Federal Agency and Organization Element to Which Report is Submitted: 4900

Federal Grant or Other Identifying Number Assigned by Agency: 1532133

Project Title: **MRI: Development of an Urban-Scale Instrument for Interdisciplinary Research**

PD/PI Name: **Charles E Catlett, Principal Investigator**

Peter H Beckman, Co-Principal Investigator

Kathleen Cagney, Co-Principal Investigator

Michael E Papka, Co-Principal Investigator

Daniel B Work, Co-Principal Investigator

Recipient Organization: **University of Chicago**

Project/Grant Period: 10/01/2015 - 09/30/2020

Reporting Period: 10/01/2018 - 09/30/2019

Submitting Official (if other than PD\PI): N/A

Submission Date: **September 11, 2019**

The following report was submitted through Research.Gov and has been formatted for public dissemination.

Table of Contents

1 Accomplishments

1.1 What are the major goals of the project?

During the no-cost extension of FY19 (and planned for FY20) the major goals, layered atop the original grant goals (below), include:

a. Users and Data: Completion of data dissemination (bulk download and near-real-time-API) services.
b. Traditional Sensing: Completion of air quality sensor characterization and evaluation.
c. Intelligent Sensing: Integration and testing of machine learning codes in edge processors ("at the edge") for pedestrian and vehicle counts and for flood detection.
d. Deployment and System Evaluation: Detailed evaluation of system functionality and reliability of deployed units spanning 3 hardware/software generations of the Waggle platform and Integration of Insights into System Design.

This annual report will focus on these four objectives along with updates on the overall project objectives below that occurred during FY19. Please see FY18 annual report for detail on these overall project objectives and accomplishments prior to FY19.

Overall Project Goals: Create a new form of cyberinfrastructure comprising an "instrument" with three specific **use modalities**:

1. Data Users. Access to high spatio-temporal measurements for a major urban area, including environmental, air quality, and indicators for various forms of activity (e.g., pedestrian or vehicle traffic flows). All data to be open and free, in multiple forms including longitudinal as well as near real time data through web portals, workflow applications, and application programming interfaces.
2. Sensor and Hardware Developers. Enable rapid evaluation and test, at urban scale, of new sensing and other hardware components.
3. Intelligent Infrastructure Developers. Support "software defined" measurements through remotely programmable edge (in-situ) computation, applying machine learning and other techniques to process images, sound, and other data and report on extracted information (e.g., flooding, traffic flows, cloud cover).

Based on the open source, modular "Waggle" hardware/software platform developed by Argonne National Laboratory, the University of Chicago, and Northwestern University, the instrument goals also included supporting these use modalities through cyberinfrastructure demonstrating **critical properties**:

4. Scaling and Resilience. Ability to scale to thousands of devices, providing reliable operation, data collection, and management approaches (and costs) that scale sub-linearly with system size, including policy that can be embedded in, and propagated through, the cyberinfrastructure.
5. Replicability. Adaptable technology to enable similar instruments to be readily constructed, at various scales (smaller and larger than the Chicago instrument) in other venues and contexts (urban and otherwise).
6. Policy and Practice to Support Cyberinfrastructure in the Public Way. Policy and process for embedding cyberinfrastructure *in public spaces*, with appropriate governance, policy, documentation, and public engagement.

1.2 What was accomplished under these goals?

1.2.1 Major Activities:

1.2.1.1 Users and Data

AoT data is accumulated on internal servers running the Cassandra database. At midnight UTC time the previous 24h of data is pushed to a public download server with a web interface that allows users to select among different packaging sizes, from "all data" (nearly 300GB) to "yesterday's data." Every five minutes the previous 30 minutes of data is pushed to an API server running on commercial AWS systems, enabling support for application development or near-real-time mirroring of the data by other groups. All data is free and open/public.

Data API. Worked with user community to refine API functionality, scalability/performance, and reliability. Developed automated processes for fault detection using commercial cloud (AWS) services.

Data Tools. Developed and published data manipulation tools such as extracting only certain sensor values or data from a subset of nodes.

Data Management. Established per-project data download and API functionality for different deployments, organized by project (e.g. partners in other cities).

Third Party Data Integration. Created import functions and code for non-AoT platforms to allow for sensor data to be reported from third party devices, using AoT data schema, for integrated analysis in concert with AoT data. This allows, for instance, data from PM2.5 devices such as the low-cost Airbox (Location-Aware Sensing System) project with over 4500 sensors in Taiwan, to be combined with AoT PM2.5 measurements. In Chicago several science teams studying air quality and human well-being (education performance, health, behavior, etc.) have deployed dozens of low-cost Airbox units, and these now can effectively expand the AoT PM2.5 footprint.

