Carnegie Mellon University
Response to the
Request for Information on the
Development of an Artificial Intelligence (AI) Action Plan
from the
Networking and Information Technology Research and Development (NITRD)
National Coordination Office (NCO), National Science Foundation (NSF)
On behalf of the Office of Science and Technology Policy (OSTP)

March 15, 2025

<u>Institutional Contacts</u>
Theresa S. Mayer, Vice President for Research, [redacted]
Matthew Christian, Assistant Vice President for Strategic Research Initiatives,
[redacted]

I. Introduction

Carnegie Mellon University (CMU) has led and shaped the trajectory of artificial intelligence since its inception, even before the 1956 Dartmouth Conference. From foundational contributions in the early days of AI through our seminal work in machine learning, robotics, natural language processing, and automated discovery, CMU's work has supported the United States. Our leadership in the national interest continues through work that seeks to utilize AI to accelerate scientific discovery, provide reliable and secure autonomous devices for national security and industry, and transition knowledge from the university to the marketplace and workforce. Our work on technology transfer and workforce development create and expand industries and grow an AI-enabled workforce in the Pittsburgh region and across the nation.

The US has been a global leader in materials science, engineering, and biomedicine since the end of World War II. This sustained advantage has been due to partnerships between the government and universities and the translation of those efforts into industry. While industry R&D has advanced significantly and in some cases exceeds the abilities of non-profit partners, it continues to benefit from both the fundamental discovery of science and the training of future scientists who apprentice in universities and government funded laboratories. To maintain US leadership in science, engineering, and medicine, there needs to be a national investment in AI and autonomous research methodologies that enable rapid cycling between computational models and real world use of AI through significant advances in robotics to offset labor shortages, autonomous vehicles and embodied AI, and accelerated laboratory science.

The bottlenecks in generating value from AI systems are increasingly being driven by human constraints: how people can incorporate them into complex work, trust and verify their output, avoid unintended consequences, and use them to enhance critical thinking and innovation by humans. The United States and AI are at an inflection point. With GPU-driven growth in the power of AI, to remain competitive and thrive we must leverage the power of AI in domains as diverse as scientific discovery, manufacturing, farming, and education with an across-the-board approach.

Work on and national investment in AI directly address and mitigate the risks of the US falling behind in AI research, AI education, AI workforce and economic growth and prosperity, and useful and safe AI deployment in a wide range of applications and industries. We address: domains in which AI is critical to US leadership and success, core needs in developing AI capabilities, the need for energy to support growth in AI, ensuring AI growth and deployment are well managed for benefit, and critical growth in the workforce for an AI-enabled nation.

This white paper encourages nationwide efforts to ensure long term U.S leadership and competitiveness through systematic efforts and partnership among government, industry, and higher education to make the leap to fully embodied AI, i.e. the movement from computational models into real world applications such as autonomous experimentation, vehicles, and robotics.

II. Why: AI In the Real World Requires Action Now

Why Embodied AI and Why Now

The rapid advancement of artificial intelligence, particularly large language models (LLMs), has revolutionized information processing, decision support, and simulation-based research. However, the true potential of AI remains constrained when limited to computation and modeling alone. Without direct interaction with the physical world, AI lacks the ability to autonomously test hypotheses, execute precise mechanical tasks, and adapt to dynamic real-world conditions. A national investment in embodied AI—intelligent robotics and autonomous systems such as vehicles, machines, and manufacturing technologies that interact with the physical environment—will bridge this critical gap, enabling AI to go beyond theoretical and computational insights and into tangible action. By integrating AI with the physical world through robotics and embodied AI, we can create systems that not only analyze and predict but also act: build, refine, and optimize real-world processes to augment human endeavors.

This investment is essential to accelerating discovery, enhancing manufacturing, and addressing labor shortages that inherently limit scalability in critical industries. Our nation faces growing labor shortages in highly manual, critical sectors like specialty agriculture, advanced manufacturing, logistics, and infrastructure maintenance. We cannot be competitive without the labor to support these industries. Human labor, while skilled, is constrained by availability, cost, and physical limitations, making it a bottleneck in high-demand sectors. There are also places humans cannot, or should not, go (dangerous situations like wildfires, combat, etc.). Embodied AI can dramatically increase productivity, reduce operational and safety risks, and expand the frontiers of innovation by performing complex tasks autonomously and reliably. From self-improving factories to AI-driven scientific experimentation, the integration of AI computation with robotics will not only mitigate current limitations but also future-proof national industries against labor shortages and supply chain vulnerabilities. By strategically funding research, development, and deployment of embodied AI, the nation can secure its position at the forefront of technological and economic progress, ensuring long-term competitiveness in an increasingly automated world.

