March 2021

Mapping India's AI Potential

CSET Data Brief

CENTER *for* SECURITY *and* EMERGING TECHNOLOGY

AUTHORS
Husanjot Chahal
Sara Abdulla
Jonathan Murdick
Ilya Rahkovsky

Executive Summary

India is a vital strategic partner for the United States and a nation with considerable potential in artificial intelligence. As the Indian government moves forward with its plans to invigorate AI-based innovation and build AI-ready infrastructure, understanding the full scope of its AI capabilities—today and in the near future—becomes increasingly important. Drawing from a variety of original CSET datasets and primary source data, this paper presents an overview of India's AI potential along five categories pertinent to the development of AI: 1) talent, 2) research, 3) patents, 4) AI companies and investments, and 5) cloud computing.

Key Takeaways:

- Talent: Talent is a key factor for success for India. The stronger the talent pipeline, the better the research output, intellectual property (IP) creation, and the ability to attract equity investments. India produces almost twice as many master's level engineering graduates as the United States, second only to China, yet less than one-third as many PhDs as the United States. Weaknesses in India's higher education sector limit its ability to continue training a highly skilled AI workforce at scale, and causes Indian students to predominantly pursue PhDs in foreign countries, especially in the United States.

- Research: India has a vast AI research community and is the fourth largest producer of AI-relevant scholarly papers since 2010. However, Indian AI researchers are less likely to collaborate with foreign authors in comparison to the other top 10 AI research-producing countries, to the detriment of the quality and impact of their research. Established AI research fields in India like speech and pattern recognition offer opportunities for Indian researchers to increase their outreach, especially with their counterparts in the United States; Indian AI researchers coauthor papers five times more frequently with U.S. researchers than any other country.

- Patents: Patents are conceptualized as a measure of innovation. Even though the number of Indian-owned AI patents is very small in comparison to its AI research output, India ranks in the top 10 AI patent producing countries, having witnessed high growth in AI patent applications since 2012 driven by a rapid increase in AI-related inventions. The four largest categories for AI patents in India are personal devices and computing, business, telecommunications, and life sciences, which are collectively associated with more than 70 percent of India's AI patents and indicate that Indian innovators have focused on applying AI to areas of traditional strength. India has come a long way in AI patenting in the past two decades, as its patent system has been updated and companies have started to see the benefit of using patents to protect their innovations. That said, India still has a long road ahead to catch up with China and the United States, currently the dominant sources of AI patents.

- AI Companies and Investments: Private-market investment in India's 361 privately-owned AI companies witnessed a steady growth between 2015 and 2018, after which it nearly tripled in dollar value. The percentage jump in India's estimated investments in 2019 alone was higher than that of any country, including the United States and China. Among foreign investors in the Indian market, U.S. investors far outpace Chinese investors. Over half of all Indian companies applying AI to its product are active in business analytics, medicine, finance, and sales, retail and customer relations, and they attract more than 60 percent all equity funding given to Indian AI companies.

- Cloud Computing: Since India does not have a domestic manufacturing capacity for AI chips, market cloud computing spending is a useful proxy for the country's capacity to support its AI computing needs. India lags in cloud computing with a cloud adoption rate that ranks below the world average. But India also has one of the fastest growing cloud markets with rising demands for computing power that is expected to grow at double-digit rates. In order to make progress in this domain, a pragmatic and cost-effective path forward for India would be to build a domestic centralized AI-specific cloud infrastructure facility to facilitate the work of a broader spectrum of stakeholders in its AI ecosystem. Notably, the Indian AI strategy is moving in this direction.

Across all of these categories, the prevalence of India-U.S. interactions and widespread evidence of existing ties are hard to miss. Moving forward, any successful partnership between the two countries will rely on their ability to provide a healthy environment to educate and train future AI innovators, and provide low cost and accessible cloud infrastructure to advance innovation, while at the same time further strengthening research collaborations, IP cooperation, and investor ties.

Table of Contents

Introduction

As one of the fastest growing major economies and the second most populous country in the world, India has a significant stake in the development of artificial intelligence globally.[1] The Indian government believes the country's thriving tech community, robust information technology (IT) ecosystem, and growing economy could position it among the leaders in AI.[2] These expectations hold merit. India's workforce holds the highest average share of AI skills represented among its top 50 skills than any other country in the world—roughly 2.6 times the global average.[3] It has experienced one of the fastest growths in AI hiring from 2015 to 2019.[4] India has a vibrant technology market and the third largest startup ecosystem globally, which added 1,300 startups in 2019 alone.[5]

However, India also has significant impediments to making AI progress. It lacks the facilities to support large scale experimental test beds, an enabling data ecosystem that facilitates access to intelligent data, and sufficient high-quality talent working on cutting-edge AI algorithms.[6] India's existing compute capabilities are at best "islands of modest excellence with capabilities of a few petaflops," not suitable for AI workloads at scale.[7]

