

Tracking China's Growing AI Innovative Capacity

Using Publications, Patents, and People

Sagar Baviskar, Lee Branstetter, Cameron Drayton, Doug Hanley, Eduard Hovy, Lei Li, Xibao Li, Prasanna Tambe, Jiarui Wu, Liuji Wu, Huiyan Zhang

Preliminary and Incomplete

September 12, 2026

Abstract

The innovative capabilities of some Chinese universities, research institutes, and firms in the domain of AI have grown quickly and reached high levels. However, efforts to quantify these developments and compare them to developments outside China have been clouded by the differences between Chinese and Western university and patent systems, significant policy interventions by various levels of the Chinese government seeking to boost the level of patenting, differences in Chinese and Western financing of firm innovative efforts, and the role of state-owned enterprises in the Chinese economy. It has also been hampered by the nature of AI as a general purpose technology and the inherent difficulty in drawing boundaries around “AI invention.” This paper draws upon the emerging literature on the Chinese innovation system to take these issues into account and re-evaluate what can be learned from the large increase in recent Chinese AI-related publications and patents. We conclude that Chinese organizations have developed impressive AI invention capabilities over a relatively short period of time, but the reflection of those capabilities in Chinese patent data is significantly biased by the enduring impact of increasingly well-documented policy interventions.

1. Introduction

“When a measure becomes a target, it ceases to be a good measure.”

Charles Goodhart, British economist, 1975

Abundant evidence exists of China’s growing scientific prowess and even intellectual leadership in key fields. Yet standard bibliometric analyses of article texts can struggle to ascertain where and to what extent this exists. Chinese computer scientists working for industrial firms and academic institutions publish large numbers of papers documenting their advances, but determining the true scientific value and innovativeness of Chinese publications is complicated by well-documented biases in both Chinese scientific publications and the citations in these documents that are not unique to computer science but are certainly present in it. Strong incentives to publish have led some Chinese scientists, especially at lower-tier universities, to rush publications, cut corners, and occasionally fabricate data; 60% of science article retractions have a Chinese author (Van Norden, 2025). Chu, Steinwinder, and Azoulay (2025) document a striking “home bias” in Chinese scientists’ citations of other Chinese academic publications -- a bias so strong it substantially distorts the true global impact of Chinese academic science across fields.

Similar challenges obtain in the case of Chinese patents. An expanding literature documents the strong incentives in China to patent low-quality inventions and the distortions this generates for the Chinese innovation system (Boeing et al, 2026; Branstetter et al, 2025; Cao et al, 2024; Chen et al., 2021; Dang and Motohashi, 2015; Li, 2012; Wei et al, 2021; Sun et al, 2021; Hu, 2014; Lin et al., 2025). Simply counting and comparing “AI” patents granted in the U.S. and in China is neither appropriate nor sufficient as a means of evaluating Chinese AI prowess relative to the U.S.; attempts to do so generate a number of anomalies (Fang et al, 2026; Lerner et al., 2026). The emergence of Chinese technological leadership in important fields of engineering is undeniable, but, as with papers, standard NLP approaches to the analyses of patent texts can obscure as much as they reveal.

Despite these challenges, publications and patents reveal important truths about the advance of China’s growing AI capabilities that cannot be obtained elsewhere. Properly used, these data not only permit broad international comparisons but also illustrate how and where in the technological product space Chinese institutions and firms are generating major advances. The key is to separate the signal contained in documentation of real advances in academic science (by papers) and industrial technology (by patents) from the noise generated by low-quality patents and papers. This is difficult to accomplish in practice using standard NLP tools because of the role patents and papers play in China in signaling compliance with and support of government policy priorities. Authors of these documents can manipulate patent and paper text to reflect the “novel” fields in which authorities wish to see greater scientific and R&D effort. Yet, the appearance of key words and phrases and the apparent elaboration of novel-sounding concepts in these documents can mask a limited degree of truly substantive innovation.

How can we separate the wheat of real AI innovation documented in patents and papers from the chaff of patent and paper texts deliberately constructed to make this separation difficult? In the case of papers, the normal operations of the scientific community offer one approach – a focus on papers accepted at leading international conferences (the usual approach taken in the fast-moving field of computer science) or published in leading journals. While anyone can post a working paper online and an ongoing explosion of lower-tier conferences and journals means many marginal articles are eventually “published,” an elite tier of conferences and journals focused on AI have used international committees to carefully screen and select articles for inclusion. The conferences have expanded over time as interest in AI has grown, but the acceptance rates have remained fairly stable, such that a reasonably consistent quality screening process, undertaken by credentialed experts, has been applied to all submitted papers over time. By limited our purview to papers published in those elite venues, we can create baseline indicators of scientific capability that are broadly comparable across publishing institutions worldwide.¹

The quality certification problem is far more difficult in the case of patents. Under international law, the granting of a patent is prerogative of national governments. While the TRIPs (Trade Related Intellectual Property Rights) Agreement supposedly imposes minimum standards of operation on the patent systems of WTO-signatory states, the reality is that national patenting systems vary greatly in their operation across countries and over time, and there can be enormous differences in the degree of screening for quality and novelty. However, the recent emergence of AI out of academic science, and the documented close connection between AI science, AI scientists and AI invention that persisted in the United States through at least the mid-2010s offers up the possibility of using AI publications and their authors as a source of ground truth concerning the distribution of real AI innovative capabilities across organizations, geographic areas, and time in China and elsewhere (Baviskar et al., 2026).

The approach taken in this paper will therefore focus on author and inventor identity, and the position of authors and inventors within the Chinese (and global) networks of AI co-authorship, co-invention, and co-work. As in the United States, the current coincidence of scientific opportunity and economic value in AI generates significant flows of talent between elite universities, leading firms, and disruptive start-ups in China. With some well-documented exceptions, leading AI innovation in China is generally undertaken by top executives, research directors, and/or leading engineers trained in elite university labs in China or the West. Top Chinese academic and industrial labs regularly present at top international conferences and have their work cited extensively by elite researchers outside China. While a rising tide of derivative, lower-quality academic work outside these elite circles is emerging in response to

¹ This statement may require some caveats. Chu, Steinwender, and Azoulay (2025) document a pronounced home bias in the citations of Chinese scientific papers across fields that could exaggerate the true quality of Chinese publications in AI accepted to these conferences. Set against that is the potential for bias in the opposite direction: the selection committees may have limited Chinese participation, the emergence of Chinese universities as centers of truly pathbreaking AI research is quite recent, and it could be that European and American academics are exercising some degree of (perhaps unconscious) bias against papers authored by scholars they may not personally know affiliated with institutions of which they have limited knowledge.

government directives, our work will identify the authors, academic centers, and industrial labs with strong network positions and international impact. Similarly, we will draw upon the approaches developed in Baviskar et al. (2026) to screen 8.5 million Chinese invention patents, identify those that are AI-related, and then screen Chinese AI patents on the basis of applicant/inventor status within the computer science network, the propensity of inventors to patent the invention outside China, and other factors.

2. Tracking AI patents in the U.S. and China: a first look

The earliest efforts to track AI inventions were constrained to search patent titles and abstracts for a limited set of keywords and /or screen patents on the basis of a limited number of patent classes now known in the United States as CPC codes (Cockburn, 2019). This is problematic, because the nature of AI as a general-purpose technology means that conceptually similar algorithmic tools can be applied in nearly every sector of the economy, and patents that codify such applications are often classified according to the industry of use rather than the underlying technology (Agarwal et al., 2019a, 2019b; 2025b; Agarwal, Brynjolfsson, and Korinek, 2025; Alderucci et al., 2024).² Efforts to leverage natural language processing (NLP) and machine learning (ML) techniques to analyze patent text led later researchers to identify a much larger number of patents across a broad domain of patent classes (Alderucci et al., 2024; Giczy et al., 2021, 2023; Miric et al., 2023). More recently, researchers have also used the growing capabilities of large language models (LLMs) to delve even more deeply into patent texts generated by different patent systems (Baviskar et al., 2026; Fang et al., 2026; Wu et al, 2026). In principle, LLMs hold a great advantage over earlier approaches, since they can take into account sentence structure, embeddings, and other features of text to more accurately distinguish between AI-related inventions and other categories of inventions (Ganguli et al., 2023).

As the next section will make clear, the growing power of LLMs and the widening availability to academic researchers of GPUs and TPUs designed to accelerate the use of LLMs on even large-scale data processing tasks does not eliminate the challenge of delineating the boundaries of AI invention. Modest differences in LLM prompts or training data can lead to differences in what gets classified as AI invention (Alderucci et al., 2025; Fang et al, 2026). The general-purpose nature of AI means its ultimate economic impact may be very large (Jones, 2026). That same general-purpose nature means that determining the difference between applications of real AI and software algorithms that “make decisions” (e.g., standard control system programs that have existed for decades) or “learn from data” (e.g., statistical packages used by economics graduate students) can be quite difficult. This study will deal with these challenges, in part, by using different approaches to AI definition, and demonstrating that the main conclusions of this paper remain quite robust to these differences.

We start with our “baseline” definition, which is meant to be broad and inclusive, in a manner that includes applications of well-established AI and ML ideas to specific products, services, and functions, as well as fundamental advances in the theory and generation of large language

² See also Sun and Trefler (2023), who use a more extensive set of key words and classes identified by WIPO to identify AI patents.

models. The baseline is also deliberately meant to include inventions that would have been considered AI in earlier eras, such as the 1990s and early 2000s.

The baseline definition is “Artificial intelligence refers to computer-implemented systems or methods that perform intelligent functions, including, but not limited to, learning, reasoning, perception, knowledge representation, planning, pattern recognition, prediction, autonomous decision-making, or goal-directed action.”³

Figure 1 depicts the main empirical puzzle of this paper, by graphing the numbers of AI patents by grant year according to this definition, in both the U.S. and China, stretching from 1987 through 2025. Through 2015, patents are rising over time in both countries, with the U.S. maintaining a significant lead in overall patent numbers. Then we see a phenomenal explosion of AI patenting in China that dwarfs activity in the U.S., expanding the number of grants by an order of magnitude in a handful of years. The figure seems to suggest a promethean burst of innovation in China that is leaving the United States – and the rest of the world – in the dust. If mass media speak of an AI “race,” these numbers suggest an overwhelmingly one-sided Chinese victory.

We are not the only researchers to generate pictures like Figure 1. Fang et al. (2026) also undertake a careful comparative analysis of AI patenting in the U.S. and China. Their definition of AI and the time period over which they observe patent grants differs somewhat from that used in our paper, but they generate a picture qualitatively similar to the one depicted in Figure 1. Lerner et al. (2026) examine the explosion of Chinese patenting in multiple domains of “critical technologies,” including AI, albeit with a definition of AI patenting that differs from us and from Fang et al. (2026). Their findings are also broadly consistent with Figure 1. Are all of these similar-looking graphs accurately reflecting the sudden death knell of American technological leadership and the unstoppable rise of overwhelming Chinese technological supremacy?

Our authorship team includes computer scientists with academic and industry experience in AI at elite levels in both the United States and China. As impressive as the recent Chinese wave of academic science and industrial invention in AI is, our experience does not suggest a relative accounting that is so one-sided. Can we presume that real innovative outcomes are proportional to these patent grant numbers? Or are there structural features of the Chinese innovation system that make that kind of comparison problematic?

The rest of the paper is devoted to a careful investigation of the latter question. Like other recent studies, we find strong evidence that the distribution of AI patenting across organizations is quite different in China and the United States (Fang et al., 2026; Lerner et al., 2026). We then draw heavily from the emerging literature on the Chinese innovation system, its unique structure, and its particular interlocking web of incentives, to argue that patents play very different roles in China and the United States. In particular, patents play an array of signaling roles in China that do not have a close parallel in the United States. We build on this existing

³ This definition is summarized from the textbook *Artificial Intelligence: A Modern Approach* by Stuart J. Russell and Peter Norvig

literature by leveraging the fact that the AI industry has recently emerged from breakthroughs in science, its advance depends heavily on the work of a relatively small group of inventive individuals with highly specialized human capital in both countries, and the distribution of innovative outcomes across organizations tends to closely reflect the presence within those organizations of members of this elite group. The important role of academics and ex-academics in AI invention, the high propensity of these individuals and their co-innovators to *publish* major breakthroughs as well as take out AI patents, and the role of leading computer science conferences as showcases for important advances all allow us to create a map of AI innovation drawn from these publications and compare that to the map drawn from patents. In the U.S., these maps are broadly similar. In China, they are quite different, although the degree of difference depends on the definition of AI. Similarly, we can identify an elite group of science and technology leaders in both countries, track their movement across organizations, and examine how the map of elite innovators compares to the map of patenting. Once again, the maps broadly cohere in the U.S., but not so much in China.

These stark differences suggest reasonable corrections to Figure 1 that dramatically narrow the emergent gap in innovative outcomes between the two countries suggested by facile comparisons of patent counts. These lines of reasoning also suggest that a significant amount of AI patenting in contemporary China is not necessarily informative concerning the pace and direction of real innovative activity, especially in core AI. That being said, the map of innovative activity drawn by publications, people, and plausibly corrected patents point to a rapid convergence of AI innovation capabilities across the two countries and an impressively rapid accumulation of Chinese invention that survives all of our redefinitions and adjustments. If an overwhelmingly one-sided Chinese victory in an AI race seems inconsistent with these new facts, so does any hubristic notion of an enduring and significant American “lead.”

3. Issues with defining AI in patent data

A rapidly growing number of other studies also use patent data to identify AI-related inventions. One early study that influenced our own work is Cockburn et al. (2019). These authors select patents from two U.S. Patent Classes: 706, which is the designated patent class for artificial intelligence patents, and 901, which is the designated patent class for robotics. In addition, the authors search patent titles for a small number of specific keywords, such as “neural networks,” and include patents containing these keywords regardless of the classes to which the patents are assigned. Combining these two approaches, the authors find 13,615 unique “AI” patents granted between 1990 and 2014. Webb et al. (2019) take a broadly similar approach with a focus on narrower categories of advanced technology patenting, identifying over 2,000 patents related to “machine learning” and over 4,000 patents related to “neural networks.” Sun and Treffer (2023) use a much broader set of patent classes and key words, drawing upon the efforts of the World Intellectual Property Organization (WIPO) to generate a definition of AI patents that can be applied across patent corpora. Note that many of these papers preceded the emergence of Gen AI or LLMs, and they all adopt definitions that include what would have been considered AI invention in the late 1990s or early 2000s, as well as more recent developments. We take a similar approach, in that sense, adopting a broad definition that

effectively includes a whole genealogy of technologies, some of which are much less intensively used today.

As natural language processing NLP methods advanced and diffused from computer science into social science, scholars begin using ensembles of machine learning models and larger portions of patent text to classify patents as AI or not AI. One example is Alderucci et al. (2024) identified between 52,896 and 146,952 AI-related patents granted between 1990 and 2018, depending on the stringency of their definition. These authors examined the core text of the patent (including the patent claims and the patent description), not just the title and abstract. Second, their ensemble of machine learning algorithms allowed them to detect AI-related patents even if they do not explicitly include the small number of keywords used by, for instance, Cockburn et al. (2019). Third, they did not limit their purview to a handful of patent classes – and one of the most important findings of their paper is the (very) large number of AI-related inventions that are classified into patent classes associated with the domain of application of the invention rather than the patent class traditionally associated with the “upstream” AI concepts.

On the other hand, Alderucci et al. (2024) *exclude*, by design, a number of the patents Cockburn et al. (2019) included in their designated AI patent set, and this difference is easiest to describe in the context of the “robotics” patents from class 901. Patents in this category can include “hardware” patents that represent important advances in the arms, sensors, and other components of a robot “body,” but have little to do with the algorithmic advances or software innovations that have equipped recent generations of robots with more artificial intelligence, and it is the latter categories of invention on which they seek to focus. Alderucci et al. (2024) do not exclude all patents with a hardware component, but their methods are designed to exclude those that are not somehow also related to the ability of a robot or other AI-enabled system to learn, think, and respond intelligently to its environment. Earlier versions of this paper circulated at the same time that Miric et al. (2023) and the U.S. Patent and Trademark Office (Giczy et al., 2021, 2023) were writing papers taking a similar approach.⁴

Discussion of the debate over the inclusion of patents describing robot bodies highlights both the importance of the scope of AI adopted by a study and the ambiguity over the appropriate boundaries of AI that leads so many different research teams to make different choices, all of which lead to different definitions of AI versus non-AI patents. Should a definition include robot bodies as well as brains? Should a definition include patents that describe the semiconductors on which AI runs (e.g., GPUs and TPUs)? Should it include the equipment used to make those semiconductors or the software tools used to design them? Then there is the nettlesome question of where AI begins and traditional programming ends. Keyword searches or even modern NLP methods will often classify control system software, of the kind that has existed for decades, as AI. Some of these methods will also flag the kinds of statistical software packages long used by economists and engineers, because they “learn from data” or employ optimization tools that seem similar to those used in more sophisticated algorithms. Reasonable experts can

⁴ See also Keisner et al. (2015), De Prato et al. (2018), Van Roy et al. (2019), and Mann and Puttman (2018)

come to different conclusions, and there is, as yet, no widely acknowledged standard definition or approach.

