OPINION

Kim, Jinyoung

# Foundational AI Innovation by Startups and the Role of Venture Capital

Artificial intelligence (AI) is widely recognized as a transformative technology driving structural shifts in the global economy and society in the 21st century. The domestic development of foundational AI technologies has emerged as a key national priority for many countries. An analysis of arXiv, a key platform for tracking global AI research trends, shows that the US published roughly 13 times more AI research papers than Korea between 2021 and March 2025, far exceeding the two countries' population ratio of 6.5 to 1. When narrowing the scope to private companies, the gap widened further to 17.3 times. These figures reflect the broader reality that US startups actively contribute to foundational AI research, whereas Korean startups play only a limited role in this domain. This disparity largely stems from the structure of the US venture capital (VC) market. The US VC ecosystem supplies AI startups with long-term "patient capital", enabling them to prioritize research and development (R&D) over immediate financial returns. In contrast, private-sector VC investment in Korea remains very limited in scale, reflecting a typical market failure in financing high-risk technology sectors. To address this challenge, a strategically designed, government-backed fund of funds is needed to support startups developing foundational technologies. In addition, proper incentive structures must be established to attract long-term, risk-tolerant limited partners (LPs) into the venture capital market.

## Enhancing research capabilities for foundational AI innovation

Artificial intelligence (AI) has emerged as a transformative technology that is reshaping the global economy and society in the 21st century. The advancement of domestic AI research serves as a critical foundation for securing core technologies. However, AI research and development, particularly in the deep tech segment, entails higher technological uncertainty,

* All opinions expressed in this paper represent the author's personal views and thus should not be interpreted as Korea Capital Market Institute's official position.

* Ph.D., Research Fellow, Fund & Pension Department, Tel: 82-2-3771-0645, E-mail: jykim@kcmi.re.kr

longer commercialization timelines, and substantial upfront investment in early stages, compared to conventional technologies.[1)]

Despite these challenges, strengthening domestic AI research capabilities is essential for ensuring technological sovereignty and reducing dependence on foreign-controlled critical infrastructure. In highly sensitive sectors such as defense, finance, and healthcare, overreliance on foreign technologies may increase exposure to security risks and supply chain vulnerabilities. Moreover, domestically driven research enables the effective development of AI infrastructure tailored to a nation's linguistic and cultural contexts.

Foundational AI technologies can also serve as a new growth engine, particularly through the creation of high value-added jobs. Securing these technologies can help curb the outflow of domestic AI talent and attract top-tier global researchers, thereby fostering greater interest among younger generations in science and technology. In addition, a strong domestic research ecosystem and a talent pool with deep expertise can enable the country to secure meaningful influence in the global discourse on AI ethics, regulation, and standardization. In this context, enhancing AI research capabilities is a key strategy not only for securing sustainable national competitiveness, but also for ensuring social stability and national autonomy.

## Tracking AI research trends through arXiv preprints

The pace of AI research continues to accelerate across the globe, with new findings emerging daily. A key indicator of these developments is arXiv, an open-access preprint repository where researchers can share their work prior to formal journal publication. In the AI field, in particular, arXiv serves as a primary outlet for cutting-edge research.[2)] However, since arXiv does not require peer review, it includes a significant number of early-stage or experimental papers, and in some cases, authors' institutional or national affiliations are not clearly identified. Despite its limitations, arXiv remains a valuable resource for tracking trends at the global AI research frontier.

An analysis of papers submitted to arXiv's three major AI-related categories reveals that the

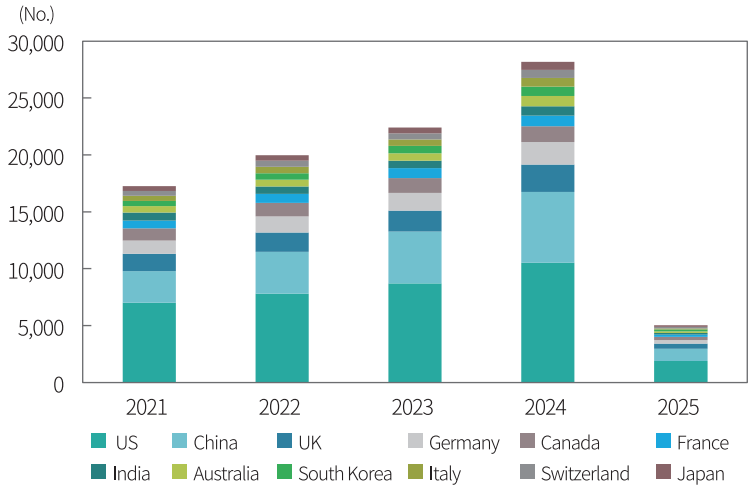
1) Deep tech encompasses a range of fields, including AI, advanced materials, drones, robotics, optoelectronics, quantum computing, climate technology, biotechnology, and blockchain.

