IOSR Journal of Research & Method in Education (IOSR-JRME)
e-ISSN: 2320–7388, p- ISSN: 2320-737x Volume 15, Issue 4 Ser. 1 (May. – June. 2025), PP 08-18
www.iosrjournals.org

Aligning the United States Land-Grant Institution Research Enterprise with Modern Computational Research Needs

Jason A. Hubbart[1,2]
[1]*Division for Land-Grant Engagement, Davis College of Agriculture and Natural Resources, School of Natural Resources and the Environment, West Virginia University, Morgantown, WV 26506, USA; jason.hubbart@mail.wvu.edu.*
[2]*West Virginia Agriculture and Forestry Experiment Station, West Virginia University, Morgantown, WV 26506, USA; jason.hubbart@mail.wvu.edu.*

Abstract:
Public research universities, particularly land-grant institutions in the United States of America (USA), face mounting fiscal pressures, aging physical infrastructure, and intensifying competition for external funding. Simultaneously, the rapid decline in the cost of high-performance computing and the rise of scalable cloud services have redefined what is possible in data-intensive research. It is presented herein that transitioning from traditional field and bench-based methodologies to computational, cloud-enabled workflows is no longer optional but strategically essential. Drawing on representative discipline examples, illustrations are provided of how researchers can leverage affordable cyberinfrastructure to generate research results faster, at lower cost, and with improved reproducibility. Across domains, examples demonstrate that modest investments in GPU nodes, research software engineering, and reproducible pipelines yield superior returns in terms of funding success, publication speed, and employability. Funding agencies are reinforcing this shift by embedding computational readiness and data-sharing expectations into proposal criteria, while industries are increasingly demanding graduates who are fluent in data science and algorithmic thinking. This paper provides a framework for institutional transformation, outlining how public universities can align infrastructure, workforce development, and academic culture with the computational imperative. By embracing computational research as a core research modality, institutions can transform resource scarcity into a competitive advantage, accelerating discovery, increasing funding leverage, and fulfilling their public mission in a data-rich century. The manuscript concludes with an implementation roadmap for principal investigators, centers, and leaders, emphasizing the urgency of building durable, scalable, and mission-aligned cyberinfrastructure as the foundation for institutional resilience and relevance.
Key Words*: Cyberinfrastructure; Computational Research; High-Performance Computing (HPC); Cloud Computing; Data Science; Reproducibility; Land-Grant Institutions; Workforce Development; Funding Agencies.*

Date of Submission: 01-07-2025 Date of Acceptance: 09-07-2025

I. Introduction

Public research universities, including land-grant institutions, particularly those in the United States of America (USA), face significant challenges to their funding models. State appropriations as a share of total revenue have fallen, on average, to levels unseen since the 1970s, while single-PI success rates at major USA agencies now hover below 20 percent [1,2]. Concurrently, the postponement of necessary maintenance for field stations and animal units imposes a multibillion-dollar burden on operational budgets. Conversely, researchers have never had such affordable access to petascale cloud nodes, open-source analytics tools, and exabyte-scale public datasets. For example, leasing a 64-GPU virtual machine or computing environment for a day now costs less than sending a graduate student into the field for a week, thereby eliminating the financial costs that once separated computational laboratories from other research types [3].

The convergence of computationally affordable research infrastructure and technologies with other research disciplines is fundamentally altering the scientific method. High resolution, high accuracy simulation is now recognized as the "third pillar" of inquiry and data-driven discovery as an emergent "fourth paradigm," co-equal with theory and experiment [4,5]. For example, hydrologists can now interrogate global satellite data archives to predict discharge in ungauged basins at 1-kilometer resolution [6]. Foresters map wildfire susceptibility with centimeter-scale LiDAR models trained on open cloud Graphics Processing Unit (GPU) stacks

[7]. Such breakthroughs result in journal-ready articles in weeks rather than field-season cycles, offering cost savings that can exceed an order of magnitude.

Funding agencies are taking notice. The United States of America (USA/US) National Science Foundation's AI Research Institutes, in partnership with the US Department of Agriculture (USDA), US Department of Energy (DOE), US National Institutes of Health (NIH), and US National Institute of Standards and Technology (NIST), now channel nine-figure investments into agriculture, climate resilience, and smart infrastructure, explicitly favoring campuses that demonstrate mature cyberinfrastructure and reproducible workflows [8]. For example, the Extreme Science and Engineering Discovery Environment (XSEDE) was a National Science Foundation (NSF)-funded virtual organization that provided researchers with access to advanced digital resources and services, like supercomputers, data storage, and visualization tools, to support computational and data-intensive research. A 2023 cross-campus analysis of XSEDE partner sites showed that each dollar invested in shared clusters leverages approximately five external grant dollars and accelerates publication output by 40 percent [9]. Given these advancements, it is not surprising that employers across sectors rate data fluency as a top hiring criterion. Ironically, fewer than one-third of non-computer science programs require substantive coding experience, underscoring an urgent workforce gap [10].

Considering the preceding context, the arguments presented in this article will emphasize that shifting from resource-intensive field and bench protocols to data-centric, computational research is not merely an increase in efficiency; it is a strategic imperative for scholarship, talent cultivation, and public relevance. Select discipline-specific transitional strategies are outlined in the following areas: Forestry, Fish and Wildlife, Animal Sciences, Agriculture, Landscape Architecture, Economics, Innovation, Civic Engagement, Workforce Development, Environmental Stewardship, and Education. For each discipline, example workflows and funding pathways are offered as springboards for initiation. Finally, an implementation framework will be provided in response to and aligned with the National Academies' call to treat cyberinfrastructure as "mission-critical" for the land-grant system. Notably, although a roadmap is not provided for all disciplines, the examples provided share many commonalities in approach that can guide other fields. Furthermore, although funding venues may shift, the fundamental changes and processes described herein are anticipated to remain constant for the foreseeable future. By embracing the computational research enterprise, universities can transform fiscal austerity into an engine of discovery, turning fieldwork research projects into cloudwork while graduating students who are prepared to lead in the data-rich century that lies ahead.

II. The Computational Imperative

The logic for pivoting toward computational research rests on three mutually reinforcing realities: (a) unprecedented cost, speed, and scope advantages; (b) clear signals from funding agencies that reward data-centric readiness; and (c) strong labor-market demand for graduates who can think algorithmically.