This is an area where national investments will have a multiplier effect, fueling new applications (e.g. in transportation, manufacturing, farming, health care, aerospace, defense) and industries by derisking research and development. Universities and small companies supported by government grants will enable the field to move forward quickly. The national investment will be amplified by private capital to move these ideas forward into commercially viable products.

Embodied AI will also require regulatory strategies. Autonomous AI systems face a special problem in that they interact with the physical world, including humans, and are currently deployed in fragmented and inconsistent ways. But, AI when applied to the physical world must abide by the laws of physics and meet safety constraints. Incorrect decisions can cause grave harm, injury and/or death. This can be especially problematic, for example, in deployment of autonomous vehicles, on the surface, in the air, underwater, and in space. Uptake of embodied AI applications will be limited unless the public can be assured that safety and security are paramount in the deployment of autonomous systems.

Thus, to advance this field we need to develop a uniform national framework: a comprehensive, rational framework for developing and deploying autonomous systems, balancing innovation and use with safety and regulatory oversight. Such a framework will provide regulatory clarity to innovators and investors, reducing the risk that a project could be upended after massive investments of time and capital. It also assures end users and stakeholders that as we move forward with embodied AI, key socio-technical tools, artifacts, and design processes were in place from inception for overall safety and to mitigate privacy risks.

How Embodied AI Advances Science and Technology

For generations, the core of discovery and innovation has been a process of generating questions and hypotheses, collecting data, analyzing it, and using the outputs to suggest new questions and hypotheses. In traditional science, this process is hard to scale. Most labs still look as they have for decades (labor intensive, error prone, slow). Even the most ambitious efforts have been limited by the ability of an individual or group to collect sufficient data and analyze it and to automate lab work. Efforts to speed scientific progress have been made by supporting increasingly large interdisciplinary teams, such as those routinely working in high energy physics. These large teams allow some parallel processes, but are ultimately rate-limited by the ability to collect, share, and communicate between individuals and to run automated experiments.

AI is already revolutionizing the scientific process, speeding discovery, and enabling breakthroughs that would otherwise be difficult, slow, or impossible. But it is limited in several ways. First, the models in use are based on incomplete or limited data sets due to the ways that we collect, store, and analyze data. Second, speeding up science and engineering requires faster cycling between computational models and real world experimentation and testing. And, third, autonomy can contribute to future automation and speed up the scientific process.

Creating Data Repositories Usable By AI: Modern scientific experiments both rely on and generate massive, complex datasets that are often beyond the capacity of traditional analysis methods. For scientific purposes, AI needs to provide well-understood uncertainties and be able to handle complex high-dimensional (and sometimes multi-modal) data at scale. Investing in AI method development for science will boost the scientific value of billion-dollar investments in scientific facilities in physics and other fields. Scientific journals, even in electronic form, are still tied to a paper medium. Discoveries move slowly through peer review and the data sets and methods require manual manipulation to combine them for use with other data sets. Review articles seeking to align seemingly similar data sets can take months and years of work.

New types of data repositories that allow sharing of data and methods in ways that can be easily read and synthesized by AI platforms will create faster iteration and hypothesis generation. The NIH and NSF already require that federally funded results be made available for free to the public, but they remain trapped in old paradigms. Creating the standards for data sharing is a first step, but in order to make a paradigm shift, these will need to be made available as a public good to US researchers, making it easy for one scientist to (re)use data created by another scientist and form the basis for a national resource. (The success in sharing data in the Human Genome Project is an early example.) This will also democratize participation in and access to science.
Data sharing websites for each relevant field would be very valuable. Just as GitHub has become the go-to website for software developers, a “Hub” for scientific data sets that serves as both data repository and search engine for discovering the data sets most relevant to a particular topic, hypothesis, or planned experiment would be a significant resource for advancing science.

In addition to collecting and sharing data in innovative ways, AI systems require improved complex modeling. AI can create sophisticated models of complex systems, such as climate patterns, biological systems and networks, astronomical phenomena, and even natural disasters. Scientists need ways to handle complex datasets with multiple types of data and need to be able to quantify multiple forms of uncertainty, both statistical and systematic. These models can be used to simulate different scenarios and make predictions, leading to a deeper understanding of these systems. Creating initiatives that focus on using AI for Physics or AI for Mathematics, for example, would fill key gaps in complex modeling – and provide an exceptional testing ground to advance the frontiers of the field and the frontiers of AI, leveraging the large and growing volume of complex, multimodal, publicly available data.

