PUBLIC SECTOR

Building the AI-Forward Research University

By Tejus Kothari, Bob Wu, J.R. Sullivan, and Daniel Goetzel

ARTICLE AUGUST 12, 2026 12 MIN READ

US research universities are the training ground for many of the PhD students who power private labs, yet these universities are not at the forefront of setting the AI agenda.

A key reason is the scale of private AI investment. The largest academic computing commitment in US history, the 10-year $457 million National Science Foundation (NSF) Leadership-Class Computing Facility (LCCF) at The University of Texas at Austin (UT Austin) is roughly what a single hyperscaler now spends in a matter of weeks.[1] Besides capital, universities face other headwinds including ongoing federal funding uncertainty, unprecedented numbers of AI-related PhDs migrating to industry, and industry dominance in frontier model development. At the same time, the federal government is rethinking how it deploys its $200 billion annual R&D portfolio: the Office of Science and Technology Policy's *Science: A New Golden Age* report has called for an increased focus on more flexible research models and new forms of collaboration across government, universities, and industry.[2]

To succeed, universities do not need to match industry's capital or attempt to compete at the cutting edge of development. In fact, many already have the capability to work across specific domains of foundational AI research, often anchored by standout faculty or a long-standing lab collaboration with an industry partner. They also have the opportunity to make gains in the broader arena of applied AI research, a rapidly expanding source of value focused on developing models across real-world systems. (See sidebar, "Which AI Strategy to Pursue?")

– Which AI Strategy to Pursue?

Not every university should pursue the same AI strategy. Institutions must first identify where they have a right to compete across the AI stack:

Foundational AI Research. The university develops AI models, with research spanning generative and foundational models, natural language processing, reinforcement learning, and probabilistic machine-learning frameworks.

- We estimate about 10 universities globally have the technical depth, capital, and scale to compete with leading private AI players at the frontier of foundational model development, while also engaging deeply at the applied AI research layer.
- Another 20–30 universities can differentiate within specific domains of the foundational stack, often anchored by standout faculty or long-standing industry collaboration. Most, if not all, will also partner at the applied layer.

Applied AI Research. The university deploys AI across real-world systems, including energy, industrial automation and robotics, drug discovery, clinical decision support, and financial risk modeling.

- Over 100 universities will compete primarily at the applied layer, leveraging domain strengths in medicine, agriculture, or aerospace.

Crucially, a focus on applied AI research is not a fallback to competing at the foundational level. Rather, it is a rapidly expanding source of value. Universities that can define a distinctive, system-level role in applying AI to real-world domains have a significant opportunity to lead, regardless of whether they compete at the frontier of foundational research.

This article proposes a four-layer framework to enable universities to position themselves as indispensable partners to industry and realize the potential value of such partnerships. (See Exhibit 1.) It's a reinforcing cycle of steps: world-class research attracts talent, talent requires compute to do their research, compute enables corporate partnerships, and partnerships fund the next research cycle. Universities that build a coherent strategy across all four layers can do what neither corporate labs nor isolated academic departments can achieve alone.

EXHIBIT 1

The Four Layers of the University AI Stack

Layer 1 **The AI-forward research lab**	Build flexible, interdisciplinary research structures with protected funding, guaranteed compute, and partnership models designed to produce field-building work
Layer 2 **Securing and deploying compute**	Assemble a diversified compute strategy that gives researchers reliable access to advanced infrastructure while managing funding concentration, rapid obsolescence, and refresh-cycle risk
Layer 3 **Reimagining corporate partnerships**	Shift from transactional sponsorships and licensing toward institution-level partnerships built around proprietary datasets, co-development, and helping companies apply AI to real challenges
Layer 4 **Winning the R&D talent competition**	Compete for top researchers through research freedom, multi-year support, compute access, flexible appointments, and stronger connections between academic labs and their alumni networks

Source: BCG analysis.

