Applying Artificial Intelligence to Broad and Deep Cyber Infrastructure

William Kramer[1], Charles Catlett[2], Alok Choudhary[3], James Brandt[4], Aaron Saxton[1], Saurabh Jha[1], Ann Gentile[4], Ravishankar Iyer[1], Peter Beckman[3], Ewa Deelman[5], Brett Bode[1], Kjellrun Olson[1], Roy Campbell[1], Volodymyr Kindratenko[1], E. A.Huerta[1], Celso Mendes[1], Zbigniew Kalbarczyk[1], Wei-keng Liao[3], Rafael Ferreira Da Silva[5], Ankit Agrawal[3], William Gropp[1]

1 – University of Illinois, 2 – University of Chicago, 3 - Northwestern University, 4 – Sandia National Laboratory, 5 – University of Southern California

Corresponding Author – William Kramer – wtkramer@illinois.edu, +1(217) 979-7577

Abstract: Cyber Infrastructure (CI) is increasingly harnessed and is expanding to support the use of artificial intelligence (AI) techniques and applications to enable smart cities, power grids, computational simulations, real-time processing in big-data experiments (LIGO), supercomputing, and other important innovations. Computation is essential to developing and using the analytics while CI networks and data facilities are central to moving and managing the immense volumes of data that enable progress and insight. Just as AI offers opportunities to revolutionize all aspects of our lives, the potential for improving the design, evolution, optimization, performance, operation, security and sustainability of the CI is equally promising. A convergent research and implementation approach is required to enable the application of *AI for the CI*, comprising experts across the spectrum of cyberinfrastructure research, design, and deployment; computer and computational sciences; and systems design and engineering.

Introduction: All domain science and commerce depend heavily on a robust Cyber Infrastructure (CI) comprising hardware and software systems across ranges of scale (from embedded devices to high performance computing (HPC)), functionalities (from sensing to computational modeling to data analytics to visualization, from data management to communications), and architectures (from centralized to distributed, from tightly coordinated to autonomous, with a wide diversity of underlying technology). All aspects of NSF's, and indeed the nation's, endeavors depend on the creation of scientific and engineering capabilities through the orchestration of CI layers and components that is only increasing as intelligence and autonomy are integrated into all layers and components. Harnessing the power and capacity of current and future CI technologies presents unique and time critical challenges that must be identified, understood, and addressed to maintain and increase the nation's leadership in scientific, industrial, and societal innovation and discovery.

CI is increasingly harnessed to support artificial intelligence[1] (AI) techniques and applications to enable innovation for smart cities, power grids, real-time processing in big-data experiments (such as LIGO), supercomputing simulations, and others. Computation is essential to developing and using the requisite analytics, while CI networks and data facilities are central to acquiring, moving and managing the immense volumes of data required. Just as AI offers promises to revolutionize all aspects of our lives, the potential for improving the design, evolution, optimization, performance, operation, security and sustainability of CI is equally promising. Following the terminology used in the NSF Big Data and Extreme Computing Workshop series we refer to AI applied to challenges of improving Cyber Infrastructure as AI for CI [1].

A convergent research approach is required to explore the application of AI for CI, comprising experts from across the spectrum of cyberinfrastructure research, design, and deployment; computer and computational sciences; and systems design and engineering. To this end, we propose to create a National AI for CI research and development effort to drive transformative exploration of CI using AI methods including areas ranging from CI performance and optimization to the concept of enabling scientists to specify high-level science objectives from which workflows might be constructed and resources (hardware, software, data) might be assigned and made available.

There are clear needs for AI innovations and implementations focused on the CI itself to build a new multidisciplinary community of AI and CI researchers. It will include people designing, implementing and supporting CI across the entire CI spectrum (from the largest to the smallest systems, from the tightest to loosest integration); and core AI/ML experts who can explore the use of AI to improve the effectiveness and efficiency of CI design and implementations.

[1] while there are differences in implementations and use of Machine Learning, Deep Learning and Artificial Intelligence, there are related for the sake of this paper we will refer to all methods as AI

Conceptual Framework: Harnessing AI to Reinvent Cyber Infrastructure: Machine Learning (ML), Deep Learning (DL), and AI have recently become important and practical alternative methods for data-driven discovery in a number of science and engineering domains including physics, agriculture, natural language processing, facial and object recognition, imaging, and automation. Recent successes have been enabled a combination of massive increases in raw processing power, the availability of exponentially increasing data for analysis and learning and successes in algorithms for applying that processing power to solving AI problems. All of these areas now rely on powerful and stable CI to apply AI methods to accomplish their domain science goals and innovation.