2) Global tech leaders have frequently shared their core research on arXiv. Google's 2017 paper, "Attention Is All You Need," which introduced the transformer architecture central to today's large language models, was initially published on arXiv and played a catalytic role in accelerating AI research.

scale and trajectory of AI research. The number of papers, which stood at 15,683 in 2021, has steadily increased each year, reaching 25,512 in 2024.[3)4)]

Figure 1 presents annual counts of AI-related research papers by country from 2021 to March 2025. For papers involving international co-authorship, each participating country is counted once per paper. The data confirm that the US and China remain the top contributors to global AI research, followed by the UK, Germany, India, and South Korea. In 2024, US-affiliated institutions contributed to 41% of the 25,512 papers submitted to arXiv, highlighting the country's continued dominance in AI research.

**Figure 1. Annual AI paper submissions on arXiv by country of affiliated institution[5)]**




Source: arXiv

AI research typically demands extended development timelines and significant resources for technology design, experimentation, and data collection and validation. As a result, universities and research institutes have long played a central role in advancing AI research. As illustrated in Figure 2, 73.1% of US-authored AI papers originate from academic or research institutions, and listed and private companies account for 13.2% and 6.3% of the output, respectively, indicating

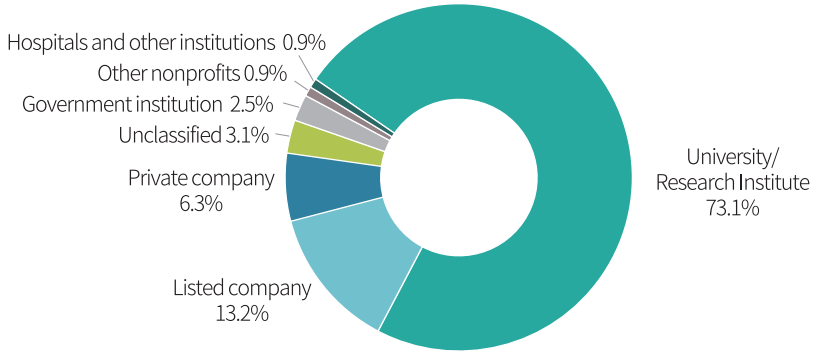
3) The three main AI-related categories on arXiv are stat.ML (statistical machine learning), cs.LG (computer science – machine learning), and cs.AI (computer science – artificial intelligence).

4) For papers with multiple versions, only the first version was counted. Papers listed under multiple categories were counted only once to avoid duplication.

5) For 2025, the analysis includes only papers published through March. Author affiliations were extracted from the first page of each paper, with data available for the majority of the 82,517 papers collected, excluding roughly 5,000 entries. The nationalities of the extracted institutions were identified using Wikipedia, the SEC's EDGAR system, and the Preqin database.

growing industry participation in AI research.

**Figure 2. Distribution of AI paper submissions by type of institution (US)**




Source: arXiv

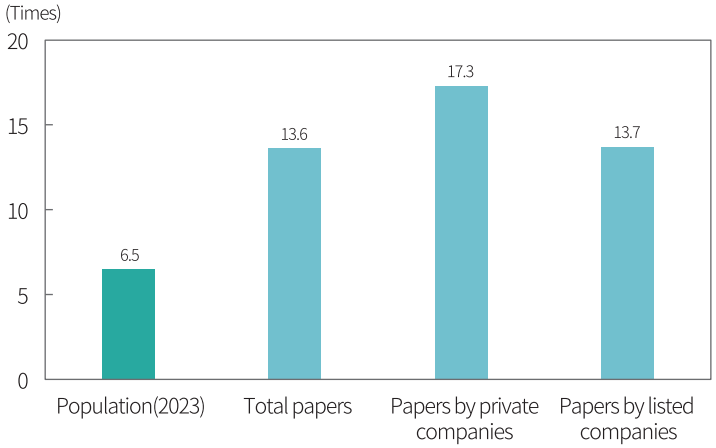
## Comparing AI research capacity: US vs. Korea

While Korea's AI research output by institution type resembles that of the US, population-adjusted comparisons reveal notable patterns. Figure 3 shows the ratio of US to Korean AI paper counts across institutional categories, based on arXiv submissions and 2023 population figures.

