<u>Cross-Cutting</u> (5 questions)

1. **<u>How should the United States achieve the goal of accelerating the pace of scientific innovation? What roles should be played by Congress, the administration, industry, civil society, and academia?</u>**

Scientific innovation does not happen without an educated, capable workforce. None of the recommendations in this document can be implemented without an education system that prioritizes strong science and technology skills. The responsibility for ensuring that the United States continues to be a leader in innovation must be shared by the public and private sector.

The US government should lead by continuing to invest federal resources in AI research and development; and govern and provide incentives to ensure its responsible use in the private sector. But government cannot single handedly ensure that the American workforce is equipped with the knowledge and skills necessary to participate in the innovation economy. The private sector, which relies on a continuous pipeline of capable workers, can play an equal role by investing in everything from early education and worker upskilling programs, to apprenticeship training models for people entering the workforce.

In addition to ensuring the US has a strong workforce, we must explore the untapped potential of AI in various fields. Ask, where can we use AI technologies to make a big difference? In what areas is AI not used much because we have not paid enough attention to them or there isn't enough access, support, or funding? For example, helping doctors in rural areas with clinical support and patient monitoring; using models to study traffic patterns and design safer roads; provide support for non-profits by predicting community needs, and analyzing community feedback to plan for things like parks, clinics, or schools.

Ensure an AI-ready workforce

a. **Start early to build the next AI-capable generation.** Early AI-focused education gives students the opportunity to think differently, ask questions, and challenge traditional ways of doing things. Programs could include exercises in solving real-world problems, and activities that spark creativity and innovation. They could also introduce important issues like the possible issues related to ensuring bias and fairness that should guide AI deployment and use.

b. **Launch a National AI Tutor Program to develop public-private partnerships to deploy adaptive AI learning platforms in K-12 and community colleges.** Spurring and supporting partnerships between experts from academia, research, and industry, gives young students an opportunity to explore different career paths in a world dominated by use of AI. Most importantly, this empowers students to feel confident about being part of the future of AI innovation, including the exciting possibilities ahead with powerful technologies like AGI.

c. **Develop a hands-on AI program for middle and high school curriculum to help students understand the application of AI in everyday life.** Sample use case for curriculum - Students develop a simple model to predict when classroom equipment (like projectors or laptops) might need maintenance. This kind of exercise gives students experience with deploying AI technologies to create practical solutions to real world problems.

d. **Support AI-based upskilling programs targeting underserved communities.** Much like owning a cell phone today, AI fluency will be a necessity as our world become increasingly AI-powered. Trust in technology is crucial if people are to use it. Fear of the unknown could cause some to resist using or benefiting from AI, thus it is important that all of society possess at least basic skills and understanding of AI technologies.

Re-ignite U.S. Scientific Leadership

a. **Create one unified ARPA – USARPA (United Strategic Advanced Research Projects Authority).** The separate ARPA agencies (DARPA, IARPA, ARPA-E, HSARPA, etc.) buy the same cloud services, vet the same vendors, and maintain separate sets of books.[1] A single umbrella "ARPA" agency that shares finance, HR, cybersecurity and super-computing while leaving mission Directorate/Program Managers independent could result in savings, allow for faster contract awards, and stronger bargaining power for AI resources (eg: chips, cloud services, etc.). This action requires no new funding as current overhead lines can be merged into a single "Shared-Services Hub." Each Directorate could maintain the current Senate-confirmed deputies.

   Creating a single agency also alleviates duplicative efforts; for example, the possible problem of near-identical research projects receiving ARPA awards. To prevent duplication, a unified registry under USARPA could be created; every new proposal could be auto-scanned by an AI system. If 70 percent overlap appears, project managers would file a one-page "Why Different?" waiver or merge efforts.

b. **Establish an Innovation or Science Score Office (For example, something like Sabermetrics used in sports analysis).** Currently, grant funding still relies on slow peer panels, not real-time performance stats.[2] To remedy this, perhaps require NSF's National Center for Science & Engineering Statistics (NCSES) to scan publications, patents, and start-ups, and then post a public Science Innovation Score at regular intervals—stats for R&D ROI. This could result in a transparent dashboard and accountability.[3]

   NSF NCSES already publishes the Science & Engineering Indicators and has authority to collect cross-agency R&D data, so no new bureaucracy is needed. A steering group of data leads from DOE, NIH and DOD could advise.

   Data pulled from existing IRS, USPTO, NIH using APIs[4] could provide reliable "early-signals" of research pay-off: IRS (start-up formation & R&D tax credits), USPTO (patent filings)[5], and NIH RePORTER[6] (paper outputs vs. grant dollars). Other feeds—e.g., SEC EDGAR for venture raises or DOI[7] DataCite for dataset DOIs.

c. **Establish deregulated innovation zones.** Lengthy federal rules deter development of bold prototypes. The federal government should work with the states to create "innovation zones," five-year sandboxes where select FDA/EPA/FAA rules are waived in exchange for live safety data feeds. Congress could provide fast-track pilot authority and the states would self-nominate and compete to receive the designation. These innovation zones could enable speedier technological breakthroughs (eg: drones, small-modular reactors, gene therapies) as well as retain jobs onshore.

d. **Partnerships between NSF and private companies could create micro-grants for frugal innovation.** Small inventors can deliver frugal innovations in a frugal manner but often lack seed cash.[8] Enabling NSF to issue smaller grants, for example $100K open-hardware grants to research and develop low-cost AI systems, biotechnologies, or climate sensors—pay only on milestone proof. This could tap under-represented talent and spread jobs beyond coastal hubs.

---

[1] https://www.gao.gov/products/gao-24-10755; https://www.gao.gov/duplication-cost-savings

[2] https://www.embo.org/documents/science_policy/peer_review_report.pdf

[3] For example, create a new dashboard called AI-to-Cure which links every federal AI-drug grant to FDA nods: spend, models, trial speed, royalties, jobs, lives saved. Small town biotech Founder John spots an open antibiotic model, reuses model code, wins SBIR award—clear ROI and public accountability in one click. John adds his derivative work back into the dashboard's pipeline, unlocking fast-track FDA mentoring under the same program umbrella. For lawmakers: clear line-of-sight from federal dollars to FDA-approved drugs, domestic jobs, and savings. For innovators: one click reveals high-ROI code and contacts, lowering entry barriers and multiplying the original investment's impact.

[4] Think of an API (Application Programming Interface) as a restaurant's take-out window — you place a specific order ("give me all cancer-related patents filed in 2024") and the kitchen (the agency database) serves just that dish (data) in a standard container (JSON file). No need to enter the kitchen or see the recipe. Most federal APIs follow the OpenAPI/Swagger standard; documentation lives on *api.data.gov*. What is an API? - https://developer.uspto.gov/api-catalog

[5] USPTO API - https://developer.uspto.gov/api-catalog

[6] NIH RePORTER API - https://api.reporter.nih.gov/

[7] DOI (Digital Object Identifier): a permanent alphanumeric code that points to a specific research output—like a book's ISBN but for research papers. https://www.doi.org/the-foundation/about-us/

[8] https://www.nature.com/articles/d41586-023-03816-7

e. **Require a human-centered AI federal grant clause.** The public fears AI will replace, not empower, workers. To alleviate these fears, federal grant applications could add a Talent Stewardship clause to require the applicant to illustrate how the project would multiply, not replace, human skill. Such a clause could build trust among the US workforce as well as lead to federal funds helping to ensure that research and development leads to workforce-upskilling outcomes.

f. **Boost evidence-based federal R&D portfolio.** The US government under-invests in research and development; for every $1 of R&D investment, the return is $5 to GDP, yet federal budgets lag.[9] Congress should shift non-defense discretionary research spending to use-inspired research (defined as Pasteur's Quadrant, a type of research that is motivated by a desire to solve practical problems or create real-world impact) and networked research, guided by the aforementioned Sabermetrics data. This could result in fiscal prudence and a growth narrative.

2. **<u>What infrastructure needs to be built to make scientists more productive, and for each type of infrastructure you recommend, what should the funding model be for the construction and operation of that infrastructure?</u>**

Use what we have, share what we build, and pay for both through the value they create.

a. Create a National Compute Exchange (NCX), a single open marketplace for compute. SEE MORE BELOW under COMPUTING.

b. Create *a* Micro-Cluster Program (similar to Event Horizon Telescope which didn't build new observatories; it networked telescopes that already existed and added precision clocks.[10]) - Seed thousands of small clusters and federate them into one grid. New super-computers cost billions and take years[11], but modest CPU / GPU nodes could rival a national lab quickly and cheaply if they were networked correctly. For every GPU a company donates to a campus, the Department of Treasury could grant a matching tax credit. Software could link them into a secure national level grid that could plug into the NCX mentioned above. Offer tax credits and partner with universities and industry for donations. Public–private match—each private dollar could trigger one federal credit; ongoing power costs could be absorbed by host campuses; grid management could be funded by NCX fees.