In recent years, India has amplified its efforts to improve its AI infrastructure. In June 2018, the Indian government released the country's National Strategy for Artificial Intelligence, which identified priority sectors for AI deployment—healthcare, agriculture, education, smart cities and mobility—and has since partnered with leading tech players (like Intel and Microsoft) to implement AI projects in these areas.[8] India's Ministry of Electronics and Information Technology constituted four committees to promote AI initiatives and develop policy frameworks for faster adoption of AI.[9] Subsequent reports released by these committees in July 2019 proposed action in areas of data; identification of AI-relevant national missions in key sectors; skills development and research and development (R&D); and cybersecurity, safety, legal, and ethical issues.[10]

Other key initiatives include a training program aiming to make AI tools and skills accessible to youth, an AI-specific cloud compute infrastructure facility, a standardization committee to bring functional uniformity in AI architectures across sectors, and various centers of research excellence for transformational AI.[11] India has been an active participant in many of the new multilateral efforts around AI, such as the Global Partnership on Artificial Intelligence, and is likely to play a significant role in global AI standard setting. The U.S. National Security Commission on Artificial Intelligence proposed creating an India-U.S. strategic tech alliance, citing the country's tech talent, strong innovation, and technical infrastructures together with its growing geopolitical challenges and a shared commitment to freedom and democratic principles.[12]

Despite India's accomplishments in AI, there is no single, holistic assessment of the country's technological capabilities in this area. In this paper, we evaluate India's potential for AI by examining its progress across five categories of indicators pertinent to AI development—talent, research, patents, companies and investments, and compute. These elements are widely regarded as crucial enablers of AI and have prominently figured in many reports evaluating a given country's individual or relative standing.[13] They also represent CSET's own active lines of research in studying AI, with most of them classified as AI foundations. Data is often identified as a precondition to train AI algorithms, but we exclude it from this report because data needs for AI are task specific, and broader proxies estimating a country's data potential more generally (population size, number of digital activities, etc.) could be misleading.[14] A comprehensive assessment of India's data management infrastructure could be a useful exercise and an area for future work.

We draw our assessments for the five categories under consideration based on data from a variety of original CSET datasets and primary data sources. The measures evaluated in this brief include:

a. Talent: computer science and engineering graduates.

b. Research: scholarly papers from academic journals and conference proceedings.

c. Patents: patent documents associated with global inventions.

d. AI Companies and Investments: equity investment flows into privately held AI companies.

e. Cloud compute: market trends on cloud computing capacity.

Graduates from higher education institutions determine the quality and availability of domestic AI workforce. Scientific publications are key indicators of a country's research strength. Patents measure the transformation of research into inventions that can be used to design new products and services. Private equity investment flows provide meaningful insight into innovation, health, and growth in the commercial sector. Computing power is the engine that runs AI.

All of these resources provide useful insights into India's AI landscape, but each has its own limitations, and even together, they remain an imperfect measure of true AI progress. For example, talent figures are based on broader proxy discipline categories, because data for STEM graduates is not consistent across countries. Not all Indian research is published, some non-English research may not be included in the journals we cover, and it is hard to judge the quality of this research. Meanwhile, innovation conceptualized as patents may not provide a comprehensive picture, since organizations may elect not to patent all of their inventions, but instead hold them as trade secrets, especially if they believe that a given country's legal system does not provide sufficiently strong protection for intellectual property.[15] Moreover, our measures for AI-related investment do not cover equity investments made in public companies or internal investments in AI by companies that are not focused on AI. Unlike the other levers, computing capacity estimates are based primarily on third-party sources, rendering less confidence in our evaluation. Therefore, we encourage readers to continue exploring other metrics beyond our assessment for a complete picture.

These limitations notwithstanding, this paper lays the foundation for a comprehensive assessment of India's capabilities in AI, including previously underexplored dimensions. It serves as a primer for anyone looking to estimate India's relative strengths and weaknesses in developing AI, specifically what it brings to the table while collaborating with AI heavyweights or in its attempt to be a leader amongst other developing economies. India is emphasizing reforms and growing quickly across most of the five AI levers. Therefore, we encourage readers to approach this analysis in view of the fast-evolving priorities and trends in Indian AI development.

The following section presents our findings across each of the five categories—talent, research, patents, AI companies and investment, and cloud computing—followed by concluding thoughts.

Findings

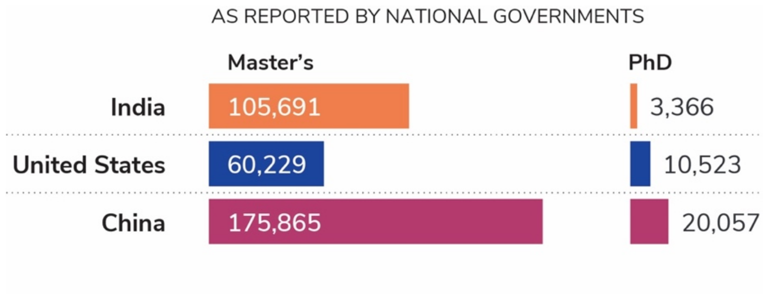
Talent

This section interprets the AI talent landscape in India through the lense of the number of master's and doctoral graduates in STEM fields.[16] The data on this is drawn from a cross-national dataset on the number of engineering or computer science graduates as reported in each country by their respective governments—namely the All India Survey on Higher Education (AISHE), Ministry of Education in China, and the National Center for Education Statistics in the United States. It includes numbers on global talent flows as reported by the United Nations Educational, Scientific and Cultural Organization (UNESCO), and data on international students in the United States as reported by the Institute of International Education.