Moving from Modeling to Experimentation: AI can automate repetitive tasks and even design and conduct sequences of experiments, in a data-driven way from previous results, freeing up scientists to focus on higher-level analysis and interpretation. This is particularly valuable in fields like materials science and drug discovery, where experimentation can be time-consuming and expensive. AI can help identify hidden connections and patterns in scientific literature and databases, leading to new insights and hypotheses and pointing the way more reliably than scientific intuition. Unlocking this potential requires an alliance of automation and AI with applied science and engineering. For instance, AI can advance the science of environmental toxicology and testing capability by combining real-world high throughput screening with AI-driven predictive modeling in a closed loop system – creating an autonomous platform for testing and designing safer chemicals. A common “language of experimentation”—the equivalent of mathematical or chemical notation for the 21st century—is required to both run experiments and precisely convey the process to other experimenters for verification and reproducibility.

To achieve the promise of AI in science and engineering, the nation requires investments to advance embodied AI into an “AI Scientific Discovery Platform” that will empower advanced AI systems – particularly large language models (LLMs) – to act as researchers (under human guidance) that propose and conduct experiments across multiple fields. In this state-of-the-art setting, an LLM could formulate a hypothesis in biology or chemistry, then immediately test it either by running simulations in a vast digital sandbox or by directing automated physical laboratory instruments. The environment seamlessly integrates high-performance computing for virtual experiments with cutting-edge robotics in real-world labs, creating a foundational playground for innovation. By allowing AI models to iterate experiments rapidly, autonomously, and reproducibly, this platform promises to accelerate breakthroughs in machine learning, biology, chemistry, physics, medicine, and beyond – pioneering a new era of cross-disciplinary scientific exploration. Such an initiative will unite AI and domain experts in a shared quest for knowledge, yielding discoveries at a speed and scale previously unimaginable.

Embodied AI in the form of shared self-driving and autonomous labs significantly reduces the costs for scientific infrastructure. The cost of infrastructure for basic science is growing rapidly. Today national laboratories, universities, and industry invest heavily in capital research instrumentation that in many cases is underutilized. Robotic laboratories enable 24/7 operation and scheduling algorithms that optimize usage of instruments. Because autonomous experimentation processes could take place anywhere, experiments can be sent to facilities based on capacity, price, or optimal instrumentation rather than the limits of where a particular researcher happens to sit. Facilities can be constructed based on relevant factors: cost of power, clusters of talent, and availability of low cost real estate.

Better translation from AI modeling to real-world experimentation opens the door to true shared user facilities. This in turn would change the economics of scientific discovery by reducing the barriers to entry and the implicit "cost per breakthrough." To do so, AI methods must be robust, well-understood, and fit within the scientific method to drive scientific discovery.

Safety, Trust, Explainability, Assurance: As noted above, advances in embodied AI also require efforts to build public trust. This in turn requires advances in explainable AI. The results of applying AI methods need to be explainable in terms of the specific domain (physics, healthcare, safety, etc.). The US must continue research into explainable AI, and V&V (verification and validation), so that the results are correct and understandable, not opaque. We need investment in models, methods, standards, tools, and best practices for ensuring AI safety and reliability across different applications and sectors, so as to build public trust in the results and safety in their use. Key domain areas that require trust and risk analysis, including domain-specific work, for the productive and accepted use of AI include, for example, medicine and healthcare, finance, and emergency response. Systems will come to be expected to both produce the "right" answer and be able to explain that answer (e.g. why a specific medical diagnosis).

This has a secondary benefit. Explainable models, experiments run in ways that are understandable by humans and machines, and highly reproducible results will increase the public trust in science. Fears of researcher bias, amplified by reports of results that cannot be replicated, undermine efforts to develop and deploy life saving cures, advanced technologies, and even national security. As we move to embodied AI, we have the opportunity to deploy systems that increase transparency into the methods and processes. Buttressed by clear standards for safety, privacy, and ethics, AI-driven systems can "show their work" in ways that allow others to provide further validation of these discoveries and systems. Working with and supporting academia and industry in setting these standards is a correct role for the government, providing assurance to the public that their tax dollars are well spent and that the outcomes are highly reliable.

We are on the verge of a paradigm shift in discovery. Done correctly, the nation will move forward rapidly in the deployment of embodied AI tools. By setting clear standards for transparent and reproducible research, we de-risk the development of new technologies, techniques, and instruments to support this science. By providing access to the facilities with which to perform this work, we enable researchers throughout the US to participate in science and provide public assurance. Accelerating discovery will, in turn, provide benefits to industry and the economy—more rapid and less expensive discovery and testing (including on a small scale prior to larger investments). It is not too strong to say that democratizing science in this way will usher in a new wave of advanced technologies, materials, and processes.