According to the World Bank, the US population in 2023 was roughly 6.5 times larger than Korea's, but the US produced 13.6 times more AI-related papers on arXiv.[6] The disparity is even greater in the category of private companies, where US firms submitted 17.3 times more papers than their Korean counterparts. For listed companies, the difference stood at 13.7 times, underscoring a particularly pronounced gap in AI research output among private firms. These figures indicate that US startups, particularly those backed by venture capital (VC), are contributors to AI research, while Korean AI startups may either lack sufficient research capacity or face structural barriers to participation in the AI research landscape.

6) https://data.worldbank.org/indicator/SP.POP.TOTL?locations=OE

**Figure 3. Comparison of population size and AI research output: The US vs. Korea**




Source: arXiv

AI research by US private companies is drawing global attention not only for its scale, but also for its qualitative impact. In its 2020 paper, OpenAI introduced the GPT-3 model, demonstrating the effectiveness of few-shot learning in large language models (LLMs).[7] This breakthrough redefined expectations around AI's generalizability and scalability. Anthropic proposed a novel approach called "Constitutional AI", aimed at improving model safety by enabling self-guided learning to reduce harmful behaviors.[8] Hugging Face, a leading open-source AI startup, enhanced model efficiency by distilling Google's BERT into a smaller, faster, high-performance variant.[9] These companies have played a pivotal role in driving LLM innovation and broadening access to the open-source AI ecosystem.

US AI startups, equipped with robust research capabilities, are producing transformative models. OpenAI's release of ChatGPT marked a turning point in the generative AI era. According to Stanford University's Human-Centered Artificial Intelligence (HAI) report, 7 of the 15 most prominent AI models released globally in 2024 originated from startups, including 3 from the US, 3 from China, and 1 from France.[10] These startups now compete directly with global tech giants such as Google, NVIDIA, Meta, and Apple, indicating their growing influence in the race for AI leadership. This indicates that the US startup ecosystem is catching up to established

7) Brown et al., 2020, Language models are few-shot learners, Advances in neural information processing systems 33, 1877-1901.

8) Bai et al., 2022, Constitutional ai: Harmlessness from ai feedback, *arXiv preprint arXiv*: 2212.08073.

9) Sanh, V., Debut, L., Chaumond, J., Wolf, T., 2019, DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, *arXiv preprint arXiv*:1910.01108.

10) Stanford University. 2025. Artificial Intelligence Index Report 2025.

companies in innovation capacity and increasingly serving as a key source of advanced technologies.

In 2025, Stanford's HAI report ranks Korea first in AI-related patents per 100,000 people and seventh in overall AI competitiveness,[11] reflecting the country's strong capabilities in applying commercialized technologies. However, Tortoise's Global AI Index places Korea 13th in the research category, exposing its weakness in foundational research and core technology development.[12] Korean startups have largely focused on integrating existing AI models into services, rather than pursuing the development of proprietary core technologies.[13]

## Venture capital as the engine of US AI startup growth

The divergence in AI research capacity between Korean and US private firms can be attributed to the scale and sophistication of the US VC ecosystem that supports AI startups. The US continues to lead the global VC market. According to KPMG, global VC investment reached $368.3 billion in 2024, with the US accounting for $209 billion or roughly 57% of the total.[14]

US leadership is particularly pronounced in VC investment in AI startups. Figures 4 and 5, based on Preqin data, illustrate both US and global VC investment trends from 2016 to 2024, measured by value and deal count. VC investment in AI peaked in 2021, driven by abundant liquidity under the quantitative easing policy, before moderating due to macroeconomic uncertainty and higher interest rates. Nonetheless, the US remained the world's most active investor in AI startups throughout the period analyzed. In 2024, the US captured nearly 66% of global AI VC funding and 36% of total deals.

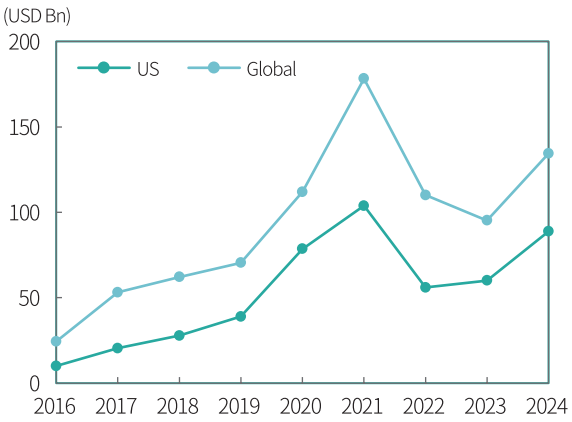
11) Stanford University. 2025. Artificial Intelligence Index Report 2025.