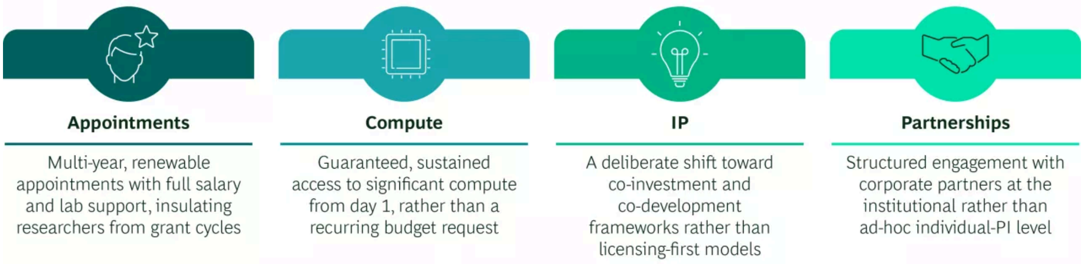
Layer 1: The AI-Forward Research Lab

The balance of AI research has decisively shifted. Industry now produces more than 90% of notable frontier models.[3] According to research published in *Science*, industry models are on average 29 times larger than academic ones. Since 2006, the number of AI research faculty has stayed essentially flat while industry hiring of AI faculty rose eightfold.[4] Universities remain the primary training ground for PhD talent, the locus of curiosity-driven research, and the custodians of unique datasets. But the old ways of doing things, built around the single principal investigator and the grant cycle, are insufficient.

The authors propose an ambitious path forward, termed the AI-Forward Research Lab, focused on interdisciplinary, field-building outputs. (See Exhibit 2.)

EXHIBIT 2

Key Dimensions of the AI-Forward Research Lab



Source: BCG analysis.
Note: PI = principal investigator.

There are three models worth watching.

Model 1: Independent Nonprofit Labs and Focused Research Organizations. A new class of research institutions, built for speed and insulated from grant cycles, is emerging outside the university. The Arc Institute, founded with $650 million in philanthropic funding in partnership with Stanford, UCSF, and UC Berkeley, provides researchers renewable eight-year appointments with full salary and lab support, freeing them from grant cycles. In a short period, the model has produced technology breakthroughs (e.g., bridge recombinases, a programmable tool that can insert, delete, or flip large sections of DNA in a single step), companies raising hundreds of millions of dollars, and public goods co-developed with industry partners such as the Evo 2 DNA foundational model, developed with Nvidia. Similarly, focused research organizations (FROs), time-limited, nonprofit entities gaining traction through NSF's newly created $1 billion X-Labs program are often able to move with greater speed and flexibility than traditional universities.[5]

Universities that become founding partners are best placed to benefit from FROs, for example by structuring pre-negotiated IP terms with aligned FRO entities, establishing predetermined slots for recurring sabbaticals and exchanges, creating post-doc externships at FROs, and negotiating master service agreements for FROs to use university facilities and equipment to perform their research. Universities that do not risk having their best researchers recruited away.

Model 2: For-Profit Labs with Strong University Ties. For-profit labs with university ties show how porous the boundary between university and commercial labs has become. The 2024 Nobel Prize in Chemistry, awarded to the University of Washington's David Baker for computational protein design and jointly to Google DeepMind's Demis Hassabis and John Jumper for AI-based protein structure prediction, shows how the leading edge of science now spans academic and corporate labs. Genesis Therapeutics (Stanford) and Xaira Therapeutics (co-founded by Baker Lab alumni) reflect the same fluidity. (See sidebar, "The Baker Lab Alumni Network.") Anthropic's co-founder and chief science officer, Jared Kaplan, remains on leave from a Johns Hopkins physics professorship.

– The Baker Lab Alumni Network

David Baker's lab at the University of Washington Institute for Protein Design (Baker was a 2024 Nobel laureate in Chemistry) offers a lens on how a leading computationally intensive academic lab produces and distributes talent. An independent review of the lab's published roster of former members, comprising more than 80 former graduate students plus a comparable cohort of postdoctoral researchers, shows:

- Just under half (roughly 45%–50%) remained in academia: faculty appointments, postdoctoral roles, and research institutes.
- Roughly one-third to 40% moved into industry, including Microsoft, Google X, and biotech firms such as Ginkgo Bioworks, Generate:Biomedicines, and Xaira Therapeutics.
- The balance went into medicine, national laboratories, and other roles.