Cost, Speed, and Scope

High-performance computing (HPC) and elastic (scalable in terms of power, storage, memory, and cost) cloud services have lowered the entry barrier that once confined intensive modeling and analytics to a handful of elite laboratories. For example, in 2000, a teraflop of processing power in US dollars (USD) cost more than $40 million; in 2025, the same capacity can be rented from a commercial provider for less than $5 per hour [3]. This price inversion means that a 48-hour GPU burst sufficient to run a continental climate and water ensemble now costs less than a week of field per diem for a graduate assistant. Speed gains are equally remarkable: hydrologists training deep-learning discharge models shorten calibration cycles from months to days [6]. Foresters mapping wildfire susceptibility can refresh statewide risk layers overnight rather than once per field season [7]. Crucially, once a workflow is containerized, marginal analyses are almost free. New scenarios require computing time but no additional travel, consumables, or animal-unit overhead.

Funding-Agency Signals

Federal and philanthropic sponsors have aligned their solicitations with this new reality. The US National Science Foundation now embeds cloud-credit supplements and reproducibility mandates in many core programs, and its National AI Research Institutes competition explicitly prioritizes proposals that leverage shared cyberinfrastructure and open pipelines [8]. The USDA's Agriculture and Food Research Initiative (AFRI) introduced a dedicated computational track in 2023, which funds machine-learning and digital-twin approaches across the food–energy–water nexus. A return on investment (ROI) study of 16 US campuses found that every institutional dollar allocated to centrally managed clusters attracted $5 to $6 in new external awards within three years and shortened the time to publication by approximately 40 % [9]. Funding agencies, in effect, use proposal review criteria to encourage universities to adopt platforms that guarantee scalability, transparency, and rapid translation.

Workforce Necessity and Institutional Relevance

Industry demands that institutions aggressively pursue computational research. Surveys of 1,200 US hiring managers showed that data fluency now ranks alongside disciplinary depth as a top screening criterion for entry-level positions [10]. However, fewer than one-third of agricultural, life-science, and environmental programs (and many others) require substantive coding or statistical computing beyond an introductory course. The mismatch poses a reputational risk: universities that fail to equip graduates to handle large data sets become less attractive to students, employers, and donors. Conversely, campuses that weave computation through the curriculum strengthen their enrollment pipelines and reinforce their public-service mandate. This may be particularly relevant for land-grant institutions, whose Morrill-era mission included, among other mandates, democratizing cutting-edge knowledge [11].

Strategic Take-Away

Collectively, decreasing computing costs, sponsor-level incentives, and labor market dynamics create a wide-open, yet potentially brief, window for early adopters to transform fiscal constraints into strategic advantages. Researchers who redesign projects around simulation, machine learning, and data reuse can publish more quickly, secure additional grants, and produce students who are immediately ready for a digitized economy. For administrators, investing in shared cyberinfrastructure and research software engineer support is no longer just discretionary overhead; it has become a necessity for maintaining competitiveness in both scholarship and workforce development. In summary, the computational imperative has now evolved into a mission imperative.

III. Discipline-Specific Pathways and Funding Channels

The computational transformation will never be a one-size-fits-all exercise. The value of computational transformation will be recognized when workflows are customized to disciplinary epistemologies, data cultures, and funding ecologies. The following text distills ten representative research domains: Forestry, Fish and Wildlife, Animal Sciences, Agriculture, Landscape Architecture, Economics, Innovation and Entrepreneurship, Civic Engagement, Workforce Development, and Environmental Stewardship and Education. It illustrates (a) how researchers can shift from costly field or bench routines to data-centric methods, (b) the minimal cyberinfrastructure and human expertise required, and (c) the most active public and private funding streams that reward such shifts. Each domain integrates peer-reviewed exemplars and cites current solicitations or program rules, demonstrating that transitioning from fieldwork to cloud-work is both technically feasible and financially strategic (Table 1).

Forestry

Legacy fire-risk mapping relies on labor-intensive fuel-load transects that cover only a fraction of the landscape. High-resolution airborne LiDAR and Sentinel-2 imagery, processed through gradient-boosting and convolutional neural networks, now predict crown-fire potential at 10-m resolution across entire states, updating after every lightning storm [7]. A 2024 pilot in Oregon (USA) demonstrated that machine-learning maps captured 91% of subsequent fire perimeters while reducing field sampling costs by 70% relative to classical transects. Minimal infrastructure consisted of a campus GPU node (4 × A100 cards) and an open-source stack (PDAL + XGBoost). The Joint Fire Science Program (USDA-DOI), NSF's Dynamic Ecosystems and Modeling for Sustainability program, and NASA's A.37 FireSense Applied Science call all explicitly solicit "advanced analytics of multi-sensor remote-sensing streams." Private timber consortia are also co-funding cluster refreshes in exchange for early access to predictive layers, creating an unrestricted revenue loop that helps universities amortize hardware [7].

Fish and Wildlife Sciences

Population viability analyses once required multi-year mark-recapture studies. Agent-based models (ABMs) calibrated against open acoustic-telemetry or camera-trap repositories can evaluate reintroduction or harvest scenarios within weeks [12]. In the Upper Missouri Basin, an ABM of pallid sturgeon integrated hydrodynamic outputs from the Community Land Model, reproducing 85 % of observed recruitment dynamics and saving the US $1.2 million in vessel time. Infrastructure demands are modest: a 32-core CPU node with 256 GB of RAM, and open-source software (NetLogo, Repast, or Mesa). Potential funding lines for such investigations include USFWS Science Applications, NSF's Biodiversity on a Changing Planet, and the Luc Hoffmann Institute's computational conservation fellowships. Because ABMs are fully replicable, investigators currently receive bonus credit under NSF's "FAIR Data" merit-review criterion, increasing proposal competitiveness [12].

Table 1. Digital Transformation across Select Land-Grant Domains: Legacy versus Computational Workflows, Impacts, Infrastructure, and Funding.