- **Measuring success:** Use cost per flops versus traditional builds, geographic spread of nodes, increase in compute-driven publications from under-resourced regions, research outputs per $X M invested.
- **Impact:** It could deliver petaflop-scale capacity for a small-scale budget, it could keep talent nationwide and could further lower prices on the NCX (more supply drops market prices further and speeds every AI-enabled discovery).

c. Create the "ShareLab USA," an Airbnb for scientific gear, a trusted booking platform for spare lab time. Microscopes, wind tunnels, and bioreactors bought with public funds often sit idle while small labs and start-ups queue or overspend to use similar tools. Institutions would post available hours, and vetted users would book and pay online. Click-through agreements could cover IP and safety. A NIST badge system could rate labs bronze/silver/gold, e.g., high-risk work allowed only in gold labs.

- **Who leads?** NIST (platform standards), NSF (pilot grants), participating universities and federal labs (list equipment).
- **Funding & sustainability:** $X M, one-off NSF pilot (could leverage CHIPS & Science Act "shared infrastructure" authority). Platform could earn a X % booking fee and optional subscription tiers; after year three no federal outlay.
- **Measuring success:** utilization rate of listed instruments, fee income returned to hosts, and growth in shared-facility citations in peer-reviewed papers.

---

[9] https://www.aei.org/economics/yes-america-underinvests-in-public-rd/; https://www.aau.edu/newsroom/leading-research-universities-report/new-research-suggests-returns-federal-investments-rd
[10] https://eventhorizontelescope.org/blog/astronomers-reveal-first-image-black-hole-heart-our-galaxy
[11] https://www.digitaltrends.com/computing/microsoft-openai-working-on-a-data-center-project/

■ **Impact:** It could turn sunk costs into revenue, cut duplication of new equipment purchases, and let every district host world-class experiments.

3. **How do we ensure appropriate design of new scientific workflow models that offload certain tasks to AI while keeping human scientists at the center of the discovery process?**

Adopt a "two-key AI workflows" standard - Powerful "self-driving labs" like the Air Force's Autonomous Research System (ARES)[12] can already design, run and re-design experiments overnight. Without built-in checkpoints, scientists risk becoming passive button-pushers, legal liability blurs, and hidden model bias can steer research off-course.[13] Adopt a Federal "Two-Key AI Workflow" standard for automated research (especially for every federally funded project). For example, the National Institute of Standards and Technology (NIST) could publish a framework that any lab can bolt onto existing platforms such as ARES. Key elements (generic across disciplines):

–Explain-Why panel — every AI suggestion must include a plain-language rationale (Explainable AI with Interpretability and Contextual Adaptability[14]) and confidence score[15].
–Two-key rule[16] — no high-risk step executes until it is approved once by the AI model and once by a human Principal Investigator (PI).
–Red-button fail-safe — any lab member can pause the run; auto-alerts the PI.
–Immutable provenance log — results and approvals are hash-stamped to a low-cost reliable ledger.[17]
–Document compliance with established standards such as IEEE 7000 (ethical process), 7001 (transparency) and 7003 (bias tests) inside its electronic lab notebook.
–Red teaming[18] for greater security.

Impact[19]:

–Faster discovery, same accountability. An overnight AI run could test several material formulations; the morning-shift scientist could accept the most promising and reject the rest. The ledger could show exactly who approved what, satisfying auditors and journal reviewers while potentially cutting trial time by X %.
–Fiscal win – Re-using existing established standards (such as IEEE) could avoid a costly rule-writing exercise and deliver immediate, auditable guard-rails.
–Global trust signal – An American workflow that is transparent by design could answer OECD calls for human-centered AI science.[20]
–By off-loading drudgery—never judgement—the U.S. could accelerate innovation without surrendering human insight, aligning with calls for "people in the cockpit."

Measuring success:

- Track what portion of all federally funded research teams start using the new workflow.
- For each project, record how often a scientist steps in to overrule the AI and why.
- Compare the calendar time from first idea to a confirmed result against today's average to see how much faster the work could move.
- Measure how long it takes a scientist to read (understand) and, if necessary, stop an AI recommendation; for example, the goal is minutes, not hours.
- Periodically (for example, every three months) an automated bias check could grade each project; for

[12]https://www.afmc.af.mil/News/Article-Display/Article/3564196/ai-research-robots-key-to-democratizing-and-revolutionizing-science-world-class/
[13]https://www.nationalacademies.org/news/2022/05/automated-research-workflows-are-speeding-pace-of-scientific-discovery-new-report-offers-recommendations-to-advance-their-development
[14]https://hellofuture.orange.com/en/for-a-contextual-approach-to-explainability/
[15]https://learn.microsoft.com/en-us/azure/ai-services/language-service/question-answering/concepts/confidence-score
[16] The "two-key" concept mirrors nuclear launch authorizations—both keys must turn; here one key is digital, one human.
[17] https://www.mdpi.com/1999-5903/13/6/143
[18] https://csrc.nist.gov/glossary/term/red_team
[19]For example, a materials-science robot could propose raising the furnace temperature to 950 °C. The notebook could auto-attach an IEEE 7001 transparency card: "Chosen because models predict 12 % higher superconducting yield." The PI could review, note that 950 °C exceeds alloy warping limits, and click "reject." AI could retrain overnight; next morning suggest 910 °C, pass both IEEE 7003 bias check and human approval – experiment proceeds. Result: hours saved, human judgment preserved, full audit trail logged.
[20] OECD, *Recommendation on AI in Science* - *https://www.oecd.org/en/publications/artificial-intelligence-in-science_a8d820bd-en.html*

example, aim for nearly all to pass on the first try.
- Track how often data sets with provenance records are cited in papers; more citations could mean better trust and reuse.
- Laboratory logs could show the number of staff hours saved because routine tasks were potentially handled by the AI assistant.
- Most important, every decision—human or machine—could be stamped with a time, a name, and a reason so that any reviewer could follow the entire chain later

4. **In order to measure the success of ASAP, we need to have objective metrics that measure the speed of scientific innovation. What metrics already exist and what ones need to be created? What information should the federal government have to understand the health and productivity of our innovation ecosystem, and what tools, processes, or institutions should be used to do so?**

Some existing metrics are:

a) Test scores of AI capabilities relative to human performance.[21]
b) Performance on knowledge tests vs. training computation.[22]
c) Patent applications filed and awarded (there is a time lag).
d) Venture capital funding.
e) Cumulative number of large-scale AI models by domain.[23]
f) Peer reviewed publications.

What metrics do we need

a) Metrics to measure 'collaboration and diffusion' across domains.
b) Metrics to measure 'societal impact.'

Problem: Studies on innovation ecosystems show that siloed knowledge is a major impediment to breakthrough innovation. Yet academic reward systems and corporate competition often keep silos intact.[24]

Recommendation: Leverage findings from innovation science. For example, the "triple helix" model in economic theory[25] posits that academia, industry, and government interactions drive innovation. Congress should explicitly adopt a triple-helix approach – measure and reward the program not just for publications or patents, but for partnerships formed. OECD reports also highlight that multi-stakeholder collaboration is key to tackling complex societal challenges.[26] The recent energy-efficient computing innovation, neuromorphic computing, is from a collaboration of neural science and computer science.[27] Therefore, include collaboration metrics (such as number of cross-sector co-authored papers) as success indicators. By aligning ASAP with evidence-based collaboration strategies, we could institutionalize a culture of cooperation.

Problem: The United States still judges science mostly by inputs (dollars, patents, papers). Yet the real race with global competitors is about how fast an idea goes through the litmus test, becomes a working product, and becomes a market success. A GAO review found that there is a lot of scope to further improve awards timeline.[28] Without a national stopwatch, delays hide, overruns grow, and Congress cannot steer resources where they matter most. Even a lightning-fast discovery is not a win if it never reaches classrooms, clinics, or the shop floor.

Recommendations: Tell a complete story: how fast we move, how well we control cost, how much good we create, and whether the underlying evidence is findable and complete.