1.2.1.2 Traditional Sensing

Gas Sensors Evaluation. During FY19 the NO_2, SO_2, and O_3 sensors were evaluated using two nodes collocated with EPA sensors. Results were published in the proceedings of the AGU, Fall 2018 meeting. Evaluation of CO data began in FY19 and results should be available by the fall 2019 AGU meeting. Preliminary findings, now that we have collected sensor data over a long enough period of time (18 months) show that the printed electrochemical gas sensors (NO_2, SO_2, O_3, CO, etc.) exhibit a predictable linear decline in sensitivity to heat generated by the on-board sampling microprocessor over the six quarters studied. This will be an important factor for longitudinal studies.

Particulate Matter. In FY18 we evaluated Alphasense ($500) and Plantower ($30) particulate matter sensors. Roughly a dozen of the first 100 nodes deployed had Alphasense, and we collaborated with Academia Sinica Taiwan to examine the over 4500 low-cost devices they had deployed using Plantower. We saw no appreciable difference in data quality (using the same EPA collocation process as described above for gas sensors) despite the significant price disparity. Moreover, multiple Alphasense devices tested as functioning properly in the final quality check at the time of manufacturing, but once they had been mounted on city poles their data was corrupted, likely due to misaligned optics caused by shocks and vibration in transit and installation. We decided in FY19, then, to install only the Plantower sensors, and the 100 units built in FY19 all have these PM sensors.

Sound Sensor. Microphones chosen for the 2018 and prior devices were evaluated as problematic in that they were overly sensitive to low frequency and wind noise. Combined with their placement at 22’ above the street, they produced little valuable sound data. The nodes manufactured in FY19 use an upgraded microphone that measures sound intensity in 10 octaves, and these will be evaluated in FY20.

Light Sensors. The sky-facing light sensors were so sensitive that they became oversaturated during daylight, effectively providing binary measurements. Nodes manufactured in FY19 included neutral density filters to address this saturation. We will evaluate these in FY20 with the new nodes, but preliminary data shows marked improvement in the data streams. With this improved light sensing capability during summer 2019 one of the student projects developed a machine learning method for correcting temperature readings based on light intensity, compensating for the heat offset produced by the Stevenson shield.

1.2.1.3 Intelligent Sensing.

Pedestrian and Vehicle Counts. Initial results from pedestrian and vehicle counting software (machine learning within the edge processors) validate that the Odroid CPU/GPU systems are adequate at 30-second samples, which are useful for longitudinal studies but less so for pedestrian and vehicle flow studies. To this end we tested these algorithms extensively on multiple platforms (Google TPU, NVIDIA TX2, ARM, and Intel hardware) during FY19 to drive the selection of upgraded processors for future AoT nodes. For pedestrian and vehicle counts we currently use the off-the-shelf Mobilenet SSD (Single Shot Detector) model pre-trained with the COCO dataset.

Vehicle Type Identification. In partnership with an industry partner, we evaluated edge ML software for determining the make of individual vehicles. We found that the birds-eye-view of AoT nodes increases the difficulty, particularly given that most training data is ground-level (person-eye) view, often of the front of the vehicle with the manufacturer logo prominent. Similarly, most pedestrian flow training data is person-eye view, with similar challenges presented to optimizing such functions using AoT birds-eye-view. A transfer learning testbed is being established at Argonne National Laboratory's main entrance with both ground-level (6-8' elevation) and AoT-level (22' elevation) cameras. Improvements from this effort will translate into software updates improving AoT measurements.

Initial evaluation of flood (standing water) detection as well as cloud cover (AoT sky camera) began in FY19, with algorithms performing reasonably well in simple cases (standing water with movement vs. still; distinct clouds vs many diffuse clouds) but require more work and testing.

Virtual Containers and Scheduling ML@Edge Codes. In FY19 we worked with ARM to prototype a virtual container-based deployment scheme for edge computing functions along with a scheduling function that would support prioritization of such functions on nodes. This approach will allow users to develop and debut edge computing software, and in turn will support automated deployment and prioritization of user codes on AoT nodes.