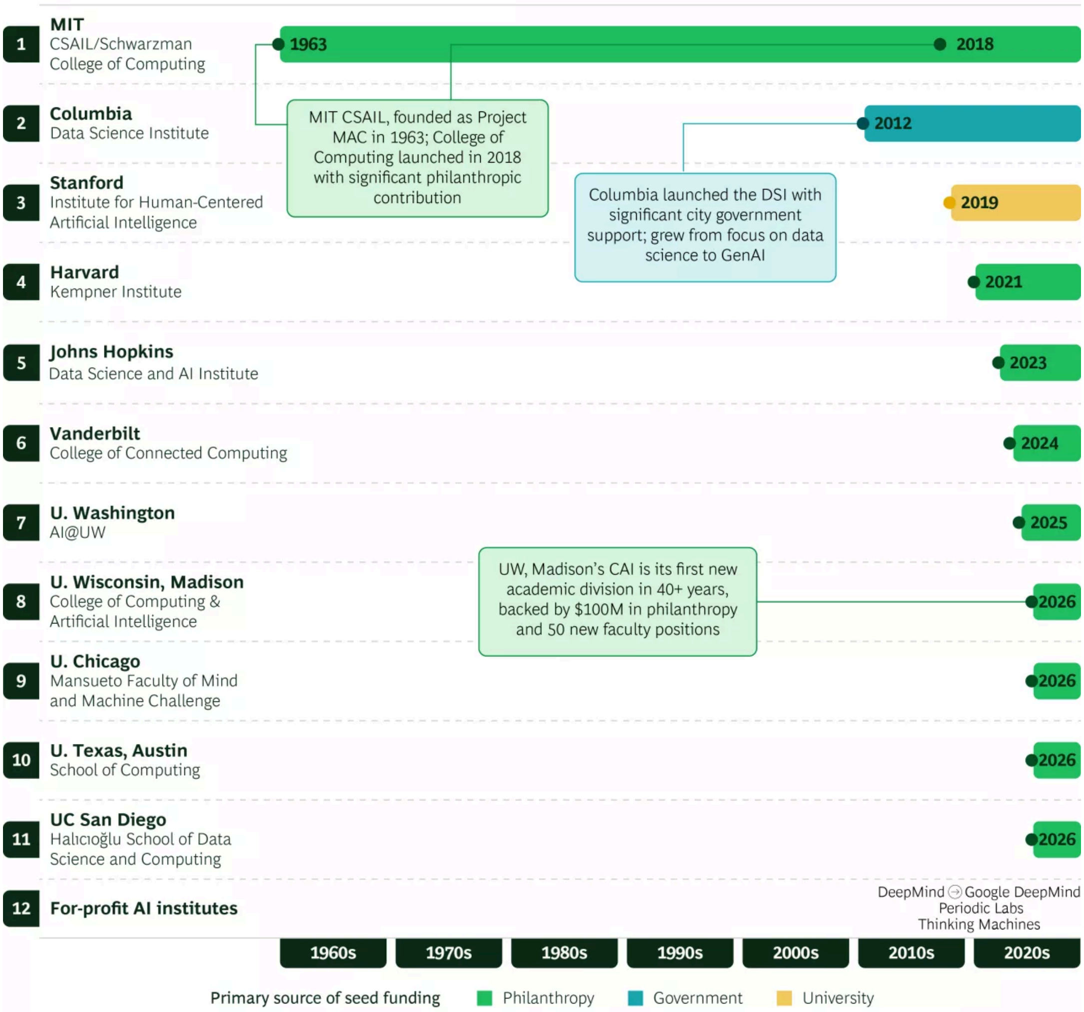
Former trainees Sergey Ovchinnikov (now MIT faculty) and Minkyung Baek (now Seoul National University faculty) are among the architects of the modern protein-structure-prediction revolution, and co-founders of Xaira Therapeutics, Outpace Bio, and Vilya all trace their scientific lineage to the lab. Baker himself has co-founded 21 biotechnology companies. Even one of the most commercially prolific academic labs in the world still sends a plurality of its PhD talent back into academia. The key is that it generates compounding networks of academic and commercial influence. Institutions can do much more to map out these networks and use them to their advantage, as they already do for school-level alumni.

Flexible leave and IP policies matter more than ever. Universities that make it easy for faculty to move between academic and commercial roles, and cleanly negotiate equity participation in ventures that emerge, can capture more value than those that force binary choices.