These examples of using CI to empower AI represent tremendous research opportunities [5]. CI is complex, vast, expensive, and can be difficult to use effectively. AI for CI endeavors will help reduce errors, decrease the time needed to rectify problems, faults and inefficiencies, and significantly decrease barriers to improving AI and other application performance.

Disease vectors are increasing at an alarming rate with the increase in travel and export/import on a global scale as evidenced by the two recent outbreaks in coronavirus and aggressive flu epidemics. Creating timely cures and modeling contagion spread requires sophisticated CI that is currently lacking (e.g., distributed sensors coupled with computation, global connections of medical providers, etc.) All aspects of enabling this acceleration hinge on our ability to efficiently build and harness the CI that is integral to all aspects of modern technology. This is but one example of current and urgent problems; improving the CI addresses many other science, research, and engineering innovations needs that will enable acceleration of productivity and decreased time to insight/solution across all disciplines. Enhancing prediction of severe weather, reducing time to discovery of new drugs and new materials, understanding our changing earth and these changes impacts on society and security, reducing energy use and increasing efficiency, understanding human behavior, understanding the universe's origin and function, and many other areas all rely on fundamentally and dramatically improving how our CI works. With the rate for hardware enabled improvements decreasing we must increasingly rely on much smarter and more effective software.

To enable the end goal of accelerating discovery, it is critical that we incorporate AI improvements in our CI starting now. The urgency comes from two parallel time-critical issues. First, AI for CI can enable important improvements to the existing and soon to be deployed CI over the next 5 years but only if we begin work immediately. Possibly even more important, new, long lasting CI architectures and systems are being designed today and it is imperative to their future performance that the new architectures incorporate AI in their CI in a comprehensive and efficient manner from the ground up.

The Human-in-the-Loop (HIL) methods prevalent in many aspects of CI operations today are too limited when dealing with complexities for even small CI implementations and may become impossible when complexity grows to national scale CI. The cost of HIL in terms of inefficient use, delayed or incomplete discovery, or simply unavailability of CI components is now too high given CI is expensive and each implementation has a limited time-span window to make an impact. The current "top-down, human-in-the-loop" approaches for its design, implementation, effective use and improvement are very limiting and prone to error. Studies [7] [8] show that human actions contribute significantly to downtime, errors, and rework in CI implementations. HIL is typically limited to being reactive (or retrospective), using a postmortem analysis of events, anomalies or performance. Data-driven AI based techniques appear the best options for making predictive approaches possible.

Ultimately, AI has opportunity to go beyond simply assisting or augmenting present CI capabilities. It will be used to explore new CI capabilities such as developing high-level scientific problem specification constructs that can discover and evaluate relevant resources (data sources, instruments, computation) and services (algorithms, models), testing promising workflows from such components, and providing end users with prototype workflows to explore a specified problem.

How might AI approaches be used to address CI challenges? First, AI can reduce the complexity of designing and operating CI. A single and/or small system may consist of millions of interacting components, whereas a larger, national-class system may contain billions of components. Similarly, the amount of code used to implement the systems ranges from millions to billions of lines. By definition, CI is highly interconnected and systems are co-dependent. This complexity makes systems that afford security and sustainability increasingly difficult to correctly design, build, and operate. Discovering services and optimizing their interaction in workflows in CI will require new, more scalable approaches than in the past.

Second, AI can reduce the delay in responding to conditions of interest. Almost all CI systems have significant limitations for data movement. To improve efficiency, workflows and overall system performance, mismatches responses to changes in conditions and delays in mitigation generation must be identified and alleviated. With the increased power of edge devices and the heterogeneity of network capabilities, emerging CI systems require complex orchestration of computation and data movement to reduce dependence on communications (e.g., reducing data volume) as well as to enable CI to act in real time despite having components that are tens of milliseconds or further apart. Concurrently, the use of embedded, wearable, or other personal devices expands the concept of CI.