Create an "Innovation Velocity Index" (Are we faster than last year?): for example, the National Center for

---

[21] https://ourworldindata.org/grapher/test-scores-ai-capabilities-relative-human-performance
[22] https://ourworldindata.org/grapher/ai-performance-knowledge-tests-vs-training-computation
[23] https://ourworldindata.org/grapher/cumulative-number-of-large-scale-ai-models-by-domain
[24] https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/06/addressing-societal-challenges-using-transdisciplinary-research_41211835/0ca0ca45-en.pdf
[25] https://www.unr.edu/business/international/blog/triple-helix-model
[26] https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/06/addressing-societal-challenges-using-transdisciplinary-research_41211835/0ca0ca45-en.pdf
[27] https://research.ibm.com/blog/what-is-neuromorphic-or-brain-inspired-computing
[28] https://www.gao.gov/assets/720/717092.pdf

Science & Engineering Statistics (NCSES) could link four time-stamped events for every federally funded project: (1) grant signed, (2) working prototype, (3) first regulatory filing, (4) first customer revenue. The median of these intervals could be published as the Innovation Velocity Index (IVI).

How will we know it works? Agencies must cut their IVI X % per year.

Create a cost-and-schedule index (Did we stay on budget?): for example, the NCSES could calculate it as (money spent ÷ money budgeted) × (days used ÷ days planned). Any score > 1 could trigger an automatic red-team review by the investigator of the funding agency.

How will we know it works? Overruns could surface early; for example, enlist an entity such as GAO to evaluate and report which programs stay on target.

Track knowledge hand-off, Go-to-market Lag (How long does knowledge sit idle?): for example, an AI tool could cross-link publication DOIs[29], patent filings, and SEC or FDA records to measure years from first paper to first revenue event.

How will we know it works? Stakeholders (such as policymakers) could see which fields (e.g., fusion, mRNA) shrink the "valley of death"[30] fastest.

Track Talent & Value Multiplier (TVM) (Was the investment worth it?): For example, NCSES could add a field that auto-pulls societal returns from existing public datasets: wages at firms that license the tech (BLS), health savings such as lives saved vs baseline (CDC), environmental gains such as CO_2 saved vs. baseline (EPA), government revenue such as federal royalty or tax receipts (Treasury). It is similar to cost-benefit analysis. Example calculation of TVM = public benefit dollars ÷ federal dollars spent, updated periodically (annually). Example calculation: A DOE battery grant costs $10 M. Three years later, the licensed startup posts $60 M payroll and EPA logs a 200 kt CO_2 cut worth $4 M in carbon-credit value. TVM = ($64 M ÷ $10 M) = 6.4.

How will we know it works? Low-TVM areas could get process audits; high-TVM areas could become blueprints for other agencies.

Create a Data Stewardship Score (DSS): Agencies require Data-Management Plans (DMPs) but rarely track whether the data are shared. Direct each grant-making agency to publish a DSS for every project. Example calculation of DSS = (% of promised datasets released × FAIR-quality grade). For example, NCSES could use AI to calculate DSS by crawling through agency-approved repositories (e.g., Figshare[31], Dryad[32]) and checking for: Findable DOIs, Accessible metadata, Interoperable formats, Reusable licences.[33]

How will we know it works? Projects scoring > X (e.g., 0.8) could unlock expedited process (renewal review); scores < Y (e.g., 0.5) could trigger data-help coaching, not penalties.

Create Null Results Transparency Quotient (NRTQ): Bias towards successful and positive outcomes in research hides negative findings, skewing meta-analyses[34]. Direct all federally funded projects to register experiments (pre-study DOIs) and publish outcomes—positive *or* null—within X months (12 months). Example calculation for NRTQ = null-papers ÷ total papers. For example, NCSES could use existing registries (e.g., ClinicalTrials.gov, Open Science Framework (OSF)[35] to compare preregistered studies with journal or pre-print outputs and then it could flag missing null publications.

---

[29] DOI (Digital Object Identifier): a permanent alphanumeric code that points to a specific research output—like a book's ISBN but for research papers. https://www.doi.org/the-foundation/about-us/
[30] "Valley of death" is a metaphorical phrase. It describes the period between a company's initial investment and when it achieves sustainable revenue or profitability. This can be a challenging time, with significant financial pressure and uncertainty
[31] https://figshare.com/
[32] https://datadryad.org/
[33] FAIR principles - https://www.nnlm.gov/guides/data-thesaurus/fair-principles
[34] The File Drawer Problem (Publication Bias) - https://web.ma.utexas.edu/users/mks/statmistakes/filedrawer.html
[35] Open Science Framework (OSF) is a non-profit online workspace where any researcher can store and share their data, code, lab notebooks, and pre-prints in one place. https://osf.io/

How will we know it works? If a research area (e.g., mRNA) scores very low on the NRTQ scale (below 0.1), it could get a gentle reminder to pick up the pace.

Potential Impact: Stakeholders (such as Congress) would be able to spot chokepoints (for example, slow FDA filings) and fix just those links. The US could achieve a competitive edge; clear velocity metrics could enable the U.S. to publicly benchmark against OECD and WIPO timelines, reinforcing America's innovation leadership story. The cost benefits of spending could be transparent; the actions could give clarity as to whether a program returns more to society than it costs.

5. **Grand challenge problems can help provide concrete direction for how to implement new innovations. What core innovations does America need that can help guide ASAP? If possible, please provide an objective quantifiable metric, such as decreasing the time it takes to bring a new drug to market from 10 years to 1 year.**

- **Problem**: Ransomware & quantum attacks can exploit US systems for months before fixes roll out[36].
  - **Recommendation**: Create a "24-Hour Zero-Day Shield" – nationwide cyber platform where AI agents could generate, test and push a patch within 24 hours of any critical flaw disclosure.
  - **Potential impact**: It could secure hospitals, grids, & critical entities; it could showcase US cyber leadership.
  - **Measure**: Mean-Time-to-Patch (MTTP) – goal < X hours (24 hours)

- **Problem**: Discovering a new high-performance material can take years; China dominates supply chains[37].
  - **Recommendation**: Create a "100-Hour Materials Pathfinder": self-driving labs must predict, make and validate any requested property (e.g., 5× stronger anode for battery) in ≤ 100 h.
  - **Potential impact**: It could create domestic alloys for chips, batteries, etc.
  - **Measure**: Idea-to-Sample Clock (ISC) – hours from hypothesis upload to verified sample. Target ≤ 100 h for X % of jobs (e.g., 80%).

- **Problem**: Every critical sector — planes, grid, medical devices, AI models — now runs on millions of lines of code. Software defects cost the U.S. ≈ $2 trillion a year in outages, recalls, ransomware and delays.[38] Existing test and patch practice cannot keep pace with ever-larger code-bases.
  - **Recommendation**: Create a National Verified & Self-Explaining Software Fabric (N-VSEF)
    - Launch a 10-year N-VSEF moon-shot: by 2035, all federally funded mission software ships with machine-checked mathematical proofs of correctness[39] and an auto-generated plain-English rationale using LLM.[40] Embed an energy-cost estimator for each function; developers must consider both correctness and efficiency (vital for AI).
    - Who: For example, NIST could partner, via open calls, with universities, U.S. cloud vendors and national labs.
  - **Potential Impact**[41]:
    - Faster breakthroughs: when software defects vanish early, prototypes could move to field tests in weeks, trimming "lab-to-market" clocks for everything from autonomous drones to vaccine-design AI.
    - Security boost: mathematically proven software quality could reduce a lot of cyber-attack.
    - Cost savings: a drop in software defect could reduce the related rework in turn could save money.
    - Global edge: it could set an American gold standard that allies could adopt.
  - **Measuring success**:
    - Bug-escape rate in operational systems: for example, it could be less than one defect per million

---

[36] https://www.gao.gov/blog/next-big-cyber-threat-could-come-quantum-computers-government-ready
[37] https://pubs.usgs.gov/periodicals/mcs2024/mcs2024.pdf
[38] https://www.it-cisq.org/the-cost-of-poor-software-quality-in-the-us-a-2020-report/
[39] *Formal verification* — mathematically proving that code meets its specification using logic solvers - https://cacm.acm.org/news/formal-software-verification-measures-up/
[40] *LLM* — large-language model; an AI system trained to generate human-like text.
[41] *Use-case story*: For example, a small startup could prove its carbon-sensor firmware correct with the N-VSEF tool-chain. The EPA reviewer could spend X hours, not X months, verifying the plain-English proof, green-lighting field deployment before the wildfire season potentially saving millions in crop losses.

lines by 2030.

  - Proof coverage: for example, it could be greater than 90 % of new critical code paths formally verified by 2028.
  - Certification time for safety-critical software: for example, it could be cut from ~24 months to less than 2 weeks by 2032.

- **Problem:** AI innovation requires great amounts of energy. Currently, U.S. data centers use 2 percent of all electricity; that number could double by 2030. A 2024 Department of Energy report on data center energy usage says, "... data centers consumed about 4.4% of total U.S. electricity in 2023 and are expected to consume approximately 6.7 to 12 percent of total U.S. electricity by 2028."[42] Some AI models use as much power as a small town just to train.[43] And many universities and labs don't have access to clean or reliable energy to support this kind of computing.[44] Without enough clean, reliable energy sources, scientific efforts cannot scale as researchers demand - especially during heat waves, blackouts, or natural disasters.