The U.S. Patent and Trademark Office sponsored a serious effort to use then-current machine learning techniques to classify the entire USPTO corpus of granted patents as AI or not. Giczy et al. (2021; 2023) describe this effort. The USPTO authors identified eight AI component technologies (e.g. natural language processing, machine learning, as well as hardware applications) and then trained a machine learning model for each of these eight components to generate their Artificial Intelligence Patent Database (AIPD). This team went on to release their data to the public, allowing subsequent researchers to investigate (and second guess) their choices. Fang et al. (2026) and Wu et al. (2026) acknowledge the effort that went into this undertaking, but each of these groups of authors devotes substantial portions of their respective papers to constructive criticism of the USPTO approach. Each offers an alternative approach with better performance. Wu et al. (2026) take the commendable next step of making their data and methods public. However, these advances do not offer an easy way around the problem of defining a boundary of AI invention. Refinements of LLMs do not tell you whether patents describing the design of GPUs should be considered “AI” or not.

Despite the effort expended by many eminent researchers to track AI patents, some readers may question whether patents (or patents alone) are sufficient to track the movement of the AI frontier. Such readers might note that industry-leading firms like OpenAI and Anthropic in the U.S. have possessed relatively few patents, even as they introduced revolutionary products.⁵ The same is true of some of their Chinese counterparts. The existing corporate patents flagged by extant algorithms are far more likely to be (often somewhat pedestrian) applications of AI to particular products or use cases rather than patents describing in detail the leaps in the latest LLMs. Sun and Trefler (2023) address this issue in their paper, and nevertheless make a strong argument that patents are useful in tracking the commercial application of AI, perhaps with a lag. We agree, but note that our paper also makes extensive use of AI publications, which may capture some of the most significant innovations that are not (yet, directly) patented in major jurisdictions.

The rapid advance of LLMs (and relatively slow publication cycles in economics and management) mean that even recent papers on AI patent classification methods run the risk of being overtaken by technology. We take a favorable view of the approaches taken in Fang et al. (2026) and Wu et al. (2026). However, we find that providing relatively simple prompts to frontier LLMs tends to generate defensibly accurate classifications of AI versus AI patents in both U.S. and Chinese patent corpora. An appendix describes our prompt engineering approach in greater detail and compares the accuracy and precision of the classifications thereby generated to a human annotated ground truth data set as well as to the results of a multi-stage approach similar to that of Fang et al. (2026)⁶. In later versions of the paper, we

⁵ Current press reports suggest that both firms, as they prepare for massive IPOs, are aggressively building out patent portfolios enabling them to protect and defend the revenue flows from their inventions.

⁶ See Baviskar et al., 2026.

expect to introduce alternative methodological approaches. However, we anticipate that the main results of our paper will prove robust.

4. China's patent system: an international comparative perspective

Why should any inventor in any country take out a patent? In principle, patents provide a government-enforced monopoly right to an inventor for a fixed period of time; in return for receipt of this temporary right, the inventor permanently discloses the essence of her invention, allowing future inventors to build on it or recombine it with other ideas. The required disclosure raises the risk that rivals could learn from but then “invent around” the patent, such that the patent holder could not credibly sue for patent infringement. A vast literature on the actual operation of patent systems across multiple countries and many centuries of operation, a full review of which is well beyond the scope of this paper, suggests that this theoretical view of the workings of the patent system is often not perfectly aligned with empirical reality, either in the United States or elsewhere (Jaffe and Lerner, 2004; Shapiro, 2001; Merges, 1999; Cohen, Nelson, and Walsh, 2000). Nevertheless, this view remains a useful benchmark.

Note that there is no “global patent” – inventors must obtain a patent grant from every jurisdiction in which they seek intellectual property protection. For this reason, innovation scholars have long used the propensity for inventors to take out patents for the same invention in multiple countries as a strong signal of the invention's true value (Lanjouw and Schankerman, 1999, 2004). In this context, it is useful to note that Chinese inventors' average propensity to patent their domestic inventions in other countries is strikingly and persistently low – on the order of 3.5% (Lerner et al, 2026). The rates for American, Korea, Japanese, and German inventors are 40%, 20%, 40%, and 60% respectively. These stark and persistent differences are one reason Western scholars have questioned the quality of Chinese invention (Branstetter, Hanley, and Zhang, 2025). As the world's largest exporter of goods and a key hub in the global trading system, Chinese firms would seem to have a strong incentive to patent truly valuable inventions abroad. Low rates of foreign patenting are consistent with the view that large numbers of Chinese patents do not, in fact, actually describe inventions with significant value outside of China.⁷

Other striking differences between China's patent system and the U.S. system include the frequency of patent litigation and the distribution of patent damage award sizes. While recent reforms have raised legal guidelines for damage awards, the average compensation awarded to inventors whose inventions have been infringed has historically been less than 1% of U.S. levels. Chinese courts can order infringers to stop the production and sale of an infringing product (a so-called patent injunction), but these injunctions are not always effectively enforced. Small commodity producers can simply cease operations and begin again under a different

⁷ It is conceivable that distinctive features of the Chinese economy make some technologies extremely valuable in China, but only in China. For example, the high degree of separation between the Chinese internet and the global internet or the particular structure and governance of China's national electric grid really do create significant opportunities for Chinese market specific investments that have limited value elsewhere. That being said, talk of a U.S.-China AI race is typically focused on the market for AI-related technology outside of China itself. Technologies so specific to China that they are nearly unusable elsewhere would not figure prominently in this “race.”

name. The low infringement rates and difficulties with injunction enforcement may explain why the rate of patent litigation in China peaked years ago at 0.13% and has dropped since even as the numbers of patents have grown (Zhang, 2022). By comparison, the litigation rate in the U.S. is reported to be around 4% (Bessen and Meurer, 2005).

China's modern patent system, established in 1985, was closely modeled on Japan's patent system, which, in turn, had been closely modeled on Germany's 19th century patent system, many features of which survive in contemporary Germany. Like its German and Japanese institutional ancestors, China's system offers protection for incremental innovations through so-called utility models that are not subject to examination for novelty and, in principle, afford much narrower legal monopoly power over the protected invention than is provided to holders of broader invention patents, which are more analogous to the patents that exist in the United States. In modern Germany and Japan, the use of utility models has declined sharply in recent decades, reflecting the emergence of large numbers of German and Japanese firms as innovators at the global technology frontier who no longer need nor demand intellectual property protection for minor, incremental innovations. In striking contrast, Chinese inventors' use of utility models exploded in recent decades, dramatically outstripping impressive growth in invention patents. Academic studies of utility models tended to confirm the widespread view that these generally document very incremental inventions (Branstetter, Hanley, and Zhang, 2025; Zhang, 2022).

In recent years, the Chinese government has issued strong guidance to inventors to sharply reduce the number of utility models and emphasize invention patents instead. The extensive subsidies and incentives to patent, discussed below, were also extensively reengineered to disincentivize the pursuit of utility models and reward the receipt of invention patent grants. The fact that Chinese patenting remains dominated by the use of utility models would seem to further strengthen the view of large numbers of patented inventions as representing incremental or lower quality inventions.

5. The rise and fall of direct Chinese patent subsidies and the shift to AI in 2017

Beginning in the early 2000s, China's central government began prioritizing the development of indigenous innovative capability. At the same time, local governments began what became a kind of subsidy war, competing with one another to address the central government imperative to strengthen "indigenous innovation." In a pioneering study, Li (2012) describes how an early pilot program in Shanghai led to rapid imitation by other provincial governments, who used generous patent subsidies and Five-Year Plan targets to drive rapid growth in patent applications. Eberhardt, Helmers, and Yu (2017) emphasize the central government's concern that WTO accession (and the required strengthening of China's patent system) would result in foreign multinationals building impenetrable fortresses of patents inside China that would stunt the development of local firms. The central government's response was an effort to encourage domestic patenting as a way of countering this foreign threat. This led to provincial efforts to set up "Patent Special Funds," subsidize patent applications, confer special tax privileges upon firms that patented, and adopt promotion targets for government officials tied to local patent filing. Research institutes and universities were directed to contribute to the construction of a defensive indigenous patent wall. Prud'homme (2012) describes how top-down central

government targets are disaggregated into municipal quotas, effectively forcing local officials to establish generous fee-reimbursement schemes in order to hit regional innovation targets. Lei, Sun, and Wright (2012, 2013) discuss the oft-cited 2006 Medium-to-Long-Term Plan (MLP), a high-level document that directly and forcefully articulated the importance of indigenous innovation and the 2008 National Intellectual Property Strategy, as policy drivers of later patent subsidies. They go on to conduct a quasi-experimental study of municipal patent subsidy increases, finding that cash subsidies increased local application counts without increasing underlying patent claims or real inventions. Applicants simply “sliced” their inventions into smaller, narrower patent applications in order to harvest more subsidy payouts per filing. All of these papers argue that patent subsidies were probably distortionary, disconnecting patent filings from real innovation. By 2008, all 31 provincial-level administrative divisions and hundreds of tier-1 and tier-2 cities had established competing cash-back programs. In some jurisdictions, the payouts on offer actually exceeded the legal cost of patent filing. This kicked off a surge in high-volume, low-quality patent filings that the government would seek to forcefully reverse by the end of the 2010s.

Direct cash payouts were not the only reason for the patent surge. The Chinese corporate income tax system offered huge tax breaks for so-called High and New Technology Enterprises (HNTes), and the possession of patents was an important component of qualifying for these reduced tax rates. Local official promotions, university rankings, academic tenure, and corporate eligibility for state procurement were all tied directly to patent volume, with relatively less attention paid to underlying invention quality. China rapidly became the world’s largest grantor of patents, but the distortions in the system were increasingly visible to experts inside and outside of China. Some organizations filed thousands of low-quality applications simply to “harvest” subsidies. Many of the invention patent applicants never requested a patent examination, which meant the applications never went into force. Other applicants stopped paying maintenance fees at the earliest opportunity, reflecting the limited utility of the invention. Patent offices at home and abroad faced growing backlogs of low-quality filings from Chinese inventors.

Academic studies such as Dang and Motohashi (2015), Hu, Zhang, and Zhao (2017), and Fisch, Block, and Sandner (2016) all provided evidence that subsidies, rather than real invention, were a major driver of China’s patent surge. The last of these three papers focused on the dramatic increase in patenting by Chinese universities and concluded that the patents thereby generated had low commercialization rates and high post-grant abandonment rates. Gu (2021) and Lin, Ding, and Chen (2025) also questioned the value of the surge in patenting by Chinese universities that grew out of the policy interventions of this era and provided evidence suggesting the value was low and perhaps even declining over time. The latter paper went so far as to refer to the surge of Chinese university patenting as a “patent bubble.” By the end of the 2010s, the rising tide of patent applications and the ensuing examination backlog lengthened the “pendency lags” for Chinese patents – that is the gap between initial application and final patent grant – meaning that patent grants would not occur until two or more years after the initial application. Contested or complex applications could have taken three to five years to process. This means that the tidal wave of patent applications induced by China’s era of

extensive patent subsidies was still showing up in patent grants throughout the early 2020s and inevitably clouds our interpretation of the AI patent explosion so clearly visible in Figure 1.

The prior literature has established the existence of generous policy support for patenting in priority areas, but when were these tools focused on AI? The 13th Five-Year Plan noted AI, but only as one of a long list of domains of interest. Similarly, the Made in China 2025 program, announced with much fanfare (and immediately seized upon by Western policymakers as evidence of WTO-illegal support of domestic high-tech industry) mentioned AI among many other technologies, but did not single it out for any kind of special emphasis. Recent academic work such as Khanal et al. (2025) and Zriek et al. (2026) point to the promulgation of the Next Generation Artificial Intelligence Development Plan (NGAIDP) by the State Council in July 2017 as marking the major turning point in Chinese government policy. In this document, China's supreme executive unit, the State Council, emphasized AI in a way it never had before, explicitly laying out ambitious targets for domestic development of an AI industry and short timelines under which these targets were meant to be achieved. The literature suggests that two external shocks played a role in motivating this important policy shift. First, the election of the first Trump Administration signalled an important shift in Sino-American relations and the onset of economic and technological decoupling on both sides.⁸ Second, the widely publicized defeat of South Korean Go grandmaster Lee Sedol by Google's AlphaGo program in 2016 alerted Chinese officials to the existence of massive leaps in AI capability by U.S.-based firms. The official Chinese reaction to this was apparently a kind of "Sputnik moment." Top leaders understood that AI was rapidly advancing and that, under the new Trump Administration, Chinese access to American AI technology was far from guaranteed. The new top priority placed on AI in the NGAIDP was continued in the 14th Five-Year Plan (2021-2025) and the new 15th Five-Year Plan. Note that the timing of this shift coincides almost exactly with the major trend shifts in Chinese and U.S. AI patenting.

As Khanal et al. (2025) report, the NGAIDP targeted a closing of the AI technology gap between China's companies the global frontier by 2020, an achievement of indigenous breakthroughs in some dimensions of basic AI research by 2025, and rise of China's AI sector as the world's technological leader by 2030. As was the case with earlier rounds of Chinese technology policy initiatives, these central directives quickly induced policy responses at the local level. Textual analysis undertaken by Khanal et al. (2025) show that more than 6,000 local-level policies mentioned AI within their texts and more than 500 local-level policies that directly refer to AI in their titles were launched. Local officials were now under pressure to contribute to national goals specifically focused on AI, and their patent targets increasingly had an explicit AI focus. Khanal et al. (2025) document that local governments quickly committed billions of U.S. dollars worth of financial resources to policies designed to support achievement of these goals within their regions.

Beraja et al. (2024) use Chinese data sources to quantify a large increase in investment by China's government-linked VC firms in the late 2010s. Both investment amounts and numbers of targeted firms increase after the promulgation of the NGAIDP. By 2024, these entities had

⁸ See Han et al. (2025) for a broader study of Sino-U.S. technology decoupling.

invested U.S. \$184 billion in over 9,600 firms through more than 20,000 transactions. Firm possession of AI patents was seen as both an attractor of these funds and a short-run indicator of effective investment. Wei et al. (2023) affirm that state-directed VC funds use indicators like patent filings and grants to justify capital deployment.

Just as the Chinese government was shifting its policy focus to AI, the technological barrier to entry in AI patenting fell sharply. Baviskar et al. (2026) document how the introduction of “automated AI” software design tools like TensorFlow, PyTorch, and Auto ML suddenly enabled software engineers with training at a masters degree level to undertake the kind of AI engineering that had heretofore required Ph.D.-level training under an elite academic AI expert to accomplish. Baviskar et al. (2026) show that these tools were rapidly adopted in the U.S., leading to an explosion of AI software patents that explicitly cited the papers that accompanied the introduction of these tools, linked to the open source code for these tools, or explicitly acknowledged these tools in the text of the patent. Chinese software engineers also took advantage of the availability of these tools through open source repositories, quickly taking up use of TensorFlow, PyTorch, Auto ML, and homegrown alternatives like Baidu’s PaddlePaddle. While few outside China’s expanding circle of world class experts could use these tools to advance the global state of the art in fundamental AI science, it was now easier than ever to patent application of standard neural network techniques to particular business uses.

Even as the emphasis on AI was massively ramping up, Chinese officials were also forced to contend with the tidal wave of low-quality patent applications induced by their constellation of patent subsidies. By the early 2020s, the Chinese government sought to engineer a rapid reversal of the accumulation of low-quality patent applications and emphasize patent quality over pure quantity. The challenge, of course, is that patent quality is an attribute that academic experts have struggled to measure for decades and is all but impossible to accurately assess *ex ante*. Inventors often take out patents early in the research and development process, because they do not want rivals’ patents to block them from appropriating the gains generated by their R&D. However, this also means patent applications are submitted before the inventors themselves really know the true economic value of the underlying inventions. That true value is often not evident until years later, when products and services based on the patents are competing in the marketplace, and clear winners emerge.

In 2021, the Chinese National Intellectual Property Administration (CNIPA) published its “Notice on Further Strictly Regulating Patent Application Behavior.” This document called for a near-immediate cancellation of direct subsidies for filing patent applications or requesting patent examinations. It also called for subsidies for patent grants (i.e., payouts made upon successful grant of a patent rather than at the time of application) to be limited to 50% of actual filing costs and it prohibited local governments from subsidizing patent maintenance fees or patent attorney fees. These declines were phased in over the next several years. CNIPA called for the complete cessation of local government financial aid for patent grants and licensing by 2025. Needless to say, these official actions amount to an implicit acceptance of the criticism of the overly generous patent subsidy policies that preceded the 2021 policy reversal.

However, subsidies for patenting did not actually end with CNIPA’s 2021 Notice. While the directive banned *direct* patent subsidies after 2025, China’s tax system, its approach to funding

universities, its *hukou* system of residency permits, its government-directed venture capital system, and its government procurement system all maintained implicit subsidies for certain kinds of patenting activity. Given the magnitude and importance of these indirect and implicit subsidies, discussed in further detail in the next section, it is challenging to know just how much the incentives for “patenting as signalling” have really changed, especially in strategic domains like AI.

6. Implicit patent subsidies in the 2020s.

One of the most important implicit subsidies for patenting in contemporary China arises through the designation of Chinese companies as high and new technology enterprises (HNTes). Enterprises which receive an HNTe designation cut their corporate tax rates from 25% (the standard rate) to 15%. The designation has to be maintained through periodic requalification on a three-year cycle. Possession of patents is an important part of qualification, as is maintenance of a threshold level of R&D spending. Chen et al. (2021) describe the program and present compelling evidence of widespread misclassification of R&D spending by firms seeking to qualify.⁹ Early administration of this tax break is also regarded as one of the largest drivers of low-quality patent filings by corporations. The total number of designated HNTes grew quickly from 2020 through 2024, exceeding 500,000 by the middle of the 2020s. Growth has slowed since 2024. As part of its recent crackdown on junk patents, the Chinese government launched nationwide audits targeting HNTes, seeking to revoke the status and claw back taxes from firms that relied on junk patents or shell IP startups to qualify.