Third, AI can reduce "Time-to-discovery" and "Time-to-actionable insights" in CI operations. Data sizes and complexities limit human's ability to discover anomalies, inefficiencies, bottlenecks, faults, security vulnerabilities, and cyber-attacks on time-scales appropriate for taking corrective actions. AI methodologies can efficiently extract such insights from large, complex data. However, as CI workflows and services become more complex it will be essential to improve our ability to create systems that behave in a way that can be explained to ensure confidence in taking action upon AI-based results.

Specifically, CI that is used across applications that require resources ranging from the farthest edge devices to the largest of facilities, from dedicated ultra-scale instruments to the broad range of distributed devices, from stationary to mobile and autonomous devices and systems. Brief review of a few exemplar use cases help explain the challenges and objectives of AI for CI, adding new use cases as more insight is gained. These exemplars include (i) mid- and large-scale data and computational facilities, (ii) CI for large scale observing instruments, (iii) CI for automated and robotic devices, and (iv) CI for widely distributed instruments whether it is in cities or national scale. Social science considerations for CI are shared among these cases. Each exemplars use has unique challenges but it is also possible to identify commonality that can be applied to all or are easily adapted for different CI implementations. We call these exemplars "Families of CI". These similarities will be conducive to developing common algorithmic approaches, interfaces, and protocols which will likely also extend to many other areas including those necessary in complex industrial scenarios. These families represent opportunities for AI for CI but are not exhaustive, and indeed are expected to be expanded.

Research Themes for Exemplar Families of Scientific CI: *Computational and Data Systems*: CI for computational and data analysis systems and facilities come in all sizes from a modest handful of compute nodes with a Network File System (NFS) to support common storage to complex High Performance Compute (HPC) systems consisting of tens of thousands of compute nodes and tightly integrated subsystems for supporting large scale data buffering and file I/O, to global scale cloud computing facilities. Across the spectrum of size, configuration and workload, C&D CI systems share the challenge of being used inefficiently. Application developers struggle determining the best algorithms to optimally utilize the increasing diversity of architectural features. These challenges are only getting worse with the ever-increasing complexity of all compute, network and data technologies. Across all such facilities this not only translates into waste in terms of energy, idle resources, and human capital, but also restricts the rate and quality of insight and output and extends the time to solution of some of humanity's most pressing problems.

In order to optimize the utilization of our compute and data related resources and minimize the time to solution of the problems being run on these resources, the technology industry and consumers have been adding CI in the form of instrumentation, communication, and analysis infrastructure for collecting and analyzing detailed information about all aspects of operation. To date this additional CI increases the burden on support staff and users of the technology to understand the complex interactions represented by the information and has been used largely for diagnosing problems reported and performance bottlenecks. Due to the high volume (tens of TB/day per system) and high dimensionality (thousands of different instrumentation points) of data, humans are only able to use a small subset of the data for analysis and are often unable to identify root causes of problems. Additionally, delays between problem occurrence and diagnosis for all but the most trivial problems limit the utility of current CI systems with respect to automated feedback and optimization of run time resources and configuration [10].

Research Themes for Exemplar Families of Scientific CI: *CI for Large Instruments:* The CI demands for next generation instruments, such as the Large Synoptic Survey Telescope (LSST) [3], are expected to exceed existing CI capabilities [11]. Upgrades to existing facilities, e.g., the advanced Laser Interferometer Gravitational-wave Observatory (LIGO) [12], and the next generation Large Hadron Collider [13] already demand the use of CI that are beyond in-house, tailored solutions for core-data analyses, requiring the use of the Open Science Grid, and other HPC platforms such as XSEDE and Blue Waters [14] [15]. There is an urgent need to rethink the existing CI paradigm to cope with the volume and speed of data production and analysis of large-scale facilities.

The convergence of AI with cloud, mid- and top-tier HPC centers has rapidly emerged as an alternative to handle the disruptive effects of the big-data revolution in strategic NSF- and DOE-funded investments. While the design and training of AI algorithms requires novel cyber-tools, such as distributed training schemes in HPC platforms, once these algorithms are fully trained, they enable real-time inference studies using minimal resources. Convergence between existing CI and AI-HPC inspired CI solutions, with expanded AI for CI technology, could boost and enrich the scientific capabilities of current and next-generation big-data experiments [16]. This approach will leverage the expertise of system architects, resource managers, and software developers to use AI to automate diagnostics tools that optimize the performance and throughput of CI resources. AI for CI will also carry out data-driven analyses, informed by the high-dimensionality of the in-situ system and application telemetry data, to form logical hypotheses about the data features essential to understanding relationships among architectures, behavioral characteristics of applications, and performance, even in the absence of labelled associated application runs and conditions, alleviating the need for relying on limited benchmarks and regression testing that is common practice today.