There is no single way to define AI talent, but a common approach is to consider people with graduate degrees in computer science and computer engineering as a proxy for it.[17] However, there are measurement challenges, especially when it comes to cross-country comparisons given a lack of uniformity in the way different countries classify disciplines and report numbers.[18] Despite these limitations, we can compare the number of graduates across countries if we broadly look at all engineering graduates for the time frame that all countries of concern have provided data for.

Table 1 presents the total number of 2016–17 graduates in engineering at the master's and doctoral level broadly across India, China, and the United States, as reported by their national governments. The total number of bachelor's graduates in engineering for 2016–17 in India were 894,437 or roughly seven times the number in United States that total 133,790 for the same time period.

Table 1: Engineering graduates, 2016–17.



Source: All India Survey on Higher Education, Ministry of Education in China, and the National Center for Education Statistics in the United States.

These numbers help us get a sense of the AI talent pool produced by each of these countries. Overall, China had the highest number of engineering graduates in 2017—both at the master's and doctoral level. It is interesting to note that even though India produces a large pool of bachelor's and master's students in comparison to the United States, its number of doctoral graduates is much smaller. While in the United States, engineering graduates with master's degrees outnumbered those with doctoral degrees nearly 6-to-1. In India, the number for master's graduates was 31 times greater than the number of graduates with doctoral degrees.

Table 2 offers a closer view into this quantitative disparity at the two degree levels by exploring the number and proportion of Indian master's and doctoral graduates in certain AI-relevant fields from 2015 to 2019.

Table 2: Indian master's and doctoral graduates in AI-related fields, 2015–19.[19]

	Total graduates	Master's		PhD	
		NUMBER	% OF TOTAL	NUMBER	% OF TOTAL
Computer Engineering	113,659	111,236	97.9%	2,372	2.1%
Electrical Engineering	53,283	51,155	96.0%	2,128	4.0%
Electronics Engineering	99,150	96,938	97.8%	2,212	2.2%
Information Technology	10,148	9,967	98.2%	141	1.4%
IT & Computer	467,164	454,804	97.4%	2,072	0.4%
Mechanical Engineering	67,548	64,894	96.1%	2,654	3.9%

Source: All India Survey on Higher Education.

Among all Indian graduate students, doctoral students comprise a significantly smaller proportion of the cumulative engineering graduates at the master's, MPhil and doctoral levels. Across all major AI-related fields under study, the share of doctoral students remains under 4 percent of the total engineering graduates. Such low proportions are due to several factors, including absence of sufficient research centers or departments in most Indian universities, shortage of qualified teachers, lack of student access to financial support, complex regulatory norms, and concerns regarding the quality of India's higher education sector.[20] According to AISHE's 2018–19 report, only 34.9 percent of all higher education institutions in India run master's or MPhil programs and only 2.5 percent of them run doctoral programs.[21]

Due to the lack of alternatives within India's resource constrained higher education system, an increasing number of Indian students have moved abroad to pursue their graduate studies. The total number of Indian students outside of India increased from 118,924 in 2003 to 181,872 in 2013.[22] In 2016, about 278,383 students were pursuing post-secondary education in foreign countries. This accounted for almost 7 percent of India's cumulative graduate enrollment.[23] Table 3 illustrates the top five destinations for Indian students moving abroad in 2016.

Table 3: Top five destinations for Indian students abroad, 2016.

	SHARE OF STUDENTS MOVING ABROAD
1. United States	**36.2%**
2. Australia	19.5%
3. Canada	9.3%
4. United Kingdom	5.2%
5. Germany	4.1%

Source: UNESCO Institute for Statistics, Global flow of tertiary-level students.

The United States stands out as the most popular choice, with over 36 percent of Indian students travelling to the country to pursue higher education. Canada was the third most favored choice in 2016, attracting roughly 9 percent of Indian graduates—but the growth for Canada in recent years has been much higher.[24]

Indian students account for nearly 14 percent of international students in the United States. India is the second largest country of origin for foreign students studying in the United States, China is the first. However, when looking specifically at mathematics, computer science, and engineering talent, the number of international Indian students in the United States has doubled in the past decade, outpacing their Chinese counterparts for six years out of 10 since 2009 (see Table C and Figure D in the Appendix).