Category	Legacy → Digital upgrade (brief)	Key impact	Core infra-structure	Main funding lines*	Select Citation
Forestry	Transect fuel loads → LiDAR + Sentinel-2 + GBM/CNN	91 % fires predicted; field cost ↓ 70 %	4×A100 GPU node	JFSP, NSF DEMS, NASA FireSense, timber consortia	[7]
Fish & Wildlife	Mark–recapture → ABM w/ telemetry & cam-trap data	85 % recruitment reproduced; US$1.2 M vessel cost saved	32-core CPU, 256 GB RAM	USFWS SA, NSF BCP, Luc Hoffmann	[12]
Animal Sci.	Human welfare scoring → YOLOv8 vision on video/thermal/IMU	F1 > 0.93; labor ↓ 60 %; 24×7	RTX 4090 or burst-cloud GPU	USDA AFRI-PLF, USCDI, FFAR Digital Ag	[13]
Agriculture	Micro-plots → APSIM + Bayesian DL emulators + PlanetScope	2 seasons faster release; N-trial cost ↓ US$340 ha¹	Shared cluster or AI-Ag cloud credits	USDA/NSF AI-Ag Inst.., state soybean boards	[14]
Landscape Arch.	Scale models → Generative design over GIS layers	Design labor ↓ 75 %; retention ↑ 14 %	High-RAM WS or elastic CPU	NEA RfC, EPA WISE	[15,16]
Economics	Single-run CGE → Ensemble CGE sweeps	4000 shocks overnight (128 cores)	Campus HPC cluster	NSF SciS, DOE EW-Nexus, WB CC&DR	[17]
Innov. & Entr.	Delphi panels → NLP on global patent corpus (Spark)	Ag-robotics convergence shown 2y early	Cloud Spark via NSF CloudBank	NSF PFI-TT, USDA SBIR I, regional EDA	[18]
Civic Engage.	Extension pamphlets → Python/Plotly public dashboards	Counties save ≈ US$600k; publishable insights	JupyterHub on cluster	USDA RHSE, NIH HEAL, HUD T4H	[19]
Workforce Dev.	Ad-hoc skills → RSE corps + micro-credentials	Cluster use ↑ 10× post-hire	Salaried RSEs; existing HPC	NSF IUSE, DOL Apprentice BA	[10]
Env. Steward-Ed.	Field trips only → Digital-twin walk-throughs + LA	Vehicle miles ↓ 40 %; scores ↑ 8 pp	LMS plug-ins + modest GPU	NOAA CRET, NASA AIST-Ed, ED IES LA	[20]

*Abbreviations: JFSP = Joint Fire Science Program, NSF DEMS = NSF Dynamic Ecosystems & Modeling for Sustainability, USFWS SA = US Fish & Wildlife Service Science Applications, USCDI = US Council on Dairy Innovation, FFAR = Foundation for Food & Agriculture Research, NEA RfC = National Endowment for the Arts "Design for Climate", EPA WISE = Water Innovation, Science & Engagement, NSF SciS = Science of Science Program, DOE EW-Nexus = Energy–Water Nexus, WB CC&DR = World Bank Country Climate & Development Reports, NSF PFI-TT = Partnerships for Innovation, Technology Translation, HUD T4H = Tech4Housing, DOL Apprentice BA = Apprenticeship Building America.

Animal Sciences

Precision livestock farming (PLF) utilizes computer vision and deep learning to analyze continuous video, thermal, and inertial data, thereby replacing periodic human scoring of welfare indicators. A multi-site dairy study utilized a YOLOv8 pipeline to detect lying, ruminating, and heat-stress postures, achieving F1 scores greater than 0.93 while reducing labor by 60% and enabling 24/7 surveillance [13]. The necessary hardware, as an example, could include a single RTX 4090 workstation running on-farm or a burstable cloud GPU, which costs less than a year of undergraduate labor previously budgeted for scoring. USDA NIFA's Agriculture and Food Research Initiative (AFRI) now maintains a PLF focus area, and the private US Council on Dairy Innovation funds data-centric welfare research with up to a 50% university match. Such projects also currently qualify for the Foundation for Food & Agriculture Research's (FFAR) Digital Ag Rapid Cycle grants, in which computational readiness is a primary review factor.

Agriculture

Experiment stations historically relied on dozens of micro-plots per cultivar compared to as many (or more) treatment combinations. The Agricultural Production Systems Simulator (APSIM), coupled with Bayesian deep-learning yield emulators, now replaces half of those plots, shaving two growing seasons from cultivar release time [14]. In Iowa, integrating APSIM with PlanetScope imagery and soil grids reduced nitrogen rate trial costs by USD 340 per hectare, while maintaining prediction error below 5%. Essential infrastructure is again lightweight, a shared cluster or commercial cloud credits provided under the USDA/NSF AI Institute for Agricultural Systems, whose RFA explicitly lists "APS-powered digital twins." Matching state soybean boards have co-sponsored GPU cycles, creating a public–private co-investment model that shielded the university from full depreciation risk [14].

Landscape Architecture

Planning stormwater retrofits with architectural awareness traditionally required the use of scale models or full-sized pilot installations. Generative-design algorithms, running with integrated GIS layers of topography, parcel boundaries, and imperviousness, now automatically produce thousands of bio-retention layouts, optimizing for runoff, cost, and aesthetic coherence [15,16]. In suburban Maryland, the approach reduced design labor by 75% and produced a 14% improvement in volumetric retention relative to conventional computer-aided drafting (CAD) workflows. Computational research methods are increasingly integral to high-performance landscape architecture. In Singapore's Bishan-Ang Mo Kio Park, designers ran 1-D and 2-D MIKE flood-routing models to test channel geometry and floodplain widths before replacing a concrete canal with a sinuous, ecologically functional river corridor, a process detailed in a comparative nature-based-solutions study of Singapore and Lisbon [21]. At New York City's Freshkills Park, planners combined fine-scale GIS layers with ecosystem services valuation models to quantify the ecological and economic benefits of converting an 890-hectare landfill into a regional green infrastructure system [22]. Designers are also leveraging generative tools. Huang*, et al.* [23] employed Grasshopper-based multi-objective optimization to place rain gardens and permeable pavements in a "sponge-city" micro-renewal project, thereby reducing peak runoff while minimizing costs. Additionally, Orsi's prescriptive agent-based model rearranged built and green plots, allowing residents to remain centrally located yet close to nature [24]. Among other sources, funding is currently available through the National Endowment for the Arts Research Labs (Design for Climate) and the EPA's Water Innovation, Science, and Engagement (WISE) program, both of which require community-scale data products. In terms of computing infrastructure, a high-memory workstation or elastic cloud CPU node is often all that is necessary [15].