Examples of Ways Embodied AI Will Further Health, Security and the Economy

Healthcare: AI can be used to develop new diagnostic tools, personalize treatments, and accelerate drug discovery, leading to improved patient outcomes and reduced national healthcare costs (e.g. as a percent of GDP). Private sector investments are incented to follow basic breakthroughs in biomedical research led by public investment in basic science, where the academic sector excels. AI-driven organizations can contribute to the discoveries in, and resulting economies to, the healthcare space, even separate from medical school facilities. Examples at CMU include work on precision cancer diagnostics and treatments and translational neuroscience.

Transformational catalysis: AI and robotics hold immense promise for discovering and developing new catalysts that can revolutionize industrial and energy sectors by enhancing efficiency, sustainability, and cost-effectiveness. Traditional catalyst discovery is a slow, labor-intensive process that relies on trial-and-error experimentation and computational modeling, often taking years to yield viable results. By integrating AI-driven simulations with autonomous robotic laboratories, we can exponentially accelerate the identification and optimization of novel catalytic materials. For example, the Open Catalysis Project utilizes machine-learned potentials to design and predict materials properties spanning 55 elements with near quantum chemical accuracy at under 1% of the cost. The largest models contain 200M+ parameters necessitating state-of-the-art GPUs for training and inference. Increasing access to GPUs beyond what any one entity currently holds will open up new possibilities. Advances in robotic systems that can autonomously synthesize, test, and refine catalysts at unprecedented speeds will not only streamline real world results, but also serve as a feedback loop improving the data in the models. This synergy enables the rapid development of catalysts for cleaner energy production, carbon capture, and more efficient chemical manufacturing, ultimately reducing reliance on scarce resources and minimizing environmental impact. Investing in AI-powered catalyst discovery has the potential to drive transformative advances in renewable energy, green chemistry, and next-generation manufacturing.

Materials Science: Via data-driven discovery, AI can help design new materials with specific desired properties, leading to advancements in areas like energy storage, electronics, and construction, hence more versatile, more use-specific, and more cost-effective solutions. For example, current work at CMU uses AI to advance discovery and development of new alloys designed to withstand extreme environments, from deep-sea exploration to space travel and next-generation energy systems. Traditional metallurgy relies on incremental improvements and extensive testing cycles, limiting the speed of innovation. By leveraging AI-driven simulations, robotic experimentation, and additive manufacturing, we can rapidly design, test, and refine novel alloy compositions with enhanced strength, heat resistance, and corrosion tolerance. Additive manufacturing allows for precise, layer-by-layer fabrication of complex materials, enabling the creation of alloys with optimized microstructures that traditional methods cannot achieve. AI-driven analysis of vast material property datasets can predict new high-performance alloy formulations, while robotic systems can autonomously synthesize and evaluate them in real time. This convergence of AI, robotics, and advanced manufacturing (embodied AI) will unlock groundbreaking materials critical for aerospace, nuclear fusion, and other high-stakes applications where conventional materials fail.

Causal Inference: AI helps scientists establish causal relationships from observational data, which is crucial for understanding complex phenomena and developing effective intervention. AI for Causation has proven extremely successful as a method to help discover genetic drivers of cancer (and flowering for plants that have gone into space), to categorize autistic vs. neurotypical brains, to find which educational approaches help students learn more, and many other important scientific breakthroughs. Recently, Causal Machine Learning was the crucial ingredient in building an image generating LLM that is more accurate than popular platforms like OpenAI or Google, but is a small fraction of their size.

Mathematics: Like scientific fields where advances have been limited by our inability to scale problems beyond a small number of individuals, in Mathematics there exist significant opportunities to promote the development, adoption, and use of computational technologies that support and enhance mathematical reasoning. Investment in and development of AI tools for mathematical reasoning will enable society at large to engage in mathematics and mathematical thought. This requires building a common framework and platform for computer verification of proofs based on the LEAN platform.

The benefits go far beyond the field of Mathematics. For instance, current widely used large-language models such as ChatGPT use transformer architectures. They are based on simple principles and their performance was observed to steadily improve with their size. While such models perform well at a number of tasks, their ability for logical reasoning and planning is rather limited and it is thus critical to invest in developing advanced reasoning capabilities for future AI. This requires advancing the understanding of how knowledge is represented by AI models. The next advances in AI itself will depend on how the relations between objects that are based on logic, programming and mathematics are incorporated into the models.