1.2.1.4 Deployment and Evaluation

Cameras and Imagery. Evaluation of image data revealed that even after 6-12 months the build-up of dust on the inside of the downward-facing camera enclosure lens was unacceptably high, rendering images of limited use. The manufacturing line was stopped in December 2018 while a camera shroud was designed, tested, and integrated into the assembly process. This delayed deployment for several months as prototypes were developed and tested and then production parts were manufactured and integrated into the manufacturing assembly guide.

Moisture Vulnerability. During FY18, many nodes experienced moisture issues including condensation on the upward-facing camera lens, which was corrected by adding a second layer of Plexiglas and Goretex vents to the processor and controller enclosure. Moisture on both the environmental sensor board (in the Stevenson shield) and the light sensor board (in the processor enclosure) caused incorrect readings from these boards, which we corrected with firmware modifications. In new nodes manufactured in FY19 we also (a) improved the conformal coating process for sensor boards in the Stevenson shield and (b) added conformal coating for the sensor boards in the processor enclosure.

Node Controller Filesystem Resilience. During FY19 the Node Controller system (Odroid C1+ single board Linux computer) was modified to use a read-only filesystem, addressing corruption issues with key configuration files, seen in prior generations of nodes.

Power Reliability and Recovery from Outages. A significant number of node outages were diagnosed to be caused by loss of power. In some cases, the wiring in the traffic control cabinets (where AoT power circuits originate) had been disconnected or the circuit breakers tripped. The cabinets are accessed by multiple City of Chicago and commercial providers, not all of whom are familiar with the AoT project. Working with our primary installation partner (Chicago Department of Transportation (CDOT), Department of Electrical Operations (DEO)) we developed initial briefing strategies and are now moving to (a) more visibly label the circuits within the cabinet, with contact information; and (b) move to a new power connector on the nodes, which we designed with our manufacturing partner, with an LED such that node power can be verified without opening the cabinets or checking the wiring. The labeling has already resulted in contact from other city departments (unfamiliar with the AoT project) to coordinate outages during pole replacements. This caused an interruption in the installation of the 100 units in FY19 after the first 20 units were deployed. Installation is scheduled to resume in late September and should be completed by the end of 2019.

Cellular Connectivity. Although AoT is deployed in a major metropolitan area using a commercial cellular provider we have documented significant portions of the 600 sqkm City of Chicago where cellular signal is inadequate to support node connectivity, resulting in nodes that appear to be failed but in fact are simply offline. New nodes in FY20 will take this into consideration, in that we will avoid deployments in poor coverage areas.

Alternate Installation Partnership. Concurrently we have established a partnership with a company deploying 5G cellular towers in Chicago (and similar installations in other US cities). The company has agreed to deploy nodes at no cost during FY20 and are open to providing no-cost access to their fiber backhaul as well. This partnership began in FY19 with a deployment of 7 AoT nodes on partner company's private towers and will be expanded to as many as 30 nodes in FY20. We anticipate that power reliability will be considerably higher with these commercial deployment partners relative to the power reliability of the installations using City of Chicago power.

Debugging Systems in the Public Way. The challenge to deploying science equipment in the public way, where physical access for routine maintenance and debugging is impractical, is exacerbated by the security architecture of the nodes, whereby there are no open network ports. That is, nodes must "phone home" and if there are issues such as cellular coverage or booting of the node controller CPU (to which the 3G modem is connected) then the node is unreachable and cannot be debugged without retrieving it. In addition to the node controller CPU console cable used for benchtop debugging and manufacturing quality assurance, the updated WagMan (controller board) in the 2019 nodes also has a console cable. Summer students in 2019 prototyped additional visual status features planned for future nodes, where an LED-equipped terminator will be attached to each console cable and used to provide error and status codes visible from the street. The WagMan console cable also supports integrating devices such as wireless-equipped microcontrollers that could be battery-powered and used to (a) report, for instance, diagnostic information associated with loss of power, and (b) allow remote interactive debugging.

Node Resilience to Power Outages. AoT nodes use Odroid single board computers, an open source hardware device produced by Hardkernel in South Korea. The Odroid C1+ is used for the Node Controller Linux system (handles data management, security, configuration, etc.) and the Odroid C4 (with GPUs) is used for all edge machine learning (with cameras and microphone directly connected). We observed some cases where a power outage would cause the Odroid C1+ Node Controller system to erase the Wagman (Waggle controller board, see [1]) microprocessor. Modifications to the Wagman board for version 4 (nodes built and deployed in FY19) addressed this issue, however we continue to work with Hardkernel (Odroid manufacturer in South Korea) to understand random C1+ reboots that randomly hang.