Economics

Wei and Aaheim [17] conducted a systematic review analyzing the applications of computable general equilibrium (CGE) models in climate change adaptation at the Center for International Climate Research (CICERO) in Oslo, Norway. They reviewed 97 peer-reviewed studies, concluding that while CGE models extensively evaluate planned adaptation measures, strategies deliberately enacted through policies, they insufficiently address autonomous adaptation, or the spontaneous, market-driven responses by economic actors without direct policy intervention. Recognizing this limitation, Wei and Aaheim emphasized the need for integrated modeling frameworks that better capture both planned and autonomous adaptation responses. Building on their insights, recent research teams, including a Department of Energy (DOE)-funded group at a western US land-grant university, expanded the use of large-ensemble CGE modeling techniques. Utilizing advanced computational clusters with 128 cores, this team efficiently processed 4,000 shock matrices overnight, a task traditionally spanning several months on conventional desktop systems. This methodological advancement enhances the capability to manage and quantify uncertainties in climate impact assessments, particularly in relation to agricultural and energy policy implications. Such innovative computational efforts are increasingly supported through funding from organizations such as the National Science Foundation's Science of Science: Discovery, Communication, and Impact initiative, the DOE's Energy-Water Nexus program, and the World Bank's Country Climate and Development Reports. These developments collectively indicate significant progress toward more comprehensive and policy-relevant climate-economic modeling through sophisticated, high-performance computational approaches.

Innovation and Entrepreneurship

A significant methodological advancement for entrepreneurial-ecosystem scholars was introduced by applying natural language processing (NLP) to global patent databases, effectively replacing the traditionally used Delphi panels [18]. While Delphi panels rely on expert consensus and can be slow and subjective, NLP rapidly processes extensive unstructured patent data, objectively identifying emergent technology clusters. This innovative approach enables researchers and policymakers to efficiently track technological trends, facilitating timely strategic decisions and fostering targeted investments in entrepreneurial innovation. A 2024 study of agricultural-robotics patents revealed regional convergence dynamics two years before the official Organization for Economic Co-operation and Development (OECD) statistics, which guided venture capital placement. The OECD is an international organization comprising member countries that are committed to democratic and market economy principles. The organization collects, analyzes, and publishes comprehensive statistics on economic, social, and environmental aspects. The infrastructure for this initiative included an Apache Spark cluster (a widely used open-source software framework designed for big-data processing and analytics) on a commercial cloud, accessible via US NSF-funded CloudBank credits through the NSF Partnerships for Innovation–Technology Translation (PFI-TT) program. The USDA SBIR Phase I now also accepts patent-text analytics as an eligible feasibility study. Corporate co-investment often follows, as regional economic-development authorities subscribe to the resulting dashboards, generating non-federal indirect revenue.

Civic Engagement

The land-grant extension model is a federally supported initiative historically focused on connecting universities with local communities to deliver practical knowledge, resources, and technical assistance.

Traditionally, extension efforts have involved distributing seed packets, providing advice to local farmers, conducting educational workshops, and participating in numerous community activities. However, this model has evolved toward digitalization and data-driven methods, exemplified by the emergence of Community-Engaged Data Science (CEDS) teams. Olvera et al. [19] illustrated this approach by describing how CEDS embedded student data ambassadors within Ohio counties to collaboratively develop interactive dashboards, specifically mapping critical regional issues such as opioid overdose trends and broadband accessibility gaps using Python and Plotly visualization tools. This shift toward interactive dashboards enabled counties to make informed decisions based on dynamic, real-time data, resulting in approximately $600,000 in savings from external consulting fees and the generation of publishable social-science insights. Notably, the infrastructure required for these initiatives is minimal, relying primarily on accessible, browser-based platforms such as JupyterHub, connected to university computing clusters. Recognizing the significant impact and efficiency of these data-driven community tools, major funding programs such as USDA NIFA's Rural Health and Safety Education, NIH's HEALing Communities, and HUD's Tech4Housing Cooperative Agreements currently prioritize proposals that incorporate computationally driven approaches, offering extension faculty with computational expertise a distinct competitive advantage.

Workforce Development

Human capital has emerged as a critical bottleneck in the effective utilization of high-performance computing (HPC) and artificial intelligence (AI) resources within academic research environments. To address this challenge, institutions are adopting a dual-faceted strategy: establishing dedicated Research Software Engineer (RSE) teams and implementing stackable micro-credential programs focused on HPC, AI, and reproducible research practices. Chen, Ghafoor and Impagliazzo [10] highlighted a case where a Midwestern land-grant university significantly increased its HPC cluster utilization tenfold by hiring four RSEs and integrating a mandatory one-credit "Data Notebook" course across all STEM majors. This course emphasized essential skills in data management, coding, and reproducibility, thereby enhancing students' computational competencies. The RSEs provided critical support in developing and maintaining research software, ensuring that computational tools met rigorous standards for reliability and reproducibility. This approach not only optimized the use of existing computational infrastructure but also fostered a culture of best practices in research software development. Moreover, funding agencies such as the US. National Science Foundation's Improving Undergraduate STEM Education (IUSE) program and the US. Department of Labor's Apprenticeship Building America initiative have recognized the value of this model. These agencies offer financial support, often covering up to 50% of early-career RSE salaries when matched by institutional funds, thereby incentivizing the integration of RSEs into research teams. By investing in human capital through Research Support and Educational (RSE) programs and targeted educational initiatives, institutions can enhance research productivity, ensure software sustainability, and better prepare students for the evolving demands of computational research.

Environmental Stewardship and Education

The integration of learning analytics platforms with satellite data is revolutionizing conservation education by personalizing curricula and enhancing field readiness. Students now engage in digital twin walkthroughs, virtual simulations of real-world habitats, before participating in fieldwork, allowing them to anticipate environmental conditions and challenges [20]. This approach has led to significant improvements; for instance, a large-enrollment environmental science course reported a 40% reduction in vehicle miles traveled for field activities and an eight-percentage-point increase in concept inventory scores. Implementing such systems requires minimal resources: existing campus learning management system plug-ins and a modest GPU capable of real-time image classification. Funding opportunities to support these initiatives currently include, but are not limited to, NOAA's Coastal Resilience Education and Training, NASA's Advanced Information Systems Technology (education track), and the US. Department of Education's Institute of Education Sciences Learning Analytics for STEM program. Proposals that emphasize reproducible data pipelines and the development of open data repositories are particularly competitive, reflecting the broader shift towards computational proficiency in environmental education. This evolution not only enhances student learning outcomes but also contributes to more sustainable and efficient educational practices.

