SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS

3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
Phone +1-215-382-9800
Fax +1-215-386-7999
www.siam.org · siam@siam.org

The Society for Industrial and Applied Mathematics (SIAM) is an international community of 14,000 individual members. Almost 500 academic, manufacturing, research and development, service and consulting organizations, government, and military organizations worldwide are institutional members. SIAM was incorporated in 1952 as a nonprofit organization to convey useful mathematical knowledge to other professionals who could implement mathematical theory for practical, industrial, or scientific use. Through publications, research, and community, the mission of SIAM is to build cooperation between mathematics and the worlds of science and technology. Our response to the Request for Information (RFI) on Frontiers in AI for Science, Security, and Technology (FASST) Initiative is below.

1. Data

Investment is needed in research and development to create an integrated suite of data lifecycle methods and tools, informed by the specific needs of DOE scientific communities. Techniques that lead to data harmonization and normalization, such as data cleaning, validation, and error correction must be developed to overcome the natural heterogeneity of data sources which is in inherent conflict with the need for data fusion. Open research questions remain around analyzing highly distributed data sources, enabling data discovery and integration, tracking data provenance, coping with sampling biases and heterogeneity, ensuring data integrity, privacy, security, and sharing, and visualizing massive datasets.

Many scientific areas, ranging from materials science to astronomy to biology, have very few datasets with labeled training data. Many current AI algorithms, for example, deep neural networks, cannot perform well when trained on these limited samples. New developments in foundational AI, such as few-shot learning, self-supervised learning, mixed-scale dense networks, and meta-learning techniques, need to be developed to enable important scientific AI applications

Computational science has always been an interdisciplinary endeavor, drawing on mathematics, computer science, and application domains. With the growth in the importance of data science and AI the field is becoming even broader. Interdisciplinary education and research has long been challenging due to the departmental structure of universities. But recent years have seen a steady growth in cross-campus programs that facilitate interdisciplinary education, collaboration, and research. In addition, the DOE National Laboratories, academia and industry all have specific strengths to benefit AI research and development. Many of the most pressing scientific and societal challenges can only be solved through the efforts of multidisciplinary, multi-institutional teams. The federal government should create incentives for interdisciplinary education and research collaboration that engage National Laboratories, academia, and industry.

Investments should be made in fundamental research to develop AI technologies that can advance DOE science. This could include efforts through existing programs and new focused opportunities that enable research collaborations among applied mathematicians, computer scientists, and computational scientists. Possible examples include: • Effective ML training for small or sparse datasets; • Incorporation of physical models into AI structure; • Verification, validation, and uncertainty quantification; • Real-time decision making; • AI methods for management of computational resources and workflows; • AI

3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
Phone +1-215-382-9800
Fax +1-215-386-7999
www.siam.org · siam@siam.org

methods for planning, monitoring and running experiments; • Interpretable models and algorithms; • Hypothesis generation from data.

2. Compute

A facility or facilities should be established to ensure that researchers have access to experimental non-traditional hardware that will inform a holistic codesign cycle between manufacturers and users of transformative computing technologies. As the non-traditional computing hardware industry evolves and the roles of different enterprises change, the government should remain open to establishing new types of collaborations and relationships with both existing and emerging industrial partners. These new devices will likely be initially employed as accelerators, so their ability to integrate with existing HPC platforms must be part of the design process.

Investments are needed in research and development collaborations between computational scientists and computer vendors to ensure development of future energy efficient compute platforms that meet the needs of the computational science community. The high performance computing market is dominated by AI and Cloud, which create challenges and opportunities for scientific computing and applied math. In particular, our reliance on double-precision arithmetic is no longer sustainable and requires new mixed-precision algorithms to take advantage of new hardware and to ensure computational results remain reliable and predictive in the presence of lower precision from energy-efficient hardware. To address these challenges, investments are needed into algorithms and methodologies. In addition, incentives are needed to encourage vendors to consider the needs of scientific application beyond their own markets. In this emerging environment, DOE should continue to leverage its successful history of investments in holistic co-design collaborations aimed at advancing the scientific computing market.

An explicit program should be established to leverage and continue to advance the community software created by the Exascale Computing Project (ECP). This body of software developed under ECP will underpin the community for many years to come and will provide a cost-effective way to develop, support, port, and maintain a host of advanced applications. But this software will not be able to address continually evolving needs without further investment. DOE should oversee a research and development program to continue to extend this functionality, adapt it to new computers, and apply it to new application areas.

3. Models

Investments are needed in the research, software tools, and system management tools needed to enable complex workflows that combine simulation with machine learning. The computational science community should leverage recent progress in AI and ML in industry and identify AI and ML research directions that will enable new and faster scientific and engineering capabilities. Many existing applications involve the use of ML models as very fast "surrogates" (approximate replacements) for more expensive simulation codes.

Incorporation of physical models into AI structure. The current theory and practice of AI struggles to incorporate physical models and constraints. Formulations of new AI algorithms structured to include

SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS

3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
Phone +1-215-382-9800
Fax +1-215-386-7999
www.siam.org · siam@siam.org

physical models, rather than having to "learn" the physics, will allow the designs of new classes of AI applications optimized exactly for the scientific tasks that require them. To do this will require advances in such formalisms as projection operators that enforce physical principles, data structures that encode symmetries and constraints, and the construction of physical priors and their injection into mathematical models.

4. Applications

In some cases, AI and ML capabilities developed by industry can be applied to science, but in many cases, significant research will be required to make these capabilities usable in the scientific domain. In particular, DOE should focus on the unique needs associated with scientific applications and with the needs of DOE's broader missions. These investments must both leverage current exascale and edge computing technologies and be able to take advantage of future computer architectures as they are developed.