Research Themes for Exemplar Families of Scientific CI: *CI for Automation and Robotics*: AI based cyber-physical autonomous agents that directly assist or interact with humans are becoming ubiquitous [17] [18] [19] [20]. However, field-deployment of trustworthy, safe, and performant autonomous agents has the following two CI interrelated challenges. First, the lack of automated processes for design and development of an end-to-end field-deployable autonomous agent CI [21] [22]. Addressing the system challenges in design and development is critically important as evident from the delays in deployment of self-driving cars despite the showcase of a working concept of a functional AV in DARPA 2007 Urban Challenge [23] 13 years ago. Towards this goal, new and innovative automated processes must be created for operating systems and computational kernels that meet energy, latency and deadline requirements, and methods/protocols for testing, failure mitigation, recovery, and safety. Moreover, these problems are increasingly more challenging in the context of autonomous swarms that are geared towards execution of shared tasks (e.g., search and rescue missions) as they require methods for remote management (from a base station) and coordination of geo-distributed agents (edge devices) in uncertain human-centric environments. The second challenge is the lack of data, code, and knowledge sharing across the field of automation and robotics [24] [25]. There is a significant commonality across autonomous agents in the field of automation and robotics, such as self-driving vehicles, manufacturing/industrial/surgical robots, and delivery bots. This commonality spans CI requirements and specifications (e.g., remote management and coordination), meeting real-time deadlines for executing actions, and maintaining safety. Although agents operate at different levels of autonomy and perform significantly different tasks (e.g., human transportation vs. package delivery), the significant commonality among them can be leveraged by investment in standards to create reusable design, development, and deployment patterns to maximize safety, efficiency, and benefit to society.

In order to test ML and AI algorithms for perception, decision, and control of autonomous agents, industry has developed several open-source comprehensive testbeds [26] and [27] that utilize real-time physics-based simulators to support software-in-the-loop, hardware-in-the-loop, and human-in-the-loop methods. This allows developers to iterate on architecture, design and implementation of the autonomous CI system and generate large amounts of data on execution time of computational tasks, their safety, and resilience.

Unique AI opportunities for automation/robotics: AI approaches are promising as the current methods for design and development of safe CI for robotics and automation are fragile as it depends on painstakingly tuned heuristics and partial solutions with insufficient safety assurance. AI-techniques can leverage simulation and field datasets as well as domain knowledge to enable automatic design space exploration and development of CI components that are performant, safe and resilient.

Research Themes for Exemplar Families of Scientific CI: *Edge Systems and Distributed Intelligent Observatories*: NSF has supported the creation of distributed observatories ranging in scale from urban (e.g., Array of Things (AoT) in Chicago [28]) to regional (e.g., WIFIRE in Southern California [29]) to continental (e.g, National Ecological Observatory Network (NEON) [30]). At each of these scales there is need for measurements outside of the capabilities of traditional sensors, aimed at sensing factors ranging from vehicle or pedestrian flows in cities or the migratory patterns of wildlife. Desired capabilities include the ability to modify sampling rates to suit current conditions (e.g., low sampling rates when nothing out of the ordinary is going on and higher sampling rates surrounding events or conditions of interest that resolve desired detail) and the ability to aggregate and anonymize data to levels appropriate to types of information being collected (e.g., anonymous pedestrian and traffic flows vs. identification in the case of collisions and altercations).

Industry has begun a transition to monitoring and control using Internet-of-Things (IoT) technologies and is rapidly adopting edge computation to embed AI algorithms. Whether in factories, electrical and gas distribution networks, or urban measurement instruments, edge computation supports the creation of “software-defined sensors” that, like software-defined networks, are remotely programmed to allow flexibility and evolution of their measurement capabilities.