Strategic Implications

Across the above ten domains, three patterns repeat. First, virtual experimentation and big-data analytics displace the most expensive stages of field or bench work, freeing scarce dollars for student stipends or additional projects. Second, shared cyberinfrastructure, whether an on-premises GPU node or cloud credits, scales across disciplines, lowering marginal cost and flattening equity gaps between resource-rich and resource-poor departments. Third, sponsors are already embedding computational readiness into their review criteria, effectively rewarding institutions that adopt reproducible, data-centric workflows. The implications for principal investigators are clear: a modest upfront investment in code containers, research software engineer support, and training converts into faster publications, higher hit rates at funding agencies, and demonstrably job-ready graduates. Land-grant administrators, in turn, gain a compelling narrative for legislators and donors in favor of computational infrastructure.

IV. Mapping the Funding Landscape

Federal research funding is tightening, resulting in reduced resources for traditional disciplines. Regardless, the rapid growth of data, interdisciplinary needs, and global competition make it essential to integrate computational methods across nearly all fields for scientific relevance and funding success. Institutions and researchers that do not adapt risk falling behind, despite limited budgets. A successful shift toward computational research could be supported through a mix of federal, philanthropic, and industry investments that increasingly favor data-centric readiness. While current and future funding environments and competitions may evolve, the strategies outlined here are expected to remain relevant for the foreseeable future.

Federal Baseline

In fiscal year 2024, the United States National Science Foundation (NSF) requested $11.355 billion, reflecting a 19% increase from the prior year, yet funding success rates remained highly competitive across directorates, with the Directorate for Biological Sciences (BIO) maintaining a 19% success rate and divisions such as Biological Infrastructure (DBI) falling as low as 13% [25]. Similarly, the US. Department of Agriculture's National Institute of Food and Agriculture (NIFA) allocated $445 million to the Agriculture and Food Research Initiative (AFRI), with $300 million earmarked for the Foundational and Applied Science (FAS) program, which explicitly prioritizes projects leveraging advanced analytics and artificial intelligence to address key societal challenges in areas such as food security, environmental sustainability, and public health [26]. Given this heightened competition and the strategic focus on computational methodologies, funding agencies are increasingly favoring proposals that deliver greater insight per dollar by integrating reproducible data pipelines, open data repositories, and advanced analytics, positioning computational approaches as a critical differentiator in securing competitive federal research funding.

Dedicated Computational Programs

Federal agencies have developed high-profile funding programs explicitly designed with robust cyberinfrastructure capabilities in mind. For example, the NSF, USDA, DOE, NIH, and NIST jointly established the National AI Research Institutes, offering seven-year awards of up to $ 20 million (USD) each. Three institutes selected for 2024 specifically target agriculture–food systems, climate resilience, and workforce development, emphasizing open, scalable data pipelines [8]. The NIH's Bridge2AI initiative was developed to distribute $130 million over four years to create comprehensive, "AI-ready" biomedical datasets that adhere strictly to the FAIR (findable, accessible, interoperable, reusable) data criteria [27]. DOE's Exascale Computing and Earthshot programs prioritize large-scale ensemble simulations, which require documented high-performance computing (HPC) capabilities, as indicated by cluster utilization metrics similar to those reported by previous authors [9]. Many other funding sources exist; researchers should proactively engage with institutional, state, and federal research support offices to identify and align with these valuable opportunities.

Philanthropic Catalysts

Private foundations are increasingly treating open cyberinfrastructure as a public good, recognizing its role in enabling scientific discovery, education, and innovation across disciplines. Schmidt Futures exemplifies this shift with its $148 million "AI in Science" fellowship, supporting postdoctoral researchers who apply machine learning outside traditional computer science domains, aiming to accelerate discovery in fields such as biology, chemistry, and climate science [28]. Similarly, the Gordon and Betty Moore Foundation's Data-Driven Discovery and Moore Inventor Fellows programs collectively provided approximately $34 million through 2025, targeting early-career scientists and inventors developing novel computational tools, data curation methods, and scientific software to advance research impact [29]. Complementing these efforts, the Alfred P. Sloan Foundation allocates around $80 million annually through its Digital Infrastructure and Exemplary Pathways initiatives, much of which supports open-source software, data stewardship, and cyberinfrastructure development at minority-serving institutions [30]. These investments reflect a broader recognition that open, interoperable, and accessible digital infrastructure, combined with support for human capital, is foundational to equitable, reproducible, and scalable science. Furthermore, these programs emphasize community-centered development models, ensuring that tools, data, and knowledge generated are accessible to a broad range of researchers, educators, and communities, thereby amplifying the societal return on philanthropic investment.

Industry and Cluster-Access Models

Campus *community clusters* have emerged as a financially sustainable and academically enriching model for shared research computing infrastructure. By offering tiered corporate memberships that provide priority job queues, exclusive research collaboration opportunities, and student capstone partnerships, institutions can recover 15–20% of their annual refresh costs while simultaneously enhancing experiential learning opportunities for students. A longitudinal return on investment (ROI) analysis by Stewart et al. [9] of XSEDE partner campuses showed that every dollar invested in shared-cluster expenses leveraged approximately US $5 in external awards, significantly increased interdisciplinary collaboration, and reduced time-to-publication by 40%, effectively accelerating research dissemination and impact. Supporting these findings, a recent study at Purdue University demonstrated a strong correlation between institutional investments in campus high-performance computing (HPC) infrastructure and growth in financial, academic, and reputational outputs [31]. Collectively, these studies reinforce that investments in campus community clusters not only offer economic efficiencies but also position institutions to foster research innovation, accelerate knowledge dissemination, and strengthen industry partnerships.

Strategic Take-Away

Federal solicitations, philanthropy, and corporate cost-sharing now incorporate computational expectations into research, development, and education. Investigators documenting containerized workflows, reproducible code, and cluster access signal lower project risk and higher translational velocity, attributes increasingly rewarded in competitive funding environments. Demonstrating detailed plans for data management and analysis pipelines enhances scalability and collaboration potential. Institutions that leverage shared computational facilities effectively demonstrate a commitment to resource optimization and interdisciplinary engagement, significantly increasing proposal competitiveness and appeal, thereby positioning projects for accelerated review and success in securing diverse funding streams.

V. Implementation Roadmap

Translating computational opportunities into lasting capacity requires a phased strategy that integrates finance, IT, faculty affairs, and external relations. This process includes auditing current needs, piloting high-impact projects with research software engineers, adopting a cost-sharing model for community clusters, and embedding data science into curricula and promotional standards. Continuous monitoring and ROI assessments ensure alignment with institutional goals, enhance research productivity, accelerate funding success, and establish cyberinfrastructure as a core strategic asset (Figure 1).

