multiply the Nation's investment in research instruments, convergent science requires an even more ambitious scientific and technical integration. SAGE architecture and specific design choices will be guided by an open community of scientists focusing on the resilience of individuals, communities, cities, and regions to perturbations such as extreme weather as well as to longitudinal change such as the impact of urbanization on ecological diversity, or the impact of agriculture and the food chain on antibiotic resistance.

Ultimately, the convergence science driving SAGE is to understand why a given policy, investment, or intervention produces resilience in one setting--whether a person, family, community, city, or region--but not in others. Today scientists observe outcomes and can measure underlying conditions, events, or trends, but the interdependencies of these coupled, complex systems at multiple scales represent a hidden layer of mechanisms and "levers" that cannot be plumbed with a single instrument or within a single discipline.

Workshop Scope and Goals

Following plenary discussions and presentations, four working groups met to flesh out plans for an NSF Mid-Scale Research Infrastructure Proposal. Each group met in several sessions, then group leaders presented findings and facilitated plenary discussions during the closing session of the workshop.

Working Group Leaders, Charters, Highlights

Leaders and Charters

A. Convergent Science \ (J. Olds, GMU; D. Work, Vanderbilt; M. Berman, UChicago)
Charter: Outline multiple convergent science questions and their testable hypotheses.

B. Cyberinfrastructure (Pete Beckman, NU; Nicola Ferrier, UChicago)
Charter: Outline hardware and software development required to build and operate SAGE.

C. Broader Impacts (V. Taylor, UChicago; S. Williams, UChicago; N. Papka, NIU)
Charter: Outline broadening participation, training, education, and data dissemination to diverse scientific and other communities to provide a persistent national data resource.

D. Management, Construction, Operation (C. Catlett, UChicago; G. Ricart, US-IGNITE)
Charter: Outline project work plans, schedules, risks, infrastructure expansion and replication, operational requirements, and sustainable operations and support plan.

Highlights

Convergent Science

Convergent science questions are supported by the SAGE network instruments that are reconfigurable and event driven, incorporate many factors that influence resilience through poorly understood mechanisms, and are motivated by the goal of understanding national resilience to environmental, societal, economic, and health disruptions. The SAGE network presents an opportunity to crack open the black box of resilience and its emergence from neighborhood to national scales and extract rule sets that guide the overarching resilience theory. Each of the three science questions outlined in the following table drive at the hypothesis that the SAGE network will enable us to delineate why regions of the country exhibit resilience and others do not.

Science question	State of the Art	Hypothesis	SAGE's role
How do severe weather events affect community well-being?	Relevant data are sampled at different scales Human impacts not characterizable with ARM	Weather's influence on community well-being is driven by key characteristics of the weather event, the built environment, and human behavior.	• ARM, HPWREN, Waggle combo • More frequent snapshots of community and conditions pre-, post-, and during event • New ARM sensors provide high fidelity weather and air pollutant data
How can we assess cognitive load of a city without a survey instrument?	Survey instruments used, characterization of human behavior indoors, where most time is spent, is limited	A city's cognitive load can be evaluated with no survey instrument but with data on human characteristics such as walking speed, posture and other observables identified with SAGE.	• Provision of high frequency images, weather and ambient condition data
How does human activity affect the microbiome?	NEON gives DNA timeseries at specified intervals but no one has used these data to explain metagenomics. DNA is manually sampled	High fidelity, high frequency sampling can provide links between anthropogenic soil inputs, changes in the microbiome, and corresponding environmental effects.	• NEON input can trigger SAGE that higher fidelity sampling is needed

Cyberinfrastructure

Establishing the cyberinfrastructure for SAGE requires examining the following questions.

- What is a viable route towards integrating mobile units?
- Edge computations must be portable. How can they be developed?
- How might devices be controlled or actuated?
- How would experiments be created?

C. Catlett (cec@uchicago.edu) and P. Beckman (beckman@anl.gov)

- What would be the process for a scientist to add their own sensor to a Waggle?

There are several routes for usage of the infrastructure SAGE would establish. For example, users could access training data from an instrument and develop edge computations that could be run periodically. An example is to detect nocturnal bat calls. These training data would be published into an open component library. To do this, a virtualized development and test bed is needed. Furthermore, a slice of the network could be saved for dedicated access to allow a group of nodes to be used for experiments that develop edge code and benchmarks. In addition, the sensor data would be open and available for use and if-then-then-that statements could be used to trigger sensor activation on data values. It would also be possible to schedule an edge computation from the library for a set of nodes, using a scheduler and resource manager for the edge. New training data sets could be published and AWS or other cloud computation could be linked to the data stream to trigger actions, start predictive steering, or commence edge analysis. Instruments could be the subject of workshops and training programs.