As we have already noted, however, patent quality is hard to assess at the time of application or even upon initial grant. As a matter of practice, the government uses a scoring system that provides additional points for invention patent grants in state-supported strategic fields (such as AI). The shift to a scoring system that prioritizes patenting in nationally designated priority fields could have plausibly contributed to the collectively large increase in AI patenting that we observe among inventing firms which each individually hold small numbers of patents. Even a modest increment in AI-patenting could help these smaller entities qualify for or maintain status as an HNTe. AI was first named as a targeted technology (among many others) in the mid-2010s, but the NGAIDP elevated it to a central priority starting in 2017, and that centrality was reaffirmed in the 14th and 15th Five-Year Plans. Firms were also incentivized to use invention patents in priority areas, such as AI, as collateral for bank loans, since this is seen as providing “market” evidence that the patents actually have value. In China’s state-owned banking system, banks are also incentivized to make loans based on the possession of patents in priority areas. Under this system, possession of patents in strategic domains makes it easier to get bank loans at favorable rates. Finally, the elevation of AI as an increasingly central national priority also led more of China’s “Government Guidance Funds” (government-directed venture capital funds) to invest in “early stage” AI firms. As we have already noted, Beraja et al. (2024) show that by 2023, these funds had invested in 9,623 unique firms with investment totaling \$184 billion. Firm possession of AI patents was seen as both an attractor of these funds and a short-run indicator of effective investment (Wei et al., 2023).

⁹ That being said, Chen et al. (2021) also conclude that the policy increased innovative effort.

We noted earlier that during China's unconstrained patent subsidy period universities were highly incentivized to generate large numbers of low-quality patent applications. Professors could secure promotion, tenure, and salary increases through easily generated patent applications, and university funding could also be increased through generating patent applications, especially in fields prioritized by the government. We have already cited a series of peer-reviewed articles that have assessed the patents thereby generated as a “patent bubble” (Fisch, Block, and Sandner, 2016; Gu, 202; Lin, Ding, and Chen, 2025). As the government sought to shift from its former emphasis on patent quantity to patent quality, it also sought to shift the incentives confronting university professors. In the patent policy reform era of the early 2020s, the Ministry of Education and the Ministry of Science and Technology pushed universities to move from a funding and promotion system based on patent quantity to a regime based on patent “commercialization.” Invention patents could still count toward tenure and promotion for faculty in applied research and engineering disciplines, so long as they were licensed, transferred or commercialized. Without this evidence, invention patents could play, at best, a supplementary role in promotion and tenure decisions.

However, it is easy to see how the commercialization requirement could be gamed by China's growing legions of engineering and applied science professors seeking promotion and tenure in a brutally competitive academic marketplace, especially when we also consider how academic incentives intersect with the HNTE tax concession. A professor could transfer an invention patent to an SME seeking to qualify as an HNTE. Instead of money, the professor and her university could receive an equity interest in a startup, which could be assigned a range of values. The professor/university has now demonstrated “commercialization,” and the SME has acquired an invention patent that can count toward HNTE status. The SME could also use the patent as collateral for a concessionary loan from a state-owned bank, providing further evidence of commercialization. Active collaboration with university labs and government research institutes is an activity that independently counts toward HNTE status, so both the SME and the university professor have a strong incentive to play up the value of a marginal patent and then take the benefits the state then confers on both. These incentives apply, in principle, to patents taken out in any field. However, as AI has become increasingly elevated by the regime as a national priority, AI patents become increasingly useful as a pathway to unlocking the benefits provided by current policy interventions. As a general-purpose technology, AI has the additional benefit of being plausibly related to the revenues of a wide range of businesses, in a way that modern nuclear fission reaction technologies, for instance, are not. With the advent and widespread adoption of automated AI software tools like TensorFlow, AI patents were increasingly easy to write.

University professors may have an additional incentive to coinvent with their students as a way of navigating China's deepening youth unemployment crisis. Since COVID, officially reported youth unemployment rates have remained at high levels as millions of Chinese students graduate into an economy generating rapidly expanding levels of exports but an insufficient number of new jobs. The fixation of the regime on patents – especially patents in technologies designated as national priorities – yields an advantage in the job search to students who are named as coinventors on patents with their professors. Inventor status suggests a student has not only mastered subject matter material learned in classes but can apply it to real problems; it

can raise the probability of corporate employment and lead to a higher starting salary. Students willing to assign their patents to the firms that employ them can help these firms qualify for HNTE status, which confers an additional benefit to the firm and therefore an increased willingness on the part of the firm to hire students who are also inventors. State owned enterprises have formal evaluation systems that provide explicit points for patent inventor status. Government-affiliated research institutes are also more likely to hire students who are also inventors. In addition, Chinese citizens live under the so-called hukou system of local residency requirements. They are not allowed to move anywhere in their own country and seek employment. Instead, any citizen seeking to move to another city must obtain a hukou (official residency status) from the local jurisdiction in order to be eligible for the full range of local social services and benefits. Many of China's leading cities have an explicit allowance in their hukou systems that confers extra credit toward hukou status for individuals with patent grants. The list of cities with such systems includes Shenzhen, Guangzhou, Nanjing, Suzhou, and Shanghai. Beijing awards credit if the patent is assigned to an HNTE or generates verifiable investment or revenue. Over the course of the 14th Five-Year Plan (2021-2025), local officials and state-owned enterprises were judged not on the basis of patent applications but on the number of "high-value invention patents per 10,000 population." The official definition of "high value" specifically notes invention patents in a strategic emerging industry, of which AI was certainly one. This timing is likely related to the striking AI patent explosion in Figure 1.

7. The emergence of AI inventive capabilities: lessons from the U.S.

The preceding three sections document an extensive array of Chinese policy initiatives that dramatically expanded the incentive to patent, especially in technological domains designated as priorities for the state, even when the individual agents would have a far more limited incentive to patent in an undistorted market for intellectual property. While the incentives may be less generous and more focused after 2021, they remained significant, and were increasingly focused on AI. Despite the best efforts of state agencies to emphasize high value patenting over raw application counts in the presence of these incentives, the intrinsic difficulty of assigning accurate economic value to patents when they are first granted would appear to provide ample opportunity for opportunistic actors in China to game the system. Yet we also know, thanks to the international commercial success of some Chinese companies introducing frontier-level AI products and the scientific prominence of China-based academic experts, that strong AI engineering capabilities really do exist. But where? The likelihood that Chinese AI patent grant numbers inflate and obscure true innovative capabilities does not, by itself, help us distinguish between the patenting entities whose patents reflect real innovative capabilities and those whose AI patents are more appearance than substance.

To help identify the kinds of Chinese entities with real AI capabilities, we first turn to two streams of literature in economics that describe the emergence of new industries and identify some of the characteristics of the successful entrants whose innovations ultimately shape the technological trajectory of the entire industry. A key contributor to the first stream is Steven Klepper, who undertook a careful study of the industrial evolution of key manufacturing industries in the early 20th century United States to formulate theoretical models that captured important elements of that evolution. Klepper (1996) observed that these industries went

through similar phases – a phase of rapid entry and radical experimentation, followed by a “shakeout,” in which the number of incumbents dramatically declined as dominant firms with better products and approaches emerged outcompeted their rivals. Once the ultimate winners settled on their designs, they shifted from a focus on product innovation to refinement, cost reduction, and process innovation, achieving an enduring advantage that ensured their survival of the “shakeout” and their long-run success as one of a limited number of incumbents refining their successful approach rather than trying radically different new approaches. In later work, Klepper observed that industries often featured early entrants who experienced initial success. Disputes about different technological directions within these companies often led clusters of leaders to “spin off” new companies for the initial market leader, many of which also went on to great success (Klepper, 2007, 2010). Klepper was able to construct a number of “industrial genealogies” tracing out lines of descent from the initial market leaders to the majority of eventually successful incumbents that weathered the shakeout. Similar linkages arose in more recent, postwar episodes like the emergence of the modern semiconductor industry (Klepper, 2010). Klepper’s work advanced the hypothesis that the initial market leaders became a kind of school in which ambitious innovators acquired the best possible familiarity with the new technology and its application. Graduates of these schools were best positioned for success as the industry evolved.

The second stream of literature focuses on the modern phenomenon of industries that arise out of recently discovered science. The industries on which Klepper initially focused were generally ones for whom the key scientific discoveries were already decades old at the time of the industrial “takeoff.” Lynne Zucker, Michael Darby, and their coauthors examined the role of star scientists and their students in seeding the emergence of a new industry based on science so new that its techniques and approaches had not yet been codified into university curricula or embodied in scientific equipment (Zucker and Darby, 1996; Zucker et al., 1998a, 2002). Focusing on the biotechnology industry, these scholars found that successful early entrants had to absorb the new techniques by hiring the star scientists or their doctoral students in whom the critical new knowledge was embodied. Collaboration with these star scientists and their students was strongly correlated with the hiring firms’ patenting, product development, employment, and other measures of commercial success (Zucker et al., 1998a, b, 2002). Over time, the centrality of these scientists and their students as a channel of knowledge diffusion weakens. This happened in part because new discoveries were incorporated into standard science and engineering curricula, substantially expanding the circle of individuals with the required knowledge to participate in invention at or close to the technological frontier. It also happened because new scientific equipment or capital goods embody the technology that was once the preserve of leading academic labs.

In the United States, the history of artificial intelligence as a domain of scientific inquiry dates back at least to the 1950s. Initial expectations of rapid advance gave way to decades of disappointment. By the end of the 20th century, visions of HAL, the hyperintelligent (and murderous) computer savant featured in the 1960s landmark film *2001: A Space Odyssey* seemed as ludicrously overoptimistic as the film’s envisioning of expanding colonies on the Moon. However, over the course of the 1990s, the rapid expansion of computational power and memory storage and the growing availability of very large machine-readable data sets enabled

algorithms to engage with increasing effectiveness in certain kinds of “machine learning” by applying fairly simple statistical procedures.¹⁰ The pace of advance quickened in the first decade of the 21st century, leading to a growing range of profitable commercial applications of machine learning, and by the early 2010s, the world’s top academic labs were generating multilayered neural networks that could solve complex, multidimensional problems like classifying visual images with an accuracy comparable to that of humans – but at immensely greater speed. AI-related patent applications and the commercial entry of AI products and AI startups accelerated significantly over the course of this decade (Alderucci et al., 2024). The rise of large language models in the early 2020s led to even further acceleration of patenting and firm entry. By this point, the AI industry had clearly “taken off.”

Baviskar et al. (2026) present evidence suggesting that the relatively brief history of the modern AI takeoff appears to include at least 2 periods within it. An initial period, running roughly through 2015, is one that bears a strong resemblance to the early biotechnology industry studied by Zucker and Darby. During this period, the emergence of fundamentally new AI science created a strong rationale for firms to hire the graduate students of star academic scientists in order to access their frontier knowledge. Baviskar et al. (2026) document a strong statistical relationship between the hiring of these elite academic experts and subsequent AI patenting by the hiring firms. They also document a weaker, but still generally statistically significant relationship between the hiring of elite academic experts and firm’s productivity. These statistical relationships are robust to controls for the presence of other kinds of AI talent within the firm and they extend across multiple industry segments. However, these statistical relationships generally collapse into insignificance after 2015. Why?

Baviskar et al. (2026) explain the stark shift after 2015 by pointing to two different but complementary shifts already documented by industry observers. First, a growing array of software design tools, such as TensorFlow, PyTorch, and AutoML, enabled midlevel software engineers without specialized academic training to broadly apply machine learning and AI approaches that hitherto would have required Ph.D.-level training in a top academic lab (Rock, 2019). These tools began to diffuse in 2015 and spread rapidly throughout the industry, becoming the latest historical example of the tendency for scientific discoveries to become embodied in scientific instruments and capital goods. The second shift was a pronounced movement of the locus of scientific leadership in some key fields of AI from academic institutions to industrial labs. At the beginning of the 2010s, the most highly cited papers in computer science were emerging from academic labs. By the middle of the 2010s, the most highly cited papers were emerging from industrial labs, and this phenomenon has continued into the present (Farber and Tampakis, 2024; Gnewuch et al., 2025; Klyman et al., 2024).

This shift in scientific leadership had multiple causes. Recognizing the commercial potential of deep learning approaches to AI, firms aggressively hired some of the best and brightest students – and even faculty -- out of top Ph.D. programs, acquiring a critical mass of internal AI

¹⁰ Early AI-based patents are visible in the data from the late 1980s. In the industry space, one early application was in quantitative hedge funds, where ML techniques were used to identify short-lived asset pricing anomalies against which hedge funds could trade. A misapplication of this technology led to the bankruptcy of Long-Term Capital Management in the late 1990s.

talent well-versed in frontier methods (Hagendorff and Meding, 2021; Hartman and Henkel, 2020; Jurowetzki et al., 2021). This critical mass of talent was paired with immense compute and data resources no academic lab could match (Klyman et al., 2024). To put it simply, top industrial labs were no longer dependent on academic institutions for the best new ideas. They were creating their own. As a consequence, ambitious young AI entrepreneurs were no longer absolutely required to complete their training in an elite AI-focused doctoral program with a top mentor; they could instead complete their training in one of the top labs, potentially achieving even greater access to the knowledge frontier in that setting.¹¹ A culture of relatively open publication at many of the top industrial labs enables us to track this shift of scientific leadership through the lens of highly cited papers presented at leading conferences. The large-scale influx of (formerly) academic talent into leading firms appears to have carried with it a propensity to publish breakthrough innovations even when the same firms are also engaged in energetic patenting of AI-related inventions. This, too, is not new. Stern (2004) and others have documented a strong desire to publish on the part of top corporate R&D talent recruited from elite Ph.D. programs, where they were socialized into the open publication norms of science.¹² Initially documented in the biopharma sector, this appears to be – for the moment – a feature of AI as well.

Linking these well-documented developments back to the economics literature, Baviskar et al. (2026) suggest that the rapidly evolving AI sector has recently emerged from a “Zucker/Darby” era in which successful invention relies on ties to top academic labs and has now entered a “Klepper” era in which leading industrial entrants are an even more important generator of the elite human capital and innovative ideas that will drive the next cycle of industry growth.¹³ These insights suggest a way of predicting the entrants with the strongest inventive capabilities. Through 2015 or so, organizations that have hired elite graduate talent are likely to have an edge. After 2015, the application of standard ML and AI techniques to particular products and services is increasingly commoditized through the diffusion of software design tools while frontier innovation is increasingly driven by a small number of leading industrial labs and the “alumni” who have spun off from them to found a growing wave of AI startups. In both cases, a

¹¹ We note that leading incumbents and more recent startups continue to hire graduates of top computer science Ph.D. programs, and the competition for this limited pool of talent has led to impressively high salaries for new recruits. That being said, these graduates are now hired as much for their raw talent as for the access they bring to new ideas.

¹² Stern (2004) finds the norms of science are so deeply embedded that new industry hires appear willing to accept lower salaries in order to retain the right to publish some of their work.

¹³ It is probably more accurate to characterize the emerging new era as an era in which recent science continues to drive industrial invention at the technological frontier, but the generation of that science moves, at least in part, from universities to firms. It is also not apparent that a Klepper-style “shakeout” will leave the industry with a limited number of large, dominant AI incumbents. An alternative equilibrium could look something like the modern biopharma sector, in which large incumbents specialize in turning successful ideas into products, but these incumbents permanently outsource upstream innovation to a marketplace of independent firms who specialize in the generation of ideas that are then licensed downstream to the aforementioned incumbents. We could see this emergence across the industry space, with leading firms outsourcing the creation of industry-specific AI tools to an upstream cluster of industry-focused firms who license out their tools.

limited network of elite human capital drives successful innovation; the nature of the network evolves over time.

We believe the same patterns are likely to characterize China's most successful AI inventors. China's leading tech incumbents, such as Tencent, Baidu, Alibaba, and Bytedance, have increasingly incorporated AI into their expanding web of product and service offerings and have key leaders whose career paths began in China's own elite universities and followed paths reminiscent of their Silicon Valley-based predecessors. China's more recent wave of AI startups, which are effectively competing with the offerings of Google, OpenAI, and Anthropic, usually have origin stories similar to those of their U.S.-based counterparts, drawing tech talent from China's elite universities and business vision and managerial experience from the executive ranks of earlier waves of tech incumbents. By focusing on the individuals that lie within these academic and industrial genealogies, and tracking their movement across geographic space, time, and organizational boundaries, we are able to identify those entities whose patents are likely to reflect the most impactful inventions. This is especially true if we narrow our focus from an inclusive definition of AI that includes new but very industry-specific applications of mature techniques to more fundamental innovations in "core AI."

8. Tracking AI publications

If AI technology remains closely connected to science, the locus of scientific leadership has moved from academic institutions to industrial labs, and leading industrial labs retain at least some degree of the scientific world's historic commitment to publishing, then careful study of the relevant scientific literature can provide a window into the identities of the true innovators in AI, and especially in "core AI." This window is subject to its own limitations, but these limitations are different from those impacting patents, especially Chinese patents.

We therefore focus on presentations in the top international conferences across 10 subfields of AI. Note that we focus on conferences rather than journals – in a fast-moving field like computer science, conferences are generally preferred outlets to "archival" journals. Our categorization of venues across or 10 AI sub-domains is rooted in three pillars of historical and modern academic prestige spanning a 35-year window:

1. **Core Rank & Institutional Stature:** Elite venues map directly to the CORE A* rating (the highest tier designated by the Computing Research and Education Association). Gold and Silver tiers correspond to high-prestige CORE A and highly stable CORE B venues, respectively.
2. **Citation Velocity & Long-Term Impact:** Elite and Gold venues maintain the highest Google Scholar h5-indices and historical citation volumes within their respective fields, indicating they host the most impactful papers.
3. **Selectivity & Peer-Review Rigor:** Elite venues maintain relatively low historical acceptance rates (typically 15%-23%) and undergo multi-stage, double-blind peer-review cycles overseen by the world's foremost area chairs.