1.2.2 Key outcomes or other achievements:

In addition to the activities and results noted above:

Supporting a Broad Range of Science and Community Measurement Initiatives. Node location selection is based on requests through an *open process available to scientists, community groups, city organizations, and the general public*. Input from these sources, combined with explicit community feedback through public meetings and discussions with local city council members, and overall project strategies such as broad spatial coverage, measurement of representative microclimates and urban forms, and availability of adequate cellular network coverage. The following projects have designated node locations to date, ranging from a single node to a dozen or more nodes:

(a) Chicago Department of Planning and Development; Chicago Department of Transportation. Chicago North Branch Framework nodes measure the environment and traffic to understand the impact of redistricting and new development.

(b) Atmospheric Sciences Community (per AoT workshops). North-South lakefront and East-West transects support studying the "lake effect" on weather and air quality as well as urban land use impact on regional weather.

(c) Chicago Department of Public Health. Lower West Side (Pilsen, McKinley Park, Little Village) has many known air pollutant sources and is the site of an EPA superfund project. These nodes aim to test AoT node density and placement strategies to measure air pollutants.

(d) Northwestern University Growing Convergence Research project (NSF CBET 1848683; Convergence: RAISE: Systems Approaches for Vulnerability Evaluation and Urban Resilience). Chatham is one of Chicago’s most flood-prone neighborhoods. Scientists at Argonne, Northwestern, and UChicago are developing flood detection capabilities using AoT image analysis and wireless moisture sensors.

(e) Ravenswood Theater. Measuring pedestrian flow and use of public spaces with community events aimed at addressing racism.

(f) Pullman National Monument Preservation Society. Measuring traffic changes resulting from the creation of the Pullman National Monument.

(g) Chicago Department of Fleets and Facilities Management. Testing new methods to measure pedestrian and vehicle flow and related impacts (noise, congestion, air pollutants, etc.) in the Chicago Loop, in and near public buildings, and along the River Walk.

(h) UChicago Population Research Center. Measuring air quality and noise in neighborhoods of the hundreds of participants in an NIH-funded study on poverty, aging, and health.

(i) General Public. Several dozen locations were requested by individual residents and community groups.

(j) Chicago Department of Public Health, UChicago *COMPASS* health project, and the UChicago/UIC NIH-funded Chicago Center for Health and Environment *(CACHET)*. Measurements to understand urban environmental impacts on health.

(k) Illinois Department of Transportation and Chicago Metropolitan Agency for Planning require rail crossing delay data to prioritize investments in transportation infrastructure.

(l) The Bronzeville Community Smart Grid (IIT, ComEd) will leverage Exelon’s partnership with Argonne to investigate AI-based energy optimization and failure prediction.

(m) Ten Chicago departments and agencies: Chicago’s Vision Zero program seeking to eliminate traffic-related fatalities by 2022. Nodes are installed at 30 of the city’s most dangerous intersections and corridors. Additional nodes along Ashland avenue aim to measure the impact of rapid future bus transit systems.

(n) Chicago Public Schools (Lane Technical H.S.) have worked with the AoT team to train over 600 students to develop sensor-based science projects, and use nodes near the school in conjunction with the course teaching students to build and use wireless sensor technologies.

<u>Interest from Science Groups in other Cities</u>. As of early FY19 there were six partners in as many cities (UW/Seattle, City of Portland, Panasonic/Denver, Stanford/Menlo Park, SyracuseU/Syracuse, and UNC/Chapel Hill). Each had between 3 and 10 nodes. Without exception these partners have experienced delays due to the pace of approvals and actions in their cities. We concluded that the partnership program needed to be re-structured to ensure that the Chicago team would have the resources to provide more hands-on assistance, which was not possible without impact to the Chicago AoT project. During FY19 we moved from a “cost per node” to a “cost per project plus cost per node” model, which will be evaluated in FY20. In this period, additional partners joined including Vanderbilt/Nashville, GaTech/Atlanta, UUtah/Salt Lake City, UBristol/Bristol (UK), CSIRO/Melbourne (AU), and NCHU/Taichung City (TW). Additional interest and discussions regarding partnerships in FY20 are ongoing with universities in Dallas, Houston, San Diego, Boston, Vancouver, and Chattanooga (where the Waggle platform is deployed), as well as

Taipei (TW), Phnom Penh (Cambodia), Hong Kong, Tel Aviv (Israel), Copenhagen, Seoul, Daejen (Korea), Busan (Korea), and Tokyo.