Finally, it is important to consider computer science convergent questions such as the following:

- How does lightweight training improve a model for a specific deployment, and what is the path for that?
- How can transfer learning be applied to edge, software-defined sensors?
- For unsupervised learning, how can additional (fused) data sets provide the accuracy/feedback/boost?
- Resource management / OS / scheduling
- Trust / security / privacy
- How to engage students in help labeling
- ML architecture evaluation (Google TPU, etc)

Broader Impacts and Broadening Participation

Universities and National Laboratories are key potential SAGE users. Their engagement could be achieved through providing repositories at different institutions, assigning data sets a DOI, and developing model toolkits to standardize data and training across institutions. Data literacy teach-ins could be useful tools in addition to curating data sets for teaching. These could be provided for capstone projects such as in the Chicago Studies program at the University of Chicago. An array of things summer camp at Argonne and workshops to train social science majors in data science research methodology could also be valuable.

It is important to expand the educational community and the public's interaction with SAGE which can be achieved through educational initiatives, enabling these groups to work with simple data sets, to run small examples on SAGE training nodes, and, potentially, to run experiments on actual SAGE nodes. To better characterize these routes, details on data products and the pipeline for computational science are needed. Examples of broader outreach include through public and cultural institutions such as the Lincoln Park Zoo which three million tourists visit annually. Other cities are home to peer institutions that also engage audiences not interlinked with universities such as parents and young children. The Zoo engages with underserved communities' middle and high school students to teach wildlife biology research techniques. SAGE could provide new data to efforts like these that would allow for hypothesis testing.

Management, Construction, and Operation

SAGE construction and management will be guided by lessons learned in prior deployments including NEON and AoT. This will include key activities such as per-sensor commissioning based on evaluation from sensor experts and domain scientists (such as were done with AoT gas sensors, resulting in multiple publications in the proceedings of the annual meeting of the American Geophysical Union (AGU), and design reviews by scientists and technologists at multiple points in the design-test-evaluate-adjust-deploy stages (some of which loop). These lessons will also go into a rigorous risk register and project execution plan effort as part of the proposal preparation process.

Sustainable operations will entail a combination of local support (within a given instrument, or within a given AoT partner city) and central services (extending the current AoT operations activities). We estimate that local support will be modest incremental cost relative to the extant operations in instruments such as NEON, ARM, HPWREN, UWI, or any of the AoT partner cities (the majority of which will have only 6-12 nodes). However, central costs for node configuration management, cybersecurity, data services, and technical support will be in the range of $1-1.2M per year.

Multiple sustainable operations models were discussed, including industry programs through existing partner organizations such as City Tech and US-IGNITE, and an annual fee for common services, assessed on a per-node basis. A purely subscription-based effort, at $1M/yr across 500 nodes would be $2k/yr/node, which was viewed as too high.

Finally, node upgrades were discussed as part of the ongoing operations, in that an assumed mean-time-between-failures of 5 years would mean that 20% of the nodes should be replaced or refurbished annually. To this end, the new Waggle design will be modifiable (as are today's "ugly boxes" used in prairies, rooftops, and other non-urban locations) rather than sealed (as are the current AoT nodes). The target for refurbish/upgrade cost is $1000 per node.

All of these factors will be integrated into a sustainable operations plan, which will evolve over the course of the construction project.

Agenda

April 29: Plenary

Introduction: Welcome, Workshop Outcomes

SAGE Vision: AI@Edge as an Enabling Architecture

Speaker: Rick Stevens, Argonne National Laboratory and the University of Chicago

Plenary 1: Convergent Science Driving Measurement Capabilities

Moderator: Jim Olds, GMU
Purpose: Begin to Form a set of science questions that will drive SAGE, with testable hypotheses that require SAGE research infrastructure to measure.
Key Questions:

- What science does today's instruments support (for me)
- What unique new measurements will AI@Edge enable
- What testable hypothesis might we investigate with these new measurements
- Describe the national-scale experiment that could leverage AI@Edge in NEON, AoT, ARM, HPWREN, UWIN, etc. to test this hypothesis.