The Arts and Design: Embodied AI presents a transformative opportunity for the Arts and Design, positioning them as a strategic driver of economic growth, workforce development, and national innovation in AI. While AI has already revolutionized creative practices through generative tools in writing, design, and composition, the integration of embodied AI— intelligent systems interacting with real-world materials through robotics and automation—unlocks new frontiers for artistic production and industrial application. Autonomous fabrication, AI-assisted sculpting, and robotic performance technologies allow artists and designers to engage with materials and forms beyond human capabilities, expanding creative expression while fostering new interdisciplinary collaborations, including those that support large industries. As examples, AI is already showing progress in design in areas from architecture to furniture to apparel, contributing across the US economy and building US leadership in this area. As the arts and cultural industries contribute over $1 trillion to the US economy and serve as a key asset in global cultural influence, investing in embodied AI ensures that American artists, designers, and creative technologists remain at the forefront of this technological shift. Furthermore, as AI reshapes industries at a rapid pace, cross-trained professionals with expertise in both AI and the arts and design will be in high demand, strengthening the nation’s creative workforce and reinforcing the arts and design as an essential pillar of American innovation and competitiveness.

III. What: Rapidly Expanding US AI Capabilities

The United States stands at a critical juncture in the global race for AI leadership. To maintain our technological edge and ensure national security, we must urgently address the bottlenecks hindering AI innovation and deployment. There exists an urgent need for a comprehensive strategy focused on rapid expansion of AI capabilities, particularly in critical sectors like science, engineering, and national defense and notably in embodied AI, where real time response is often critical.

Overcoming the Compute Barrier

Academic researchers in the United States face significant barriers in accessing the computational resources needed to train and deploy AI models for real-world applications. While large industry enjoys rapid progress due to its ability to leverage massive computing infrastructure, universities and small enterprises struggle with resource limitations, despite being essential for long-term AI innovation. This imbalance creates two critical challenges: first, industry continues to rely on academic research for fundamental breakthroughs that eventually fuel commercial success, yet without adequate resources, this pipeline of innovation is at risk. Second, as industry research becomes increasingly proprietary, it restricts broader access to AI advancements, limiting reproducibility and hindering scalable applications that benefit society at large.

To bridge this gap, the nation must invest in a robust AI infrastructure that provides academic and public sector researchers with state-of-the-art computational capabilities. Establishing large-scale computing facilities—modeled after national laboratories—will ensure that research groups and consortia have access to the necessary compute power to train advanced AI models for fields such as healthcare, national defense, and manufacturing. Expanding programs like the National AI Research Resource (NAIRR) pilot initiative is a proven step in this direction, as it has already demonstrated the feasibility of democratizing AI resources. Furthermore, targeted investments in domain-specific computing, such as automated, cloud-based science labs, will enable AI to make meaningful contributions to high-impact scientific disciplines and speed scientific progress through reproducibility.

However, expanding computational access alone is not enough. To maximize efficiency and ensure national investments are well utilized, there must be parallel investments in GPU-centric scheduling, resource allocation, and diagnostic tools to prevent wasted and energy and cost inefficient compute cycles. Additionally, addressing AI’s limitations in reasoning and planning requires dedicated efforts to build specialized datasets tailored to real-world applications. A "National AI and Digital Commons Initiative" could serve as a foundation for open-source AI resources, including curated datasets and advanced AI models designed for scientific discovery. By coupling enhanced computing access with investments in fundamental AI research—particularly in hybrid models that integrate symbolic reasoning, logic, and mathematics—the United States can create AI systems that are computationally powerful and capable of intelligent, real-world decisions.

IV. How: Powering the US Future through AI

Data Centers, Energy Consumption, and Efficiency: Rapidly increasing demand for AI and AI-powered solutions, especially with arrays of GPUs, is driving an exponential growth in demand for computing and a resulting demand for power to drive and cool large data centers. We describe some key thrusts that are imperative for the US to continue to lead in the development of core AI and of its applications and to do so in a cost-effective way, to ensure the costs don't outweigh the benefits.

Resources: We must ensure access to robust, scalable, and secure cyberinfrastructure for academic AI fundamental and applied research. This includes critical resources such as high-performance computing (HPC), data storage, networking, and cloud computing, including at supercomputing centers, ensuring that US universities and research institutions remain globally competitive in AI innovation.

Measurement: to optimize energy consumption and cost of AI workloads, we first need to measure them. We must develop and enforce standardized energy reporting metrics that emphasize location-based costs and include a comprehensive look at all metrics: energy input, efficiency, output, water usage (for cooling), carbon emissions, and computational efficiency (e.g. Performance per Watt).