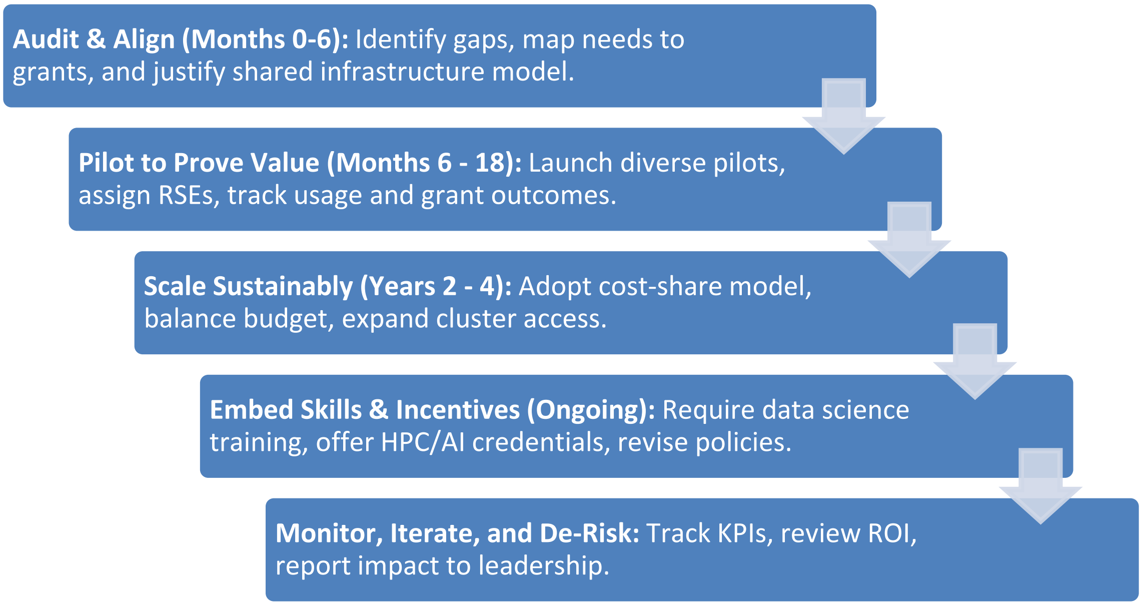

Figure 1. An example implementation roadmap including a timeline for building durable computational capacity. Timing and time periods may vary depending on institutional capacity and need.

Audit and Align (Months 0 – 6)

Organizations should begin by conducting a comprehensive gap analysis, including inventorying existing computing workloads, storage profiles, and GPU demand, then overlaying those data with recent proposal reviews to pinpoint where inadequate cyberinfrastructure drew criticism. They should then cross-tabulate these requirements against public grant calendars projected 24 months forward, such as NSF AI Institutes and USDA AFRI Foundational calls, to forecast future demand. This audit phase has previously shown that approximately 60 to 70 percent of anticipated usage can be satisfied by a single, shared mid-tier cluster augmented with cloud-burst credits, obviating the need for multiple discipline-specific servers. Empirical studies reinforced the value of this data-driven approach: Stewart et al. [9] showed that each dollar invested in shared XSEDE-style resources returned about five dollars in external awards and shortened time-to-publication by 40 percent, validating the economic rationale for consolidating infrastructure. Complementing this, Smith [31] demonstrated that balanced investment in on-campus HPC hardware and support personnel significantly enhanced institutional research outputs, confirming that strategic, evidence-based planning directly improves both productivity and competitiveness.

Pilot to Prove Value (Months 6 – 18)

During the 6- to 18-month pilot phase, three to five high-visibility research projects should be launched (this may vary by institutional capacity and need). For example, one each in the life sciences, physical sciences, socio-economic analysis, and digital humanities, on the new shared cluster. Each research team could receive fractional support from a Research Software Engineer (RSE) plus a modest allotment of cloud-burst credits, ensuring that code is containerized and reproducible from the outset. At Indiana University (USA), a comparable initiative led to a tenfold increase in core-hour utilization and a 38 percent rise in annual grant volume, outcomes directly linked to RSE-led containerization and workflow hardening [32]. Complementary evidence from an extensive survey of container technologies in high-performance computing revealed that systematic container adoption enhances portability while incurring negligible performance overhead, thereby accelerating time-to-science across heterogeneous environments [33]. Institutions captured pilot metrics, core-hour consumption, job-queue wait times, grant dollars secured, and publications produced in quarterly dashboards; these data streams underpinned compelling ROI narratives, sustained cabinet-level enthusiasm, and informed subsequent investments in both hardware expansions and additional RSE personnel.

Scale Sustainably (Years 2 – 4)

To escape the inefficiencies of one-off, grant-funded server purchases, institutions could consider implementing a community-cluster cost-share model in which faculty contribute start-up or sponsored-research dollars, central IT matches those funds, and industry affiliates purchase queue-priority memberships. As noted previously, Stewart et al. [9] quantified the impact of this model at XSEDE partner campuses, reporting that every US $1 invested yielded approximately US $5 in new external awards and shortened median time-to-publication by 40 percent. Similar benefits were observed at the national scale: the XSEDE program's shared cyberinfrastructure lowered per-project computing costs and accelerated scientific output compared with siloed resources, underscoring the leverage gained through pooled investment [34]. Financial modeling across multiple land-grant campuses now recommends a 40/40/20 spending split for hardware procurement, human capital (Research Software Engineers and data stewards), and continuous training, respectively, because this ratio maximizes return on investment while sustaining a skilled support ecosystem that keeps the cluster fully utilized.

Embed Skills and Incentives

Ultimately, hardware investments may deliver limited returns until institutions pair them with systematic human-capital development. In response to this gap, several universities have mandated entry-level, credit-bearing data science courses for all majors. They then offered stackable certificates in high-performance computing (HPC) and artificial intelligence. Within two graduating cohorts, programs that adopted this scaffold saw data-centric job placements rise by approximately 15 percent, a trend echoed in Communications of the ACM, where Chen et al. [10] documented how a competency-based HPC curriculum produced graduates who were immediately productive in real-world research settings. To cement cultural change, promotion and tenure policies also evolved. An eLife consensus paper urged institutions to credit curated data sets, FAIR-compliant workflows, and reusable open-source code equally with traditional articles, arguing that such research products accelerate discovery and widen access to knowledge [35]. Embedding micro-credentials and revising incentive structures thus transformed underutilized compute resources into a vibrant, talent-driven ecosystem that sustains both research excellence and workforce readiness.