Model 3: New Schools of Computing and AI Institutes. Increasingly, universities are turning toward creating entirely new schools dedicated to exploring the foundation and frontier of data science and computing. (See Exhibit 3.) Johns Hopkins is building the Data Science and AI Institute (DSAI) with 80 new tenure-track faculty positions as part of a broader expansion, which is expected to generate $1.6 billion in economic impact. Vanderbilt, the University of Chicago, and others are following. Federal funding is evolving toward large-scale funding and nimble, team-based efforts: NSF's X-Labs funding call and the Department of Energy's Genesis Mission both signal a preference for cross-sector collaborations, often led by full-time teams over traditional single-PI grants and research centers.[6]

EXHIBIT 3

Timeline of Formation of New Schools of Computing and AI Institutes



Source: BCG analysis.
Note: CSAIL = Computer Science and Artificial Intelligence Laboratory. DSI = Data Science Institute. CAI = College of Computing and Artificial Intelligence.

At the same time, IP disputes in AI research partnerships are rising. Who owns a model trained on university data with corporate compute is unsettled in law and practice. Therefore, it is important that universities establish clear frameworks before entering these arrangements. We expect a likely shift toward co-investment and co-development partnerships rather than traditional licensing-first models.

Layer 2: Securing and Deploying Compute

The scale of the gap is considerable. Big Tech's four largest hyperscalers (Amazon, Google, Meta, and Microsoft) spent approximately $410 billion on AI infrastructure capital expenditure in 2025, with projections approaching $785 billion and about $1 trillion in 2026 and 2027, respectively.[7] Additionally, US-based AI companies raised approximately $286 billion in private investment in 2025, of which California-based entities alone accounted for $218 billion (75%+ of US total).[8] University compute budgets are orders of magnitude smaller than industry's. The strategic question is not whether to close this gap, but how to make university research indispensable despite it. Three models are emerging:

- **Model 1: The Single-Institution Powerhouse.** UT Austin's approach has been to become a single-institution powerhouse by assembling the most concentrated academic computing infrastructure through the $457 million NSF LCCF grant and partnerships with Dell and Nvidia.[9] The Horizon supercomputer entered production in 2026 and will offer roughly 100x the AI performance of Frontera, until now the leading US academic supercomputer.

- **Model 2: The State-Convened Consortium.** New York's Empire AI is the most ambitious multi-university compute consortium in the country, backed by over $500 million in public and private funding, including up to $340 million in state capital.[10] The consortium includes 10 institutions (Columbia, Cornell, NYU, RPI, SUNY, CUNY, University of Rochester, RIT, Flatiron Institute, and Icahn School of Medicine at Mount Sinai) and has more than 320 active users across more than 120 projects. Its Beta system, featuring Nvidia's Blackwell chips and slated to come online in 2026, will be 11 times more powerful than the current Alpha system. California has followed with CalCompute, housed in the UC system. The Empire AI model has also seeded adjacent talent efforts such as the Empire AI postdoctoral fellowship at Cornell Tech, funded with $10.5 million from Bloomberg co-founder Tom Secunda.[11]

- **Model 3: Federal Supercomputing.** The Department of Energy's Genesis Mission and NSF's LCCF represent a new model for allocating supercomputing time to academic researchers. DoE's proposal to offer federal lands for private data center development in exchange for shared compute access points to an emerging model that states could replicate within their research universities.[12]

Leaders should be aware that a compute investment dependent on a single funding source is inherently fragile; diversified models leveraging multiple funding sources tend to be more resilient. This is especially the case in times of federal budget uncertainty. Universities should also account for capital expenditure cycle risk: compute investments made today reflect current GPU architectures, and there is limited clarity on how institutions will refresh infrastructure at the pace that technology is advancing. A $100 million cluster that becomes obsolete within three years before generating its intended research output represents significant stranded-asset risk. Compute access is also increasingly a recruiting tool, with commitments to compute becoming a standard element of faculty startup package negotiations.

Layer 3: Reimagining Corporate Partnerships

Many corporate partnership offices were built for an era in which companies sought brand affiliation, sponsored hackathons, and recruited graduates. Two important developments are prompting university leaders to reevaluate how they approach these partnerships.