We also need to measure the load of various aspects of AI, from training AI models (widely recognized as energy-intensive) to the inference phase where these models perform real-world tasks on a much larger scale every day and in real time for embodied AI. We need to develop clear frameworks to quantify consumption, especially across various hardware configurations and across real world industry use cases and workflows (healthcare, finance, transportation, manufacturing, defense.…)

Demand: without a view towards optimization, the challenges and costs of producing energy for AI will grow far faster than we can keep up. We must enact comprehensive legislation to use existing sources, where possible, to address AI's energy demands and promote efficiency. Enablers include interagency collaboration among DoE, NIST and EPA, and across sources of power, grid modernization, incentives for renewable energy use by data centers, incentives to build energy efficient AI chips and systems, wise investment (including nuclear), reporting mechanisms, and so forth. Interagency collaboration can also contribute to the evaluation of the efficacy of AI systems.

Grid and Data Centers: The US electric grid is increasingly straining under the weight of rapidly growing data center energy demands in part due to its fragmented, antiquated nature. The grid lacks high-voltage transmission lines, has inefficient permitting processes, and significant regional concentration of load, which collectively force utilities to rely on outdated and expensive (and environmentally harmful) energy sources. The US should undertake a comprehensive approach to grid modernization and decentralization, including upgrading transmission lines, deploying smart grid technologies such as virtual power plants, integrating energy storage systems to improve efficiency and manage demand-response strategies, locating energy sources near data centers, and optimizing centers with advanced cooling systems and power management strategies. Smart design of data centers, through reduction of e-waste and cooling demands, also improves cost efficiency.

V. Ethics and Safety: Navigating the Complexities of Responsible Innovation

Artificial intelligence and embodied AI are rapidly transforming society, presenting both unprecedented opportunities and significant ethical and trust challenges. We emphasize the need for proactive, interdisciplinary approaches that prioritize accountability, transparency, and human-centered design, towards public trust and productive deployment and adoption.

Ensuring AI Trustworthiness and Safety

AI systems, particularly those deployed in high-stakes and embodied environments such as medical, security, and military applications, often lack sufficient accountability and validation measures. This leads to risks of bias, inaccuracies, privacy breaches, and a lack of trust among users and the public. AI systems, particularly in mission critical applications, must be trustworthy, robust, and secure. The lack of understanding of black box solutions poses vulnerabilities and limits their applicability and reproducibility in high-stakes scenarios.

Security and Privacy: Integration of AI in various domains, especially those interacting with the public and physical/embodied applications, raises significant concerns regarding data privacy and security. We must create policies and mechanisms in distributed architectures, encryption techniques, cybersecurity, privacy-enhancing technologies, socio-technical tools, and *ad incipio* design processes to ensure compliance with data protection regulations, as well as mitigating risks and creating public trust, while fostering collaborative AI innovation across sectors and applications. Investments in secure infrastructure and interoperability standards, as well as in testing and evaluation for security will enable trustworthy AI development in key fields such as healthcare, finance, and national security and defense, balancing data privacy and trust with technological advancement and application deployment.

AI Engineering: Solutions can be found by investing in the development of AI engineering as an interdisciplinary approach, increasing rigor by focusing on frameworks, tools, systems, and processes for developing and deploying AI in real-world contexts, particularly in high-stakes scenarios, such as use by warfighters and operators. At the same time, we must seek to include human-centered AI standards emphasizing user testing, interdisciplinary evaluations, and human oversight mechanisms. Similarly we should be establishing national standards and specifications for AI interoperability and security, ensuring that different AI systems can connect and work interactively as needed._AI engineering must include:

- Establishing Robust Accountability Frameworks: Implement mandatory accountability and validation measures for AI outputs, including rigorous testing, auditing, and documentation of training data and system limitations.

- Promoting Transparency: Require AI developers to disclose system limitations, document training data sources, and integrate human review mechanisms.
- Independent Audits: AI models should be auditable by independent researchers to assess their fairness, reliability, and long-term impact.
- Human Factors, i.e. an understanding of the humans interacting with the AI systems, and including working conditions, cognitive and emotional factors, expert vs. routine use, interactions for decision making (in various fields). Human factors work derives insights from human-computer interaction, psychology, labor studies, and ethics, and a host of additional factors for embodied AI systems.