Table 1 provides a comprehensive list of the venues we use to track publications. We also measure citations received by articles published in these venues from subsequent articles published in the same set of venues. We are able to access these data through collaboration

with the Elsevier research group based in Amsterdam, which has generously provided us with access to the Scopus database. The article-level data contain information on the affiliation of each author, allowing us to connect publications to Chinese and American universities, corporations, and government-affiliated research institutes. We use these data in a number of ways. First, we note whether a patenting entity in either country has ever presented a paper in any of these venues. Second, we set a modest threshold of 2,000 cumulative citations received, and flag any university in either country whose presence at these conferences falls below that threshold. In the U.S. ranking of institutions, the marginal university that exceeds this threshold is the University of Miami. The university that falls just below it is Oklahoma State University. Given the historic concentration of pathbreaking innovations (and innovators) in a small cluster of elite universities on both sides of the Pacific, we feel confident in downweighting the impact of patents emerging from institutions below this threshold.¹⁴

While the first threshold of 2,000 citations sets a reasonable minimum standard that establishes a given university possesses research active faculty in fields related to AI, extensive evidence suggests that the economic value of innovation is likely to be clustered in the most elite institutions. So, we set a second, higher threshold of 30,000 cumulative citations. In the U.S. ranking, the marginal institution that just exceeds this threshold is the University of Central Florida. The institution that falls just below it is the University of Wisconsin at Madison.

The fast pace of advance in computer science, especially those subfields linked to AI, imply that important recent papers quickly accumulate large numbers of citations and much “older knowledge” is quickly rendered effectively obsolete. That being said, our use of cumulative citations inevitably privileges U.S. departments with a longer publishing track record, since citations to their papers have had longer to arrive. In future versions of this paper, we can track citations received over periods of fixed length as a way of limiting this measurement bias.¹⁵

Having measured the scientific quality and impact of universities across subfields of AI, we can now highlight a critical difference between university patenting in AI between the U.S. and China. In the U.S., university patenting is quite limited, and this holds regardless of which particular definition of AI we adopt in our classification algorithms. As an illustrative example, Table 2 lists the total amount of Core AI patenting across all U.S. universities from 1990-2024 and the fraction of that patenting that is concentrated in the top 20 institutions. The table also undertakes the same exercise for Chinese universities. While the most highly ranked Chinese universities also tend to stand out above the rest in terms of patents, a very large fraction of Chinese AI patents granted to universities are held by institutions that fall below our minimum citation threshold (and would be ranked similarly below the University of Central Florida or lesser-known institutions in the United States). The fraction of total university-held AI patents

¹⁴ Future versions of the paper will provide rankings of Chinese and U.S. universities and companies according to these publication citation counts, along with the citation counts themselves.

¹⁵ We also intend to screen computer science publications originating in China for evidence of the same “home bias” in citation patterns identified by Chu et al. (2025) in chemistry.

held by the top 20 institutions is significantly smaller than the U.S., and we see large numbers of core AI patents awarded to universities with a very limited publication or citation profile.

9. Tracking AI experts

The second way in which we use the publication data is as a means of identifying the top innovator *individuals*. We hypothesize that some top innovators may establish the quality of their human capital by coauthoring a highly cited paper presented at a major conference, but then may take their talent and human capital into a new organization with a very different approach to scientific publication. The new entities may be more inclined to patent than to publish, but the presence of a critical mass of top innovators in that organization would suggest that the potential to advance the state of the art exists within that organization, even if it publishes few (or no) papers.

Baviskar et al. (2026) provides a detailed explanation of our use of Elsevier Scopus data to identify the world's top academic AI experts across the same 10 subfields of AI. We then engaged in a manual, time-consuming process of using university and lab websites, academics' CVs, and the ProQuest database to link these experts to their doctoral and post-doctoral advisees. Next, we used access to the Revelio database, based on scraped LinkedIn profiles, to track the movement of these advisees across organizational boundaries. Matching these data to the Compustat database, which tracks publicly traded firms in the United States, we created a firm-year level stock of elite AI graduates for each firm that hired at least one of these graduates, adjusting the stock to reflect both entry and exit. The presence of an elite graduate within a firm represents a separate measure of connection to frontier scientific knowledge, even if the individual does not publish while affiliated with the hiring firm. Firms that include a member of this elite innovator network are considered more likely to generate patents that actually advance the AI technological frontier. Our database includes the career histories of more than 10,000 individuals. We are in the process of extending this set of experts to include additional experts who may not have been mentored by an elite academic computer scientists but were nevertheless coauthors or coinventors of multiple highly impactful papers or patents after the mid-2010s, when the locus of scientific research leadership in AI in some fields of AI began to shift from academic to industrial research labs.

China presents a more challenging environment in which to track the emergence and movement of elite AI talent. LinkedIn is not widely used, nor is there a Chinese language equivalent that plays the role LinkedIn does in the global software industry outside of China. Given these constraints, candidate lists for Chinese AI experts were compiled from the leadership teams of top Chinese AI companies, keynote speakers at AI conferences held in China, members of AI-related professional committees of the China Computer Federation (CCF), and the AI 2000 Global AI Scholars Ranking by AMiner. AMiner, formerly known as ArnetMiner, is a 20-year-old research project originally launched by Chinese star computer science professor Tang Jie with funding from China's National Science Foundation. Over time, AMiner came to index millions of researchers and more than 200 million articles across a wide range of scholarly fields.¹⁶

¹⁶ Tang Jie later acquired global fame as the founder of [Z.AI](#) (formerly known as Zhipu), which is the first LLM startup to go public anywhere in the world. Its market valuation peaked at around U.S. \$128 billion.

Academic identifiers and affiliations of experts were inferred through web searches and the integration of information from multiple online sources. Expert affiliations were further expanded using arXiv papers linked to each expert through their academic identifiers. A parsing system was developed to extract author affiliations and email addresses from the LaTeX source files of arXiv papers. The start and end years of each affiliation can then be inferred based on the publication dates of the corresponding papers.

The final Chinese AI expert network currently contains more than 3,000 experts. For each expert, aliases, academic identifiers (e.g., Google Scholar ID, ORCID, and OpenReview ID), affiliations with start and end years, and expertise distributions are included when applicable. This allows us to identify entities that have hired a member of the Chinese AI expert network as well as those who have never been connected to any of the Chinese experts. Private firms, SOEs, and other entities connected to this network are judged more likely to generate commercially valuable AI patents that could meaningfully advance the technological state of the art.

10. A second look at AI patents in China and the U.S.

In previous sections, we reviewed the growing literature on the incentives woven into the Chinese tax, finance, and higher education systems that undeniably have led important categories of Chinese patent applicants to file large numbers of low-quality patent applications in the recent past. We argued that, despite the weakening of direct incentives to file such patents, the indirect incentives to do so remain strong. We have also argued that the historically close relationship between relatively recent AI science and AI invention enables us to use AI publications in competitive international conferences as a separate source of information on AI innovation capabilities. Finally, we have used important findings from studies of the emergence of new industries to argue for the disproportionate importance of a core network of key innovators who carry frontier knowledge from leading academic labs and pioneering firms to later entrants. We have argued that the most important advances in commercial application of AI on both sides of the Pacific are likely to be driven by entities affiliated with members of this network, we have identified thousands of individuals in both China and the U.S. (and other Western economies) that are likely to be part of this network, and we have measured – as comprehensively as we can – where these individuals have worked to date.

Using all this information, we now seek to decompose AI patenting in both the U.S. and China into groups of AI patents associated with different categories of patent applicants. We do so, using a narrower definition of AI patenting than that used at the beginning of the paper. In this section, we focus on “core AI” patenting in both countries. We prompt our LLMs with a narrower definition, that seeks to eliminate the large numbers of boundary cases we encounter in both patent systems when we employ prompts specifically designed to capture the full range of AI applications to specific industries and use cases.¹⁷ Of course, this comes at a cost. It screens

He is also known as the mentor of Yang Zhilin, his former student, who went on to found Moonshot, another high-profile Chinese LLM startup worth billions.

¹⁷ We use this definition of Core AI: “Algorithms, methods, and technologies on foundational framework or core capabilities of AI, including theoretical foundations of AI, visual perception, speech recognition,

out a large number of applied, industry-focused or product-focused patents that may have significant economic value.

Figure 3 illustrates China's AI patent explosion from 2018 through 2025 (patent grant years), using this narrower definition. While the numbers of patents are much smaller, the same general shape appears here as in Figure 1 – sharp growth and a pronounced divergence from the level of patent grants at the USPTO over the same period, using the same algorithmic definition.

However, Figure 3 also breaks the core Chinese patents into categories based on the identity of the patent applicant, which is the Chinese patent system's equivalent of the USPTO "assignee," using a stacked area graph. The dark red segment at the top denotes "small" applicants (entities with fewer than 10 cumulative core AI patents by the end of 2024) that have never employed a member of the Chinese AI expert network. Each applicant is individually small, but, as a group, they account for a collectively nontrivial fraction of total core AI patents, and we consider this group to be highly likely to seek patent grants for the purpose of qualifying as an HTNE.

The next segment, colored red orange, represents core AI patents taken out by Chinese universities that fall below our second threshold of 30,000 cumulative elite conference citations. This is not a particularly stringent threshold; as we have already noted, it omits universities with a lower publication and citation profile than the University of Central Florida in the U.S. With apologies to any alumni of that institution who may be reading this paper, the University of Central Florida is an R1 institution and a leading center of engineering education in the Sunshine State, but it is not internationally recognized as a major center of AI research. Given that, it is hardly surprising that very few core AI patents are assigned to University of Central Florida. What is surprising is that tens of thousands of core AI patents are assigned to Chinese universities with similar (or weaker) publication profiles at major international AI conferences.

The third (orange) segment represents core AI patents taken out by state-owned enterprises that have never employed a member of the Chinese core AI network. Taken together, the segments associated with small entities that are almost certainly not innovative startups, second-tier (or weaker) universities, and SOEs who have never employed an elite expert account for an extremely large fraction of core AI patenting and a large fraction of growth in core AI patenting. However, it seems unlikely that these categories of patent applicants are the ones driving significant advances in the technological state of the art.

The yellow segment illustrates core AI patenting by Chinese universities that lie above our threshold. The very best of these universities, such as Tsinghua University, are internationally recognized as possessing computer science departments that now rival those of MIT, Stanford, or Carnegie Mellon in terms of elite publications and the generation of dynamic start-ups, and the quality of its best core AI patents should not be discounted. That being said, even within

understanding language, reasoning, planning, classification, recognizing patterns, prediction, executing actions, taking feedbacks and rewards to improve itself."

this elite group, the number of publications slopes downward sharply as we move from the top-ranked institution to the bottom, suggesting differences in AI inventive capability.

The green segment illustrates core AI patenting by Chinese firms (not SOEs) with significant total numbers of AI patenting. It includes well-known, globally respected Chinese tech firms such as Alibaba, Tencent, Baidu, and Huawei, as well as large numbers of firms that are considerably less well known in the West. The light blue segment illustrates core AI patents taken out by foreign companies in China. These add to the totals, but cannot be interpreted as evidence of innovative strength by China-based entities; most of these patents are invented abroad or with the input of research staff based outside China. The dark blue segment illustrates patents taken out by SOEs that have employed at least one member of the Chinese expert network. The gray segment illustrates patents taken out by a diverse set of entities that cannot be classified into the previous groups – this fraction is small relative to the total.

Overlaid on top of this graph is a purple line that reflects total core AI patents granted by the USPTO over the same years. As can be easily seen, this purple line looks small relative to total core AI patenting in China. However, if we focus solely on patents generated by the Chinese entities most likely to possess real AI invention capabilities and the ability to turn those inventions into real products or services – top universities and large private firms, respectively – the numbers appear much closer to U.S. patenting levels. There is impressive evidence of Chinese growth in inventive capability, to be sure, but the picture suggested by this more selective approach is one of rapid convergence and something closer to parity in the mid-2020s rather than an explosion of invention that leaves the U.S. lagging far behind.

Figure 4 takes a similar approach to the disaggregation of core AI patents granted by the USPTO into different categories, using a stacked area graph. The dark red dotted line superimposes onto this graph the total Chinese level of core AI patent grants, summing across all categories of applicants. The pink dotted line superimposes onto this graph the level of Chinese core AI patents granted only to Chinese companies with more than 10 such patents by the end.

We can immediately see that the breakdown of core AI patenting across organizational types is very different from that observed in the Chinese data. The vast majority of patent grants are concentrated in either U.S.-based firms or multinationals based in Western economies outside the United States that individually have more than 10 patent grants by the end of our sample window. U.S.-based universities – even those above our citation threshold – account for very few patents. Small enterprises, a category that includes some startups employing members of the North American AI expert network, also account for a very small fraction of total core AI patenting. Other entities – a component that reflects government agencies, national labs, and a variety of other organizations – are visible, but quite small. What is also immediately clear in this graph is that the AI patenting of large Chinese firms in China is broadly similar in magnitude to the combined totals of American and non-U.S. Western MNC patenting in the U.S. The former grows more quickly over time and exceeds the latter by the end of our sample, but the margin by which it does so is not large – at least, not yet.

It is also interesting to note the striking differences in university patenting across the two countries. Even the very top-ranked North American universities whose science has disproportionately driven AI advance and whose graduates feature prominently in the global industry's most successful firms (outside China, at least) take out relatively small numbers of AI patents. MIT, Carnegie Mellon, and the University of California system are awarded 107, 87, and 85 patents, respectively, by the end of our data window. This is dramatically smaller than the set of patents awarded to what many would consider to be China's top institutions, such as the Chinese Academy of Sciences, Tsinghua University, Zhejiang University, the University of Science and Technology, Peking University, and Shanghai Jiao Tong University, which were awarded 3,157; 1,571; 1,373; 737; and 640 patents, respectively. Computer scientists would now rank China's top computer science programs as peers of the best North American universities. Some might even say they are somewhat better, at least in the AI domain. However, few objective members of the research community would conclude that the Chinese institutions are ten times better or 30 times better than their North American peers, as a simple comparison of patent totals might imply. Even for universities that are at the very top of China's academic hierarchy, patent volumes appear to reflect not just innovative capabilities but also policy incentives that do not exist for their North American peers.

11. International patent grants as a screen for patent quality

As we have already noted, innovation scholars have long used the propensity for inventors to take out patents for the same invention in multiple countries as a strong signal of the invention's true value (Lanjouw and Schankerman, 2004). In this context, it is useful to note that Chinese inventors' average propensity to patent their domestic inventions in other countries is strikingly and persistently low – on the order of 3.5% (Lerner et al, 2026). The historical rates for American, Korea, Japanese, and German inventors are roughly 40%, 20%, 40%, and 60% respectively. These stark and persistent differences are one reason Western some scholars have questioned the quality of Chinese invention (Branstetter, Hanley, and Zhang, 2025). As the world's largest exporter of goods and a key hub in the global trading system, Chinese firms would seem to have a strong incentive to patent truly valuable inventions abroad. Low rates of foreign patenting are consistent with the view that the large numbers of Chinese patents do not, in fact, actually describe inventions with significant value outside of China.

It is conceivable that distinctive features of the Chinese economy make some technologies extremely valuable in China, but only in China. For example, the high degree of separation between the Chinese internet and the global internet or the particular structure and governance of China's national electric grid really do create significant opportunities for Chinese market specific investments that have limited value elsewhere. That being said, talk of a U.S.-China AI race is typically focused on the market for AI-related technology outside of China itself. Technologies so specific to China that they are nearly unusable elsewhere may not figure prominently in this "race."

Guided by these considerations, we go back to our database of core AI patents and we retain only patents taken out by any category of applicant or assignee in either patent system that have been granted in more than one country. Figures 5a and 5b illustrate the new stacked area

graph for China-granted patents over the 2018-2024 patent explosion period; figures 6a and 6b construct the new stacked area graphs for USPTO-granted patents over the same time period.

Figure 5a appears substantially different from Figure 3. The total number of Chinese core AI patent grants falls by more than 90 percent and the relative importance of different applicants is considerably altered when we limit our focus to patents granted in multiple jurisdictions. Both lower tier and elite Chinese universities are much less likely to patent abroad than at home. These important sources of Chinese AI patents shrink dramatically once we screen on international patenting. State-owned enterprises are also much less likely to patent abroad, whether or not they employ members of the Chinese AI expert network. SOE patenting declines so much as to be virtually invisible in the new graph. The most important domestic patent grantees are now large private Chinese companies. In sharp contrast to the striking decline in the importance of all categories of domestic patent grantees, foreign companies in China now emerge as the most important source of China-granted core AI patents, once we limit our purview to patents granted in multiple countries. The purple line overlaid on top of the graph measures total USPTO-granted core AI patents also issued in at least one other country as a basis of comparison. For the first time in our set of comparative figures, the USPTO patent grants substantially exceed those granted in China. In fact, if we subtract the patenting of foreign companies from the Chinese total, then a significant gap opens up between 2018 and 2022 in favor of the USPTO totals. This is occurring over the very same years in which purely domestic core AI patenting in China is skyrocketing. Figure 5b represents the same picture with Chinese patents granted to foreign applicants stripped out. This makes the small scale of Chinese multicountry patenting even more obvious.