- **AI companies need and often cannot easily acquire or replicate the curated datasets that universities hold.** While especially common in universities with affiliated hospital systems (e.g., datasets on longitudinal health studies and genomic databases), academic datasets span industries and include decades of data on topics from agricultural field trials to robotics testing, as well as detailed data on infrastructure and energy systems and more. There is no proven, replicable model for extracting full strategic value from these assets. The institution that develops one, covering governance, pricing, access terms, and model ownership, can seize durable competitive advantage.
- **Manufacturers, Fortune 500 companies, and local corporations are seeking academic AI expertise for product design, pipeline, and deployment.** For example, Indiana-based Eli Lilly and Purdue University announced a $250 million partnership in 2025 to accelerate AI-

powered drug discovery and AI-enabled manufacturing scale-up.[13] Universities have the potential to become preferred AI implementation partners for regional and national industry, not just model developers. (See Exhibit 4.)

EXHIBIT 4

There Is No "One-Size-Fits-All" Approach to Corporate Partnerships

Model	Best fit	Tradeoffs
Central corporate partnerships office A single front door with university-wide authority	University president will commit significant political capital, and the current model is underperforming *or* Few existing "fiefdoms" exist to resist consolidation	Strong faculty and school buy-in required Risk of centralization slowing nimble research-level deals
Hybrid consulting An internal consulting shop taking on 2–3 high-priority institutional projects per year	University president has a curated set of high-value relationships to reimagine/scale and wants to use focused wins as a proof of concept before broader rollout	Slow to scale Depends heavily on presidential sponsorship and a skilled in-house team
Federated Provision of playbooks and support for school-level teams	Strong internal teams already exist, or restructuring corporate partnership revenue is not a top institutional growth driver or priority	Hard to coordinate cross-university opportunities Dependent on quality and cooperation of existing teams

Source: BCG analysis.

There is the concern that corporate partnerships can become talent pipelines in the wrong direction: close collaboration with an industry lab can accelerate a faculty member's departure. To address this, it is important to structure partnerships to deepen engagement rather than facilitate exits. Health data, student records, and research subject information carry significant privacy obligations that need proper oversight. Finally, convincing individual researchers and schools that there is value in expanding corporate partnership scope requires both a compelling institutional vision and an incentive structure that rewards faculty who bring partners to the table for institution-wide collaborations. OpenAI, Microsoft, Salesforce, and AWS are among the companies expanding collaborations with universities through research funding, compute and model access, and support for applied AI projects.

Layer 4: Winning the R&D Talent Competition

The AI talent landscape is rapidly evolving. Roughly 70% of AI-relevant PhD graduates now take industry positions post-graduation, up from 20% two decades ago.[14] A 2025 study found the outflow is highly selective: young, highly cited researchers at leading institutions are the most likely to leave, and their research tends to show reduced novelty and impact after the move. [15] This drain of early-career, high-output talent has compounding effects on the pipeline of academic mentors, the direction of research agendas, and the capacity of universities to train the next generation. Compensation is a big driver of the exit from academia. One 2026 study found the salary gap between industry and academia has risen more than fivefold since 2001. AI researchers at large tech firms can earn $500,000–$2 million+ in total compensation versus $120,000–$150,000 base salaries for assistant professors. [16]

Federal funding uncertainty has contributed to what is a paradoxical market for talent. Admissions to PhD programs at 55 leading US research universities fell 15% for the Fall 2026 semester, and several top universities have imposed hiring freezes. [17] At the same time, demand for AI expertise has never been higher. A new $47 million NSF pilot, combining doctoral training with industry placements and joint advising, points to one way universities can better align research education with the labor market. [18] Even with such efforts, the compensation gap will remain a challenge.

Universities cannot match industry salaries. But they can compete with industry when it comes to:

- Freedom to publish openly and pursue curiosity-driven agendas unconstrained by product roadmaps or earnings calls.
- Multi-year research endowments that insulate researchers from grant cycles.
- Flexible leave policies that allow faculty to move between academic and commercial roles, including dual appointments with affiliated research entities.
- The opportunity to train and mentor the next generation.

Universities could also do much more to go beyond the school-level tracking they currently do. One step is to begin mapping their AI lab alumni networks more systematically—identifying where a lab's alumni have landed, maintaining those connections, and creating pathways for philanthropic giving and co-founding opportunities.