Many scientists are exploring ideas and developing new insights, but the field of scientific machine learning is very young. Many existing applications involve the use of ML models as very fast "surrogates" (approximate replacements) for more expensive simulation codes. The speed of these ML surrogates allows for more instances to be run, and for a larger search space to be explored which can lead to more optimized designs or better characterizations of uncertainties. Alphafold13, Google's breakthrough technology for predicting protein conformations, showcased the potential for ML to outperform traditional scientific approaches. Other potential roles for ML in science include automatically monitoring and running experiments or series of experiments, improved methods for combining simulation results and experimental data, and the generation of new hypotheses from data. Quite likely, the most important ideas have not yet been thought of. AI may also dramatically change the software development process, which might make it much easier to build new computational science tools and applications.

Investment is needed to define and develop a common software stack for data science edge computing resources at scientific user facilities. This will include development and testing of new paradigms for integrating real-time information from sensors with edge computation, enabling realtime predictive analytics, control, and optimization in support of operation of complex infrastructure at DOE scientific user facilities. The software infrastructure should support all aspects of scientific user facility operations, from generic operation and analysis services to specialized software applications specific to each facility.

Many facilities in the DOE portfolio and beyond generate vast quantities of heterogeneous data that must be analyzed in real-time. Further work is required to achieve sub-microsecond decision times for applications such as particle physics, real-time particle accelerator or fusion reactor control, or the kinds of processing required for applications such as radio astronomy. Real-time decision making can also be used to optimize the use of scientific experiments and facilities and to manage key systems like the power grid.

Many AI models focus on prediction accuracy without a focus on explainability. For scientific applications, prediction without insight is seldom enough. A key area of exploration is the development

3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
Phone +1-215-382-9800
Fax +1-215-386-7999
www.siam.org · siam@siam.org

of interpretable models and algorithms that provide insight on why and how a model is producing a particular output. This requires investment into sparsification techniques for explainable AI models.

Artificial intelligence and machine learning have enormous potential to impact the processes of scientific research, but broad investments in mathematics and computing will be required to realize these opportunities.

Investments should be made to develop partnerships between computational and applications scientists to customize and apply AI capabilities to new science areas. As with all areas of science, active collaboration between science domain experts and computational experts will ensure that the best possible computational capabilities can be developed and brought to bear on challenging scientific problems. The SciDAC program provides an attractive model for this kind of investment.

5. Workforce

Creating new workforce pipelines that expose students to scientific computing earlier in their careers is vital to expand the pool of practitioners. DOE's Workforce Development for Teachers and Scientists (WDTS) has been successful at collaborating with the national laboratories to engage with educators and undergraduate, graduate, and postdoctoral students. However, with the changing landscape in scientific computing, more investment is needed to address the growing workforce issues at all levels. Exposure to DOE has proven to be an effective way to inspire students at all levels to pursue careers in the field and ensure they the requisite skills. The national labs provide these opportunities but are a finite resource; additional programs at educational institutions to attract and train the next generation of the CSE workforce will help to alleviate the current shortages.

DOE should invest in programs at diverse educational institutions to increase the reach of workforce development programs. One model DOE could follow is the National Science Foundation Research Experiences for Undergraduates, which funds students to work on research projects at the host institution. DOE should also consider additional K–12 engagement to ensure students are aware of and interested in CSE fields as they make college decisions. To support the workforce needs of the future, students need to be exposed early and often to DOE, and partnering with other institutions will multiply the impact of this effort. Over the past decades, workforce development programs have been created that aim to build a skilled CSE workforce. A principal example is the DOE Computational Science Graduate Fellowship16, which has enabled students to pursue a multidisciplinary program of education in CSE coupled with practical experiences at the DOE laboratories. Graduates of this program have pursued professional careers in industry, academia, and at federal laboratories. More recently, the ECP Broadening Participation Initiative 17, including the Sustainable Research Pathways program18, 19, has successfully brought hundreds of faculty and students from underrepresented communities to the DOE national laboratories for summer research experiences and HPC training. Continuation and expansion of these programs, and the development of additional programs modeled after them, will be essential for building the diverse CSE workforce needed to address the challenges and take advantage of the opportunities discussed in this report.

The federal government should expand and broaden investments in workforce development in computational science. Areas of focus should include pathways for underrepresented communities,

3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
Phone +1-215-382-9800
Fax +1-215-386-7999
www.siam.org · siam@siam.org

retraining opportunities for existing workers, and academic programs with strong interdisciplinary elements.

DOE should continue to support, and look for opportunities to expand, the Computational Science Graduate Fellowship. The CSGF program has been integral in ensuring early career researchers have the skills and experiences they need to effectively contribute to DOE. As the future of scientific computing continues to evolve, the program should be supported in expanding to new areas, such as quantum computing and artificial intelligence for science.

Many of the most pressing scientific and societal challenges can only be solved through the efforts of multidisciplinary teams. The federal government should create incentives for interdisciplinary education and collaboration at both the graduate and undergraduate levels.

6. Governance

A facility or facilities should be established that will ensure that computational science researchers have access to emerging hardware, programming models, and heterogeneous systems to enable assessment of their utility for scientific applications. With the plethora of emerging computing designs, it is difficult to know which are the most promising, how they can be best exploited, and which application areas are best suited to each. A program to provide broad access to new technologies, and to assist new users, would allow the community to identify the most promising approaches. A similar program in the late 1980s and early 1990s was essential for the successful transition from vector computers to parallel computers.

Sincerely,

Sven Leyffer	Suzanne L. Weekes	Alejandro Aceves
President	Chief Executive Officer	Vice President for Science Policy