Deployment Framework. Autonomous and embodied AI systems face a special problem in that they interact with the physical world, including humans, and are currently deployed in fragmented and inconsistent ways. This can be especially problematic, for example, in deployment of autonomous vehicles, on the surface, in the air, underwater, and in space where failed control systems could have devastating outcomes for humans.

We must:

- Develop a Uniform National Framework: a comprehensive, rational framework for deploying embodied systems, including autonomous vehicles and agents, balancing innovation with safety and regulatory oversight.
- Encourage Responsible Testing: Create guidelines that encourage investment and responsible testing of embodied systems, autonomous vehicles, and agents.
- Support the development of socio-technical tools, artifacts, and design processes to mitigate privacy risks and ensure that the AI is not driven to undesirable outcomes due to unintended priorities or underlying issues in data sources.
- Mandate Human-Centered AI and Design: The bottlenecks in generating value from AI systems are increasingly being driven by human constraints: how people can incorporate them into complex work, trust and verify their output, avoid unintended consequences, and use them to enhance human critical thinking and innovation. We must require human-centered design (from the outset) and human oversight of AI systems, prioritizing user testing and interdisciplinary evaluations, including in the field. We need “AI that works.”
- Research on the Human Factors of AI Work: Support interdisciplinary research examining the human labor involved in AI systems, including data annotation, model testing, and content moderation, as well as interactions with embodied AI.

Liability and Accountability: There is ambiguity in liability and accountability when AI-assisted decisions lead to adverse outcomes, making it unclear if the responsibility lies with, for example, the clinician or the model developer. This uncertainty creates obstacles in adoption and misalignments among stakeholders. Regulatory frameworks and guidelines, for example based on proportional level of control, are needed to clarify accountability for AI-assisted decisions, specifying who is responsible in adverse cases. These guidelines would provide clear liability structures, making AI model adoption more feasible by aligning the interests of stakeholders, including delivery (e.g. health) systems, autonomous system/embodied AI providers, developers, and end users.

Example: Balancing Bioethics and Safety in Embodied AI for Biotechnology
As embodied AI becomes increasingly capable of autonomously generating and testing biological hypotheses, it is critical to establish ethical and safety frameworks that prevent misuse while enabling innovation. Unlike traditional AI, embodied AI in biotechnology operates within a closed-loop system where autonomous experimentation refines AI models in real-time. This presents unique risks, as failures in accountability, transparency, and ethical oversight could lead to unintended consequences. To mitigate these risks, responsible AI principles must be embedded from the earliest stages of product development. Current approaches often address fairness, accountability, transparency, and ethics (FATE) reactively, after problems have emerged. Instead, AI development should integrate ethical responsibility into ideation and problem formulation, ensuring that risks are anticipated and mitigated proactively rather than relying on *post hoc* regulatory intervention. As noted earlier, doing this correctly and transparently will increase the public trust in both process and outcomes and in the scientific community as well.

A key policy priority for embodied AI in biotechnology is the implementation of built-in safeguards such as model alignment, watermarking, and unlearning techniques to prevent malicious applications. For generations, we have relied on inefficient, biased, and often nonexistent external oversight. New approaches are required because AI systems operate autonomously in scientific discovery—and the development of such approaches will positively impact any regulated industry. Standardized risk assessment frameworks are necessary, including dynamic benchmarks for AI-generated biological entities, proactive security testing (such as red teaming to identify vulnerabilities), and clear safety evaluation protocols before AI-driven embodied biotechnologies are deployed. Additionally, as AI-driven biomedical research advances, real-time monitoring mechanisms—such as multi-agent defense systems—may be required to prevent unintended or malicious outcomes. These policies would reduce the likelihood of AI being misused to create harmful biological agents, bypass safety screenings, or inadvertently generate hazardous compounds.

However, balancing safety and innovation is essential to avoid overregulation that stifles progress. While strong safeguards are necessary, excessive restrictions could slow beneficial AI applications in biomedicine, limiting advances in drug discovery, personalized medicine, and disease prevention. Transparency is essential and regulations that threaten to expose potential trade secrets may hinder open collaboration and transparency in scientific research, driving key advancements into proprietary, closed environments. Therefore, ethical AI governance must strike a balance—ensuring safety without undermining the transformative potential of embodied AI in biotechnology. By integrating responsible AI principles from the outset, establishing adaptive risk frameworks, and fostering a regulatory environment that supports both security and innovation, we can harness the power of embodied AI to drive ethical and groundbreaking scientific progress.