Panelists: Jim Olds, GMU, Christina Negri, UChicago, Scott Collis, ANL, Dan Work, Vanderbilt, Ilkay Altintas, SDSC, Kathleen Cagney, UChicago
Panel Question: What Convergent Questions Cut Across Many Scales? Disciplines?

Plenary 2: Architecture of a National Infrastructure

Moderator: Pete Beckman, NU
Purpose: Establish a common vision for exactly what SAGE will look like and how it will provide measurements to examine the convergent science questions.
Format: Short talks each outlining vision for some facet of SAGE

- Vision: An Integrated Research Infrastructure (Pete Beckman, NU)
- Resilient Edge Computing Hardware and New Sensors (Rajesh Sankaran, ANL)
- Programming Environment and Data Pipeline (Ilkay Altintas, SDSC)
- ML Inference at the Edge and Training in the Cloud (Nicola Ferrier, ANL/UChicago)

Vendor Lunch Round Table: NVIDIA, ARM, HPE

Moderator: Nicola Ferrier, Northwestern
Purpose: Overview of AI@Edge Benchmarks and Use Cases possible with emerging (2019) and forecasted (2022) AI@Edge hardware technologies

Plenary 3: SAGE Partner Instruments: How Will AI@Edge Be Integrated

Moderator: Charlie Catlett, UChicago
Purpose: Preliminary descriptions of what new measurements, sensors, AI@Edge capabilities, will be required for each instrument (what measurement and AI@Edge features are required).
Format: 8-minute talks- how will AI@Edge/Waggle be integrated to augment each instrument

NEON (Eugene Kelly, CSU); AoT (Charlie Catlett , UChicago); HPWREN (Ilkay Altintas, SDSC); ARM (Scott Collis, NU); UWIN (Seth Magle, Lincoln Park Zoo)

Plenary Session 4: Round Table on Bringing the Vision & Science Together

Moderator: Dan Work, Vanderbilt
Purpose: Review the science discussions, instrument architecture, and vision - What are the key challenges? How do our science challenges relate to NSF Big Ideas? Convergent Science? What are we missing?.
Format: Panel Discussion and Round Table
Panelists: Dan Work (Vanderbilt), Jim Olds (GMU), Valerie Taylor (UChicago), Nicki Hickmon (ANL), Pete Beckman (NU)

Post-Plenary: Working Group Leaders and Scribes Planning Session

April 30: Working Group Sessions

Plenary: Big Picture and Working Group Outline Review

Purpose: Identify topics to add to working group discussions (look for missing pieces)

Working Groups Session 1: Survey, Inventory, Outline

Purpose: Refine and finalize outline of opportunities and strategies; big questions, plans, etc.

Plenary: Update from Working groups

Purpose: Read-out and discussion from working groups, identify cross-topic opportunities.

All-Hands Update, Closing Discussions, and Assignments

Participants

Altintas de Callafon, Ilkay	San Diego Supercomputer Center
Anderson, Jason	University of Chicago
B. Dunn, Jennifer	Northwestern University
Beckman, Pete	Northwestern University / ANL
Berman, Marc	University of Chicago
Bruning, Eric	Texas Tech University
Cagney, Kathleen	University of Chicago
Catlett, Charlie	University of Chicago / ANL
Chalamalasetti, Sai Rahul	HPE
Christian, Theresa	Exelon Corporation
Collis, Scott	Argonne National Laboratory / NU
Ferrier, Nicola J.	University of Chicago
Forgione, Vince	University of Chicago
Hickmon, Nicki	Argonne National Laboratory
Kelly, Eugene	NEON / Colorado State University (by video)
Magle, Seth	Lincoln Park Zoo
Nutter, Mark	ARM
O'Brien, Dan	Northeastern University / BARI
Olds, Jim	George Mason University
Papka, Mike	Northern Illinois University / ANL
Potosnak, Mark	DePaul University
Rheinheimer, Randal	Los Alamos National Laboratory
Ricart, Glenn	US-IGNITE
Sankaran, Rajesh	Argonne National Laboratory
Shahkarami, Sean	University of Chicago
Taylor, Valerie E.	Argonne National Laboratory / University of Chicago (by video)
Thiruvathukal, George K.	Loyola University Chicago
Van Hensbergen, Eric	ARM
Williams, Stacie	University of Chicago Libraries
Work, Dan	Vanderbilt University