Integrating Data from Lower-Cost Measurement Platforms. While in FY18 and FY19 we continued to prototype microWaggle devices (see FY18 annual report), we also explored how data from third party sensor platforms might be combined with AoT data. This would mean that AoT tutorials, APIs, and analysis tools would be available to these additional user communities for analysis of their data. Of particular importance is the difficult of designing and deploying a reliable data platform such as AoT has done with the Waggle platform. With many sensor projects this is an afterthought, where the design of a simple sensor device (without edge computing or other AoT functions) is quite simple. We collaborated with the Location-Aware Sensing System (LASS) project at Academia Sinica in Taipei, Taiwan to support a summer student from the LASS team in 2019. The student completed the preliminary data translation service so that LASS "Airbox" PM2.5 measurements could be integrated with AoT data (and APIs, tutorials, and tools).

Support for Alternate Edge Computing Hardware. After evaluation of multiple machine learning hardware devices, we collaborated with ARM to begin porting Waggle/AoT node software to their CPI hardware, exploiting the new WagMan controller board's compute-agnostic design.

Machine Learning User Code. In partnership with CPU companies (see "Intelligent Sensing" activities and accomplishments above) we now have a prototype system that will (a) improve edge machine learning performance by factors of 30-50 depending on algorithms (based on benchmarks performed in FY19), and (b) streamline edge compute development by partners, enabling them to develop and test in virtual systems and then simplify the system by which AoT deploys those containers to AoT nodes. We also began testing a scheduling capability for these containers, to be tested in FY20.

Machine Learning: Training Library. Currently, AoT nodes record and save one image from both the street-facing and sky-facing cameras every 30 minutes, along with 1 minute of audio. This data is processed using quality check filters, discarding images that are too dark or audio that has no signal (due to no noise). We have established a data use agreement that allows any scientist to obtain the entire training set.

High School Curriculum. In four years, the curriculum, "Lane of Things," renamed "School of Things" (SofT) in 2018, has trained some 600 high school students at Chicago's Lane Technical High School. The SofT team submitted a NSF Smart and Connected Communities planning grant proposal in September 2019 to build on lessons learned regarding scaling and replicating to other schools. This planning effort will explore the potential for (a) integrating with CSforALL curricula, and in particular the AP course "Computer Science Principles" taught nationally including at the over 100 Chicago Public High Schools, and (b) packaging for continuing education in the form of 2-3 day workshops to be taught in suburban and rural communities. Industry partners include Underwriters Laboratory (online education) and American Infrastructure (operates over 300 post office facilities and working with their tenant, the US Postal Service, to develop community programs).

1.3 What opportunities for training and professional development has the project provided?

Nearly 20 undergraduate students from the US and abroad participated in the summer program led by AoT team members Catlett, Beckman, and Sankaran in 2019. Additionally, a team of 3 students from the University of Chicago's Computational Analysis and Public Policy (CAPP) master's degree program began their capstone project in FY19 (to be completed in FY20) using AoT system data (nodes report sensor as well as internal system data) to develop predictive models for node or sensor failure.

Collaborators at Northwestern are using AoT data in concert with satellite data (Tropomi, provides air quality measurements at 5x8km resolution once daily) to evaluate AoT Ozone and other gas sensors, and preparing for the TEMPO satellite (estimated to be ready in 2021) which will provide hourly 3x5km resolution data.

A team of five students at Vanderbilt University (three graduate students and two undergraduate students; supervised by co-PI D. Work) are installing four nodes at Vanderbilt University in Nashville, TN as part of the university's MoveVU mobility strategic plan. The team is also developing new automated and unsupervised data cleaning methods based on an approach known as tensor factorization. The approach imputes missing data and corrects erroneous measurements that are inherently present in large-scale sensor networks. The methods have been applied to Chicago AoT data and cross-validated with *National Oceanic and Atmospheric Administration* (NOAA) data where applicable, improving agreement with the NOAA temperature readings by up to 30%. This work is part of a NSF CMMI grant (#1727785; Signatures and Barcodes: Data-driven Understanding of Transportation System Performance during Extreme Events; PI Dan Work).

The SofT high school curriculum discussed earlier has trained 600 students in four years. In FY19 the team worked with teachers from six Chicago public schools (middle- and high-school) to follow up on summer 2018 training using four Saturday check-in workshops. They discovered that resource constraints were significant (schools could not afford the hardware, even at $50/kit) as was the difficulty of integrating an 8-week program into existing curricula. To this end the team submitted a planning grant proposal to the NSF Smart and Connected Communities program in September 2019. The LofT/SofT program has been supported over four years by Motorola Solutions Foundation with a total of $200k of funding.