Monitor, Iterate, and De-Risk

To keep cyberinfrastructure performance tightly aligned with institutional goals, campuses should consider establishing a joint Faculty, IT Advisory Board charged with evidence-based oversight. The board could track a concise set of key performance indicators, which might include cluster-core utilization hours, external research dollars captured, Research Software Engineer (RSE) tickets resolved, and peer-reviewed publications that acknowledge the system, among other criteria. When a metric softens, the board could recalibrate queue priorities, shift staffing, or recommend targeted hardware upgrades before minor inefficiencies escalate into costly bottlenecks. The groups could also commission independent return-on-investment studies that apply the International Integrated Reporting (IR) Framework, translating scientific outputs and workforce training into financial terms that resonate with trustees and legislators. A recent IR analysis of the national XSEDE program showed every federal dollar invested generated between US $4.7 billion and $22.7 billion in societal benefit, a finding that helped secure bipartisan budget support [32]. By cycling transparently through audit, pilot, scale, and embed stages, while continuously refining KPIs, universities could transform one-time server purchases into a resilient, self-reinforcing ecosystem of grant wins, reproducible science, lower carbon intensity, and graduates that are market-ready for data-intensive careers.

VI. Conclusions

The accelerating affordability and accessibility of advanced computing infrastructure have eliminated many of the historical barriers that once limited computational research to elite institutions. For land-grant universities and other public research entities facing declining appropriations, deferred maintenance, and escalating competition for federal funding, this technological shift represents not just an opportunity but a strategic imperative. As illustrated across various domains, from forestry and agriculture to economics and civic engagement, the transition from fieldwork to cloudwork enables researchers to produce high-impact insights more quickly, at lower cost, and with greater scalability. Institutions that invest in shared cyberinfrastructure, research software engineering support, and reproducible data pipelines consistently experience higher grant success rates, shorter time-to-publication, and stronger graduate employability. These gains, however, are not merely technical; they are institutional in nature. Cyberinfrastructure must now be regarded as mission-critical, on par with experimental farms, wet labs, and county extension offices. The evidence is clear: computational readiness translates directly into research productivity, funding competitiveness, and labor market alignment.

Nevertheless, the challenge extends beyond hardware and software. To fully capitalize on the computational research paradigm, universities must embed data fluency across their curricula, revise promotion and tenure metrics to value open science, reusable code, and realign extension and workforce development initiatives toward digital engagement and training. Institutions that act decisively can secure founder advantages, build interdisciplinary capacity, and reassert their relevance in a data-driven economy. Those that delay risk obsolescence as sponsors increasingly demand reproducibility, transparency, and translational velocity. The future of research, and by extension, the public mission of higher education, depends on a systematic pivot toward computational capability. This is not a passing trend but a structural transformation that will define institutional viability for decades to come. By turning financial constraint into strategic leverage, universities can revitalize their land-grant mission, accelerate discovery, and graduate a new generation of leaders prepared to thrive in the era of data-intensive science.

Funding

This work was supported by the USDA National Institute of Food and Agriculture McIntire-Stennis accession number 7003934 and the West Virginia Agricultural and Forestry Experiment Station. The results presented may not reflect the sponsors' views, and no official endorsement should be inferred. The funders had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript.

Acknowledgements

The author is grateful to reviewers whose constructive comments improved the quality and scope of the article.

References

[1] Holbrook, K.A.; Sanberg, P.R. Understanding the High Cost of Success in University Research. *Technology & Innovation* **2013**, *15*, 269–280, doi:10.3727/194982413x13790020922068.

[2] Gavazzi, S.M. The land-grant mission in the 21st century: promises made and promises to be kept. *Animal Frontiers* **2020**, *10*, 6–9, doi:10.1093/af/vfaa016.

[3] Thomas, R.; Cholia, S. Interactive Supercomputing With Jupyter. *Computing in Science & Engineering* **2021**, *23*, 93–98, doi:10.1109/mcse.2021.3059037.

[4] Blei, D.M.; Smyth, P. Science and data science. *Proceedings of the National Academy of Sciences* **2017**, *114*, 8689–8692, doi:10.1073/pnas.1702076114.

[5] Carleo, G.; Cirac, I.; Cranmer, K.; Daudet, L.; Schuld, M.; Tishby, N.; Vogt-Maranto, L.; Zdeborová, L. Machine learning and the physical sciences. *Reviews of Modern Physics* **2019**, *91*, doi:10.1103/revmodphys.91.045002.

[6] Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204, doi:10.1038/s41586-019-0912-1.

[7] Moghim, S.; Mehrabi, M. Wildfire assessment using machine learning algorithms in different regions. *Fire Ecology* **2024**, *20*, doi:10.1186/s42408-024-00335-2.

[8] Donlon, J.J. The National Artificial Intelligence Research Institutes program and its significance to a prosperous future. *AI Magazine* **2024**, *45*, 6–14, doi:10.1002/aaai.12153.

[9] Stewart, C.A.; Costa, C.M.; Wernert, J.A.; Snapp-Childs, W.; Bland, M.; Blood, P.; Campbell, T.; Couvares, P.; Fischer, J.; Hancock, D.Y.; et al. Use of accounting concepts to study research: return on investment in XSEDE, a US cyberinfrastructure service. *Scientometrics* **2023**, *128*, 3225–3255, doi:10.1007/s11192-022-04539-8.

[10] Chen, J.; Ghafoor, S.; Impagliazzo, J. Producing competent HPC graduates. *Communications of the ACM* **2022**, *65*, 56–65, doi:10.1145/3538878.

[11] National Academies of Sciences, E.a.M. *Enhancing coordination and collaboration across the land-Grant system*; 0309691079; National Academies of Sciences, Engineering, Medicine: 2022.

[12] Neil, E.; Carrella, E.; Bailey, R. Integrating agent-based models into the ensemble ecosystem modelling framework: A rewilding case study at the Knepp Estate, UK. *Ecological Solutions and Evidence* **2025**, *6*, doi:10.1002/2688-8319.70022.