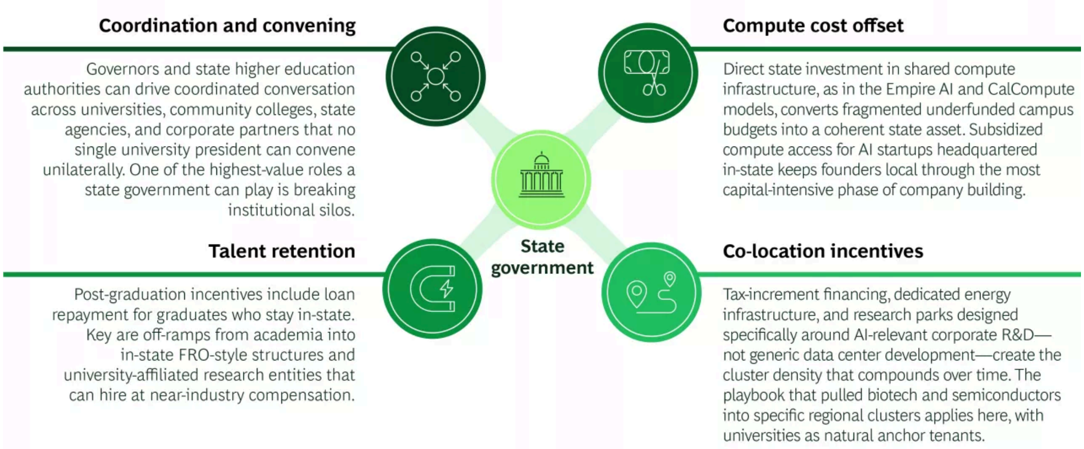
Another development worth noting is that China now produces significantly more AI-relevant PhDs than the US and is responsible for 74.2% of all AI patents granted worldwide. [19] The Arena Leaderboard spread between the best American and Chinese AI models had narrowed to 2.7% as of March 2026. [20] University leaders must navigate an environment where international research collaboration is simultaneously necessary for scientific progress and more complicated than it has been in decades.

Scaling with State Partners

The strategic choices about AI sit largely within universities, but they are difficult to execute at scale without state partnership. Increasingly, states are moving beyond general innovation strategies to capitalize on shared research infrastructure and organize universities, companies, and public agencies around AI. New York's Empire AI, Pennsylvania's Keystone AI + Quantum Factory, and the University of Florida's HiPerGator AI supercomputer illustrate the range of models emerging, from multi-university compute consortia to concentrated investments in flagship institutions and statewide shared infrastructure. University leaders can engage proactively in shaping these investments through active collaboration with their state partners. (See Exhibit 5.)

EXHIBIT 5

States Can Play an Active Role in Supporting the University AI Stack



Source: BCG analysis.
Note: FRO = focused research organization.

Seven Actions to Turn Ambition into Action

Every research university must decide where it can lead, what capabilities it must build, and which partnerships it needs to succeed through a cohesive, integrated approach to AI. (See Exhibit 6.)

EXHIBIT 6

Seven Actions University Leaders Can Take Now

1	**Appoint a senior leader with cross-cutting authority.** This individual should have a mandate to coordinate across research, IT, corporate relations, the provost's office, and state government negotiations.
2	**Develop a data strategy framework.** Start with a single high-value dataset and work through governance, pricing, access, and model ownership in a real negotiation.
3	**Join or build a compute consortium.** Empire AI, CalCompute, and NAIRR offer templates. These partnerships can lead to multi-institution datasets and co-development with AI companies and VCs.
4	**Conduct a corporate AI stack audit.** Map your position across all four layers. Identify where different offices are pursuing parallel negotiations with the same companies.
5	**Restructure the corporate partnership model.** Select the right governance model (central, hybrid, federated) based on the institution's starting point.
6	**Pursue FRO partnerships proactively.** As government funders invest in FRO structures (e.g., NSF X-Labs), secure founding-partner status through pre-negotiated IP terms, recurring talent exchanges, and facility-use agreements.
7	**Invest in talent retention, not just attraction.** Flexible leave, multi-year endowments, sabbaticals to work at FROs, and guaranteed compute access may matter more than headline salaries. Begin mapping AI lab alumni diaspora now.