  Recommendation: Accelerate clean and reliable energy infrastructure and invest in "Green Power Zones." Establish renewable energy hubs and microgrids to support research institutions and innovation districts.

    –Modernize Campus Energy Systems - Fund clean energy upgrades and storage solutions for universities, labs, and public research centers. Create a "Science-Grade Energy" Certification Program for research campuses and public institutions.
    –Incentivize Energy Efficiency - Support R&D in low-energy tools, buildings, and infrastructure through grants and public-private partnerships. Include Clean Research Zones in national infrastructure and energy investment legislation.
    –Enable Smart Grid Access - Deploy nationwide smart metering and demand forecasting technologies to stabilize energy for critical scientific facilities.
    –Ensure Resilient Power Supply - Back energy storage initiatives to prevent outages and enable 24/7 research continuity.

[42] https://www.energy.gov/articles/doe-releases-new-report-evaluating-increase-electricity-demand-data-centers
[43] https://www.contrary.com/foundations-and-frontiers/ai-inference
[44] https://edtechmagazine.com/higher/article/2024/11/universities-invest-high-performance-computing-support-ai-education

Pillars

(5 overarching questions: DATA, COMPUTING, AI, COLLABORATION, PROCESS; with multiple sub questions under each of the 5 overarching questions.)

DATA Pillar

6. **How can America build the world's most powerful scientific DATA ecosystem to accelerate American science?**

America could build the world's most powerful scientific data ecosystem by knitting three, mutually reinforcing moves:

- A National Self-Describing Data Fabric (NSDF) – so every dataset, wherever it lives, could speak the same "grammar," could be discovered in seconds and could pay its own upkeep.
- A National Science Data Cloud (NSDC) – a common "warehouse district in the cloud" where any researcher can run heavy analytics without buying hardware.
- A Certified Synthetic-Data Marketplace – a safe, legal fuel pump for teams who need realistic data but are blocked by privacy or trade-secret walls.

Together these moves could deliver availability (the Cloud), interoperability & traceability (the Fabric), and legal access to sensitive signals (the Synthetic-Data exchange). The Fabric could tell every byte *how to speak*; the Cloud could give every scientist *somewhere to work*; and the Synthetic-Data Marketplace could supply *legal, high-octane fuel* to anyone with an idea—together compressing discovery cycles, cutting waste, and keeping the U.S. at the front of global science.

Build a National Self-Describing Data Fabric (NSDF)[45] - Breakthrough innovation now depends on mixing many datasets (such as genes, materials, climate), yet a lot of U.S. research files sit in private format — slowing reuse, duplicating grants, and inviting leaks.[46] To alleviate these issues, launch the NSDF grand project by doing the following:

One "Data-Highway Code." Publish an open, machine-readable national schema (think "common grammar") (such as JSON[47]) that every federally funded dataset could export next to the raw files.

Plug-and-play "Gateway Pods." Provide small, container-based (such as Docker[48]) appliances that any lab could drop onto its existing server or cloud account; the pod[49] could speak the common schema, could run queries inside a protected sandbox, and could record every access.

Micro-Royalty Ledger. Each gateway could write usage events to a permissioned distributed ledger[50]; a fee (such as fraction-of-a-cent) per query could automatically flow back to the data owner, covering storage costs and rewarding sharing.

Explainable AI index. A national search engine could continuously review registered datasets and publish a plain-language summary—generated by interpretable AI (such as large language model LLM)—so any stakeholders (such as students or policymakers) could understand why the data matters before requesting access.

The NSDF is analogous to how the mesh of the interstate highways as used by the trucking industry: every lab keeps its cargo, but one set of open rules lets anyone find it, pay the toll, and trace the route.

---

[45] **Data Mesh** - Every team looks after its own data but promises to share it in a standard way. Local ownership, knowing who the expert is, keeping data close to the people who create it. https://www.ibm.com/think/topics/data-mesh. **Data Fabric -** A smart blanket of software that knows where every file lives, translates formats on-the-fly, and enforces the rules. Global search, common security, one-click movement of code to data. https://www.ibm.com/think/topics/data-fabric. How Data Mesh and Data Fabric work together - Think of highways and drivers. The mesh is like each city maintaining its own stretch of road; the fabric is the Interstate map, signage, and toll tags that let you drive from Miami to Seattle without stopping to ask directions. A quick example, a university materials lab "owns" its alloy test results (mesh). A start-up designing lightweight airplane parts could ask the national fabric, "Any alloys that stay strong above 800 °C?" The fabric could check permissions, spin up a secure compute spot inside the lab's server, run the start-up's code, and return the answer—no heavy data copies, and the lab could earn a small royalty.

[46] Open Science by Design" report, published by the National Academies in 2018, addresses the importance of open science and data sharing, and highlights challenges related to making data findable, accessible, interoperable, and reusable (FAIR principles)

[47] https://www.json.org/json-en.html

[48] https://www.docker.com/

[49] https://kubernetes.io/docs/concepts/workloads/pods/

[50] https://hyperledger-fabric.readthedocs.io/en/latest/whatis.html#permissioned-vs-permissionless-blockchains

A permanent Data Fabric Program Office could be housed at the National Institute of Standards & Technology (NIST). The Data Fabric Program Office will manage and define the design and architecture patterns to guide the initiative. The decision on Data Mesh or Data Fabric or Data Weave[51] falls within the ownership of the Data Fabric Program Office. As part of defining standards, it is important for the Data Fabric Program Office to consider Data Vault and Data Minimization concepts to set the content on how the data gets used, and who gets access to the data. The standards definition should also include details on what data gets loaded where, for example what datasets get loaded to data.gov initiative, and what attributes within the datasets gets anonymized or tokenized to maintain the privacy and confidentiality.

Measuring Success: Median search-to-analysis time (for example, less than 2 seconds by 2028); Duplicate-project spend cut (for example, cut by 25 %) (it could be tracked by global-unique dataset hashes[52]); Royalty dividend ratio (greater than 1 means that the system pays its own bills).

Impact:

Quicker breakthroughs. Think of doctors hunting for a new cancer pill: instead of waiting months for scattered files, they could mix gene data and drug results overnight—like searching all the cookbooks at once to find the best recipe.

Security. Every file could get a tamper-proof "package-tracker" code. If someone tries to change a number, the system could ring an alarm.

Big savings for taxpayers. When one lab has already run an expensive test, others could reuse those results instead of paying to redo it—potentially avoiding billions of dollars in repeat work over the next decade.

Academic incentives. Citations and royalty income; null-result datasets could be rewarded the same way as positive results.

Establish a National Science Data Cloud[53] that is open to all domains and to all researchers similar to the European Open Science Cloud (EOSC). US federal entities such as the National Library of Medicine, the National Institutes of Health, as well as professional associations such as IEEE, APS, and AAAS host data repositories (e.g., IEEE DataPort). But there is no central cloud for sharing this info among US entities. On the other hand, the European Union supports the EOSC with the objective of providing, *researchers and innovators in Europe with an open and trusted multi-disciplinary environment where they can publish, find and reuse data, tools and services for research and innovation. Through this environment, EOSC aims to mobilise, align and scale resources across Europe to accelerate open science, higher productivity and increased reproducibility and trust in research*. The US should establish a similar cloud infrastructure.

Create a Certified Synthetic-Data Marketplace

Problem: Real-world files such as medical records or factory logs are locked behind privacy or intellectual-property rules; small labs and start-ups cannot legally touch them, stalling AI training and replication.

Recommendation:

- Accredit "data foundries." A standards body (for example, one hosted at NIST) could certify vendors that can turn raw, sensitive data into synthetic twins that keep the statistical patterns but remove personal or proprietary details.
- Common quality & bias seals. Use standards such as IEEE so every synthetic batch ships with a "nutrition label" covering accuracy, privacy, and demographic & statistical representation.
- Tokenized exchange. List certified datasets on a micro-fee marketplace; buyers pay a few cents per query, and proceeds flow back to the original data owners. Everything clears through the same permissioned ledger used by the NSDF, so use is traceable and audit-ready.

Potential Impact

- Open fuel for innovation – a cancer-drug start-up could train on real-lifelike patient histories in hours

[51] AI & Future Technologies, The 3 pillars that support AI's data deep undercurrents - https://www.thomsonreuters.com/en-us/posts/technology/pillars-ai-undercurrents/

[52] https://www.ledger.com/academy/glossary/transaction-hash

[53] FYI China has a Data Exchange - https://www.cigionline.org/articles/data-marketplaces-and-governance-lessons-from-china/

instead of chasing waivers for years.
- Level playing field – rural colleges or small manufacturers could get the same data power as large institutes or corporations.
- Privacy preserved – because no real identities are ever shared, compliance with HIPAA, FERPA and trade-secret law could be baked in from the start.