[13] Pesenti Rossi, G.; Dalla Costa, E.; Barbieri, S.; Minero, M.; Canali, E. A systematic review on the application of precision livestock farming technologies to detect lying, rest and sleep behavior in dairy calves. *Frontiers in Veterinary Science* **2024**, *11*, doi:10.3389/fvets.2024.1477731.

[14] McNunn, G.; Heaton, E.; Archontoulis, S.; Licht, M.; Vanloocke, A. Using a Crop Modeling Framework for Precision Cost-Benefit Analysis of Variable Seeding and Nitrogen Application Rates. *Frontiers in Sustainable Food Systems* **2019**, *3*, doi:10.3389/fsufs.2019.00108.

[15] Jiang, F.; Ma, J.; Webster, C.J.; Li, X.; Gan, V.J.L. Building layout generation using site-embedded GAN model. *Automation in Construction* **2023**, *151*, 104888, doi:10.1016/j.autcon.2023.104888.

[16] Chen, R.; Zhao, J.; Yao, X.; Jiang, S.; He, Y.; Bao, B.; Luo, X.; Xu, S.; Wang, C. Generative Design of Outdoor Green Spaces Based on Generative Adversarial Networks. *Buildings* **2023**, *13*, 1083, doi:10.3390/buildings13041083.

[17] Wei, T.; Aaheim, A. Climate change adaptation based on computable general equilibrium models – a systematic review. *International Journal of Climate Change Strategies and Management* **2023**, *15*, 561–576, doi:10.1108/ijccsm-03-2022-0031.

[18] Paul, J.; Ueno, A.; Dennis, C.; Alamanos, E.; Curtis, L.; Foroudi, P.; Kacprzak, A.; Kunz, W.H.; Liu, J.; Marvi, R.; et al. Digital transformation: A multidisciplinary perspective and future research agenda. *International Journal of Consumer Studies* **2024**, *48*, doi:10.1111/ijcs.13015.

[19] Olvera, R.G.; Plagens, C.; Ellison, S.; Klingler, K.; Kuntz, A.K.; Chase, R.P. Community-Engaged Data Science (CEDS): A Case Study of Working with Communities to Use Data to Inform Change. *Journal of Community Health* **2024**, *49*, 1062–1072, doi:10.1007/s10900-024-01377-y.

[20] Samsul, S.A.; Yahaya, N.; Abuhassna, H. Education big data and learning analytics: a bibliometric analysis. *Humanities and Social Sciences Communications* **2023**, *10*, doi:10.1057/s41599-023-02176-x.

[21] Cui, M.; Ferreira, F.; Fung, T.K.; Matos, J.S. Tale of Two Cities: How Nature-Based Solutions Help Create Adaptive and Resilient Urban Water Management Practices in Singapore and Lisbon. *Sustainability* **2021**, *13*, 10427, doi:10.3390/su131810427.

[22] Vincevica-Gaile, Z.; Burlakovs, J.; Fonteina-Kazeka, M.; Wdowin, M.; Hanc, E.; Rudovica, V.; Krievans, M.; Grinfelde, I.; Siltumens, K.; Kriipsalu, M.; et al. Case Study-Based Integrated Assessment of Former Waste Disposal Sites Transformed to Green Space in Terms of Ecosystem Services and Land Assets Recovery. *Sustainability* **2023**, *15*, 3256, doi:10.3390/su15043256.

[23] Huang, Z.; Sun, Y.; Fan, Y.; Guan, R.; Zhang, H.; Zhao, L.; Zhang, B. Toward Urban Micro-Renewal: Integrating “BMP-Plan” and “LID-Design” for Enhanced Stormwater Control—A Case Study. *Water* **2025**, *17*, 992, doi:10.3390/w17070992.

[24] Orsi, F. Centrally located yet close to nature: A prescriptive agent-based model for urban design. *Computers, Environment and Urban Systems* **2019**, *73*, 157–170, doi:10.1016/j.compenvurbsys.2018.10.001.

[25] NSF. National Science Foundation, BIO Funding Rates. Available online: https://www.nsf.gov/bio/funding-rates.com (accessed on July 1, 2025).

[26] NIFA. National Institute of Food and Agriculture, Agriculture and Food Research Initiative (AFRI). Available online: https://www.nifa.usda.gov/grants/programs/agriculture-food-research-initiative (accessed on July 1, 2025).

[27] NCCIH. 2022 Press Releases. Available online: https://www.nccih.nih.gov/news/press-releases/2022 (accessed on July 1, 2025).

[28] Schmidt-Sciences. Schmidt Sciences: Advancing Discovery Through Research and Technology. Available online: https://www.schmidtsciences.org/ (accessed on July 1, 2025).

[29] TMF. The Moore Foundation: Moore Inventor Fellows. Available online: https://www.moore.org/initiative-strategy-detail?initiativeId=moore-inventor-fellows (accessed on July 1, 2025).

[30] APSF. Alfred, P. Sloan Foundation, 2024–2025 Exemplary Pathways Grantees. Available online: https://sloan.org/programs/higher-education/exemplary-pathways/2024-2025-exemplary-pathways-grantees (accessed on July 1, 2025).

[31] Smith, P. The Value Proposition of Campus High Performance Computing Facilities to Institutional Productivity: A Production Function Model. *SN Computer Science* **2024**, *5*, doi:10.1007/s42979-024-02888-0.

[32] Snapp-Childs, W.G.; Hart, D.L.; Costa, C.M.; Wernert, J.A.; Jankowski, H.E.; Towns, J.W.; Stewart, C.A. Evaluating Return on Investment for Cyberinfrastructure Using the International Integrated Reporting <IR> Framework. *SN Computer Science* **2024**, *5*, doi:10.1007/s42979-024-02889-z.

[33] Keller Tesser, R.; Borin, E. Containers in HPC: a survey. *The Journal of Supercomputing* **2023**, *79*, 5759–5827, doi:10.1007/s11227-022-04848-y.

[34] Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G.D.; et al. XSEDE: Accelerating Scientific Discovery. *Computing in Science & Engineering* **2014**, *16*, 62–74, doi:10.1109/mcse.2014.80.

[35] Kohrs, F.E.; Auer, S.; Bannach-Brown, A.; Fiedler, S.; Haven, T.L.; Heise, V.; Holman, C.; Azevedo, F.; Bernard, R.; Bleier, A.; et al. Eleven strategies for making reproducible research and open science training the norm at research institutions. *eLife* **2023**, *12*, doi:10.7554/elife.89736.