Unique AI opportunities for edge systems include: AI-at-the-edge can provide support for multiple types of CI needs, for instance enabling instruments to enhance communication conserving low-frequency samples with anomaly detection to trigger changes in sampling rates appropriate to resolve events or conditions of interest. For example, cameras deployed in cities need not stream and save (with significant privacy implications) video to central servers if their embedded intelligence can both extract routine measurements (e.g., pedestrian flows) and adapt to events of interest (such as a vehicle collision). As new CI capabilities move from prototypes to production scale, CI services in observatories such as NEON [31], AI will be increasingly critical to the operation and optimization of distributed components with high degrees of heterogeneity, diverse services, and in some cases intermittent connectivity.

Social Science Influence on CI: The social, economic, and behavioral science communities have significant experience working with sensitive data, maybe more than do the other NSF-supported sciences and have grappled with the ethics of the use of “big data” [32]. The advent of smartphones, wearables, in-home devices, and location-based services has brought CI squarely into social science research. The pace and scale of the use of such CI in the social sciences are such that the privacy, security, and ethics of the CI, and moreover of the use of AI within these devices, require a central rather than supporting role for social science in the use of AI for CI.

Equally important is the impact of AI for CI from the point of view of stakeholder trust--whether the correctness of results, the security of systems embedded in the public way and in homes, or the ethical application of AI in the use of associated data. Without trust, there will be no adoption. Initiatives must address increasing autonomy in order to reduce latency for making high-quality decisions in high-consequence and/or complex scenarios with the goal of increased performance, efficiency, security and safety. CI is a place where domain scientists make high consequence decisions and getting those decision-makers and their stakeholders to trust AI is going to be difficult unless its explainable and transparent in some way. In particular, there is a need to explore the partnership and trust between the AI-decision making tools and the humans for high risk tasks.

AI for CI research areas include low- through high-risk scenarios. Low-risk scenarios without significant incoming bias can provide a good testbed for testing acceptance and assessing confidence in results. For example, HPC scheduling decisions to improve performance can be low-risk as long as the affected applications are not high-consequence. Assessing criteria for tradeoffs in explainability-transparency-acceptance-latency change higher risk scenarios (e.g., more urgent computations, autonomous vehicles) will be critical to acceptance of the new methods.

Common Challenge Themes: The research challenges identified across the exemplar CI families all share some common themes: 1) CI systems are so complex that only domain experts can make reasonable and repeatable optimization decisions and the complexity is growing exponentially, 2) decision making should be performed with minimal lag time for optimum performance but the consequences of those decisions, in terms of human welfare and/or monetary cost, can be high, 3) processing telemetry data is a “big data” problem in itself with requirements for low latency time to solution in the presence of perhaps incomplete data, 4) individual data values are bounded but the high dimensionality and variability in aggregate makes full ensemble processing, within effective latency bounds, currently impossible, 5) social aspects of trusting the judgment of a machine must be overcome for significant deployment, 6) workload and resource scheduling, swarm coordination/management, resiliency, energy efficiency, ease of use by developers and end user, and 6) lack of labeled datasets as well as the high-cost of generating labeled datasets for training.

Across CI families, AI approaches will enable domain experts to automate the design space exploration of performant, safe and reliable CI, and unlock new use across the domains of science, health, agriculture, manufacturing, energy, and transportation by combining field datasets, simulation techniques and domain knowledge. We believe that these similarities will be conducive to the development of common algorithmic approaches (including explainability), interfaces, and protocols which will likely also extend to areas such as weather forecasting using drones/aircraft and more.

Conclusion: A National AI-for-CI effort that organizes a community of both CI and AI experts and stakeholders to focus on identifying the challenges and opportunities for AI to enhance and eventually re-invent the CI is a critical national need. Equally important is AI for CI efforts including leaders from some of the largest and most complex of NSF's CI projects and operating observatories and facilities to bring a diverse set of stakeholders and experts together from both a disciplinary standpoint and a human standpoint. The ubiquitous and foundational changes to be empowered by applying AI to CI require new perspectives from new communities that challenge traditional thinking embedded in nearly four decades of NSF-supported CI. Simply put, only such a bold approach can yield insights and new concepts that are required to fully exploit AI in achieving the promise of increasingly broad, powerful, and intelligent CI in the coming decade

REFERENCES CITED

[1] BDEC2, "Big Data and Extreme-Scale Computing 2," 2019. [Online]. Available: https://www.exascale.org/bdec/meeting/sandiego. [Accessed 27 January 2020].