- **What standards and protocols should be established to ensure interoperability of scientific datasets across disciplines?**

  Clarify "Standards" - We distinguish between technical standards such as those established by Standards Developing Organizations (SDOs) like IEEE, NIST, ISO, or W3C, which define file formats, protocols, APIs, and ontologies; and normative standards such as community-adopted principles like FAIR (Findable, Accessible, Interoperable, Reusable) or CARE for Indigenous data, which shape data ethics, sharing practices, and usage expectations. Both types are critical to enabling seamless data interoperability across scientific fields.

  Recommendations:

  1. Adopt FAIR Data Principles - Require all federally funded datasets to adhere to FAIR principles, with metadata made machine-actionable. Federal agencies should build internal capacity to evaluate FAIR compliance (e.g., metadata validators, semantic schema tools).

  2. Mandate Use of Persistent Identifiers[54] - Use DOIs (Digital Object Identifiers) for datasets and ORCID iDs for researchers.

  Use RRIDs (Research Resource Identifiers) for key materials, instruments, and protocols. Persistent Identifiers help us keep track of who did what, where, and when. A vital piece of research integrity in today's digital world. They ensure that knowledge doesn't get lost and that people and data can be reliably connected over time.

  3. Data pipelines of the proposed National Science Data Cloud could be secured end-to-end with salted-tag guardrails and continuous encryption auditing similar to the framework outlined in this paper[55].

  4. Standardize Data Formats and Metadata Schemas - Promote domain-specific standards: For example, Biomedicine: HL7 FHIR, BioSchemas, OMOP CDM or Physics/Engineering: HDF5, JSON-LD, IEEE 2791. Require schema.org-compliant metadata across repositories for semantic discovery. Schema.org-compliant metadata is just a standardized way to label your data so machines can find and make sense of it.

  5. Use OpenAPI + JSON-LD for Machine-Readable Access - All scientific datasets should be exposed via RESTful APIs documented using OpenAPI and enriched with JSON-LD for cross-dataset linking.

  6. Develop a U.S. Scientific Interoperability Profile Registry - Create a federated metadata interoperability registry hosted by NIST or NSF, with domain-specific schema mappings and controlled vocabularies (aligned with NIH, DOE, and NOAA).

- **What are the biggest blockers preventing researchers from sharing high-value scientific data today? What technological solutions could allow researchers to analyze sensitive data without compromising privacy?**

  Regulatory frameworks like HIPAA and GDPR create additional compliance requirements that many research institutions struggle to navigate. Many institutions lack data management systems that can handle complex, heterogeneous datasets while maintaining security standards. Data quality and standardization challenges further complicate sharing efforts.

  Several emerging technologies offer promising approaches to enable sensitive data analysis while maintaining

[54] Persistent Identifiers help keep track of who did what, where, and when. A vital piece of research integrity in today's digital world. They ensure that knowledge doesn't get lost and that people and data can be reliably connected over time.
[55] https://doi.org/10.32996/jcsts.2025.7.1.2

privacy protections.

Federated Learning and Homomorphic Encryption provide advanced methods for collaborative data analysis. Federated learning allows distributed computation across multiple data holders without requiring centralized data collection. Instead of moving sensitive data to a central location, computational algorithms are distributed to where data resides, with only aggregated results shared between participants. Homomorphic encryption enhances this approach by enabling mathematical operations on encrypted data, allowing analysis to proceed without ever exposing the underlying sensitive information.

Secure Multi-Party Computation and Differential Privacy offer complementary privacy-preserving approaches. These methods enable multiple organizations to jointly analyze their combined datasets while keeping individual data sources protected. Differential privacy adds carefully calibrated noise to query results, preventing the identification of individual data points while preserving overall statistical patterns.

Trusted Execution Environments (TEEs) provide hardware-based security solutions for sensitive data processing.

Synthetic Data Generation represents a different approach to privacy-preserving research collaboration. This technology creates artificial datasets that maintain the statistical properties of original data without containing actual sensitive information.

- **What new data infrastructure is needed to handle the scale and complexity of emerging scientific data?**

  See above discussion of National Science Data Cloud.

  Standardize data formats and meta data. To make data easier to share and reuse, researchers and data scientists should align with the NIH recommended FAIR data principles (Findability, Accessibility, Interoperability, and Reusability).

  National AI Cyber Shield for Science Infrastructure & federal systems - create a real-time, AI-powered cybersecurity defense system[56] to protect America's scientific infrastructure, federal systems, university research systems, biotech companies, and energy R&D networks. Implementing a zero-trust, encryption-layered design, similar to the framework as outlined in an article in the *Journal of Computer Science and Technology Studies,*[57] would provide guardrails against prompt injection, vector poisoning, and jailbreak attacks.

  What It Does:
  - Uses AI for continuous threat detection, anomaly detection, and autonomous response across high-value research environments.
  - Focuses on IP theft prevention, ransomware mitigation, and integrity of experimental data.
  - Provides threat intelligence sharing between public institutions and trusted private partners in real-time.

  Why It's Needed:
  - Scientific data is now a geopolitical asset—especially in biotech, quantum, AI, and energy.
  - Universities and startups often lack the resources for state-of-the-art cyber defense.
  - Protects the very foundation ASAP is trying to accelerate.

  Congress' Role:
  - Fund the creation of a National Science Cyber Command Center (NSC3), led by DHS, NSF, and NIST.
  - Establish cyber standards and AI-based certification protocols for federally funded labs and institutions.

[56] U.S. Government Accountability Office (GAO) finds the National Cybersecurity Strategy lacking in strength, with agencies still limited in protecting federal systems and data. Critical infrastructure sectors remain exposed to the risk of major cyber disruptions. https://www.gao.gov/blog/what-are-biggest-challenges-federal-cybersecurity-high-risk-update

[57] https://doi.org/10.32996/jcsts.2025.7.1.2

- Mandate breach reporting and coordinated incident response for research-focused organizations.

Private Sector Incentives:
- Cybersecurity R&D tax credits for AI models, tools, or services deployed in scientific environments.
- Federal matching grants for companies securing AI training data or federated research models.
- Procurement preference for security vendors offering transparency and AI-based intrusion detection systems for academic and public labs.

Create a National Science Data Grid and offer compute credits to institutions - Develop a federated AI compute and data-sharing infrastructure for national labs, universities, and startups working on scientific innovation. Democratize access to high-performance computing and curated datasets for scientific discovery.

Create Autonomous Science Labs: Develop "self-driving laboratories" that use AI and robotics to autonomously test hypotheses, run experiments, and refine theories.
- Goal: 10x the pace of physical experimentation in material science, chemistry, and bioengineering.
- Congressional Role: Establish pilot programs through ARPA-H, ARPA-E, or NIST.

- **How can we balance data privacy and security with open access to scientific data?**

  No answer.

- **How can we create sustainable funding models for data infrastructure and maintenance, and how can we improve academic incentives to reward researchers who contribute to datasets?**

  No answer.

COMPUTING Pillar

7. **What does the U.S. need to do to ensure its researchers have access to enough COMPUTING resources to power new breakthroughs?**

The United States is currently operating under the premise that the only way to innovate is by having access to large models. How do we know there is not a better method? If we ensure higher quality of data, you may not need large quantities of data? Is massive amounts of data the goal? We must ensure that USG also incentivizes research that does not require as many resources.

Create a National Compute Exchange (NCX): Transform compute into a commodity.[58] Compute inequality is hindering American scientific leadership**.** AI models run on data the way jet engines run on fuel (a commodity). Yet real-world datasets are locked behind HIPAA, FERPA, trade-secret, and copyright walls. Small labs and start-ups struggle to obtain high-quality, legally usable data.

Cutting-edge research now rises or falls on access to powerful computers, yet America's GPU clusters and cloud offerings are scattered, priced in "take-it-or-leave-it" lists, and often sit idle between projects. This drives up costs for small labs and forces agencies to over-buy hardware. The result is wasted capacity, inflated costs, and slower science—exactly the opening America's competitors are exploiting.

Direct NSF, DOE, and NIST to launch a National Compute Exchange (NCX)—one transparent, real-time marketplace where universities, startups, federal labs, and private clouds can post or rent spare GPU and CPU hours (like travelers now enjoy on Airbnb or investors on Nasdaq). By letting supply and demand meet in a single window, the NCX would slash prices, keep taxpayer-funded machines fully used, and open top-tier computing to community colleges and minority-serving institutions.