Of the total annual engineering doctoral graduates in the United States (averaged across 2015-2017), Indian doctoral graduates comprised 9 percent of the total, or 783 in number.[25] This number represented 21 percent of the total annual number of engineering doctoral students graduating in India averaged across the same time period.[26]

Talent from India in the United States:

- Indians account for 71.7 percent of approved H1B visas issued annually by the U.S. Citizenship and Immigration Services (USCIS).[27]
- Indian students comprised 13 percent of all U.S. doctoral recipients on temporary visas between 2000–17, second only to China. 93 percent of these doctorates were in the science and engineering fields.[28]
- Intention-to-stay rates (or percentage of international students intending to stay in the United States after completing their degrees) for Indian graduate students are among the highest, exceeding 90 percent.[29]
- First-generation immigrants from India comprised 11 percent of the founders of the top 50 AI startups in the United States—the largest share among all immigrant founders.[30]

The lack of strong doctoral programs in India limits the country's ability to continue training its own highly skilled domestic AI workforce, which will potentially impact its ability to be more innovative in AI. The stronger the talent pipeline the better the research output, intellectual property (IP) creation, and the ability to attract funding for startups and businesses. This is a key factor for success for India.

The Indian government recognizes its current limitations and inability to fully accommodate the talent pool at home. Consequently, it does not see itself as suffering from a brain drain requiring it to take policy steps to induce return migration of its talent. As a prominent Indian official reportedly noted, "brain drain" is better than a "brain in the drain."[31]

The development of infrastructure needed to train its AI talent is a long-term endeavor that India should push harder to achieve. In the meantime, India can benefit from strengthening its ties with its tech talent diaspora, which already plays a significant role in the U.S. tech sector. This can also serve as a bridge between the two countries, transferring skills, creating avenues for research collaboration, and other externalities such as investment flows, inducing incentives for higher and technical education, and setting up of new firms and subsidiaries of multinational companies.[32]

In reference to talent flows to the United States, India's Minister of External Affairs, S. Jaishankar remarked in September 2020 that supporting Indian talent in the United States "is a win-win for both countries."[33] India creates a huge AI talent pool, but an inability to provide the right environment to nurture it will be a loss overall—for India, the United States, and the global AI community.

Research

In this section we explore AI research in India by examining research papers published by Indian authors in areas related to AI.[34] Raw research output is a very simple measure of research capability and suffices for this quick introduction; however, other indicators can and should be explored to fully understand a country's research potential. We compare India's public-facing scientific research output with other countries and explore the top collaborators with Indian scientists across all AI-related topics over the past decade. The data for this section draws from global academic journals and conference proceedings in Web of Science, China National Knowledge Infrastructure, Dimensions, Microsoft Academic Graph, and arXiv from 2010 to 2019. We estimate that this merged dataset of scholarly literature holdings collectively captures over 90 percent of the world's scholarly output during this period with 109.8 million unique research papers in Chinese, English, French, German, Japanese, Portuguese, Spanish, and other languages.[35]

India is one of the top producers of AI scholarly research publications over the past decade. Between 2010 and 2019, we identified a total of 84,384 AI papers that have at least one author affiliated with India.[36] Table 4 presents the top 10 countries by production of AI-relevant scholarly papers in 2010–19. While China and the United States are clear leaders, Indian AI-relevant research production ranks fourth, closely behind the United Kingdom and ahead of Germany and Japan.

Table 4: Top 10 countries by production of AI-relevant scholarly papers, 2010–19.

		NUMBER OF AI PAPERS
1.	China	471,726
2.	United States	310,562
3.	United Kingdom	88,573
4.	**India**	**84,384**
5.	Germany	81,780
6.	Japan	76,491
7.	France	61,197
8.	Canada	49,583
9.	Italy	48,441
10.	Spain	47,508

Source: CSET merged corpus of scholarly literature, as of January 15, 2021.

The progress of Indian AI research is all the more impressive given that India's expenditure on R&D totaling $49 billion in 2015 is much smaller than some advanced economies, like Japan at $170 billion and Germany at $132 billion (2017 figures).[37] India's rank slips further down versus others when looking at R&D as a share of its gross domestic product (GDP) or gross expenditure on R&D per researcher.[38] This means that India's scientific research in AI is most likely underfunded in comparison to many countries, but tends to outperform in terms of the volume of research production.[39]

Indian AI scholars tend to produce research more independently and are less likely to collaborate with researchers from other countries. Of the 84,384 AI research papers produced by Indian authors, only 16 percent of the papers had at least one non-Indian coauthor.[40] This share is the lowest among all the top 10 AI research producing countries. While factors impeding international coauthorship by Indian scholars is a topic that requires further study, some of the barriers indicated by Indian scientists in surveys include lack of funding for international work, differing academic standards, bias against scholars from economically developing countries, bureaucratic impediments, and lack of institutional support.[41]

It is widely believed that international collaboration benefits research quality and impact.[42] Therefore, it is not surprising to see that India's rank slips down from the top 10 countries to 15th when we evaluate its AI research based on citation counts.[43]

Our findings indicate that Indian AI research collaborations are the highest with scholars from the United States, as shown in Table 5, with a total of 5,339 coauthored AI papers between the two since 2010.[44]

Table 5: Top 10 countries by AI research collaborations with India, 2010–19.