This dramatically different picture requires some additional context. Chinese inventors choosing to file for patent protection in other countries under the auspices of the Patent Cooperation Treaty (PCT) can delay their international patenting applications to individual foreign patent offices for up to 30 months and still retain their original Chinese priority date. Given the popularity of this filing strategy with Chinese firms, we should expect foreign patent grants to lag domestic ones by more than 2 years. This suggests that the last two years of data shown in Figures 5a and 5b should be viewed skeptically. That reality notwithstanding, we see clearly that the vast majority of Chinese core AI patent grants have no corresponding patent grant in any other country.

Figure 6a presents the counterpart to Figure 4, which examines core AI patents granted by the USPTO over the same period. This time, we only track patents granted in more than one country. Patent grants drop by more than 50%, but far less than China's 90%+ decline. Noncorporate patentees and small firms shrink somewhat in importance, but they were not significant to begin with. Large U.S. firms shrink in importance relative to non-U.S. Western firms (who nearly all take out patents in their domestic markets before patenting in the U.S.). Chinese firms also loom somewhat larger here than they did in Figure 4, since they are all patenting in China as well as in the U.S. However, the main takeaway from Figure 6a is the implied strength of "Western" core AI patenting relative to Chinese core AI patenting. As in Figure 4, lines representing total Chinese multi-country patent grants and those assigned to

Chinese companies are superimposed on Figure 6, suggesting an impressively rapid accumulation of core AI inventive capabilities, but one that is still far from overshadowing the inventive capabilities of the West. The figure also challenges the common narrative of an exclusively Sino-American AI race by highlighting the significant AI invention capabilities of firms based in European and Asian economies that have historically been important trading partners and defense allies of the United States. The most visibly successful LLMs may be developed in the U.S. and China, but the ability to apply AI technology, broadly defined, to industrial products and services is apparently well developed outside those two countries. Figure 6b presents the same picture, with patents granted to Chinese assignees stripped out, to sharpen the focus on the relative strength of Western multicountry patenting.

12. Textual analysis based approaches to patent quality.

The rapid growth of computing power and natural language processing (NLP) techniques led Kelly et al. (2021) to use new approaches to patent novelty. Prior work had focused on patent citations that were subject to well-known biases and limitations. Kelly et al. (2021) proposed to measure patent novelty by documenting the textual dissimilarity of a patent to previously granted patents and its textual similarity to subsequently granted patents. Patents that were simultaneously dissimilar from previous granted patents but similar to subsequently granted patents received higher scores. They were viewed as introducing new ideas that deeply influenced subsequent work. More formally, the KPST score is defined as

$$KPST_p = \frac{\frac{1}{|F(p)|} \sum_{j \in F(p)} s_{pj}}{\frac{1}{|B(p)|} \sum_{j \in B(p)} s_{pj}}$$

Where s_{pj} denotes the textual similarity between patents p and j . $F(p)$ denotes the set of patents granted after patent p , and $B(p)$ denotes the set of patents granted before patent p . Kelly et al. (2021) focused on five-year forward and five-year backward similarities. Lerner et al. (2026) calculate KPST scores and a variant they label KPST-XN, which computes scores calculated only within national patent corpora. In other words, KPST-XN scores only compare Chinese patents to other Chinese patents. Lerner et al. (2026) calculate these scores using the full text of the Chinese patents, and they conclude that explosive Chinese patent growth does not appear to be correlated with substantive declines in novelty, as indicated by KPST or KPST-XN scores across the critical technologies examined in their paper.

We take another look at KPST-XN scores for our core AI patents, narrowing the time windows to three years rather than five, in order to measure the evolution of KPST-XN scores over at least part of the core AI patent explosion that emerges after 2018. We also use only titles and abstracts for this calculation and therefore view the evolution of KPST-XN depicted in Figure 7 for different categories of Chinese patent applicants, as preliminary and subject to revision. Our approach is motivated partly by concern that patents designed to signal compliance with government policy directives might include text deliberately and consciously constructed to indicate “novelty” in ways that impact conventionally calculated KPST scores.

Across all categories of patent applicants, we see sharp declines in measured novelty toward the end of our sample period, as the core AI patent explosion unfolds – a different finding from that of Lerner et al. (2026). If we look at average measures of novelty over the whole sample period, large Chinese companies and universities above our citation threshold generate patents rated as more novel than other categories of Chinese inventors, strengthening our view that these categories of patent applicants are likely to be generating more substantive inventions. SOEs that employ identified members of our Chinese AI expert network appear to generate more novel patents than those without. Universities below our paper citation threshold appear to generate substantially less novel patents than those above the paper citation threshold.

For many of the more important categories of applicants, we see a common pattern over time – rising patent novelty from the mid-2010 through the end of that decade, followed by a sharp decline. Chinese AI patenting was extremely sparse until the 2010s and growth really took in the middle of that decade. In its early years, the flood of AI patents may have looked quite novel, since their text was different from what had come before but increasingly similar to what came after. As the patent explosion went into overdrive after 2020, AI patents were, perhaps mechanically, calculated to be more similar to what immediately preceded them, driving measured patent novelty sharply down. We are currently looking into the patent-level numbers to determine whether this interpretation is consistent with the disaggregated data. In any case, the sharp fall in measured novelty appears to call into further question use of raw Chinese AI patent counts as an accurate indicator of novel invention.

13. LDA / topic modeling and centers of AI expertise in China (NOTE: this section is still incomplete)

The preceding sections argued that much domestic Chinese AI patenting may be uninformative regarding that nation's underlying distribution of substantive AI innovations that advance the state of the art and used information on AI academic publications (especially in leading international conferences) as an alternative data window into Chinese advances in AI technology. Standard clustering algorithms allow us to use the text of publications generated by members of the Chinese expert network to identify the subfields of AI in which these individuals are actively publishing. We use standard LDA to generate 100 latent topics based on the arXiv AI papers collected. We then use GPT-5 to name these topics and merge them into 52 canonical topics. We manually categorize the canonical topics into broader research areas. The AI topic structure diagram provided in Figure 8 illustrates the mapping from latent topic IDs to canonical topics, as well as the overall topic hierarchy.

The bar chart shown in Figure 9 is generated based on the author-topic distributions of experts in the Chinese AI expert network. Each author-topic distribution is obtained by averaging the paper-topic distributions across all AI papers authored by that expert. For each topic, we then calculate its mean probability across all experts and rank the topics accordingly. The top 20 topics are selected as the research domains with the highest average topic weights among Chinese AI experts.

The general picture that emerges from this analysis is one of a research community that spans leading universities and major enterprises, holds a growing research presence across many productive subfields of AI, exerts increasing scientific influence, as evidenced by rising academic citations, and exhibits particularly impressive strength in domains of AI related to image classification and analysis. Future versions of this paper will compare the distribution of research themes evidenced by total domestic patenting and Chinese patents taken out in more than one country to the distribution obtained here from academic publications. It will also compare these distributions to those evidenced by the academic publications of U.S.-based experts and the distribution of topics in USPTO-granted patents.

14. Conclusions and lessons for the literature.

The explosion of Chinese patenting across AI and other “critical technology” domains over the past decade has been highlighted by leading researchers and viewed with alarm by many Western industrialists and politicians (Fang et al., 2026; Lerner et al., 2026). At the same time that these patent numbers have been growing, an emerging literature – much of it authored by Chinese scholars – has meticulously documented an array of Chinese public policies that could plausibly induce Chinese inventors to obtain far more patent grants than might be optimal under the traditional cost-benefit analyses Western firms have generally followed. While the most distortionary of these policies have been recently ended or reduced in magnitude, doubts remain about the degree to which patent counts might inflate underlying innovation.

Conventional empirical approaches to the systematic assessment of the quality of Chinese patents have been hampered by data limitations. The processes that generate patent citations in China work differently than in the U.S (Lerner et al., 2026), complicating evaluation based solely on patent citations. Following influential work by Kogan et al. (2017), researchers have attempted to infer the value of Chinese patents by examining stock price reactions to news of patent grants (Fang et al., 2026), but publicly traded companies account for a surprisingly small fraction of these patents and Lerner et al. (2026) have pointed out that widespread government intervention in Chinese equity markets casts doubt on the accuracy of these price signals. Some scholars have used the text of Chinese patents to infer changes over time in the average novelty of Chinese patent grants. Different scholars have come to different conclusions regarding what these textual changes imply about the evolution of Chinese patent novelty in recent years (Lerner et al., 2026; Boeing et al., 2024; Branstetter, Hanley, and Zhang, 2025). If patents are deliberately being written in such a manner as to suggest significant innovation in technological domains prioritized by the government, then it is not clear text-based approaches will prove dispositive. In the absence of definitive evidence (“ground truth”) to the contrary, many well-regarded researchers have chosen to take the exploding patent counts at face value (Fang et al., 2026; Lerner et al., 2026).

In this paper, we exploit the fact that much modern AI invention has closely built on relatively recent advances in AI science. As in the United States, the current coincidence of scientific opportunity and economic value in AI in contemporary China generates significant flows of talent between elite universities, leading firms, and disruptive start-ups. Leading AI innovation in China is often undertaken by research directors and leading engineers trained in elite

university labs in China and/or the West. Top Chinese academic and industrial labs regularly present at top international conferences and their work is cited extensively by elite researchers outside China. If we presume that these international conferences apply a consistent quality screen to submitted papers undertaken by expert evaluators, regardless of the national origin of the authors, then patterns of AI scientific publications across geographies, time, and organizational boundaries can be informative regarding the underlying distribution of AI inventive capabilities. In other words, the unusually close connection between AI science and invention and the role of the international scientific community as an evaluator of research quality provides us with an alternative source of ground truth. Using publications, we can identify many of the leading innovators in both countries. By mapping the movement of these innovators across organizational boundaries, we can categorize patenting entities based on their connection to the expert network, as we have measured it, in both countries. This provides us with a new basis for making judgments about the quality of patents generated by these different organizations.

Leveraging this information, we present evidence that a large and growing fraction of Chinese AI patents are awarded to entities with plausibly strong incentives to patent marginal ideas and plausibly limited capacity to engage in AI innovation anywhere close to the technological frontier. Discounting the patents awarded to these entities eliminates most of the measured gap between Chinese and U.S. AI patenting. Building on decades of research using the propensity to patent inventions in more than one country as a measure of quality, we then retain only patents granted in more than one country. This adjustment flips the picture, shrinking Chinese patent grant totals by more than 90%, and leaving Western corporate patent holders with a significant and enduring lead.

Publication data from elite international conferences clearly showcases the rapid growth of a Chinese community of frontier computer scientists with formidable innovative capabilities. Members of this impressive community are now powering a new cohort of startups whose products are keeping pace with the performance of much better funded American AI labs. Other Chinese companies are effectively generating AI inventions (including those further from the frontier) focused on particular use cases, products, and services with global scope. These developments are undeniable and arguably benefit the global user base, whose application of AI technologies will determine its impact on productivity growth at the global level. However, this paper demonstrates that simple counts of AI patents are highly imperfect indicators of this advance. Under current policy settings, that is likely to be true for years to come.

Bibliography

Agrawal, A. (2006). Engaging the inventor: Exploring licensing strategies for university inventions and the role of latent knowledge. *Strategic Management Journal*, 27(1):63–79.

Agrawal, A., Gans, J., and Goldfarb, A. (2019a). *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.

Agrawal, A., Gans, J. S., and Goldfarb, A. (2019b). Artificial intelligence: The ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives*, 33(2):31–50.

Agrawal, A. and Henderson, R. (2002). Putting patents in context: Exploring knowledge transfer from MIT.

Agrawal, A., McHale, J., and Oettl, A. (2017). How stars matter: Recruiting and peer effects in evolutionary biology. *Research Policy*, 46(4):853–867.

Agrawal, A. K., Brynjolfsson, E., and Korinek, A., editors (2025a). *The Economics of Transformative AI*. University of Chicago Press, Chicago.

Agrawal, A. K., Gans, J. S., and Goldfarb, A. (2025b). Genius on demand: The value of transformative artificial intelligence. In Agrawal, A. K., Brynjolfsson, E., and Korinek, A., editors, *The Economics of Transformative AI*, chapter 1. University of Chicago Press, Chicago.

Alderucci, D., Baviskar, S., Branstetter, L., Goldschlag, N., Hovy, E., Runge, A., Tambe, P., and Zolas, N., (2024), “Quantifying the Impact of AI on Productivity and Labor Demand: Evidence from U.S. Census Microdata,” paper presented at NBER Summer Institute.

Arora, A., Belenzon, S., and Sheer, L. (2021). Knowledge spillovers and corporate investment in scientific research. *American Economic Review*, 111(3):871–898.

Azoulay, P., Fons-Rosen, C., and Graff Zivin, J. S. (2019). Does science advance one funeral at a time? *American Economic Review*, 109(8):2889–2920.

Azoulay, P., Graff Zivin, J. S., and Wang, J. (2010). Superstar extinction. *The Quarterly Journal of Economics*, 125(2):549–589.

Babina, T. (2026). Understanding firms’ ai efforts and their economic impact. NBER Working Paper 35123, National Bureau of Economic Research.

Babina, T., Fedyk, A., He, A., and Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151:103745.

Baily, M., Bosworth, B., Goolsbee, A., and Syverson, C. (2023). Prospects for a productivity boom: Ai and beyond. *Brookings Papers on Economic Activity*.

Barrios, J. M., Fujii, B. C., Lisowsky, P., and Minnis, M. (2025). Measurement matters:

Financial reporting and productivity. NBER Working Paper 34536, National Bureau of Economic Research.

Baviskar, S., Branstetter, L., Drayton, C., Fujiy, B., Hovy, E., Tambe, P., and Wu, L., (2026). "Olympus Has Fallen: The Declining Role of Academic Science in AI Innovation," Working Paper.

Barry, Chris, Ronen Arad, Landan Ansell, Meredith Cartier, and HyeYun Lee (2017): "2017 Patent Litigation Study: change on the horizon?"

Belderbos, Rene, Jinhyuck Park, and Martin Carree (2021): "Do R&D investments in weak IPR countries destroy market value? The role of internal linkages," *Strategic Management Journal*, 42, 1401–1431.

Beraja, M., Yang, D., and Yuchtman, N., (2023). "Data-intensive Innovation and the State: Evidence from AI Firms in China." *Review of Economic Studies*, 90, pp. 1701-1723.

Beraja, M., Peng, W., Yang, D., and Yuchtman, N. (2024). "Government as Venture Capitalists in AI," NBER WP 32701.

Bessen, James and Michael J Meurer (2005): "Lessons for patent policy from empirical research on patent litigation," *Lewis & Clark L. Rev.*, 9, 1.

Beukel, Karin and Minyuan Zhao (2018): "IP litigation is local, but those who litigate are global," *Journal of International Business Policy*, 1, 53–70.

Boeing, Philipp, Loren Brandt, Ruochen Dai, Kevin Lim, and Bettina Peters (2024): "The anatomy of Chinese innovation: Insights on patent quality and ownership," Tech. rep., ZEW Discussion Papers.

Branstetter, Lee, Hanley, Douglas, and Zhang, Huiyan, (2025). "Unleashing the Dragon: The Case for Patent Reform in China," Working Paper.

Bresnahan, T. F. and Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics*, 65(1):83–108.

Brynjolfsson, E., Li, D., and Raymond, L. (2023). Generative ai at work. Technical Report 31161, NBER.

Brynjolfsson, E. and McAfee, A. (2014). *The Second Machine Age*. W. W. Norton.

Brynjolfsson, E., Rock, D., and Syverson, C. (2019). Artificial intelligence and the modern productivity paradox. In *The Economics of Artificial Intelligence: An Agenda*, pages 23–57. University of Chicago Press.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suarez Serrato, and Daniel Yi Xu (2021): "Notching R&D investment with corporate income tax cuts in China," *American Economic Review*, 111, 2065–2100.

Chu, Shimin, Steinwender, Claudia, and Azoulay, Pierre, (2025), "Paper Tiger? Chinese Science and Home Bias in Citations," NBER Working Paper 32468.

- Cockburn, I., Henderson, R., and Stern, S. (2019). The impact of artificial intelligence on innovation. In Agrawal, A., Gans, J., and Goldfarb, A., editors, *The Economics of Artificial Intelligence: An Agenda*, pages 115–146. University of Chicago Press.
- Cohen, W., Nelson, R., and Walsh, J. (2000). Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (Or Not). NBER WP 7552.
- Crane, D. (1972). *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. University of Chicago Press, Chicago, IL.
- Dang, Jianwei and Kazuyuki Motohashi (2015): “Patent statistics: A good indicator for innovation in China? Patent subsidy program impacts on patent quality,” *China Economic Review*, 35, 137–155.
- David, P. A. (1991). The computer and the dynamo: The modern productivity paradox in a not-too-distant mirror. In *Technology and Productivity: The Challenge for Economic Policy*, pages 315–348. OECD.
- De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, 135(2):561–644.
- de Solla Price, D. J. (1963). *Little Science, Big Science*. Columbia University Press, New York.
- Dy.vre, D. and Seager, P. (2023). Linking 70 years of uspto patents to compustat. Working Paper.
- Eberhardt, M., Helmers, C., and Yu, Z. (2017). “What Drives Patenting in China?” *Research Policy*, 46(5), pp. 917-986.
- Economist, The (2020): “Chinese inventiveness shows the weakness of the law,”
- Fang, H., Gu, X., Yan, H., and Zhu, W. (2026). AI Patents in the United states and China: Measurement, organization, and knowledge flows. NBER Working Paper 35022, National Bureau of Economic Research.
- Farber, M. and Tampakis, L. (2024). Analyzing the impact of companies on ai research based on publications. *Scientometrics*, 129:31–63.
- Fisch, C., Block, J., and Sandner, P. (2016). “Chinese University Patents: Quantity, Quality, and the Role of Subsidy Programs,” *The Journal of Technology Transfer*, 41 (1), pp. 60-84.
- Fleming, L., Greene, H., Li, G.-C., Marx, M., and Yao, D. (2019). Government-funded research increasingly fuels innovation. *Science*, 364(6446):1139–1141.
- Gaetani, R. and Li Bergolis, V. (2015). The economic effects of scientific shocks.
- Giczy, A. V., Pairolo, N. A., and Toole, A. A. (2021). Identifying artificial intelligence (ai) invention: A novel ai patent dataset. Economic Working Paper 2021-2, USPTO Office of the Chief Economist.