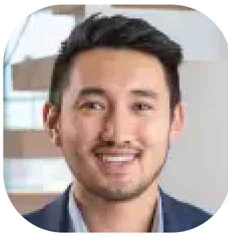
Source: BCG analysis.
Note: FRO = focused research organization.

The window to shape that position is narrowing. Talent, corporate partnerships, philanthropic capital, and public investment are already concentrating around institutions that have a clear AI ambition they are executing. University leaders that act now will be better able to attract and retain talent, maintain influence, and set the direction of research.

Authors

Tejus Kothari

Managing Director & Partner
Chicago



Bob Wu

Partner
Boston

J.R. Sullivan

Manager, BCG Vantage
Washington, DC

Daniel Goetzel

Harvard Reimagining the Economy Fellow; President, Chalant Strategies

ABOUT BOSTON CONSULTING GROUP

Boston Consulting Group bridges the gap between ambition and outcomes for the world's leading companies and organizations. We are built for this era of unprecedented change — bringing strategic clarity rooted in over 60 years of deep domain knowledge, combined with applied AI shaped by our practitioners. BCG works shoulder-to-shoulder with CEOs across industries and geographies to deliver transformative impact at scale: stronger returns, transferred capabilities, and change that sticks. For more information, visit bcg.com.

1 Data Center Dynamics, “National Science Foundation to Invest $457m in Computing Facility to House Horizon Supercomputer,” July 30, 2024.

2 White House Office of Science and Technology Policy, *Science: A New Golden Age*, July 21, 2026.

3 Stanford Institute for Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report 2026*.

4 Nur Ahmed, Muntasir Wahed, and Neil C. Thompson, “The Growing Influence of Industry in AI Research,” *Science* 379, no. 6635, March 2, 2023.

5 US National Science Foundation, “NSF Announces $1.5B NSF X-Labs Initiative to Pursue Generational Breakthrough Science Efforts,” May 14, 2026.

6 The Genesis Mission is a Department of Energy initiative, launched by executive order in November 2025, that aims to embed AI across DoE's national labs and research programs to accelerate scientific discovery. Its first 278 projects were selected for award negotiations in July 2026. Of these, 168 projects are university-led.

7 Hugh Son, "Moody's Says 'Unprecedented' AI Spending Threatens Credit Quality of Amazon, Meta, Alphabet and others", CNBC, July 24, 2026.

8 Stanford Institute for Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report 2026*.

9 *Data Center Dynamics*, "National Science Foundation to Invest $457m in Computing Facility to House Horizon Supercomputer," July 30, 2024.

10 New York State Office of the Governor, "Governor Hochul Announces $40M to Launch Empire AI Beta Supercomputer," June 26, 2025.

11 Cornell Tech, "Tom Secunda Gives $10.5M to Cornell Tech for Empire AI Research," March 5, 2025.

12 US Department of Energy, "DOE Identifies 16 Federal Sites Across the Country for Data Center and AI Infrastructure Development," April 3, 2025.

13 Wes Mills, "A Year of Advancing the Lilly-Purdue $250M Research Partnership to Accelerate Medicine Discovery and Manufacturing," *Purdue University News*, May 11, 2026.

14 Stanford Institute for Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report 2026*.

15 Ufuk Akcigit, Craig A. Chikis, Emin Dinlersoz, and Nathan Goldschlag, "The Great AI Talent Migration: Why Universities Are Losing the Future of Innovation," Centre for Economic Policy Research, April 11, 2026.

16 Ufuk Akcigit, Craig A. Chikis, Emin Dinlersoz, and Nathan Goldschlag, "Attention (And Money) Is All You Need: Why Universities Are Struggling to Keep AI Talent", NBER Working Paper no. 34964, March 2026.

17 Vimal Patel, "Decline of Ph.D. Admissions Could Imperil a 'Generation of New Talent,'" *The New York Times*, July 6, 2026.

18 US National Science Foundation, "NSF Partners with Universities and Industry on Pilot Initiative for Four-Year Ph.D. Programs with Real-World Research Placements," July 29, 2026.

19 The White House Council of Economic Advisors, *AI Talent Report*, January 2025; Stanford Institute for Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report 2026*.

20 Stanford Institute for Human-Centered Artificial Intelligence, *Artificial Intelligence Index Report 2026*.