12) Tortoise. The Global AI Index. https://www.tortoisemedia.com/data/global-ai

13) AI Times, January 21, 2025, "Korea's AI Industry: Policy Missteps and Weak R&D Commitment from Companies"; Yonhap News Agency, April 3, 2023, "Is the Boom in GPT-Based AI Services Sustainable? Concerns over Technology and Data Dependency."

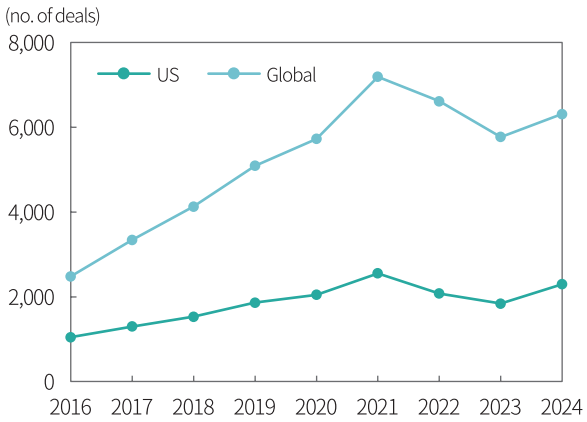
14) KPMG. 2024. Venture Pulse Q4 2024.

**Figure 4. VC investment in AI startups**




Source: Preqin

**Figure 5. No. of VC investment deals in AI startups**




Source: Preqin

By contrast, Stanford's HAI report indicates that Korea faces significant constraints in mobilizing private capital needed to build a globally competitive AI startup ecosystem. As of 2024, Korea's private-sector AI investment stood at only $1.33 billion, roughly 1% of the $109.08 billion invested in the US. According to the World Bank, Korea's 2023 GDP was around 6% of the US level, but its AI-related private investment falls far short of this proportion.[15] This divergence reflects challenges in supplying private capital to high-risk technology sectors with limited exit opportunities. As such, insufficient financing for AI startups could undermine Korea's ability to conduct AI research and secure core technologies over the medium to long term.

## Future challenges in fostering Korea's AI startup ecosystem

Among various factors for evaluating AI competitiveness, Korea continues to underperform in AI research and foundational technology development—areas where Korean startup involvement remains limited compared to global peers. Given its status as a national strategic technology, AI requires sustained and structured policy support.

As a deep tech field, AI carries far-reaching societal impact but involves high uncertainty and long development cycles before commercialization. As these characteristics require high upfront investment and lengthy periods for technology validation and market development, patient capital, especially in the form of VC, is critical for supporting deep tech startups.

15) https://data.worldbank.org/indicator/NY.GDP.MKTP.CD

Notably, VC investment in deep tech differs structurally from conventional VC models funding tech startups. Unlike conventional tech startups, deep tech firms require significantly more time to generate initial revenue, intensifying the risk of market failures as private-sector investors remain reluctant to commit capital.

Considering the unique nature of deep tech development, government support plays a pivotal role in fostering deep tech startups. In particular, government-backed funds of funds should serve as catalytic capital to address market failures, and be designed to consistently channel long-term, large-scale funding into foundational technology startups. In addition, to effectively achieve the strategic goals of a government-backed fund of funds, VC managers with expertise in different investment stages—such as early and later-stage financing—should be appropriately allocated.

Furthermore, deep tech-focused funds should carefully choose the types of LPs they bring in, ensuring they align with the long-term, high-risk nature of deep tech investing. As LPs typically provide a substantial share of a fund's capital, VC managers cannot overlook the risk appetite and investment preferences of their LPs. In the US, family offices, corporate venture capital (CVC) funds, university endowments, foundations, and defined contribution pension plans are among the key participants in venture capital funds investing in early-stage startups, driven by their higher risk tolerance in pursuit of outsized returns.[16] These LPs are considered well-suited for deep tech investments, which involve long time horizons before commercialization. To ensure a sustained, long-term capital supply for startups developing core technologies of strategic importance, Korea should build the institutional framework and incentive structure to attract such LPs to the domestic VC ecosystem.

16) The high-risk, high-return profile is a common feature of the VC market. In particular, seed and Series A investments, which carry the greatest uncertainty, recorded annualized returns of 25.5% and 19.2%, respectively. Later-stage investments, including Series B, C, and D or beyond, yielded lower returns of 12.8%, 13.1%, and 14.5%, respectively. This reflects the typical VC investment pattern where risk and return both decline with later-stage investments. Source: PitchBook, 2024, Analyzing VC Returns by Series and Alternative Exit Strategies.