Think of an NCX as an Airbnb for computing where any spare GPU can be rented. While Strategy 5 of the 2023 National AI R&D Strategic Plan wisely calls for "expanded access" to high-performance computing and creates new computing infrastructures, it never explains how thousands of scattered GPU clusters and cloud spot markets will become truly affordable and discoverable to every researcher.

Clear price signals would also guide future CHIPS-Act investments to the places America genuinely needs new capacity. A lightweight, open-source API "wrapper" would let any cloud or super-computer publish its spare cycles without changing its internal architecture; jobs (eg: AI model training) could slide from one provider to the next without code rewrites. Every transaction would carry a tiny fee that flows back to the owner as a micro-royalty, turning sunk hardware costs into a revenue stream. In short, the NCX would turn today's idle computing resources into an engine of faster, cheaper, and more inclusive innovation—delivering a win for fiscal responsibility, economic competitiveness, and scientific discovery alike.

Every federal lab, university center, and willing cloud vendor installs and publishes willing (or unused) CPU/GPU hours to a shared marketplace—no changes to their internal systems. Researchers bid or pre-book time; owners are paid automatically. Top-level research now depends on short bursts of GPU power, yet government clusters idle between grants and cloud prices stay high. Small colleges, start-ups, and many agency programs pay more—or wait longer—than they should. Certified Synthetic data can overcome the restrictions of privacy and IP. In a way, the data generators can get real or synthetic data certified (for privacy, bias, quality, IP, etc. using IEEE standards) and trade in an exchange as a commodity. E.g., a start-up that lacks real-world medical records could buy a certified batch of HIPAA-safe patient data the same way it rents two hours of GPU time.

In short, the NCX would turn idle computing resources into an engine of faster, cheaper, and more inclusive innovation—delivering a win for fiscal responsibility, economic competitiveness, and scientific discovery alike.

Funding & sustainability: A one-year, $X M NSF start-up grant could cover software and cybersecurity. After launch, a X % transaction fee could fund operations.

---

[58] China has a Data Exchange - https://www.cigionline.org/articles/data-marketplaces-and-governance-lessons-from-china/

Measuring success: track median cost per GPU/CPU hour, utilization of listed clusters, and the share of bookings going to under-resourced institutions. Real-time price data highlights capacity gaps, guiding CHIPS Act investments against competitors.
Impact - Transparent bidding could drive X % price drops, it could let agencies stop over-buying hardware and could widen access to frontier compute—at minimal new cost to taxpayers.

- **How should we construct public-private partnerships for public sector computing infrastructure that increase availability and reduce cost for high-performance computing and AI?**

Throughout this document we discuss the power of public/private partnerships as important to advancing AI innovation - everything from ensuring that we're educating, training, and upskilling our workforce to creating shared resources that enhance access to the tools and resources necessary for research and development. For example, see answer below – PROCESS pillar question – regarding leveraging AI to advance discoveries in the healthcare domain. Some additional suggestions below.

Adopt a Hybrid, Multi-Cloud Architecture for public sector agencies – Public sector agencies such as Federal labs and defense agencies juggle two tough demands: massive bursts of computing power for AI and simulations, and strict security for sensitive or classified data. Buying enough on-premises machines to cover every spike is expensive; moving everything to one public-cloud vendor risks lock-in and security gaps.

Adopt "best of both worlds" approach which uses both public cloud for scalability and on-premises federal high performance computing resources for security & control.

The U.S. government - specifically research agencies (that deal with sensitive or classified data), national labs (that run big AI or scientific workloads or temporary, compute-heavy projects), and the Department of Defense - should adopt a hybrid, multi-cloud architecture.

This approach could let public sector to move workloads around based on cost, performance, or security needs. It could also help them handle spikes in computing demand without overpaying for resources they rarely use. This could help the public sector dynamically scale up or down as and when needed, keep data safe, avoid being stuck with one cloud provider, and save money by only paying for what they use.

Reaffirm the support for public-private decentralized standards development model - As the U.S. focuses on revitalizing America's science and technology enterprise and ensuring scientific progress and technology innovation, the role of global technical standards has become increasingly important. The U.S. public-private partnership model for standards development has been a cornerstone of American innovation, competitiveness, and security. It leverages the expertise of industry, academia, and civil society to create flexible, technically sound standards that are globally respected and market driven.

Competing governance models can threaten the openness, transparency and democratic values embedded in this model of standardization. There are also proposed legislative challenges surrounding incorporation by reference (IBR) of standards and could compromise the funding and sustainability of Standards Developing Organizations (SDOs) and the broader standardization ecosystem, potentially hindering a foundation of the U.S. innovation ecosystem. Without robust voluntary standards produced in such a standards development model, there may be the potential to rely on prescriptive and fragmented regulatory approaches that can further hinder innovation and global market access. In an era of global technological advancements, growth and competition, ensuring the integrity and sustainability of the public-private standards development infrastructure is key, as voluntary, consensus-based standards are a critical pillar for innovation and growth

- **What role should distributed, federated, and decentralized computing models play in the scientific research ecosystem?**

  All three computing models – distributed, federated, and decentralized – play important roles in advancing scientific research. **Distributed computing** spreads tasks across multiple systems to process large amounts of data quickly. It offers high scalability and is usually controlled by a central authority, often using cloud clusters. **Federated computing** allows researchers from different institutions to work together without sharing sensitive raw data. This helps protect privacy and follow regulations. **Decentralized computing** uses peer-to-peer networks and technologies like blockchain and distributed ledger technology (DLT) to improve trust, transparency, and system resilience. It supports high fault tolerance and data sovereignty, with decision-making shared across participants. Key technologies also include secure multi-party computation for privacy and collaboration. See chart below.

Note that decentralized ML is mostly theoretical or niche with limited practical deployment due to complexity and lack of commercial drivers—though it has strong potential in AI democratization, especially when combined with Web3 ideas.

	Distributed	Federated	Decentralized
Scalability	High	Medium	Medium
Data Sovereignty	Low	High	High
Fault Tolerance	Medium	Medium	High
Governance Model	Centralized	Collaborative	Peer-to-peer
Key Technologies	HPC, cloud clusters	Federated learning, secure multiparty computation	Blockchain, IPFS, DLT

- **What benchmarking improvements do we need to understand the value provided by computing systems, and how should we best measure the strength of U.S. public sector compute against what is available in other nations?**

  Align with Global Standards: Work with organizations such as MLCommons, SPEC, and RISC-V International to shape global benchmarking standards for AI and high-performance computing. To compare with and measure against, U.S. public sector computing with other countries:

  - We need a clear and fair way to measure capabilities. One way is to create a **National Capability Index** that combines key areas like hardware performance, AI power, quantum readiness, network strength, and access to data.
  - We should also look at how easily researchers can use these systems considering policies, funding, and collaboration across agencies. It's important to measure how effectively each country turns computing power into real scientific results, such as discoveries, publications, and innovations. Tracking leadership in AI like model training capacity, datasets, and shared AI platforms is also critical.
  -Finally, encouraging global cooperation on benchmarking standards and open reporting will help ensure transparency and support healthy international competition.

- **What specific breakthroughs in hardware are needed to sustain accelerated scientific progress?**

  No answer.

- **How can edge computing be integrated into scientific workflows to accelerate data processing?**
  Edge computing can accelerate scientific workflows by processing data near where it's collected like in sensors, labs, or field sites allowing faster analysis, reducing the need to send large datasets to central servers, and enabling real-time decision-making. It's especially useful in remote locations, for real-time experiments, and in privacy-sensitive research. Some examples of where this is already being done (Integration with Cloud tools examples):

  – AWS IoT Greengrass / Azure IoT Hub / Google Edge TPU – Manage and sync edge workloads with cloud-based scientific pipelines.
  – Globus / iRODS – For secure, automated data movement between edge systems and research data centers.

AI Pillar

8. **What should America do to take full advantage of AI capabilities to dramatically accelerate the pace of science in both the private sector and the public sector, and what innovations should we target in the foundations of AI itself?**

To fully capitalize on AI's transformative potential, the United States must establish a national AI innovation infrastructure that unifies public and private efforts—streamlining access to compute resources, research funding, and trusted data ecosystems. Cross-agency collaborations should be deepened, enabling AI models to accelerate scientific discovery in climate modeling, drug design, materials science, and quantum simulation. Strategic investment in open foundational models—including multilingual, domain-specialized, and multimodal architectures—can democratize innovation and prevent monopolization of AI capabilities. The U.S. must also prioritize talent development and ethical AI governance, ensuring responsible deployment across sectors.

Looking ahead, the United States should lead global efforts in the development and containment of Artificial General Intelligence (AGI)—systems that match or exceed human cognitive abilities across tasks. The possibility of Artificial Superintelligence (ASI), with capabilities far beyond human understanding, warrants urgent research into alignment, interpretability, and safety. AGI/ASI could unlock radically new scientific paradigms—self-generating theories, recursive experimentation, and autonomous innovation—but only if safety keeps pace. To that end, innovation in AI alignment mechanisms, neurosymbolic reasoning, and value-centric architecture design is critical to ensure beneficial outcomes.