	NUMBER OF AI PAPERS
1. United States	5,339
2. United Kingdom	1,154
3. Australia	856
4. China	828
5. Singapore	790
6. Germany	780
7. France	749
8. South Korea	697
9. Canada	652
10. Italy	652

Source: CSET merged corpus of scholarly literature, as of January 15, 2021.

It is noteworthy that the number of India-U.S. paper collaborations are roughly five times higher than any other country that Indian AI researchers collaborate with. The 5,339 India-U.S. coauthored papers are more than half of the 10,304 paper collaborations that Indian researchers have had with the rest of the world combined. Among other factors, this could be due to established institutional partnerships between the two countries, like the Indo-U.S. Science and Technology Forum (IUSSTF) of March 2000, focused primarily on promoting scientific collaboration. For the United States, Indian authors are the eighth largest collaborators on AI research ranking closely behind Japan's 5,360 and France's 6,395 paper coauthorship figures.[45]

To get a better sense of the top operational avenues for AI research cooperation, we also examined the India-U.S. partners responsible for the most collaborations. Carnegie Mellon University in the United States, and the Indian Institutes of Technology emerge as top collaborators, having coauthored 74 papers together, followed by Harvard University and TCS Innovation Labs.[46] See Table A in the Appendix for details.

It is also worth understanding what AI fields Indian researchers focus on, as shown in Table 6A, which classifies the total AI papers produced by Indian researchers by their fields of study, along with the corresponding number of paper collaborations between India-US, India-China, and India-rest of the world (ROW).[47]

Table 6A: AI research by fields of study and collaboration type, 2010–19.[48]

NUMBER OF PAPERS AND SHARE OF TOTAL AI RESEARCH (shading scale: 5%, 10%)

	India		India-US		India-China		India-ROW	
Pattern recognition	19,265	23%	794	15%	176	21%	1,808	19%
Computer vision	12,313	15%	554	10%	126	15%	1,136	12%
Machine learning	6,253	7%	510	10%	71	9%	784	8%
Data mining	4,569	5%	270	5%	54	7%	495	5%
Natural language processing	4,071	5%	350	7%	16	2%	473	5%
Speech recognition	4,067	5%	248	5%	20	2%	370	4%
Algorithm	3,725	4%	240	4%	40	5%	411	4%
Information retrieval	2,422	3%	239	4%	15	2%	249	3%
Artificial intelligence	1,890	2%	115	2%	29	4%	261	3%
Control theory	1,798	2%	71	1%	7	1%	189	2%
Mathematical optimization	1,420	2%	150	3%	23	3%	252	3%
Real-time computing	1,366	2%	101	2%	10	1%	124	1%
Human-computer interaction	1,230	1%	131	2%	16	2%	190	2%
Data science	927	1%	90	2%	14	2%	132	1%
Unclassified	1,683	2%	57	1%	13	2%	96	1%
All other fields	17,385	21%	1,419	27%	198	24%	2,506	26%
Total	**84,384**	100%	**5,339**	100%	**828**	100%	**9,476**	100%

Source: CSET merged corpus of scholarly literature, as of January 15, 2021.

Table 6B: Average yearly growth rate in AI research and collaborations, 2010–19.[49]

GROWTH RATE FROM 2010 TO 2019

100 200% 300% 400%

	India	India-US	India-China	India-ROW
Pattern recognition	35%	166%	1,750%	216%
Computer vision	23%	181%	1,250%	179%
Machine learning	124%	233%	227%	243%
Data mining	22%	107%	530%	145%
Natural language processing	29%	308%	-	215%
Speech recognition	25%	238%	-	175%
Algorithm	63%	290%	-	284%
Information retrieval	29%	110%	-	198%
Artificial intelligence	29%	182%	280%	208%
Control theory	26%	91%	-	226%
Mathematical optimization	21%	105%	220%	100%
Real-time computing	75%	243%	-	610%
Human-computer interaction	126%	427%	-	307%
Data science	114%	215%	-	650%
Unclassified	21%	275%	-	110%
All other fields	49%	179%	386%	181%
Total	39%	180%	627%	195%

Note: This table is organized in a descending order as per number of papers in earch field.

Source: CSET merged corpus of scholarly literature, as of January 15, 2021.

Table 6A lays out the data by total paper counts for each field and their respective shares in the total. India has produced the greatest number of AI research papers in pattern recognition, comprising 23 percent of all its AI publications, followed by computer vision at 15 percent. The smallest field is data science with only 927 published papers. Research collaborations—where at least one Indian author collaborates on a paper with at least one author from the corresponding country—follow the same trend.