[2] G. H. Bauer, B. Bode, J. Enos, W. T. Kramer, S. Lathrop, C. L. Mendes and R. R. Sisneros, "Best Practices and Lessons from Deploying and Operating a Sustained-Petascale System: The Blue Waters Experience," in *Proceedings of Supercomputing'18*, Dallas, TX, 2018.

[3] LSST, "The Large Synoptic Survey Telescope," [Online]. Available: https://www.lsst.org/lsst. [Accessed 27 January 2020].

[4] MBDH, "Midwest Big Data Hub," [Online]. Available: http://midwestbigdatahub.org/. [Accessed 27 January 2020].

[5] DOE, "AI for Science Town Hall Meetings," 2019. [Online]. Available: https://www.ornl.gov/content/town-hall-artificial-intelligence. [Accessed 27 January 2020].

[6] G. Fox, "Perspectives on High-Performance Computing in a Big Data World," in *Proceedings of the 28th International Symposium on High-Performance Parallel and Distributed Computing*, Phoenix, AZ, 2019.

[7] A. Brown and D. A. Patterson, "To Err is Human.," in *Proceedings of the First Workshop on Evaluating and Architecting System DependabilitY (EASY '01),*, Gothenburg, Sweden, July 2001..

[8] A. Brown, L. C. Chung and D. A. Patterson, "Including the Human Factor in Dependability Benchmarks," in *Proceedings of the DSN Workshop on Dependability Benchmarking*, Washington, DC, 2002.

[9] "Digital Twin," 2 January 2020. [Online]. Available: https://en.wikipedia.org/wiki/Digital_twin.

[10] S. Jha, J. Brandt, A. Gentile, Z. Kalbarczyk, G. Bauer, J. Enos, M. Showerman, L. Kaplan, B. Bode, A. Greiner, A. Bonnie, M. Mason, R. K. Iyer and W. Kramer, "Holistic Measurement-Driven System Assessment," in *Proceedings of the IEEE International Conference on Cluster Computing (CLUSTER)*, Honolulu, HI, 2017.

[11] NAS, "Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020," National Academies of Sciences, Engineering, and Medicine, Washington, DC, 2016.

[12] LIGO, "Laser Interferometer Gravitational-wave Observatory," [Online]. Available: https://www.ligo.caltech.edu/page/what-is-ligo. [Accessed 27 January 2020].

[13] D. Castelvecchi, "Next-generation LHC: CERN lays out plans for €21-billion supercollider," *Nature,* Vols. 565, 410 (2019), 15 January 2019.

[14] E. A. Huerta, R. Haas, S. Jha, M. Neubauer and D. S. Katz, "Supporting High-Performance and High-Throughput Computing for Experimental Science," *Computing and Software for Big Science,* vol. 3, no. 5, 2019.

[15] E. A. Huerta, R. Haas, E. Fajardo, D. A. Katz, S. Anderson, P. Couvares, J. Willis, T. Bouvet, J. Enos, W. T. Kramer, H. W. Leong and D. Wheeler, "BOSS-LDG: A novel computational framework that brings together Blue Waters, Open Science Grid, Shifter and the LIGO Data Grid to accelerate gravitational wave discovery," in *Proceedings of the 13th IEEE International Conference on e-Science*, Auckland, 2017.

[16] E. Papka, I. Monga, J. J. Hack and S. Jones, "Facilities Integration and AI Ecosystem," [Online]. Available:

https://www.csm.ornl.gov/documents/downloads/ai/Facilities%20Integration%20and%20AI%20Ecosystem/FacilitiesIntegrationORNLOutbriefV2.pptx. [Accessed 27 January 2020].

[17] F. Codevilla, M. Müller, A. López, V. Koltun and A. Dosovitskiy, "End-to-End Driving Via Conditional Imitation Learning," in *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, 2018.

[18] A. Hussein, M. M. Gaber, E. Elyan and C. Jayne, "Imitation Learning: A Survey of Learning Methods," *ACM Computing Surveys (CSUR),* vol. 50, no. 2, pp. 1-35, 2017.

