Federal Register Notice 86 FR 46278, https://www.federalregister.gov/documents/2021/08/18/2021-17737/request-for-information-rfi-on-an-implementation-plan-for-a-national-artificial-intelligence, October 1, 2021.

Request for Information (RFI) on an Implementation Plan for a National Artificial Intelligence Research Resource: Responses

A Development Plan for a National Artificial Intelligence Research Resource

Lawrence Livermore National Laboratory
Livermore, CA USA

Technical contacts:
Brian K. Spears
AI Center of Excellence Director

James Brase
Deputy Principal Associate Director for Computing
Chris Clouse
Weapons Simulation Codes Program Director

Robert M. Sharpe
Engineering Deputy Associate Director for Research and Development

Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear Security Administration under Contract DE-AC52-07NA27344.

Executive Summary

The central purposes of the National Artificial Intelligence Research Resource should to focus US AI research strategy and to expand participation in AI research to include the broadest range of researchers possible (roadmap item A). The NAIRR should approach this using four key strategies: (1) the development of a governing consortium drawn from public and private institutions; (2) the identification by this consortium of priority research directions to guide US AI strategy; (3) the production of a set of central products for the research resource comprising common data sets, accessible compute, and common software tools; and (4) the development of partnerships that extend research collaborations between heavily resourced institutions and historically underrepresented ones. (roadmap items C, D, E). The NAIRR should be owned by the Department of Energy owing to its experience with large-scale R&D, interdisciplinary expertise (including AI work and varied applications), and access to critical resources (roadmap item B). The leadership consortium should be led by Lawrence Livermore National Laboratory given its leadership and experience in HPC and AI for applied science (roadmap items B, I)

Need for the research resource (roadmap item A)

The United States needs a coherent, directed AI strategy that both sets priority research directions and democratizes access to the tools and resources needed to advance those research directions. Currently, US AI research advances across economic sectors and national security missions according to a wide set of decoupled goals and motivations. Commercial AI addresses current business and economic needs. Research organizations, especially national laboratories, academic institutions and some commercial and nonprofit organizations, address fundamental methods and national security needs. The efforts of these capable groups need to be further enhanced by concentrating, at least partially, on key national grand challenges – particular AI tasks, methods, and applications found central to a wide range of US competitiveness and security efforts. At the same time, it is critical for future AI research to span

subject matter areas. Key applications – whether in predictive biology, advanced manufacturing, autonomous systems, and more – demand collaboration and interaction among subject matter experts from a wide variety of fields. These include computer science, engineering, physics, biology, chemistry, and applied mathematics. An additional intersection is perhaps the most critical. This is the intersection of simulation with experiment and real-world observation. As an example, at Lawrence Livermore National Laboratory, we have a large strategic effort focused on using AI to merge detailed simulation with high-precision experimental observation. We call this Cognitive Simulation, or CogSim. The ultimate endpoint is to use the combination of simulation and experiment to improve predictive models and accelerate the design of new solutions. However, without acknowledging the coming confluence of data gathered from the edge and experimentation with large-scale simulation, efforts to advance AI methods will miss a large part of the real evolutionary pressure on AI tools. This further highlights the need for a central, organizing body drawn from across the AI R&D ecosystem to identify and set this direction, making sure that efforts draw from all economic sectors, all relevant scientific disciplines, and all sources of useful data (simulation and experiment/observation). The NAIRR should serve as a body to navigate these issues.

The United States also needs a fully engaged and well-trained research and work force to advance the priority research directions. Right now, the lack of access to top tier data sets and computing systems limits full engagement in AI research to large-scale corporations, national Laboratories, and some tier 1 academic research institutions. This excludes engagement by the large talent pool residing in other institutions at tier 1 , tier 2, and those focusing on traditionally underserved groups. The US needs the aforementioned NAIRR to build large-scale collaborations across the AI R&D ecosystem focused on developing open data sets, accessible computing resources, and a common public software stack, all of which address the priority research directions established by the NAIRR. The NAIRR needs to enable establishment of critical infrastructure for the dissemination of such data sets to all levels of the AI R&D ecosystem to drive maximal participation. It needs further to enable the establishment of infrastructure for the collation and tracking of subsequent research results, and the disbursement of codified knowledge and best practices across economic sectors and security missions at a variety of skill levels.

In summary, the US needs the NAIRR to build collaboration among the disparate parties in the US AI R&D ecosystem, to identify priority research directions, to develop democratized AI research resources, and to fully engage the totality available US research talent at all levels. The effect of this effort will be to develop a leading AI research capability that outpaces the world by recruiting the effort and abilities of researchers at comprehensive levels not currently available.