Giczy, A. V., Pairolero, N. A., and Toole, A. A. (2023). Identifying artificial intelligence (ai) invention: A novel ai patent dataset. *The Journal of Technology Transfer*.

Gnewuch, M., Wahle, J., Ruas, T., and Gipp, B. (2025). Big tech-funded ai papers have higher citation impact, greater insularity, and larger recency bias.

Hagendorff, T. and Meding, K. (2021). Ethical considerations and statistical analysis of industry involvement in machine learning research. *AI & Society*.

Hampole, M., Papanikolaou, D., Schmidt, L. D. W., and Seegmiller, B. (2025). Artificial intelligence and the labor market. NBER Working Paper 33509, National Bureau of Economic Research. Revised September 2025.

Hartmann, P. and Henkel, J. (2020). The rise of corporate science in ai: Data as a strategic resource. *Academy of Management Discoveries*.

Helpman, E., editor (1998). *General Purpose Technologies and Economic Growth*. MIT Press.

37

Hu, Albert (2014): "Propensity to patent, competition and China's foreign patenting surge," in *Intellectual Property for Economic Development*, Edward Elgar Publishing.

Hu, Albert Guangzhou and Gary H Jefferson (2009): "A great wall of patents: What is behind China's recent patent explosion?" *Journal of Development Economics*, 90, 57–68.

Hu, A., Zhang, P., and Zhao, L. (2017). "China As Number One? Evidence from China's Most Recent Patenting Surge." *Journal of Development Economics*.

Hu, Wei, Tohru Yoshioka-Kobayashi, and Toshiya Watanabe (2020): "Determinants of patent infringement awards in the US, Japan, and China: A comparative analysis," *World Patent Information*, 60, 101947.

Jensen, Richard and Marie Thursby (2001): "Proofs and prototypes for sale: The licensing of university inventions," *American Economic Review*, 91, 240–259.

Jaffe, A. B. and de Rassenfosse, G. (2017). Patent citation data in social science research: Overview and best practices. *Journal of the Association for Information Science and Technology*, 68(6):1360–1374.

Jaffe, A. and Lerner, J. (2004). *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress and What to Do About It*. Princeton University Press. Princeton, N.J.

Jovanovic, B. and Rousseau, P. L. (2005). General purpose technologies. In *Handbook of Economic Growth*, volume 1B, pages 1181–1224. Elsevier.

Jurowetzki, R., Hain, D., Mateos-Garcia, J., and Stathoulopoulos, K. (2021). The privatization of ai research (-ers): Causes and potential consequences. *arXiv preprint arXiv:2102.01648*.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy (2021): “Measuring technological innovation over the long run,” *American Economic Review: Insights*, 3, 303–320.

Khanal, S., Zhang, H., & Taeihagh, A. (2025). Development of New Generation of Artificial Intelligence in China: When Beijing's Global Ambitions Meet Local Realities. *Journal of Contemporary China*, 34(151), 19–42.

Klepper, S. (1996). Entry, exit, growth, and innovation over the product life cycle. *American Economic Review*, 86(3):562–583.

Klepper, S. (2007). Disagreements, spinoffs, and the evolution of detroit as the capital of the u.s. automobile industry. *Management Science*, 53(4):616–631.

Klepper, S. (2010). The origin and growth of industry clusters: The making of silicon valley and detroit. *Journal of Urban Economics*, 67(1):15–32.

Klyman, K., Bao, A., Meinhardt, C., Cryst, E., Wald, R., Li, F.-F., and Zhang, D. (2024). Expanding academia’s role in public sector ai.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth,” *Quarterly Journal of Economics* 132(2), pp. 665-712.

Lanjouw, J. and Schankerman, M. (2004). “The Quality of Ideas: Measuring Innovation with Multiple Indicators,” *The Economic Journal*, Vol. 114 (495), pp. 441-465.

Lei, Z., Sun, Z., and Wright, B. (2013). “Patent Subsidy and Patent Filing in China,” *Fung Institute Working Paper*, U.C.-Berkeley.

Lerner, J., Narain, N., Papanikolaou, D., Seru, A., Zunda, W. X., (2026). “Chinese Sputnik Moments?” *NBER Working Paper* 35551.

Li, D., Azoulay, P., and Sampat, B. N. (2017). The applied value of public investments in biomedical research. *Science*, 356(6333):78–81.

Li, Xibao (2012): “Behind the recent surge of Chinese patenting: An institutional view,” *Research Policy*, 41, 236–249.

Lin, Fen, Waverly W. Ding, Shi Chen. 2025. “The Patent Gold Rush? An Empirical Study of Patent Bubbles in Chinese Universities (1990-2019).” *Journal of Technology Transfer* 50(4): 1602-1632
(<https://doi.org/10.1007/s10961-024-10071-z>)

Lin, J., Wu, H., and Wu, H. (2021). “Could Government Lead the Way? Evaluation of China’s Patent Subsidy Policy on Patent Quality,” *China Economic Review*, 69, 101663.

Lieberman, M. B. (1978). A literature citation study of science-technology coupling in electronics. *Proceedings of the IEEE*, 66(1):5–13.

Marquis, D. G. and Allen, T. J. (1966). Communication patterns in applied technology.

American Psychologist, 21:1052–1060.

Marx, M. and Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal*, 41(9):1572–1594.

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., and Zolas, N. (2024). Ai adoption in america: Who, what, and where. *Journal of Economics & Management Strategy*, January.

Merges, R. (1999). “As Many As Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform,” *UC Berkeley Technology Law Journal*, Vol. 14, pp. 577-615.

Merton, R. K. (1968). The matthew effect in science. *Science*, 159(3810):56–63.

Miric, M., Jia, N., and Huang, K. G. (2023). Using supervised machine learning for largescale classification in management research: The case for identifying artificial intelligence patents. *Strategic Management Journal*, 44(2):491–519.

Moser, P., Voena, A., and Waldinger, F. (2014). German Jewish emigres and US invention. *American Economic Review*, 104(10):3222–3255.

Narin, F., Hamilton, K. S., and Olivastro, D. (1997). The increasing linkage between u.s. technology and public science. *Research Policy*, 26(3):317–330.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative ai. *Science*, 381(6654):187–192.

Prud’homme, D. (2012). “Strategic Regulatory Favors in China’s Intellectual Property System,” *Journal of Intellectual Property Rights*, 17 (5), pp. 412-422.

Roach, M. and Cohen, W. M. (2013). Lens or prism? patent citations as a measure of knowledge flows from public research. *Management Science*, 59(2):504–525.

Roach, M. and Sauermann, H. (2010). A taste for science? PhD scientists’ academic orientation and self-selection into research careers in industry. *Research Policy*, 39(3):422–434.

Rock, D. (2019). Engineering value: The returns to technological talent and investments in artificial intelligence. Working Paper, Wharton School of Management.

Rosenkopf, L. and Almeida, P. (2003). Overcoming local search through alliances and mobility. *Management Science*, 49(6):751–766.

Sauermann, H. and Roach, M. (2012). Science PhD career preferences: Levels, changes, and advisor encouragement. *PLOS ONE*, 7(5):e36307.

Shapiro, Carl. (2000). “Navigating the Packet Thicket: Cross Licenses, Patent Pools, and Standard Setting,” in Jaffe, A., Lerner, J., and Stern, S., (eds.), *Innovation Policy and the Economy*, Vol 1., pp. 119-150. MIT Press, Cambridge, MA.

Singh, J. and Agrawal, A. (2011). Recruiting for ideas: How firms exploit the prior inventions

of new hires. *Management Science*, 57(1):129–150.

Stern, S. (2004). Do scientists pay to be scientists? *Management Science*, 50(6):835–853.

Susskind, D. (2020). A world without work. Allen Lane.

Sun, Zhen, Zhen Lei, Brian Wright, MarkCohen, and Taoxiong Liu (2021): “Government targets, end-of-year patenting rush and innovative performance in China,” *Nature Biotechnology*, Vol. 39, pp. 1068-1075.

Sun, R. and Treffer, D. (2022), “AI, Trade, and Creative Destruction: A First Look.” NBER WP 29980.

Sun, R. and Treffer, D. (2023). “The Impact of AI and Cross-Border Data Regulation on International Trade in Digital Services: A Large Language Model,” NBER WP 31925.

Tambe, P. (2014). Big data investment, skills, and firm value. *Management Science*, 60(6):1452–1469.

Thursby, Jerry G and Marie C Thursby (2002): “Who is selling the ivory tower? Sources of growth in university licensing,” *Management Science*, 48, 90–104.

Toole, A. A. and Czarnitzki, D. (2009). Exploring the relationship between scientist human capital and firm performance: The case of biomedical academic entrepreneurs in the SBIR program. *Management Science*, 55(1):101–114.

Van Noorden, R. (2025). “Exclusive: These Universities Have the Most Retracted Scientific Articles,” *Nature* 638(8051), 596-599.

Waldinger, F. (2010). Quality matters: The expulsion of professors and the consequences for PhD student outcomes in nazi germany. *Journal of Political Economy*, 118(4):787–831.

Webb, M., Short, N., Bloom, N., and Lerner, J. (2019). Some facts on high-tech patenting. NBER Working Paper 24793, National Bureau of Economic Research.

Wei, Shang-Jin, Jianhuan Xu, Ge Yin, and Xiaobo Zhang (2021): “Bureaucratic Bean Counting and Patent Subsidies: A Welfare Evaluation of China’s Pro-innovation Program,” Working Paper.

Xu, ShiaoXiong Brian and Hu, Guangwei (2025), Combating China’s Retraction Crisis, *Nature Human Behavior*, 9 (4), 631-634.

Yamaguchi, Kazuhiro (2010): “Japanese Patent Litigation and Its Related Statistics—Current Environment and Future Agenda—,” Notes, 2011, 2013. [8](#)

Zolas, N., Kroff, Z., Brynjolfsson, E., McElheran, K., Beede, D., Buffington, C., Foster, L., and Dinlersoz, E. (2020). Advanced technologies adoption and use by u.s. firms: Evidence from the annual business survey. NBER Working Paper 28290, National Bureau of Economic Research.

Zucker, L. G. and Darby, M. R. (1996). Star scientists and institutional transformation:

Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences*, 93(23):12709–12716.

Zucker, L. G., Darby, M. R., and Armstrong, J. S. (1998a). Geographically localized knowledge: Spillovers or markets? *Economic Inquiry*, 36(1):65–86.

Zucker, L. G., Darby, M. R., and Armstrong, J. S. (2002). Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. *Management Science*, 48(1):138–153.

Zucker, L. G., Darby, M. R., and Brewer, M. B. (1998b). Intellectual human capital and the birth of u.s. biotechnology enterprises. *American Economic Review*, 88(1):290–306.

Table 1. Tier List of Top AI Publication Venues by Sub-domain

Sub-domain	Elite Tier (A* / Flagship)	Gold Tier (A / Highly Prestigious)	Silver Tier (B / Specialized)
Machine Learning	1. Neural Information Processing Systems (NeurIPS) 2. International Conference on Machine Learning (ICML) 3. International Conference on Learning Representations (ICLR)	1. Journal of Machine Learning Research (JMLR) 2. International Conference on Artificial Intelligence and Statistics (AISTATS) 3. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 4. Machine Learning Journal (MLJ)	1. Conference on Uncertainty in Artificial Intelligence (UAI) 2. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD) 3. ACM International Conference on Web Search and Data Mining (WSDM)
NLP	1. Annual Meeting of the Association for Computational Linguistics (ACL) 2. Empirical Methods in Natural Language Processing (EMNLP) 3. North American Chapter of the Association for Computational Linguistics (NAACL)	1. Transactions of the Association for Computational Linguistics (TACL) 2. Computational Linguistics Journal (CL)	1. European Chapter of the Association for Computational Linguistics (EACL) 2. International Conference on Computational Linguistics (LREC-COLING) 3. Asia-Pacific Chapter of the Association for Computational Linguistics (AACL-IJCNLP) 4. Journal of Natural Language Engineering (JNLE)
Robotics	1. IEEE International Conference on Robotics and Automation (ICRA) 2. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 3. Conference on Robot Learning (CoRL)	1. Robotics: Science and Systems (RSS) 2. International Journal of Robotics Research (IJRR) 3. IEEE Transactions on Robotics (TRO)	1. Journal of Field Robotics (JFR) 2. Autonomous Robots Journal 3. ACM/IEEE International Conference on Human-Robot Interaction (HRI)

Sub-domain	Elite Tier (A* / Flagship)	Gold Tier (A / Highly Prestigious)	Silver Tier (B / Specialized)
Computer Vision	1. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2. IEEE/CVF International Conference on Computer Vision (ICCV) 3. European Conference on Computer Vision (ECCV)	1. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2. International Journal of Computer Vision (IJCV) 3. ACM SIGGRAPH (SIGGRAPH)	1. British Machine Vision Conference (BMVC) 2. Asian Conference on Computer Vision (ACCV) 3. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) 4. Computer Vision and Image Understanding (CVIU)
Information Retrieval	1. ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR) 2. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 3. The Web Conference (WWW)	1. ACM Transactions on Information Systems (TOIS) 2. ACM International Conference on Web Search and Data Mining (WSDM) 3. ACM International Conference on Information and Knowledge Management (CIKM)	1. European Conference on Information Retrieval (ECIR) 2. Information Processing & Management Journal (IPM) 3. Journal of the Association for Information Science and Technology (JASIST)
Speech	1. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2. International Speech Communication Association Conference (Interspeech)	1. IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP) 2. Computer Speech & Language Journal	1. Speech Communication Journal 2. IEEE Spoken Language Technology Workshop (SLT) 3. IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)

Sub-domain	Elite Tier (A* / Flagship)	Gold Tier (A / Highly Prestigious)	Silver Tier (B / Specialized)
HCI	1. ACM CHI Conference on Human Factors in Computing Systems (CHI) 2. ACM Symposium on User Interface Software and Technology (UIST)	1. ACM Transactions on Computer-Human Interaction (TOCHI) 2. ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW) 3. ACM International Joint Conference on Pervasive and Ubiquitous Computing (Ubicomp/IMWUT)	1. ACM International Conference on Intelligent User Interfaces (IUI) 2. ACM Conference on Designing Interactive Systems (DIS) 3. ACM International Conference on Mobile Human-Computer Interaction (MobileHCI)
Agents	1. International Conference on Autonomous Agents and Multiagent Systems (AAMAS) 2. International Conference on Learning Representations (ICLR)	1. Autonomous Agents and Multi-Agent Systems Journal (JAAMAS) 2. Conference on Lifelong Learning Agents (CoLLAs)	1. Agent Theories, Architectures, and Languages (ATAL) 2. IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)
Knowledge Rep.	1. International Conference on Principles of Knowledge Representation and Reasoning (KR) 2. International Semantic Web Conference (ISWC)	1. Journal of Automated Reasoning (JAR) 2. Semantic Web Journal (SWJ) 3. Extended Semantic Web Conference (ESWC)	1. Journal of Web Semantics (JOWS) 2. International Conference on Formal Ontology in Information Systems (FOIS) 3. International Web Rule Symposium (RuleML)
General AI	1. AAAI Conference on Artificial Intelligence (AAAI) 2. International Joint Conference on Artificial Intelligence (IJCAI)	1. Artificial Intelligence Journal (AIJ) 2. Journal of Artificial Intelligence Research (JAIR)	1. European Conference on Artificial Intelligence (ECAI) 2. Pacific Rim International Conference on Artificial Intelligence (PRICAI)

Table 2. Comparison of University Core AI Patenting in China Versus the U.S.

	China	US Universities Only	World ex- China
1. Any university	43,942	1,254	2,108
2. Top 20 universities	18,246	822	864
% Top 20	42.2%	65.9%	41.5%

Figure 1. The Great Divergence of AI Patents: China vs. the U.S.