Co-I Papka along with colleagues at Argonne National Laboratory have developed a big data camp for high school juniors and seniors to introduce them on how they do research with data. The camp introduces the students to a variety of different methods and tools using the Array of Things datasets. Started in 2018 with 14 participants (9 female and 5 male) for 3 days, expanding to 5 days in 2019 with 13 participants (3 female and 10 male), students get to learn first-hand what it's like to be a data scientist with data from their own backyard. The familiarly with Chicago and local weather conditions allow the students to make quick connections to the underlying data and focus attention on learning skills in mapping, transforming and filtering data to answer questions in a larger context than the individual data sources.

1.4 How have the results been disseminated to communities of interest?

Results are disseminated through talks at conferences, the project website, publications, and the media. The project website includes access to reports on community engagement and transparent policy development, project policies, project technical details, presentations, and science and education goals. The project website, http://arrayofthings.org, keeps links to all videos and major news publications about the project, and is continually deepened with detail on policy, status, and technology.

At the AoT User Workshop in August 2018 there were tutorials as well as a round table mini-tutorial related to "lessons learned" deploying cyberinfrastructure in the public way. Led by project executive director Kate Kusiak Galvin along with Chicago's current CIO (Danielle DuMerer) and former CIO (Brenna Berman), this round table reviewed aspects of the AoT deployment that had gone well, as well as reviewing "what we would do differently in retrospect."

All workshops are documented with workshop reports at the UChicago Urban Center for Computation and Data website, and specifically the link http://www.urbanccd.org/past-events

1.5 What do you plan to do during the next reporting period to accomplish the goals?

During FY20 the project will operate under a No-Cost Extension, pursuing the plans below aimed at each of the four objectives outlined in 1.1.

1.5.1 Users and Data

In FY20 we will work with the user community to evaluate and improve the documentation and features of the AoT bulk data download repository and the APIs. The bulk data download site is provided to the project by the University of Chicago and Argonne National Laboratory as part of cost-sharing and the APIs, and Plenario are implemented in Amazon Web Services.

1.5.2 Traditional Sensing

NO_2, SO_2, and O_3 were evaluated and characterized in FY18-19; CO evaluation and characterization will be completed in FY19. PM2.5 sensors will also be evaluated and characterized in FY19. New options for sensors will also be examined, given that the current set of sensors was finalized in early 2018 and new sensors have greatly improved since that time, including reductions in cost. We are also evaluating adding sensors in the next version of the platform, including CO_2, wind, and optical rain sensors.

1.5.3 Intelligent Sensing

The computer vision and hearing team comprising Argonne and UChicago scientists will continue to evaluate algorithms for pedestrian counts, vehicle counts, and flood detection, leveraging the new virtual container-based deployment methodologies to ensure that these algorithms are running on all AoT nodes. A growing number of collaborators are engaging with the team to test their algorithms on the AoT instrument and have been supportive of the container approach.

1.5.4 Deployment and System Evaluation

We will complete the “postmortem” evaluation of all AoT nodes that have failed to date, using this evaluation as input to the final design of the next iteration of the Waggle platform. This work will build on the UChicago masters student work in FY19 to develop machine learning algorithms to predict node failure using node data (nodes report not only sensor data but a significant number of telemetry measurements ranging from memory usage to CPU temperature).

2 Products

2.1 Journals or Juried Conference Papers

Silva MP, Sharma A, Budhathoki M, Jain R, Catlett CE. Neighborhood scale heat mitigation strategies using Array of Things (AoT) data in Chicago. AGUFM. 2018 Dec;2018:PA21D-0986.

Potosnak, M.J., Banerjee, P., Sankaran, R., Kotamarthi, V.R., Jacob, R.L., Beckman, P.H. and Catlett, C., 2018, December. Array of Things: Characterizing low-cost air quality sensors for a city-wide instrument. In *AGU Fall Meeting 2018*. AGU.

2.2 Audio or Video Products.

Many videos have been produced by third parties including a video done for SC17 in conjunction with an AoT-driven plenary panel and an in-depth profile of the project by BBC Click and a highlight video for Supercomputing 2017.