[19] ResearchAndMarkets, "Global Agriculture Drones and Robots Markets, 2018-2028 - Advancements in Artificial Intelligence Technology and Machine Learning Solutions," 2019. [Online]. Available: https://www.globenewswire.com/news-release/2019/05/10/1821686/0/en/Global-Agriculture-Drones-and-Robots-Markets-2018-2028-Advancements-in-Artificial-Intelligence-Technology-and-Machine-Learning-Solutions.html. [Accessed 27 January 2020].

[20] J. Edward, "Building a Smart Factory with AI and Robotics," 2018. [Online]. Available: https://www.roboticsbusinessreview.com/wp-content/uploads/2018/02/RBR_BuildingAI_WP3.pdf. [Accessed 27 January 2020].

[21] L. E. Lwakatare, A. Raj, J. Bosch, H. H. Olsson and I. Crnkovic, "A Taxonomy of Software Engineering Challenges for Machine Learning Systems: An Empirical Investigation," in *Proceedings of the International Conference on Agile Software Development*, Montreal, 2019.

[22] B. Hrvoje, M. Vukovic and C. Željka, "Requirements Engineering Challenges in Building AI-Based Complex Systems," in *Proceedings of the 27th IEEE International Requirements Engineering Conference Workshops (RE'19)*, Jeju Island, South Korea, 2019.

[23] DARPA, "DARPA 2007 Urban Challenge," 2007. [Online]. Available: https://www.darpa.mil/about-us/timeline/darpa-urban-challenge. [Accessed 27 January 2020].

[24] F. L. DaSilva and A. H. Costa, "A survey on transfer learning for multiagent reinforcement learning systems," *Journal of Artificial Intelligence Research,* vol. 64, pp. 645-703, 2019.

[25] C. Longbing, G. Weiss and S. Y. Philip, "A brief introduction to agent mining," *Autonomous Agents and Multi-Agent Systems,* vol. 25, no. 3, pp. 419-424, 2012.

[26] AirSim, "Microsoft/AirSim," Microsoft, [Online]. Available: https://github.com/microsoft/AirSim. [Accessed 27 January 2020].

[27] CARLA, "Open-source simulator for autonomous driving research," [Online]. Available: http://carla.org. [Accessed 27 January 2020].

[28] C. E. Catlett, P. H. Beckman, R. Sankarn and K. K. Galvin, "Array of things: a scientific research instrument in the public way: platform design and early lessons learned," in *Proceedings of the 2nd International Workshop on Science of Smart City Operations and Platforms Engineering (SCOPE'17)*, 2017.

[29] I. Altintas, L. Smarr, H. W. Braun and R. D. Callafon, " Hazards SEES Type 2: WIFIRE: A Scalable Data-Driven Monitoring, Dynamic Prediction and Resilience Cyberinfrastructure for Wildfires," NSF Award 1331615, 2013. [Online]. Available: https://www.nsf.gov/awardsearch/showAward?AWD_ID=1331615. [Accessed 27 January 2020].

[30] D. Schimel, W. Hargrove, F. Hoffman and J. MacMahon, "NEON: A hierarchically designed national ecological network," *Frontiers in Ecology and the Environment,* vol. 5, no. 2, pp. 59-59, 2007.

[31] NEON, "The National Ecological Observatory Network," [Online]. Available: https://www.neonscience.org. [Accessed 27 January 2020].

[32] P. Elias, P. H. Fossheim, P. Muhlberger, C. Catlett, K. A. Cagney, R. George, G. v. Halderen, T. Ausin, K. Agni, S. Maurizio, S. Hagit, P. Buleon, S. Bender, M. Martin, I. Satoh, S. Osaumu, M. Flesaoana, C. v. Zyl, M. Woolard, S. McGregor and C. Humphrey, "Research ethics and new forms of data for social and economic research.," *OECD,* no. 34, 2016.

[33] P. Beckman, C. Catlett, I. Altintas, S. Collis and E. Kelly, "Mid-Scale RI-1: SAGE: A Software-Defined Sensor Network," NSF Award 1935984, 2019. [Online]. Available: https://www.nsf.gov/awardsearch/showAward?AWD_ID=1935984&HistoricalAwards=false. [Accessed 27 January 2020].

[34] DES, "The Dark Energy Syrvey," [Online]. Available: https://www.darkenergysurvey.org. [Accessed 27 January 2020].