It is worth noting that in comparison to global trends, Indian researchers produce more papers on speech recognition versus all other AI fields, comprising 5 percent of its total AI publications, compared to the rest of the world where 47,797 speech recognition papers comprise 2 percent of the total AI publications (47,797 of 1,952,800 as of February 1, 2021). Speech recognition papers affiliated to Indian authors comprise about 9 percent of all papers on the topic produced globally. This could likely be due to the country's vast linguistic variations among its billion plus population that provides extensive opportunities to research the field.

Over time, research areas evolve with scientific breakthroughs and new trends. Table 6B presents the growth in AI-related research output from 2010 to 2019, broken down by field of study and organized in a descending order by size of field in India. The growth percentages should be seen keeping in mind that the fields up top (especially the first four) have larger paper counts and therefore more significant results versus when further down the list to smaller fields with fewer number of papers. See Table 3A for total paper counts. In India, the number of research publications linked to the general category of machine learning (ML) witnessed one of the fastest growths averaging 124 percent over the past decade. Smaller fields like human-computer interactions and data science have also expanded, growing by 126 percent and 114 percent respectively. India-ROW papers follow the same trend across smaller fields, with the categories of data science, real-time computing, and HCI growing at a high pace. For papers coauthored by Indian and U.S. scholars, other than HCI, the larger field of natural language processing witnessed a high growth of 308 percent since 2010. Overall, India-U.S. collaborations across all fields grew by 180 percent, in comparison to 195 percent for India-ROW and 39 percent for India alone.

India and China have much fewer paper coauthorships and the seemingly high growth figures in some fields could be misleading.[50] Most of the high-growth numbers for India and China collaborations are due to very small numbers of coauthored papers in 2010, translating into a significant growth jump by 2019.

As a whole, India has an advantage of building upon a huge pool of research conducted by its vibrant academic community. But without greater international collaboration, India's AI-relevant research could potentially miss opportunities to further its research quality and will fail to have a significant impact globally. AI and closely related fields where India has an established research community, like speech recognition and pattern recognition, occupy an important place in the country's scientific progression. They offer avenues for greater learning and outreach globally, as well as commercial growth as an increasing number of businesses develop applications based on this research.

Patents

This section looks at patenting as an indicator of AI-related innovation in India. We examine India's AI patent activity between 2002 to 2019 in terms of filed applications and granted patents for those IP documents first filed within India.[51] The following data is drawn from a worldwide AI-relevant patents database at CSET derived from patent categorizations by 1790 Analytics.[52] All of the applied for and granted patents are grouped into patent families, which represent all patents related to the same invention to avoid double counting.[53] Any reference to patents in the section below refer to patent families, unless otherwise specified as applications or granted patents.[54]

India has traditionally fared poorly in patent filings due to several factors, including poor infrastructure and limited resources for processing patent applications, weak IP protections, limited R&D funding, and a general lack of a patenting culture.[55] Consequently, a majority of Indian patents have been filed by nonresidents in India's patent office, amounting to roughly 68 percent of the total in 2017–18. Nearly 62 percent of these nonresident patents filed in India were from residents of four countries—the United States (31.5 percent), Japan (13.9 percent), Germany (8.6 percent), and China (8 percent).[56] Even including these nonresident patents, Indian patenting activity has been fairly moderate, and it continues to lag far behind patenting giants like China and the United States.

Keeping this in mind, India's global rank on the number of patent families related to AI was surprisingly high.[57] Table 7 lists the top 10 countries by patent families globally between 2002–19. India ranks number 8, slightly above Russia and France. It is interesting to note that, in comparison to other top 10 countries, India had no AI relevant patents prior to 2002. Therefore, it is impressive to see India get ranked in this list after having caught up in less than two decades in this domain.[58]

Table 7: Top 10 countries by AI Patent Families, 2002–19.