Approach to developing the research resource (specific answers to questions 1 and 2 of the RFI)
The NAIRR will dramatically accelerate US AI R&D effectiveness by making available a common research resource. This research resource will be built using a 4-pronged approach:

1. Develop a national consortium built of key constituencies from the US AI R&D ecosystem
2. Identify priority research directions
3. Build (horizontal) partnerships across heavily resourced institutions that produce data sets, compute resources, common software tools to serve as the central products of the research resource.
4. Build (vertical) partnerships and programs that utilize the central products of the research resource to fully recruit R&D talent at all levels for simultaneous advancement of efforts on PRD/grand challenges and development of the national AI workforce and talent reserve.

This approach will develop the central products of the research resource in alignment with national priorities while recruiting a far expanded workforce to drive progress in AI research (figure 1).

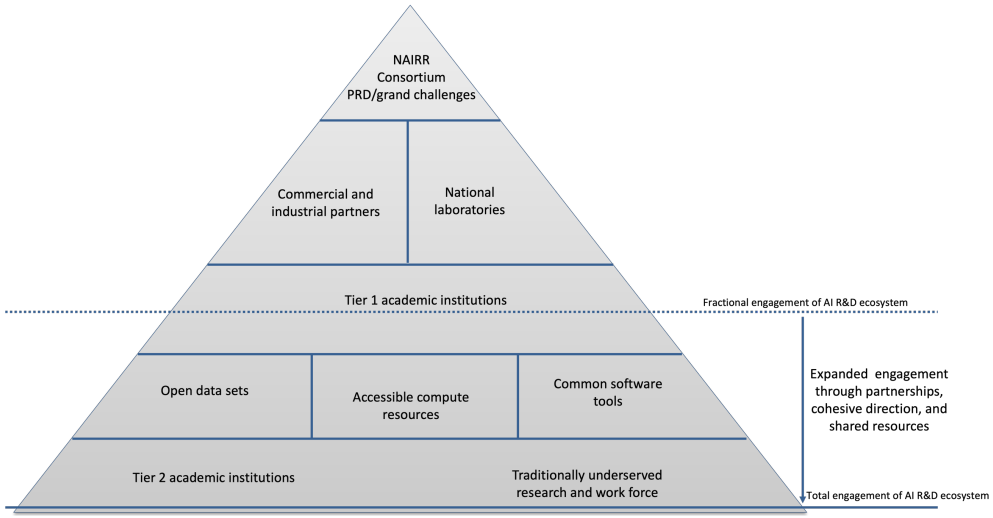

Figure 1: a model for directing US AI research, developing shared resources, and fully engaging and expanding research personnel. The NAIRR national consortium provides a cross-disciplinary leadership team guide the evolution and development of the research resource. The consortium identifies priority research directions that capture the dominant US AI needs. Partnerships among the heavily resourced (large companies, national labs, leading academic institutions) provide the central products for the research resource – open data sets, accessible compute resources, and common software tools. Partnerships across all levels use this democratized resource to engage traditionally underserved groups. These partnerships and the research resource combine to open access to AI research while accelerating progress on the priority research directions to advance US competitiveness.

<u>The NAIRR national consortium (partial answer to question 5 of the RFI)</u>

The NAIRR will develop a research consortium composed of key members of the US AI R&D ecosystem. This consortium will include commercial and industrial partners (technology, computer hardware, and computer software companies), national laboratories (national security mission and national resource stewards), and academic partners (drawn from across US academia with emphasis on traditionally underserved groups). The consortium will represent the key constituencies in US AI research. It will draw on this expertise to develop and direct the activities of the NAIRR. The central missions will be: (1) articulating the priority research directions, (2) developing partnerships aimed at producing the shared AI research resources, and (3) developing partnerships that utilize the shared resources, aimed specifically at bridging all levels in the US AI ecosystem with emphasis on historically underserved academic communities, and (4) managing the allocation of research resources.

The consortium should represent the broad range of disciplines and constituencies required for AI research. The disciplines should include expert leaders focused on fundamental AI methods and research as well as those focused on AI method development for applications. The expert leaders should further include those who operate at the intersection of AI development and application, where the applications span science, engineering, and commercial applications.

The expert leaders should be drawn from across the broad range of the US AI R&D ecosystem. This should include parties with substantial resources to contribute, specifically large companies and national laboratories. It must further include academic partners from a wide range of levels – tier 1 research universities, tier 2 schools with fewer resources, and regional schools serving smaller areas and those lacking access to AI resources (data, compute, sophisticated software).

Establishment of this cross-cutting leadership consortium will ensure that a comprehensive set of national interests are represented and that a wide range of resources can be recruited. It will further ensure that participation in the leadership of AI research is spread across all levels of the US AI research ecosystem and that the ecosystem itself is expanded to include those who often lack access.