What should we target? While today's large-language models can draft papers and design molecules, they can make mistakes, and they cannot see. False results could spread fast, and human reviewers cannot keep up.[59]

Create AI systems that can prove their own answers before a human ever sees them. Neural-Symbolic[60] Proof Loops. Combine deep-learning with automated theorem proving so every prediction could come with a machine-checkable certificate (think mathematical proof stapled to each claim). These systems could result in fewer retractions, faster trust. For example, a chemist could let the engine screen thousands of reactions overnight and wake up only to certificates that already satisfy lab-safety rules. Or, private firms could get audit-ready AI and regulators could get instant transparency.

Success metric - By 2030, ≥ X % (e.g., ≥ 80%) of federally funded AI models used in drug, materials, or climate work should issue a machine-verifiable proof of each new claim within X minutes (e.g., 10 minutes) of inference.

Invest in domain experts (i.e., scientists) - Take full advantage of the workforce - the scientists and engineers - who can build and continually enhance AI technologies. AI is not at a stage where it can replace scientists, but it is instead a very useful tool to aid them in their work. Continue to educate, train, and employ scientists in a variety of fields where AI capabilities can be applied and new AI methods can be advanced. It is impossible to silo AI research from its applications, as those very same applications inspire and motivate advances in AI capabilities that can then be applied to other areas of life.

Continue investment in science centers and research institutes[61] - These institutions provide fertile ground for collaboration and the exchange of ideas that will speed up both the application and advancement of AI methods.

- **What breakthroughs in modeling could decrease the need for expensive or slow real-world experiments? How should we best combine modeling and experiment to maximize our scientific knowledge?**

Regardless of what we do with AI, there's always going to be a role for real experiments. We need to be very

[59] https://pmc.ncbi.nlm.nih.gov/articles/PMC12045364/

[60] Neuro-symbolic AI aims to bridge the gap between the pattern recognition capabilities of neural networks and the logical reasoning abilities of symbolic AI. This integration offers the potential for more robust, explainable, and efficient AI systems. Refer - https://mitibmwatsonailab.mit.edu/category/neuro-symbolic-ai/

[61] https://www.nsf.gov/funding/opportunities/national-artificial-intelligence-research-institutes

careful with using AI to replace real experiments (or even numerical simulations). AI methods are currently not very interpretable, and AI-emulations of experiments will propagate biases in our understanding. Nevertheless, advances in AI that allow for faster/cheaper emulations of expensive simulations are essential. AI Interpretability, the ability to understand how a model is making a determination and how it might be going wrong, are critical to making AI emulations of simulations trustworthy.

Further integration of AI into data analysis of results from experiments, especially for experiments with datasets too large to analyze with traditional methods, is the best means forward. In many cases this does not require innovation on the part of AI, but instead the education and confidence of domain experts. Methods like outlier/anomaly detection, AI-assisted noise reduction, and simple AI-regression models can speed up data cleaning and the identification of interesting data points. The simpler the method, the easier it is to understand what it is doing and if it is working correctly.

Rather than cast a wide net, focus on high quality data. This activity could also address IP issues. A resource for this might be Andrew Ng of Stanford who pioneered the use of GPUs to train AI models. He is now advocating data centric AI models where one systematically engineers data to build an AI system. It would take far less data, would be environmentally more friendly and has the potential to minimize IP and copyright infringement issues.[62]

Advance Modeling to Reduce Reliance on Expensive Real-World Experiments - Accelerate the development and public access of AI-enhanced scientific models that reduce the financial and temporal burdens of experimentation. Prioritized investment in hybrid models—which combine physical laws with machine learning—can dramatically reduce cost across domains like climate modeling, biomedical research, and materials science.

Establish a National Digital Twin Infrastructure funded by NSF and DOE, where physical systems (e.g., energy grids, cell biology, supply chains) are mirrored by continuously improving AI models.

Incentivize federally funded simulation toolkits that serve as public infrastructure, similar to what TensorFlow did for AI.

Use AI to Drive Hypothesis Generation and Experimentation - AI should not only interpret scientific data but also assist in formulating and testing hypotheses. This creates a powerful feedback loop where AI proposes the next-best experiment, simulation, or question—saving years of trial and error.

Fund interdisciplinary AI-Science Integration Hubs at research institutions that merge computer science with biology, chemistry, and materials science.

Develop AI-powered "National Laboratories-in-the-Loop," where AI systems help prioritize federally funded research activities and grant disbursements based on prospective impact.

Continue the creation and essential funding of AI for Science centers at universities and research institutions. The NSF recently established two major such hubs for astronomy and astrophysics, as part of a larger program.[63] These are private-public partnerships that involve funding from outside philanthropic organizations like the Simons Foundation.

- **How can AI accelerate the generation and testing of new scientific hypotheses? How should we construct scientific research models where AI can be used to iteratively drive simulation or experimentation to achieve a particular research goal?**

The most straightforward way that AI can advance scientific experimentation is by doing more of the data analysis grunt work, i.e., data cleaning, anomaly detection, etc.

A second area of benefit would be for LLMs to accurately and succinctly summarize existing literature on a

[62] *IEEE Spectrum*, April 2022
[63] https://www.nsf.gov/news/nsf-simons-foundation-launch-2-ai-institutes-help

subject to allow scientists to better synthesize it. With the ever-increasing number of papers, it is almost impossible to stay current. If these systems were sufficiently trustworthy, i.e., they directly referenced the literature in a way that allowed users to quickly check the summary's veracity, this would speed up hypothesis generation.

A third area of interest is "symbolic regression" methods that use machine learning to construct analytical expressions from data. In principle these are more easily interpretable and less prone to overfitting (i.e., more robust). In practice they are still difficult to interpret and require a lot of finesse (like most AI methods) to produce useful results. Further work in methods like this is needed. If such a model could work reliably, it would allow researchers to quickly construct human-interpretable analytical models of relationships found in data. Predictions made by these analytical models could then be tested.

- **What bottlenecks limit development of, or access to, AI-driven research tools?**

Breaking Bottlenecks to Democratize AI-Driven Science - Currently, access to powerful AI research tools is limited by compute inequality, lack of data standardization, and fragmented infrastructure. These risks concentrating innovation and marginalizing smaller institutions, especially HBCUs, Tribal Colleges, and MSIs.

- Expand access to national compute resources via AI Compute Vouchers for under-resourced institutions.
- Require federally funded research to release structured, FAIR (Findable, Accessible, Interoperable, Reusable) datasets and pre-trained models for public use.
- Standardize model documentation and reproducibility requirements across federal R&D grants.

- **What should a democratic AI research ecosystem look like? What lessons can we learn from earlier American-led efforts such as the development of the internet?**

A democratic AI research ecosystem is one that ensures:

- Open Infrastructure, Open access to research tools, datasets, and compute resources;
- Broad participation from academia, startups, minority-serving institutions (MSIs), nonprofits, and international allies;
- Transparency and reproducibility of results;
- Ethical governance that includes civil society, labor, and impacted communities; and
- Sovereign resilience, ensuring that infrastructure is controlled by public or allied entities—not monopolized by a handful of private firms or foreign adversaries.

Lessons from the Development of the Internet - U.S. leadership in the internet era was possible because the core protocols and access mechanisms were public goods, not proprietary walled gardens.

Internet Pillar AI Research Analogy:

DARPA-funded ARPANET – Federally funded AI-HPC grid for open scientific research
NSFNET scaling & academic access – Publicly accessible AI research cloud + data commons
IETF open standards model – Transparent, consensus-driven AI governance protocols
TCP/IP as a public standard – Open-source foundational AI models and APIs
Private sector commercialization – Incentivized innovation but grounded in open infrastructure

- **What foundational innovations are needed in AI, such as in areas like interpretability, energy efficiency, and uncertainty quantification?**

Interpretability and uncertainty quantification are key to fully integrate AI into scientific workflows. For AI to be useful for scientific research, AI models should improve understanding by directing our attention to new relationships in the data and should be trustworthy, i.e., not learning from artifacts or non-representative features in the training set. In the context of AI interpretability, this means further research into understanding model latent space is needed, whether that be through forms of saliency mapping, symbolic regression, contrastive learning, or other means. Uncertainty quantification is also essential. Advances in ML inference techniques using Normalizing

Flows have been very helpful. Further work needs to be done in terms of mitigating/quantifying distributional shift, which is a major source of bias in simulation-based inference.