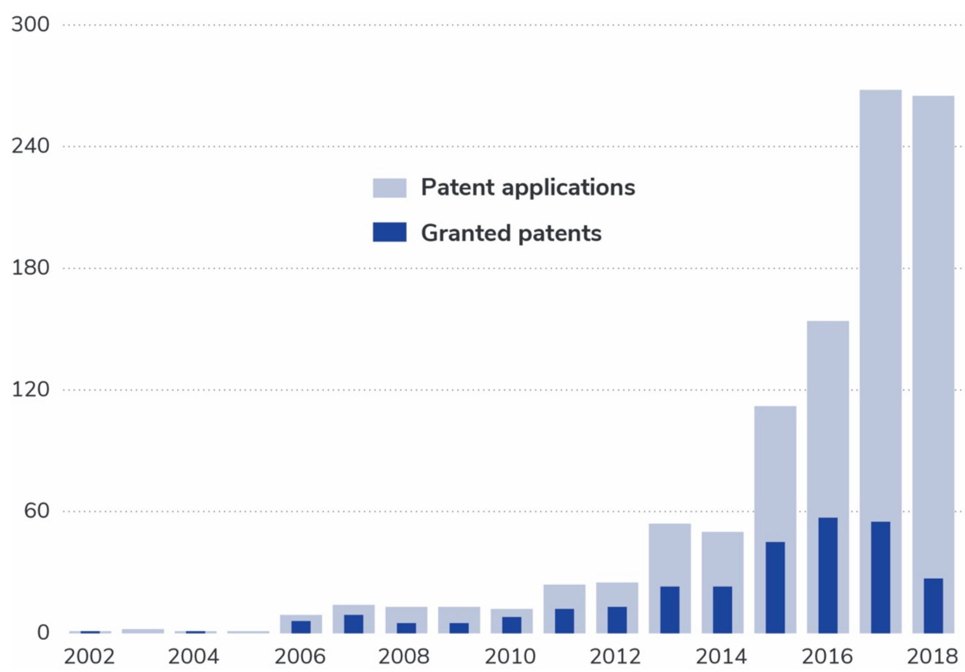
		NUMBER OF AI PATENT FAMILIES
1.	China	78,980
2.	United States	41,092
3.	Japan	9,123
4.	South Korea	7,322
5.	Germany	1,505
6.	United Kingdom	1,094
7.	Taiwan	1,089
8.	**India**	**1,037**
9.	Russia	885
10.	France	631

Source: CSET worldwide AI-relevant patents database, as of January 29, 2021.

Despite the interesting results, India's position behind Taiwan, the United Kingdom, Germany, South Korea, and Japan, is significant if seen from the perspective of patents per unit of GDP or patents relative to population, where India's rank slips further down. If patents are an output of research, then India clearly underperforms on this output especially when contrasted with the country's voluminous AI research production capacity. Still, India's underwhelming performance in the realm of AI patenting should be understood in the context of the confounding factors mentioned above, such as lack of patent litigation and IP protections, combined with the understanding that while patents cost money to file, scholarly papers are often free or low-cost to submit.

Even though India's performance in AI patents is not as impressive as its AI research, patent production in India has witnessed significant growth since 2012.[59] Tata Sons Ltd. is the top patenting assignee and most of the large players are big companies. Figure 1 shows the number of Indian AI patent applications and patent grants by year from 2002 to 2018, indicating a sharp increase in AI patent activity in recent years. In India, there were 10 times as many AI-relevant patent applications in 2018 as in 2012. This trend aligns with global AI patent production activity, which saw roughly similar growth in this time period.[60] This AI patent boom post 2012 is commonly associated with increased connectedness, more data, and greater computing power that brought new breakthroughs, including the rise of cutting-edge AI algorithms like deep learning.[61]

Figure 1: Number of Indian Patent Grants and Patent Applications, 2002–18.

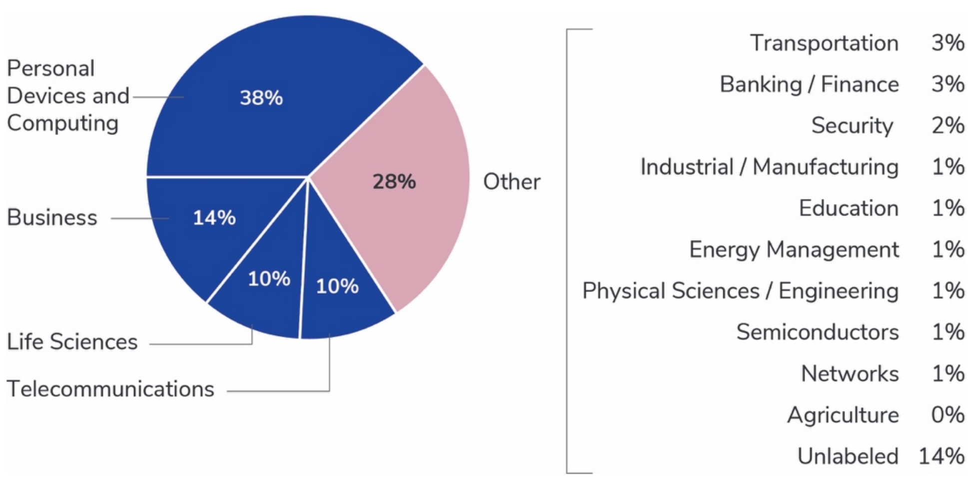


Source: CSET worldwide AI-relevant patents database, as of January, 15 2021.