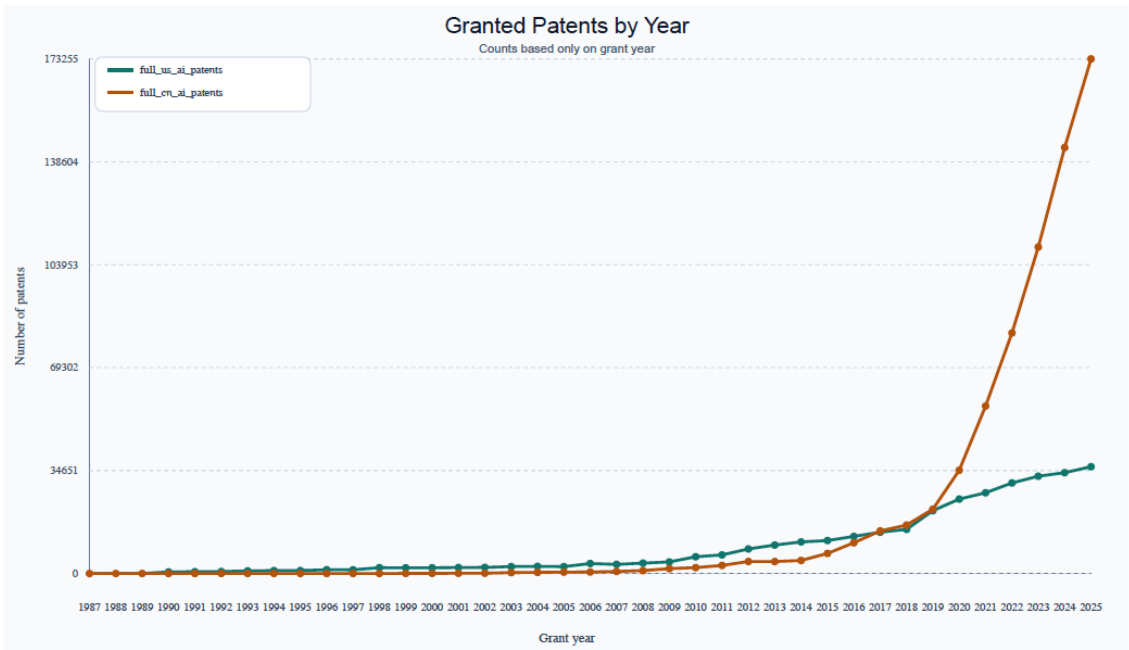


Figure 2. Focus on Core AI Patents: China vs. the U.S.

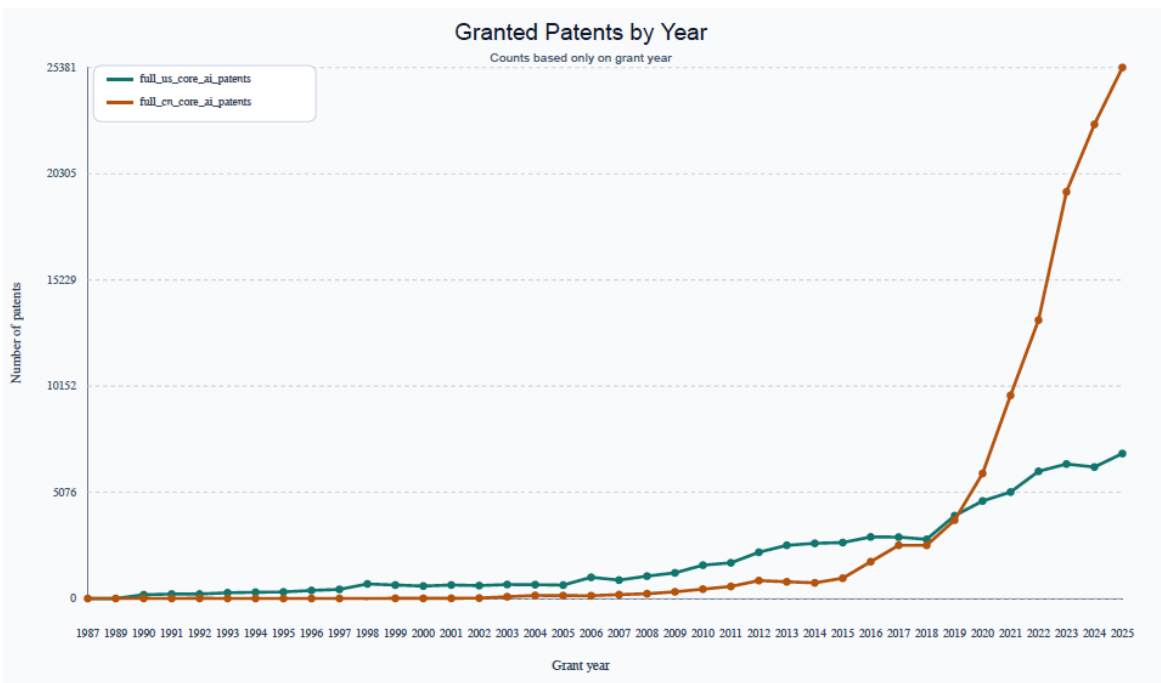


Figure 3 China’s Explosion of Core AI Patents, 2018-2025

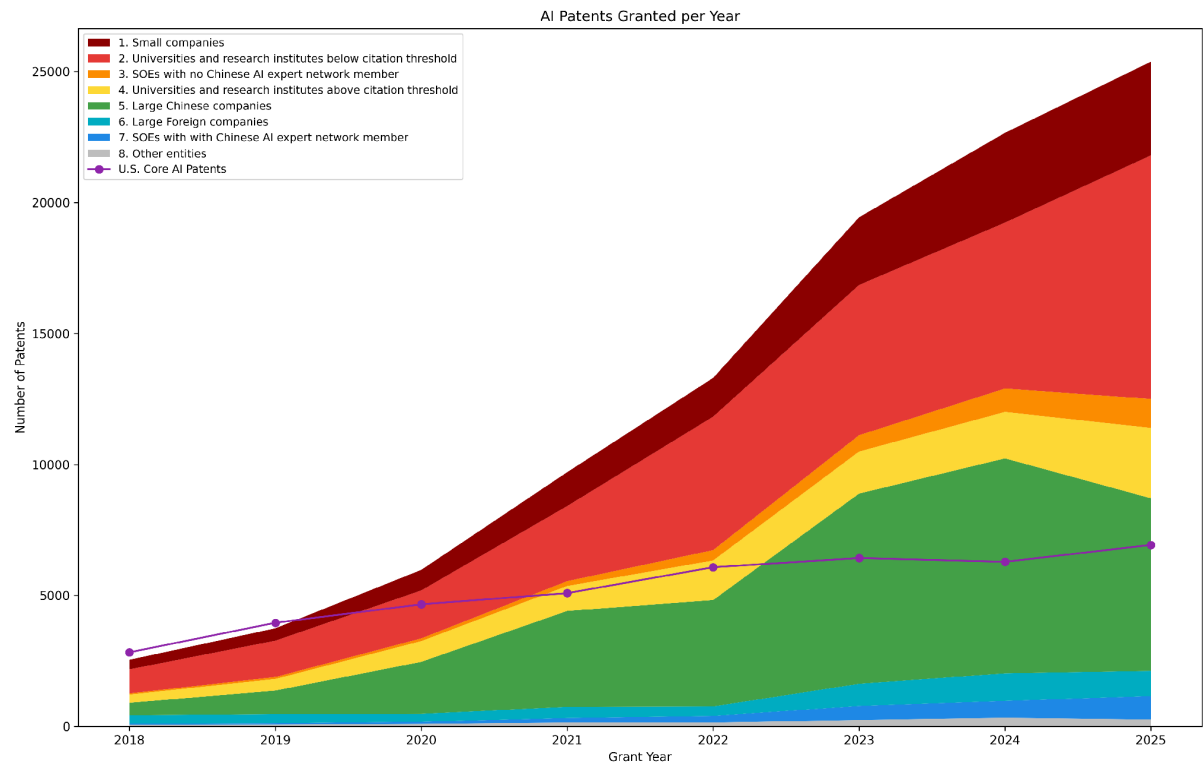


Figure 4. U.S. Growth in Core AI Patents, 2018-2025

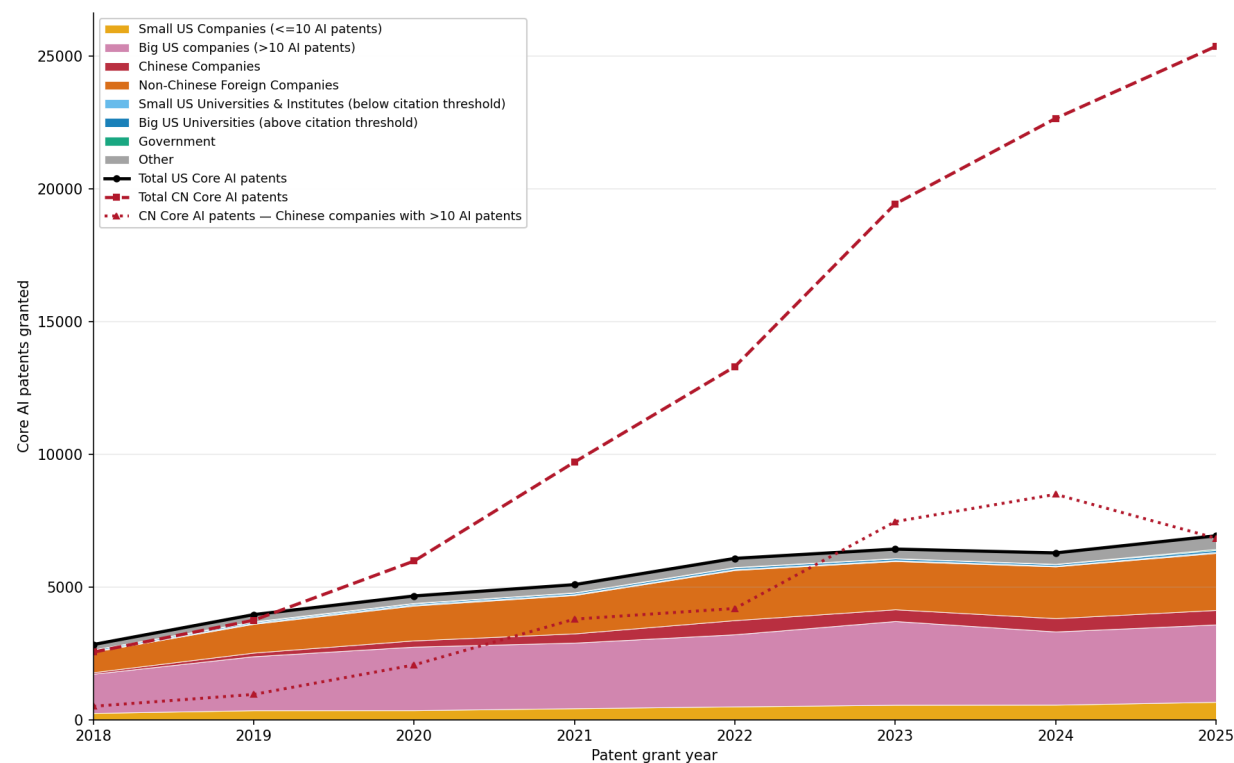


Figure 5a: Chinese Core AI Patents Granted in Multiple Countries: 2018-2025

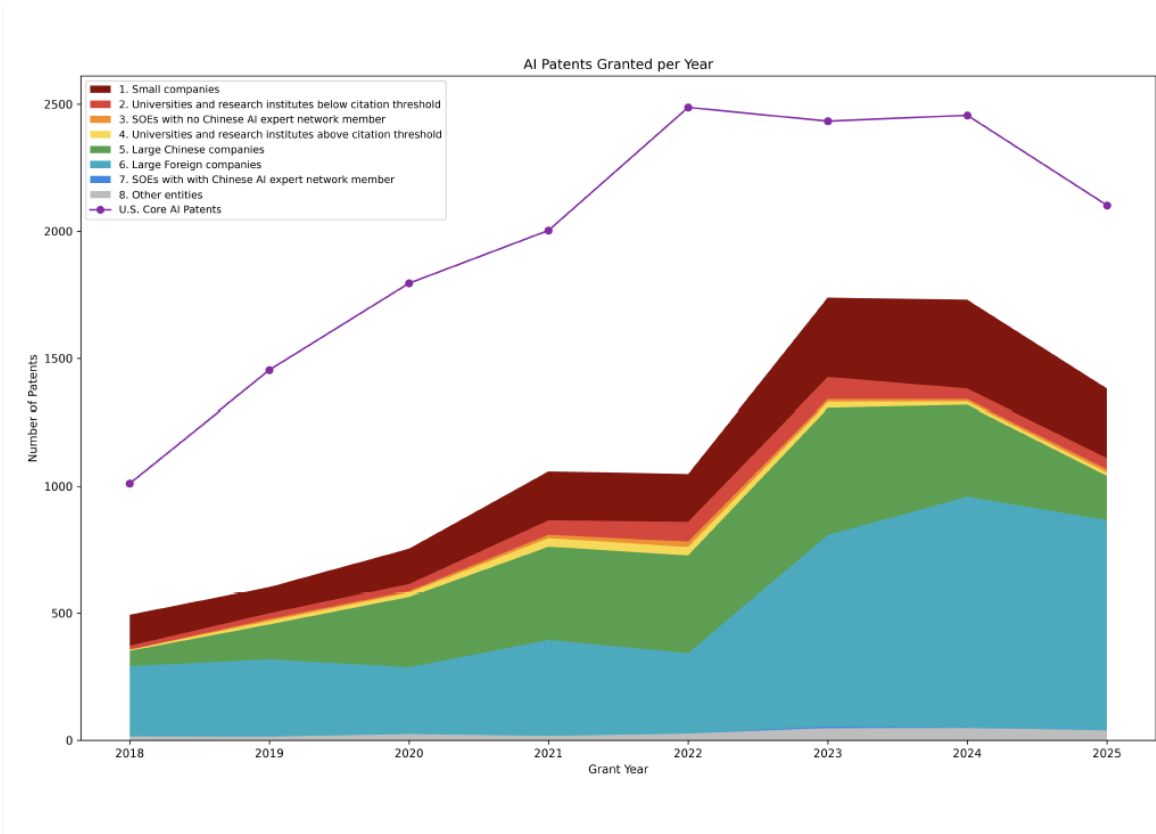


Figure 5b: Chinese Core AI Patents Granted in Multiple Countries (Domestic Applicants Only): 2018-2025

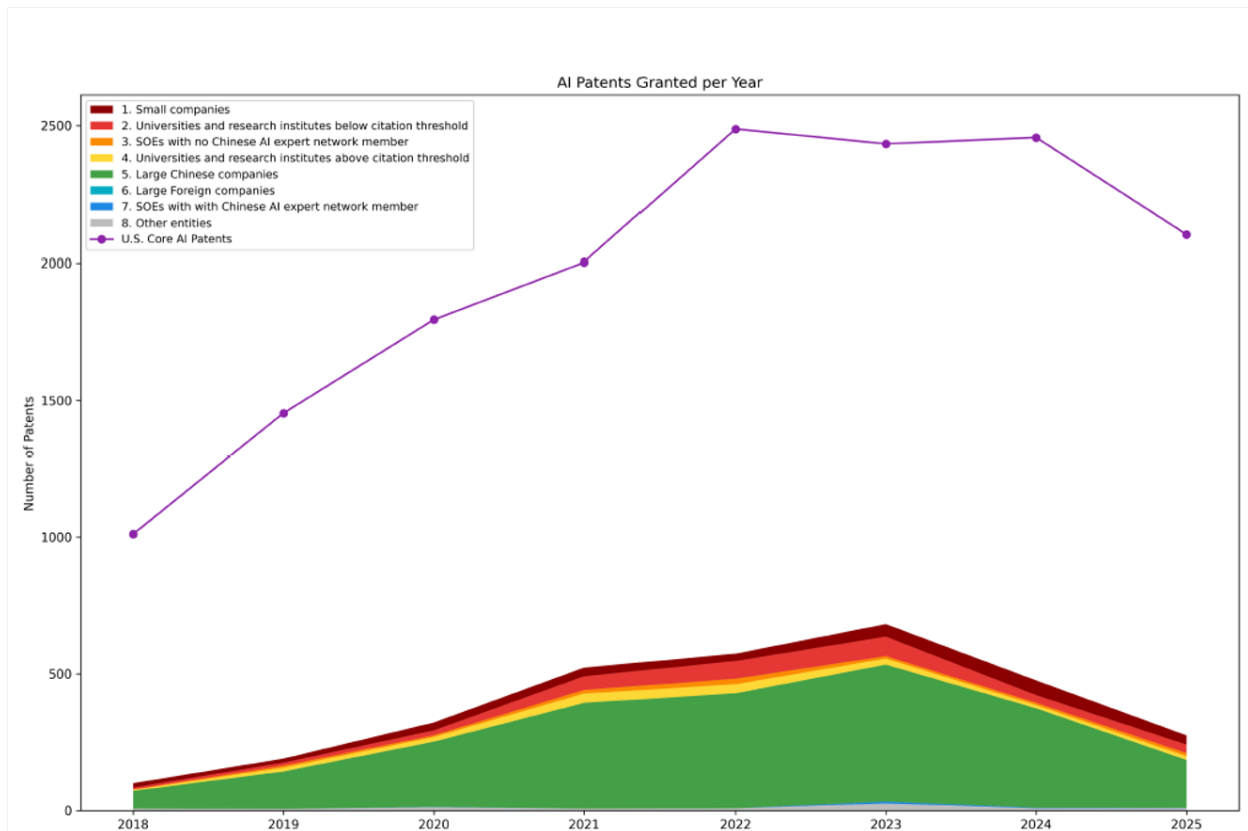


Figure 6a. USPTO Core AI Patents (Granted in Multiple Countries), 2018-2025

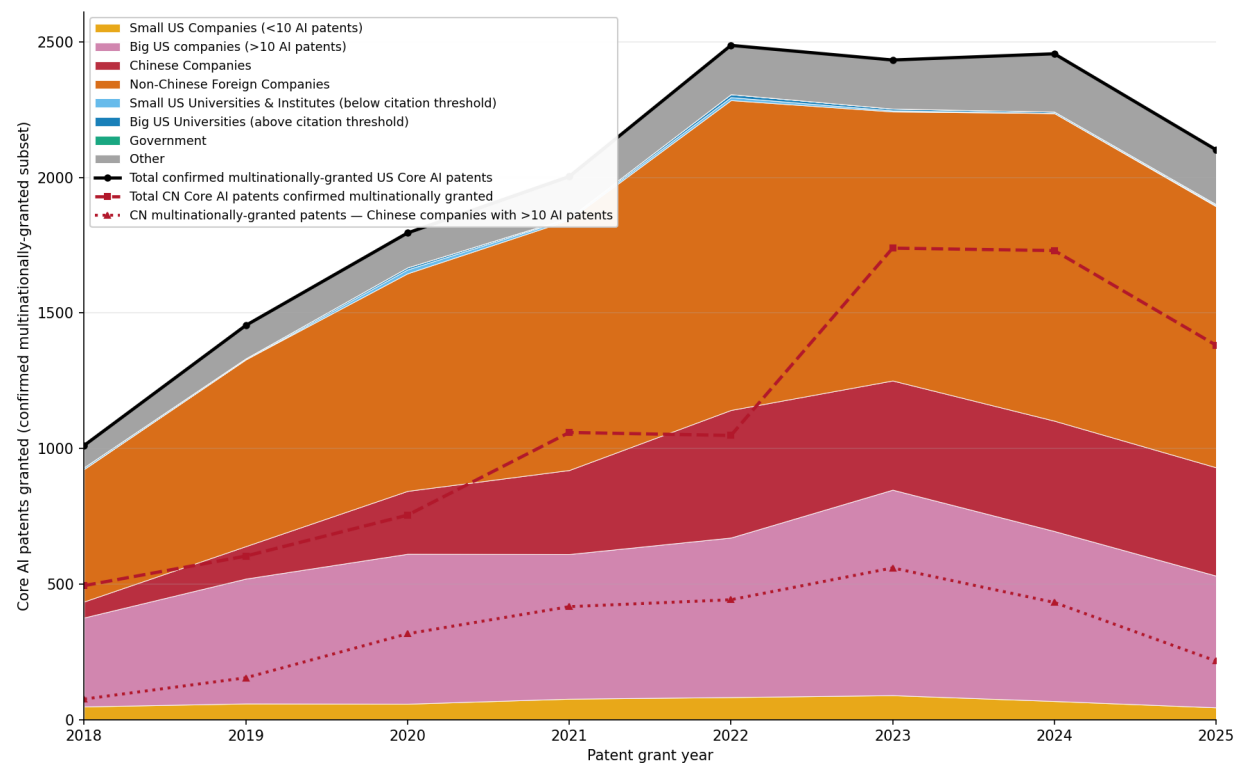


Figure 6b. USPTO Core AI Patents (Granted in Multiple Countries), 2018-2025 (Chinese Assignees Excluded)