To aid research in these core areas, Congress should continue robust federal support for scientific research across a variety of fields. AI is already an important component to work done throughout the scientific workforce. Pure AI research is important, but scientific research provides an excellent and unique proving ground for experimenting with novel methods, especially in the areas of interpretability and uncertainty quantification.

COLLABORATION Pillar

9. **How can we radically increase the scale, speed, and impact of scientific collaboration across disciplines, institutions, and sectors?**

Incentivize collaborative research. Assign points to federal funding requests. Higher points/ higher grades mean you get the funding.

. *Establish a National AI Commons* **-** Democratize access to high-performance AI models and infrastructure.

- Fund and govern a secure, interoperable AI infrastructure (compute, storage, datasets) available to startups, researchers, and nonprofits—like a digital public utility.
- Create IP-safe sandboxes where innovators can train and deploy models using sensitive or proprietary datasets under clear rules.

Establish a National AI Sandbox Network to speed up scientific breakthroughs. Create a federated, secure, high-performance set of AI environments to which researchers, startups, and public institutions would have access to advanced models, computing capabilities, and sensitive data—all within regulatory-aligned, standards-compliant sandboxes. The testbeds might target key domains such as cancer research, climate modeling, drug discovery, and real-time biosecurity response.

Congress could create a network modeled after DARPA or the NAIRR, and encourage federal agencies, academia, and industry participation via tax credits and preferred vendor status incentives, actions that could encourage private sector cloud and AI companies to donate infrastructure. NIST and the relevant standards bodies could contribute to interoperability guidelines and governance.

PROCESS Pillar

10. How can the U.S. radically accelerate the introduction of breakthrough technologies to the market without sacrificing safety and public trust?

Congress could mitigate the risks that the use of AI in scientific research presents for replicability by mandating existing recommendations from both the public and private sector. For example, see recommendation 4 of the 2024 PCAST report: *Adopt principles of responsible, transparent, and trustworthy AI use throughout all stages of the scientific research* process.[64] Also see IEEE-USA's, *How Should We Regulate AI? Practical strategies for regulation and risk management from the IEEE1012 Standard for System, Software, and Hardware Verification and Validation*.[65]

- **What foundational changes should we consider in how funding agencies sponsor research? What already works well that we should double down on?**

  No answer.

- **In what ways should intellectual property laws evolve to provide greater regulatory clarity and better facilitate AI-accelerated scientific discoveries?**

  Intellectual property laws are designed to reward those who create. If uncertainty in laws and regulations would lead a creator to believe they may not be able to reap the rewards of their efforts, innovation and ingenuity may be stifled, no acceleration of discovery.

  To ensure that US innovation continues to occur at the rates that Americans have become accustomed to, Congress should support and promote existing laws that guarantee a creator/innovator's constitutionally protected intellectual property rights.

  Congress is faced with the challenge of answering the question, 'who is the creator (inventor or author) when an AI system is used when creating?' The US Constitution[66] and US patent and copyright laws already answer this question. The constitution grants, "*Authors and Inventors the exclusive Right to their respective Writings and Discoveries.*" Authors and inventors are human beings.[67]

  The law should ensure that the right to protect IP is predictable and effective, and ensure that innovators can facilitate licensing of their ideas and are able to commercialize their AI-related inventions, innovations and materials. Predictability in IP rights laws also ensures that innovators can obtain capital which in turn fosters additional research and development. Congress should:

  - Pass the PREVAIL Act (PTAB reform) and PERA (subject matter eligibility reform bills that have been introduced in both the 118th and 119th Congress;
  - Encourage copyright reform to address copyrightability issues and privacy concerns, such as materials developed using personal data;
  - Return permanent injunctive relief to its status as an automatic remedy after a finding that the patent was infringed;
  - Restore the pre-America Invents Act 35 USC 102 grace period; and
  - Improve the standard for 35 USC 103 obviousness.

- **How can peer review be modernized to encourage faster or more rigorous scientific validation? What role**

---

64 https://bidenwhitehouse.archives.gov/wp-content/uploads/2024/04/AI-Report_Upload_29APRIL2024_SEND-2.pdf

65 *How Should We Regulate AI? Practical strategies for regulation and risk management from the IEEE1012 Standard for System, Software, and Hardware Verification and Validation*. https://ieeeusa.org/product/how-should-we-regulate-ai/

66 Article I, Section 8, Clause 8, US Constitution

67 https://www.merriam-webster.com/dictionary/author and https://dictionary.cambridge.org/us/dictionary/english/inventor

should emerging technologies play in analyzing the quality of new research?

While peer review practices vary from scientific field to scientific field, there are some common challenges that ought to be addressed. Reviewers are often not incentivized to do a thorough job. Reviewing others' work is an often unrewarding, albeit time consuming, task for scientists. Reviewers are often anonymous, thus unrecognized, and usually unpaid. Solutions to this problem will vary depending on context. Two possible solutions, which are complementary, include a) requiring those who submit manuscripts or proposals to also review the manuscripts or proposals of others, and b) recognizing reviewing as a duty of research grant recipients and compensating them for their work. The former solution would need to be implemented at the journal level and is already a part of some research grant proposal systems. The latter solution could be implemented by Congress as a stipulation to grant-awarding federal agencies like NASA, the NSF, NIH, etc.

Emerging technologies, like LLMs, can be useful to the peer review process, but also pose dangers. LLMs can be useful to reviewers by analyzing their reviews, evaluating their clarity and tone. This can improve communications between reviewers and authors. Similarly, LLMs can aid authors by anticipating possible questions a reviewer might have. Future LLMs might also be capable of summarizing works, or even reviewing works themselves. However, present-day artificial intelligence is insufficiently trustworthy to review works by itself, or to create reliable summaries. Instead, AI-reviewers risk propagating existing biases, grossly misinterpreting results, and being insufficiently scrutinizing. For the time being human domain-experts must personally review new research thoroughly and completely, without overreliance on AI-assisted summarization or review.

- **What new institutional structures should we experiment with to accelerate innovation?**

  No answer.

- **In the healthcare domain, how can artificial intelligence be strategically leveraged to accelerate processes within federal agencies? Conversely, how might these agencies deploy AI-driven solutions to streamline operations in order to support a research community that increasingly depends on rapid technological advancements?**

  The challenge is to enable real-time responses to future pandemics or national emergencies. The USG should explore:

  - Building a national AI observatory to combine data from, and share data between, federal agencies. This could enable the ability of decision makers to generate rapid decisions.
  - Creating a cross-agency rapid response fund for AI-driven projects tackling national crises.

  In general, federal healthcare agencies should establish strategic goals that:

  - Accelerate regulatory approvals.
  - Incentivize public-private partnerships to develop AI tools that predict molecule behavior, optimize clinical trials, and identify therapeutic targets.
  - Offer challenge grants or tax incentives for pharmaceutical companies sharing anonymized datasets.
  - Streamline data integration, reporting, and analysis
  - Increase federal support for biomedical research and interagency collaboration.
  - Procure and use AI systems to accelerate processes at agencies such as FDA, NIH, ARPA-H.

  Specifically, AI systems could:

  Optimize drug discovery & clinical trials - AI could simulate molecular interactions, predict toxicity, and identify promising compounds, as well as analyze historical trial data to suggest better designs, populations, and endpoints.

  Example tools: AlphaFold, BioGPT, DeepChem.
  Impact metric: Reduced timelines to Investigational New Drug (IND) approval, 40-60%.

Automate regulatory review and document analysis - Use natural language processing (NLP) to summarize and analyze clinical trial applications, public comments, and inspection reports. This could result in accelerated approval timelines and reduce backlog at FDA and CMS.

> Example tools: GPT-powered summary tools, document triage bots, semantic search.
> Impact metric: Review cycle time, # of expedited reviews processed per year.

Create an Open Science High Compute Grid for Cancer and Human centered AI Innovation, an "AI BioShield" - A public/private partnership moonshot model (PPP) that could support AI-HPC (High Performance computing) and AI model training with resources like GPUs, TPUs, and large AI models for use by researchers, startups, and public institutions or labs. Various discoveries in cancer research, genomics, and drug design are limited not because of lack of human ingenuity, but because of lack of access to a complete infrastructure. Only big labs or tech companies have access to the necessary resources or can scale needed to run advanced AI workloads. The Bioshield could enable rapid therapeutics development and reduce time-to-market to months, from years.

In the PPP model, federal agencies (NSF, NIG, DOE) could partner with cloud and AI companies to share their infrastructure as part of their environmental, social or governance (ESG) decision-making processes. In return the private partners could receive incentives such as R&D tax credits or receive preferred vendor status on federal AI projects. The relationship could be similar to how the CHIPS act formalized U.S. semiconductor investments. Oversight would come from a body of representatives from federal agencies (NIH, NSF, DOE), industry, and academia.