The priority research directions

The NAIRR consortium will identify key priority research directions or grand challenges. These challenges must span economic sectors and national security missions. They must be curated so that their cross-cutting nature fully envelops the strategic needs of the US in order to maintain international AI competitiveness. They should be posed such that success on these challenges, in the time frames established by the consortium, will clearly lead to US AI global superiority.

The challenges must be interdisciplinary in nature in order to bring the largest field of researchers to the table. They must also drive work that combines computational work, e.g. simulation and modeling, with empirical data (experiment, observation, real-world production). This will drive AI methods that are developed and tailored for interaction and impact in the physical world across economic sectors.

Examples of such challenges might include:

1. Accelerated material discovery: The problem is to discover and optimize molecules to accelerate the development of new materials and medicines to boost economic competitiveness and national security. There has been substantial progress in the application of generative AI models to suggest new concepts with specified molecular properties for small molecules, peptides, and gene sequences, speeding the process of material/bio-material design and discovery. However, these capabilities do not yet exist for complex materials (e.g. polymers) at the same level as for molecular materials, and they do not extend to complex descriptions of materials such as chemical synthesis planning and real-world production. Concerted effort on rapid, AI-driven material discovery pathways will transform US industries that delivery medications, materials, materials synthesis and production methods and more.

2. Enhanced manufacturing competitiveness: The US manufacturing sector is constantly called to deliver a faster, more flexible methodology that will deliver high-quality or novel products in fractions of the current time required. Delivering on this call requires bridging the range of processes in the production chain – material discovery and selection, component and system design, and manufacturing and quality control. We propose a global, AI-driven digital twin framework that spans these processes to shorten delivery times by exploiting recent advances in AI, computing, and manufacturing methods. Research in this area will transform US manufacturing, using AI to produce a competitive advantage in international manufacturing competition that will benefit both commercial and national security operations.

These very brief examples highlight the essential ingredients for meaningful priority research directions. They must engage partners across economic sectors, they must offer wide-ranging benefit to US competitiveness or security, they must develop tools and techniques at the intersection of AI, science, and engineering, and they must operate at an inspirational scale that captures public imagination and future research engagement.

The establishment of these priority research directions by the broad perspectives of the guiding consortium will ensure that a complete set of US AI research needs are addressed by the NAIRR. It will further ensure that the focus and benefits of the resource extend not only across economic sectors but to those groups with previously limited access to AI research opportunities.

The resource production partnerships (partial answer to question 5 of the RFI, roadmap item D)

The NAIRR will further develop multipoint partnerships with the members of the US AI R&D ecosystem who are capable of producing essential resources for the solution of the grand challenges. These essential resources fall into three categories – shareable data sets, accessible compute resources, and common software tools. The production partnerships will feature collaboration between heavily resourced institutions, perhaps a large software company, a large hardware company, and a national laboratory. The collective capabilities of such partnerships will be aimed at making progress on the grand challenges, but equal emphasis will be given to producing **shareable data sets** that others with fewer resources (colleges, regional schools, small businesses) can use to grow their own efforts. The accessibility of such data sets will directly enable all members of the US AI research ecosystem to access the essential data ingredient that is currently unavailable outside of large businesses and national laboratories. The curation of these data sets and the research motivated by them will allow the NAIRR to intentionally focus AI research on areas of national need aligned with the priority research directions.

In similar fashion, production partnerships should be developed to deliver **common software tools**. This will involve developing an AI software ecosystem that spans topics – learning frameworks, workflow software, data tools, and system-level security and management utilities. The software will be interoperable and open source, allowing access by the full spectrum of US AI R&D members, as well as vendor partners, giving them a common target software stack to support. Such a common stack will also provide a uniform interface to AI research, making educational materials universally useful and lowering the barrier to entry for researchers across the AI enterprise. Standardization of interfaces and communication protocols would make it easier to integrate novel AI optimized hardware with more traditional compute resources, widening the aperture for both vendor and researcher participation in the investigation of new hardware approaches. Without common functional tools with a low barrier to initial use and entry, the common data sets and shared compute resources (described next) are functionally useless. Researchers, especially the underserved, will require enormous amounts of time to manage, build, install, and operate software on large computers using large data sets if not equipped with a well-designed and transparent software stack. Note that this will not be a great limiter for the heavily resourced who routinely develop their own tools or purchase others. Thus, a common software tool set will provide an essential key for the underserved to unlock access to other parts of the research resource.