[35] NCSA, "Center for Artificial Intelligence Innovation," University of Illinois, [Online]. Available: http://www.ncsa.illinois.edu/about/centers/ai/. [Accessed 27 January 2020].

[36] N. Stout, "Supercomputers will start building a 3D map of the world," 5 August 2019.

[37] S. Jha, S. S. Banerjee, T. Tsai, S. K. Hari, M. Sullivan, Z. Kalbarczyk, S. Kecler and R. K. Yyer, "ML-based Fault Injection for Autonomous Vehicles: A Case for Bayesian Fault Injection," in *Proceedings of the 49th IEEE/IFIP International Conference on Dependable Systems and Networks*, Portland, 2019.

[38] R. K. Iyer, A. P. Athreya, L. Wang and R. M. Weinshilboum, "Artificial Intelligence and Pharmacogenomics: A Timely Synergy for Individualizing Medicine," *Advances in Molecular Pathology,* vol. 2, no. 1, pp. 111-118, 2019.

[39] A. P. Athreya, A. J. Gaglio, J. Cairns, K. R. Kalari, R. M. Weinshilboum, L. Wang, Z. T. Kalbarczyk and R. K. Iyer, "Machine Learning Helps Identify New Drug Mechanisms in Triple-Negative Breast Cancer," *IEEE Transactions on NanoBioscience,* vol. 17, no. 3, pp. 111-118, 2018.

[40] Y. Varatharajah, B. Berry, J. Cimbalnik, V. Kremen, J. V. Gompel, M. Stead, B. Brinkmann, R. Iyer and G. Worrell, "Integrating artificial intelligence with real-time intracranial EEG monitoring to automate interictal identification of seizure onset zones in focal epilepsy," *Journal of Neural Engineering,* vol. 15, no. 4, 2018.

[41] S. S. Banerjee, R. Jha and R. K. Iyer, "Inductive Bias-driven Reinforcement Learning For Efficient Schedules in Heterogeneous Clusters," 2019. [Online]. Available: https://arxiv.org/abs/1909.02119. [Accessed 27 January 2020].

[42] S. Jha, S. Cui, T. Xu, J. Enos, M. Showerman, M. Dalton, Z. T. Kalbarczyk, W. T. Kramer and R. K. Iyer, "Live Forensics for Distributed Storage Systems," 2019. [Online]. Available: https://arxiv.org/abs/1907.10203. [Accessed 27 January 2020].

[43] V. Satone, R. Kaur, H. Leonard, H. Iwaki, L. Sargent, S. Scholz, M. A. Nalls, A. B. Singleton, F. F. Faghri and R. H. Campbell, "Predicting Alzheimer's disease progression trajectory and clinical subtypes using machine learning," 2019. [Online]. Available: https://doi.org/10.1101/792432. [Accessed 27 January 2020].

[44] S. Hashemi, S. Yothi and R. Campbell, "TicTac: Accelerating Distributed Deep Learning with Communication Scheduling," in *Proceedings of the SysML Conference*, Stanford, CA, 2019.

[45] A. Agelastos, B. Allan, J. Brandt, P. Cassella, J. Enos, J. Fullop, A. Gentile, S. Monk, N. Naksinehaboon, J. Ogden, M. Raian, M. Showerman, J. Stevenson, N. Taerat and T. Tucker, "Lightweight Distributed Metric Service: A Scalable Infrastructure for Continuous Monitoring of Large Scale Computing Systems and Applications," in *Proceedings of Supercomputing'14*, New Orleans, 2014.

[46] O. Tuncer, E. Ater, Y. Zhang, A. Turk, J. Brandt, V. J. Leung, M. Egele and A. K. Coskun, "Diagnosing Performance Variations in HPC Applications Using Machine Learning," in *Lecture Notes in Computer Science - High Performance Computing ISC*, Springer, Cham, 2017, pp. 355-373.

[47] CUCIS, "Center for Ultra-scale Computing and Information Security," Northwestern University, [Online]. Available: http://cucis.ece.northwestern.edu/. [Accessed 27 January 2020].

[48] "JELSC Overview," 29 January 2020. [Online]. Available: https://jlesc.github.io/.

[49] GE, "Digital Twin Framework," General Electric, [Online]. Available: https://www.ge.com/research/project/digital-twin-framework. [Accessed 27 January 2020].