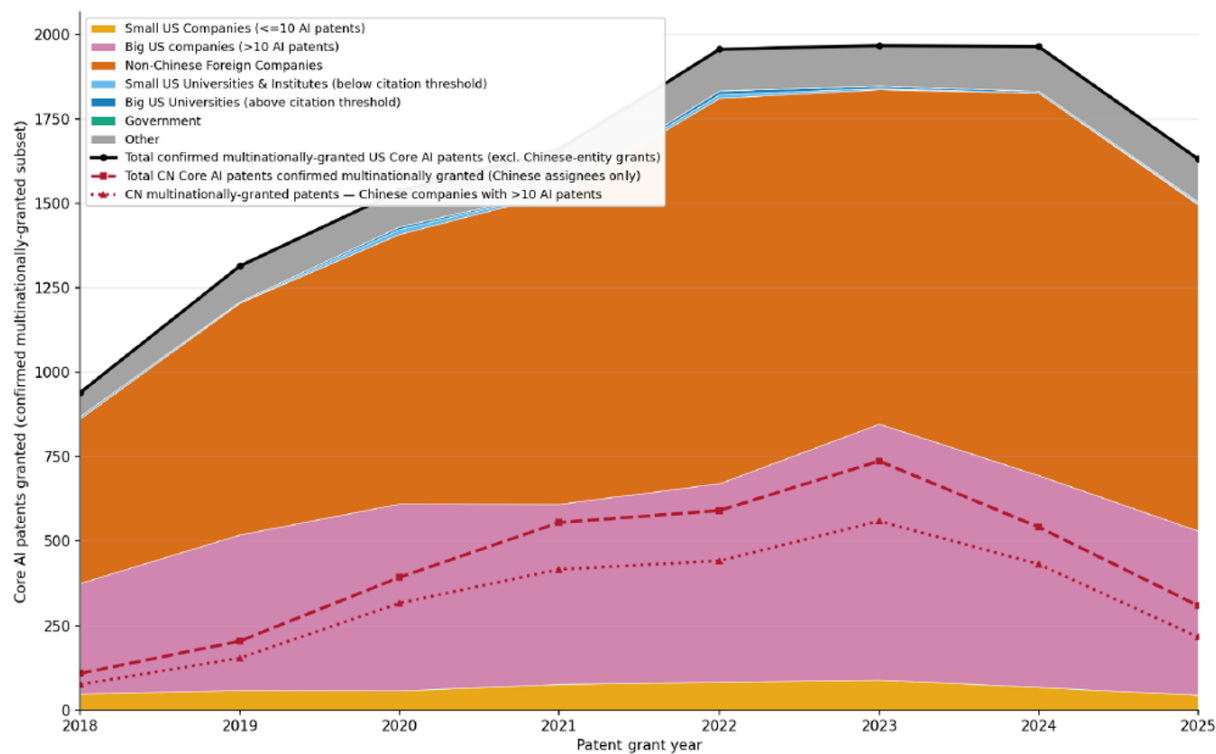


Figure 7: KPST Novelty Index / Chinese Core AI Patents 2010-2022 (Domestic Applicants)

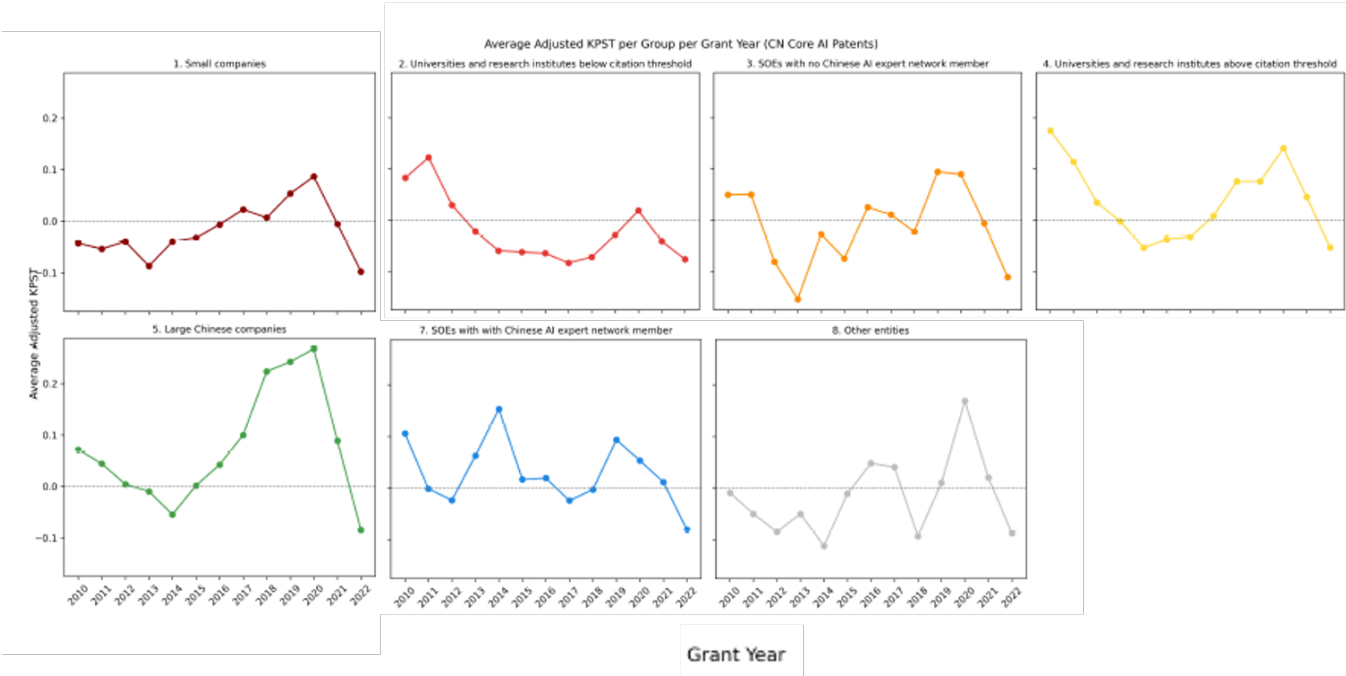
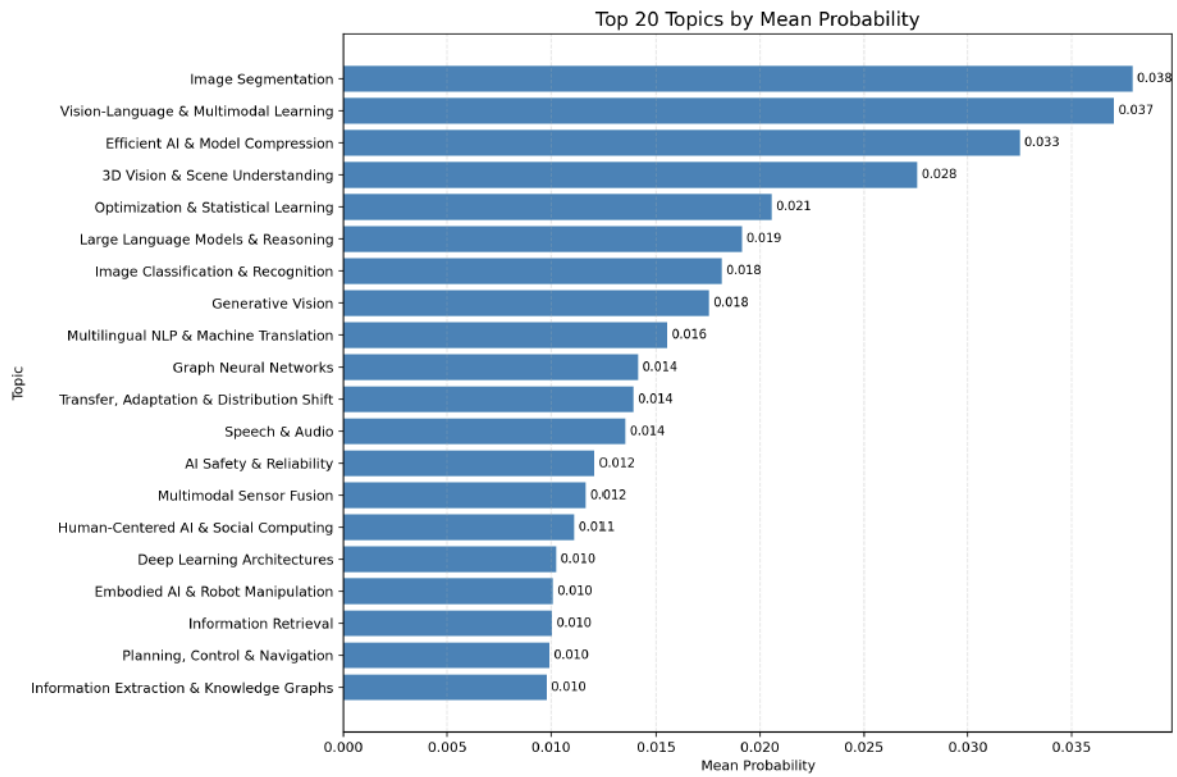


Figure 8. Topic Modeling Applied to the Chinese AI Expert Network



Figure 9. Distribution of Chinese AI Expert Specializations



Appendix

Table A1 Classification Method Evaluations

Method	General Acc	General Precision	General F1	Hard Acc	Hard Precision	Hard F1
Few-shot (12, Expert Definition)	89.75%	90.86%	89.72%	65.50%	64.79%	57.14%
Few-shot (8, Expert Definition)	89.50%	93.72%	89.09%	62.00%	65.22%	44.12%
Zero-shot (Domain List)	90.12%	89.73%	90.28%	70.00%	71.43%	62.50%
Zero-shot (Expert Definition)	90.00%	91.75%	89.90%	56.50%	52.11%	45.96%
Zero-shot (No Definition)	90.12%	91.56%	90.06%	62.00%	59.72%	53.09%
Zero-shot (Textbook Definition)	91.00%	90.10%	91.20%	56.00%	50.98%	54.17%
Core AI	—	93.96%	—	—	83.33%	—

In addition to the trained classifier, we also experimented with several LLM-based methods for patent classification and found that they performed well. These methods mainly differ in how AI is defined in the prompt and whether labeled examples are included. Below is a description of each method:

1. Zero-shot (No Definition): This is the baseline method. We directly ask the LLM to classify each patent as either "AI" or "Non-AI" based solely on its intrinsic knowledge.
2. Zero-shot (Domain List Definition): In this method, we provide the LLM with a comprehensive list of AI domains and subdomains as a reference for classification.
3. Zero-shot (Expert Definition): In this method, we provide the LLM with a concise definition of AI developed by a domain expert.
4. Zero-shot (Textbook Definition): In this method, we provide the LLM with a concise definition of AI summarized from Artificial Intelligence: A Modern Approach by Stuart Russell and Peter Norvig, one of the most widely used textbooks in artificial intelligence.
5. Few-shot (8, Expert Definition) and Few-shot (12, Expert Definition): These methods include both the expert definition and a balanced set of expert-annotated boundary cases in the prompt. The numbers 8 and 12 indicate the number of examples included in the prompt.
6. Core AI: In this method, we use a narrower definition of AI that includes only innovations related to foundational AI frameworks or core AI capabilities. We treat this as a conservative, high-precision classification method.

Description of Datasets:

We used two manually constructed datasets to evaluate the different classification methods.

The General dataset was created by students who collected AI and non-AI patents from online sources. It contains 800 patents in total, including 400 CNIPA patents and 400 USPTO patents. Among them, 404 patents are labeled as AI and 396 as non-AI. The “Hard” dataset was constructed by first sampling boundary cases on which different classification methods disagreed, and then asking a domain expert to label each patent. It contains 200 CNIPA patents, including 90 labeled as AI and 110 as non-AI.

Domain List:

AI-related technical fields include:

1. AI hardware platforms

- intelligent chips
- AI chips
- NPUs
- neural network processors
- neuromorphic chips
- brain-inspired chips
- AI accelerators
- GPU / FPGA / ASIC / SoC only when they are specifically designed for AI computation, neural networks, machine learning training, or AI inference

2. Core AI technologies

- machine learning
- traditional machine learning
- support vector machines
- decision trees
- k-nearest neighbors
- ensemble learning
- reinforcement learning
- deep learning
- neural networks
- convolutional neural networks
- recurrent neural networks
- knowledge graphs
- knowledge extraction
- knowledge fusion
- ontology-based reasoning
- brain-inspired intelligent computing
- quantum intelligent computing
- quantum machine learning
- pattern recognition
- pattern classification
- pattern clustering
- swarm intelligence
- ant colony algorithms
- particle swarm algorithms
- multi-agent intelligence

- hybrid intelligence

3. AI key application technologies

- natural language processing
- machine translation
- semantic understanding
- semantic analysis
- question answering
- intelligent speech
- speech recognition
- speech synthesis
- voiceprint recognition
- computer vision
- image recognition
- image classification
- image detection
- image segmentation
- image generation
- image enhancement when AI-based
- biometric recognition
- fingerprint recognition
- face recognition
- iris recognition
- behavioral recognition
- DNA recognition when used as computational identification or pattern recognition
- augmented reality / virtual reality when combined with AI perception, recognition, generation, or adaptive interaction
- human-computer interaction when involving intelligent perception, speech interaction, gesture recognition, body-sensing interaction, eye tracking, brain-computer interaction, or adaptive user modeling

Textbook AI Definition:

Artificial intelligence refers to computer-implemented systems or methods that perform intelligent functions including, but not limited to, learning, reasoning, perception, knowledge representation, planning, pattern recognition, prediction, autonomous decision-making, or goal-directed action.

Expert AI Definition:

AI is defined as a computational system or a machine that can think or act like human and improve its performance through data or interaction with surrounding environment and other systems.

Core AI Definition:

Algorithms, methods, and technologies on foundational framework or core capabilities of AI, including theoretical foundations of AI, visual perception, speech recognition, understanding language, reasoning, planning, classification, recognizing patterns, prediction, executing actions, taking feedbacks and rewards to improve itself.

Few-shot Examples:

idx. title (application_number) (label)

1. 应用于无线通信系统的智能调试平台系统及调试方法 (CN201010211178.2) (Non-AI)
2. 基于线性泛化回归模型的跨媒体检索方法 (CN201210171539.4) (AI)
3. 一种基于数字孪生的短中长期耦合蓄水调度方法及系统 (CN202510779274.3) (Non-AI)
4. 聚类方法、装置、计算机设备、存储介质和程序产品 (CN201711288774.9) (AI)
5. 一种导航栏推荐方法及其系统 (CN201510020029.0) (Non-AI)
6. 一种图像识别方法和装置 (CN201710536756.1) (AI)
7. 一种基于麦克风阵列波束形成算法的语音处理系统 (CN201910785555.4) (Non-AI)
8. 基于模糊C均值聚类的地铁牵引负荷运行情况对电网运行的影响分析 (CN201710019384.5) (AI)
9. 姿势适应选择 (CN201380048280.7) (Non-AI)
10. 一种本地生活资源优化与分配方法、装置以及终端设备 (CN202411963427.1) (AI)
11. 一种数据中心智能节能调度方法 (CN202310264232.7) (Non-AI)
12. 一种营运车辆挂靠行为的判断方法 (CN202010431488.9) (AI)