VI. Who: Building a US AI-Enabled Workforce

Artificial intelligence is rapidly transforming every facet of our lives, from professional landscapes to everyday interactions. AI systems rely on human expertise at multiple stages, from curating and annotating training data to evaluating and fine-tuning models.To harness its full potential and mitigate its inherent risks requires a fundamental shift in education for young people and in workforce training and education for adults. Thus the AI Action Plan should prioritize investments in education, training programs, and workforce development initiatives.

The proliferation of AI systems necessitates a workforce equipped not only to develop and maintain these technologies but also to effectively utilize and understand them. Many have noted that the current landscape is akin to a "wild west," lacking standardized training and comprehensive understanding of AI's uses and implications. The foundation of an AI-ready society lies in widespread literacy. AI literacy empowers informed decision-making, effective regulation, and ethical deployment. It enables workers to identify transformative AI applications, fosters responsible innovation, and ensures public trust.

A robust embodied AI initiative can significantly enhance workforce development by integrating AI education across disciplines rather than treating it as a standalone field. Traditional training programs often focus on computational specialists, but a truly AI-enabled workforce requires professionals across domains—scientists, engineers, doctors, and artists—to understand and apply AI in their respective fields. To achieve this, interdisciplinary education must be emphasized, ensuring that students and professionals are trained in both AI and traditional disciplines such as physics, biology, and the arts, side by side. This holistic approach fosters a workforce that can effectively use AI-driven automation, robotics, and experimentation without being forced to choose between AI and their core expertise. Institutions must invest in developing educational standards, training materials, and policies that support this integration from K-16 through graduate programs and continuing education for existing workers.

In addition to formal education, workforce development programs must focus on upskilling current professionals to ensure they can adapt to AI-driven advancements. This includes ongoing training for postdocs, faculty, and industry workers, with curricula that evolve alongside AI's rapid advancements. By embedding workers into the AI design, development, and deployment process, these programs can ensure that AI tools reflect the realities of various industries, improving functionality. The US might consider the equivalent of the "GI Bill" for reskilling and upskilling workers to use AI tools (with programs designed for upskilling while workers maintain current jobs). Employers would be provided with incentives to help their workers upskill for a 21st century AI workforce, including apprenticeships and on the job training opportunities.

Similarly, land grant universities reach every county in every state and can be harnessed to serve as AI extension centers to support AI in diverse fields (e.g. farming, manufacturing) consistent with developments in embodied AI for those industries. AI deployment is expected to cause some workers to lose jobs, while new opportunities will require workers to develop a new set of skills. A White House task force could design the infrastructure to create an AI Skills transition program, supporting American workers.

As noted earlier, use of AI to educate and train continues to accelerate, and these same ideas and technologies can and must be used to train on AI itself, for speed and scale.

Beyond workforce training, an embodied AI initiative will drive economic growth by fostering AI-driven entrepreneurship and business innovation. Programs like an "AI Co-Innovation Challenge Fund" would incentivize startups to partner with community organizations, ensuring that AI solutions address real-world challenges, meet market demand, and generate local value. Such initiatives would not only support emerging AI enterprises but also remove barriers to innovation by streamlining regulations and fostering a flexible development environment. Public funding and policy support should prioritize making AI education and workforce development accessible to all, ensuring that the US remains competitive in AI while creating broad opportunities for workers across industries. By integrating AI training with traditional disciplines, upskilling the current workforce, and fostering innovation, an embodied AI initiative can create a US future where AI enhances, rather than displaces, human expertise.

Since the 1980s, our Pittsburgh region has demonstrated how technological change impacts workers and communities and has provided a model for rapid adaptation through technology, innovation, investment, and upskilling, for continued prosperity.

VII. Conclusion

Other countries are investing heavily in AI and will rapidly deploy the technologies necessary for Embodied AI through advances in robotics and automation. The US developed these fields and currently leads, but sustained investment is essential to maintain US leadership in this critical area. China is forging ahead in this multi-trillion-dollar-a-year market that was invented in the US. They have many more deployments, targeted public and private investments, fewer regulatory constraints, and a low risk of public backlash. As such, they have taken leadership positions in key areas such as solar cells, nuclear fusion, commercial drones, electric vehicles, electric batteries, robotic manufacturing, and humanoid robots.

Investments in national AI research, education and training, and infrastructure (not just protectionism) will ensure that American technology and manufacturing companies are at the leading edge. While there is a role for federal investment, it is also critical that there is federal leadership in this area bringing together public and private entities on shared goals for a shared economic future. Strategic thinking toward US leadership in embodied AI will lead to new technologies and industries, driving economic growth and creating jobs and prosperity.

In this work, rich regional ecosystems, such as we have in Pittsburgh, are a critical driving force and resource in moving forward our nation’s leadership in AI.