Finally, the NAIRR should develop a **shared compute resource** to democratize access to computational resourced needed to exploit shareable data and common software tools. The NAIRR should draw from the recent experience of the COVID-19 High-Performance Computing Consortium and on-going efforts to develop a National Strategic Computing Reserve to support computing in response to national crises. Over 40 U.S. computing organizations, both public and private, came together to provide computing resources free of cost for projects with potential to accelerate understanding and development of treatments and interventions in the pandemic. A similar model could provide a path to providing top-tier computing and data resources to grow the ecosystem of AI R&D in the U.S. This directly addresses question 4 of the RFI. Such a shared computing model provides the final piece that will establish a truly open, accessible, democratized AI research resource. It bears repeating that the resource must include

expanded access to all three fundamental resources – data, software, and compute. Omitting any one of the three will severely limit the set of researchers who can benefit from the NAIRR.

The resource utilization partnerships

The NAIRR will develop a second type of partnership aimed at fully engaging and improving US AI R&D forces at all levels. These partnerships will focus on developing solutions to the directed grand challenges through collaboration across levels. They will be aimed at exploiting the common resources developed by the production partnerships. The efforts should have as endpoints both research solutions and advances, as well as upskilling, re-skilling, and recruitment of researchers. These partnerships must be vertical ones. Some partners should be well-resourced researchers, such as large corporations or national laboratories. They must be paired with historically underserved groups. These pairings are essential democratizing access to AI research. Simply providing underserved groups with the resources (data, software, compute) ignores the fact that progress is made through application of best practices for consuming such resources. To truly reduce the barrier to AI research, underserved institutions need access to these best research practices – methods of teaming, approaches to problems, selection of tools for tasks, optimization of workflows for problem solving, and more. These partnerships will are the final ingredient for expanding the US research ecosystem.

Ownership and Governance of the NAIRR (roadmap items B,C)

The NAIRR should be owned by a governmental agency with experience in large-scale R&D, interdisciplinary expertise (including AI work and varied applications), and access to critical resources (e.g. computing). We recommend that the Department of Energy play such a role. The experience with critical missions, wide-ranging science, world-class computation, and successful partnerships make it a prime candidate for owning the NAIRR. As already mentioned, the NAIRR should be governed by a broad consortium. Such a consortium should have a lead partner capable of steering the group. The lead partner should again have experience with a wide range of computing and science applications, including those pursued through public/private partnerships. The lead partner should also have extensive experience in AI work for a range of applications. We recommend that LLNL serve in this lead partnership capacity. LLNL is strongly coupled to DOE and has a history of global leadership in HPC, having been cited as the most dominant HPC site since 1993 by the TOP500 project. It has a long historical record of successful interdisciplinary science and engineering with a strong track record in public/private partnerships. LLNL pioneered novel risk-sharing relationships that have allowed vendor partners to aggressively innovate; a model which is now widely used across DOE. In addition, LLNL has developed a leading capability in Cognitive Simulation. This CogSim effort combines high-performance computing and simulation with high-precision experimental data using AI methods for applications ranging from drug design to fusion energy to climate modeling to national security. LLNL has created a research and oversight institute, the AI Center of Excellence to help coordinate its wide range of AI research efforts. The AI COE draws from across those efforts to offer a skilled team of technical leads with experience and mandate in developing AI vision and translating that into action via collaboration and R&D. LLNL also has considerable experience in developing and maintaining common software stacks and tools. It is the lead lab in the development of the TOSS operating system, which provides a common operating system for all commodity computing procurements made by NNSA and has made substantial contributions to HPC systems software that are in widespread use across the global HPC community. The history of LLNL, its unique CogSim AI work, and its pre-existing AI COE leadership structure make it uniquely capable of leading the NAIRR consortium.

Existing components (specific answers to question 4 for the RFI)
The DOE and LLNL bring, together with partners, a key set of initial components to help construct the NAIRR according to the plan described above.

Open data sets

1. LLNL provides a large set of open data sets for AI research spanning science missions and data scales. Hosted by the LLNL Data Science Institute as part of its Open Data Initiative
2. LLNL has a growing partnership with MLCommons to further develop open data and benchmarks

Accessible compute resources

1. LLNL has extensive experience from and contributions to COVID-19 High-Performance Computing Consortium. This provides a template for developing the shared compute portion of the NAIRR, as well as a partnership model.
2. LLNL maintains flagship class computing. LLNL machines perennially top the list of most powerful supercomputers in the world as maintained by TOP500.

Public-private partnerships

1. LLNL has an extensive history of successful public-private partnerships and sound partnerships models for doing partnered R&D. ATOM consortium is a prime example.
2. LLNL has an AI Center of Excellence that expressly develops public-private research partnerships to develop rapid advances in AI research. Their Collaboration Hub model offers a further template for developing NAIRR relationships.

Software tools

1. LLNL has developed and shared a set Cognitive Simulation tools that lay a foundation for common tools in AI for science (MaCC, DJINN, LBANN, MERLIN)
2. LLNL leads the RADIUSS project. It aims deploy a common base of foundational scientific software that would be essential to a common NAIRR software stack for use in a shared computing resource.