2.3 Websites

2.3.1 City of Chicago Public Engagement Report

http://arrayofthings.github.io/engagement-report.html

Report detailing public engagement efforts, goals, and results. Public engagement objectives were to inform residents about the instrument goals and capabilities, discuss privacy concerns, and provide details on policies. An overarching objective was to position the instrument as a community "asset" as distinct from a sterile device or a surveillance program.

2.3.2 FAQ on Privacy and Policy and Feedback

http://arrayofthings.github.io/policy-responses.html

Exhaustive (regularly updated) FAQ that addresses all of the public comments and suggestions, initially taken from public meetings, email, web forms, and multiple online options for commenting on draft policies (including the OpenGov Foundation's Madison site).

2.3.3 Node Location Request

https://docs.google.com/forms/d/e/1FAIpQLScIG6YevqzWW4d1U0eH1D2tNSQuHnlng6AMrlDE4V8Gf136qQ/viewform

Form for use by residents or others requesting that AoT place one or more nodes in a specific location or neighborhood. Allows a non-scientist to articulate a science question in terms of goals, type of measurement, and locations.

Note this form is linked from the friendlier URL of the main project site, http://arrayofthings.us

2.3.4 Plenario System Overview

https://github.com/UrbanCCD-UChicago/plenario-stream/wiki/System-Overview

Extensive system documentation for the Plenario system that will serve as a key portal for all AoT instrument data. The site is necessary to support developers wishing to create portals, analytics tools, or applications that use instrument data.

2.3.5 Project Information Portal

http://arrayofthings.github.io

Main project website with links to all policies, technical documentation, press, etc.

2.3.6 VIDEO: BBC Report on Chicago Array of Things]

http://www.bbc.com/news/av/technology-39229221/air-quality-tracker-for-chicago-to-roll-out-city-wide

An update on air quality and Array of Things, published March 2017 and following a previous BBC profile on the project:

http://www.bbc.com/news/av/technology-32511363/key-to-chicago-s-quality-of-life-the-array-of-things

2.3.7 VIDEO: AoT Technology and Privacy

https://www.youtube.com/watch?v=pFL5QNwgs6A

Short video focusing on the manufacture, technology, and privacy of the AoT instrument. The video was developed specifically to help residents understand the privacy protections built into the device and into policy.

2.3.8 Waggle Platform Software Site

https://github.com/waggle-sensor/

Developed site to document the open source software stack, and eventually the hardware specifications, for the instrument. This site is also essential to engaging the research and development community in evaluating and hardening the software stack.

3 Impacts

3.1 What is the impact on the development of the principal discipline(s) of the project?

Adoption of AoT's Waggle platform has been accelerating (see other participating organizations for a list of the dozen groups already testing nodes to support their research, and in turn providing valuable feedback to the AoT project regarding node reliability, sensors, platform, and edge computing features. In FY19 we adopted a virtual container approach to deploying and scheduling edge machine learning codes based on interactions with computer scientists wanting to use the AoT systems for research ranging from adaptive and goal-oriented computing to advanced computer vision and hearing algorithms. Lessons learned in the

AoT platform development work fed directly into a proposed mid-scale research infrastructure in partnership with the NSF NEON project and others, resulting in an FY20 award for $9M to build a next-generation platform based on AoT nodes.

3.2 What is the impact on other disciplines?

Social and Behavioral Sciences. Co-I Kathleen Cagney (UChicago) along with behavioral scientists Howard Nusbaum and Marc Berman have been working with the AoT team to use AoT sound and street activity (e.g., pedestrian flow) data to understand social cohesion and related topics ijn neighborhoods. They proposed a NSF Smart and Connected Communities (S&CC) project in September 2019 to this end. The proposed project also includes the use of AoT air quality and sound data as additional variables characterizing street segments with respect to their "naturalness" in context of a navigation application designed by Berman's group to conduct experiments on the cognitive benefits of natural versus urban settings. Dan O'Brien has similarly worked with the City of Boston to develop a NSF S&CC proposal using AoT machine learning at-the-edge capabilities to understand social interactons in public places.

Atmospheric Sciences and Air Quality. A project funded by the Department of Energy's Exascale Computing Project used AoT temperature, relative humidity, and barometric pressure data to validate and parameterize a high-resolution urban weather model (Weather Research and Forecast Model, or WRF, at $100m^2$ horizontal resolution), contributing to a publication on the use of such models to understand building performance during extreme heat events (see Products publication Jain, Rajeev and Luo, Xuan and Sever, G{\"o}khan and Hong, Tianzhen and Catlett, Charlie (2018). Representation and evolution of urban weather boundary conditions in downtown Chicago. Journal of Building Performance Simulation.)

Environmental Sciences. Northwestern University and in particular the NSF-funded Convergence: RAISE: Systems Approaches for Vulnerability Evaluation and Urban Resilience (Aaron Packman, PI, NSF 1848683) are using AoT in Chicago's Chatham neighborhood to understand and predict urban flooding.

Civil Engineering and Transportation. Co-PI Daniel Work (Vanderbilt) is working with the City of Nashville to deploy a major AoT instrument focused on mobility research (see also previous section on Training and Professional Development).

The privacy policies continue to be of interest to other projects as examples of community engagement and balancing science and privacy needs.

3.3 What is the impact on the development of human resources?

The School of Things (formerly Lane of Things) high school program, outlined earlier, is a key component to developing human resources. The curriculum has been developed by Kate Kusiak Galvin (UChicago), Douglas Pancoast, Satya Mark David and Robb Drinkwater (School of the Art Institute of Chicago), and Dan Law and Jeff Solin (Lane Tech High School). Some 600 high school students have been trained in three consecutive years, and a dozen Chicago teachers have taken a professional development workshop created by the team.

Finally, Co-PI Mike Papka created a team of undergraduate students at Northern Illinois University who developed a set of Jupyter notebooks that form a new online Python tutorial on AoT data analytics.

3.4 What is the impact on physical resources that form infrastructure?

These are covered in the supplemental documents (AoT User Workshop Report), but it is important to note that AoT is a unique urban-scale instrument that supports both measurements at unprecedented spatial and temporal resolution and edge computing experiments such as data reduction, image and sound analytics, or video analytics. As such, AoT can be seen as complementing other measurement infrastructures such as EPA sites and remote (satellite) sensing infrastructure.

3.5 What is the impact on institutional resources that form infrastructure?

The AoT project has been the impetus behind relationships between Argonne and UChicago CDOT, DOIT, and new relationships with deputy commissioners and department heads in the Departments of Public Health, Planning and Development, Buildings, Parks and Water-related agencies. During FY18 these relationships extended to county and state agencies including the Cook County Department of Environmental Controls, the Illinois EPA, and the Illinois Department of Transportation (IDOT) as well as key regional planning entities such as the Chicago Metropolitan Agency for Planning. For instance, IDOT requested that several AoT nodes be located on at-grade rail crossings so that computer scientists could analyze the traffic impact of rail crossings, with the potential to expand the program to hundreds of rail crossings throughout Illinois.

3.6 What is the impact on information resources that form infrastructure?

This year, with critical mass of data available from the first 100 nodes, multiple GIS groups have engaged from UChicago, Northern Illinois University, and the University of Illinois to explore incorporating AoT data into their GIS portals and resources. By piggybacking on the NSF-funded Plenario system (NSF-1348865, http://plenar.io) we are able to integrate sensor data with other urban data, such as 311 calls, crimes, permits, or inspection data.

3.7 What is the impact on technology transfer?

We have transferred the supply chain management and assembly/test of AoT units to a Chicagoland electronics partner, Surya Electronics (the firm that also fabricates our circuit boards). All software for the nodes, data pipeline, and central server functions is published in Github (https://github.com/waggle-sensor) and hardware specifications for circuit boards and connections will be published in FY19 once we complete detailed system evaluation of the 5th generation controller board (WagMan) through the installation of 100 units beginning December 2018.

We have established a research partnership program whereby we can support AoT projects in a sustainable fashion, with a "turnkey system" approach with centrally managed nodes irrespective of geographic location and a more extensive partnership program enabling non-US partners to operate the full system of data and management servers locally. These partnerships include agreements that require that any and all AoT projects adopt both the open/free data policy and the privacy policies.

3.8 What is the impact on society beyond science and technology?

The near weekly contacts from students and faculty interested in learning more about the project and doing data analytics or machine learning projects, and from cities and universities interested in building AoT-like infrastructure reveals an increasing need for a platform that can be rapidly deployed in cities, focused on specific questions in particular neighborhoods, to provide insight into challenges, to provide ongoing measurement of conditions, and/or to evaluate the impact of major urban development projects. Today these projects are extremely costly and tend to involve bespoke systems that are either academic-generated (not reproducible) or proprietary (not producing open data; not available to support innovation from the science community.

1. Beckman, P., et al. *Waggle: An open sensor platform for edge computing*. in *SENSORS, 2016 IEEE*. 2016. IEEE.