The question arises as to whether this rapid growth in Indian patenting reflects increased inventive activity, or results from other factors. This context is clarified by the divergence in the number of patent applications versus patent grants 2014 onwards. The sharp rise in patent applications underlines that a great deal of AI patent activity was driven by a rapid increase in AI inventions versus, for instance, clearing of backlog applications due to administrative progress. The slight dip in granted AI patents from 2017 to 2018 suggests either an increase in denials or an increase in the amount of time taken to process patent applications. The latter is more plausible considering that under the Indian patent system applicants have 48 months from application date (or priority date, whichever is earlier) to request examination of their application.[62] Hence, it is likely that the spike in 2017-18 applications may still be working its way through the Indian patent system.[63]

Consequently, the number of Indian AI patent grants are likely to grow in the coming years. Patent litigation, at least of a significant scale, is a relatively new phenomenon in India. If there is little history of litigating or enforcing patents, then companies are less likely to spend the money to file them.[64] As litigation becomes more established, the expenditure becomes more justified. Therefore, it is probable that the recent trend of increasing AI patent applications in India is also going to stay.[65]

Figure 2: AI patent families in India by application field, 2002–19.[66]



Source: CSET worldwide AI-relevant patents database, as of January 15, 2021.

Figure 2 shows the distribution of patent families across AI application fields.[67] The top application field for India's AI patents was personal devices and computing, which was associated with 568 patent families and includes subcategories like personal computers and PC applications, and affective computing or technology that enables computers to better recognize human emotions. Together, the four categories of personal devices and computing, business, telecommunications, and life sciences were associated with over 70 percent of all Indian AI patent families.[68] This large proportion is not surprising given that India is regarded as a global hub for medicine, telecommunications, customer relations, and e-commerce.

In comparison to global AI patenting trends, what stands out for India is the higher proportion of patents in business, with the application field featuring in 14 percent of India's AI patent families in comparison to about 10 percent worldwide. Another notable point is the absence of significant patents associated with transportation as a field in India. Transportation, with sub-categories like autonomous vehicles, transportation and traffic engineering, vehicle recognition, and avionics, is the most popular patent application field globally, representing 24 percent of all AI patent families.[69] In comparison, only 3 percent of India's patent filings were associated with classifications mapped to transportation as a field, accounting for a total of 43 patents overall.

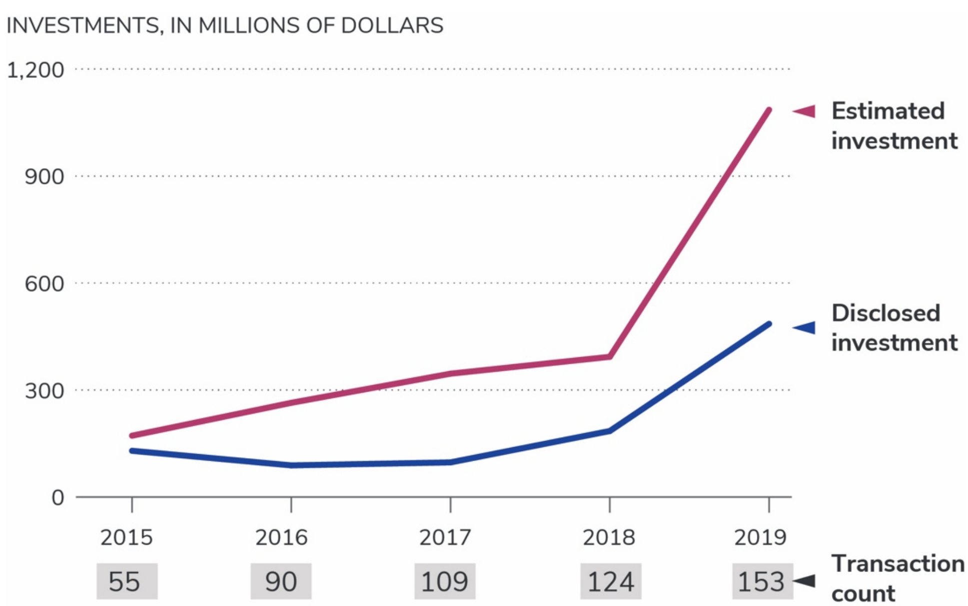
Taken together, despite bureaucratic barriers and poor protections for intellectual property, Indian patent applications have seen massive growth since 2012, which is likely to sustain. Still, the number of granted patents remains relatively low. Among all Indian AI patents, the majority are in fields like personal devices and computing, business, life sciences, and telecom, which indicates that Indian innovators have focused on applying AI to areas of traditional strength. India has come a long way in AI patenting since 2002, as its patent system has been updated and companies have started to see the benefit of using patents to protect their innovations. That said, India still has a long road ahead to catch up with China and the United States, currently the dominant sources of AI patents.

AI Companies and Investments

This paper estimates AI investment by measuring equity investments into privately held AI companies, or AI-focused companies that are not traded on a stock exchange.[70] The analysis below relies on Crunchbase data drawn from CSET's investment database containing information on companies, venture capital funding rounds, and other financial data from 2015 to 2019.[71] In total, we identified 361 privately-owned AI companies in India.

Investment in AI has increased greatly in the past five years internationally, and India is a good example of a country that has witnessed recent growth.[72] Figure 3 presents estimates of equity investment (defined as venture capital, private equity, and mergers and acquisitions) attracted by Indian AI companies from 2015 to 2019.

Figure 3: Investments in privately-held Indian AI companies, 2015–19.

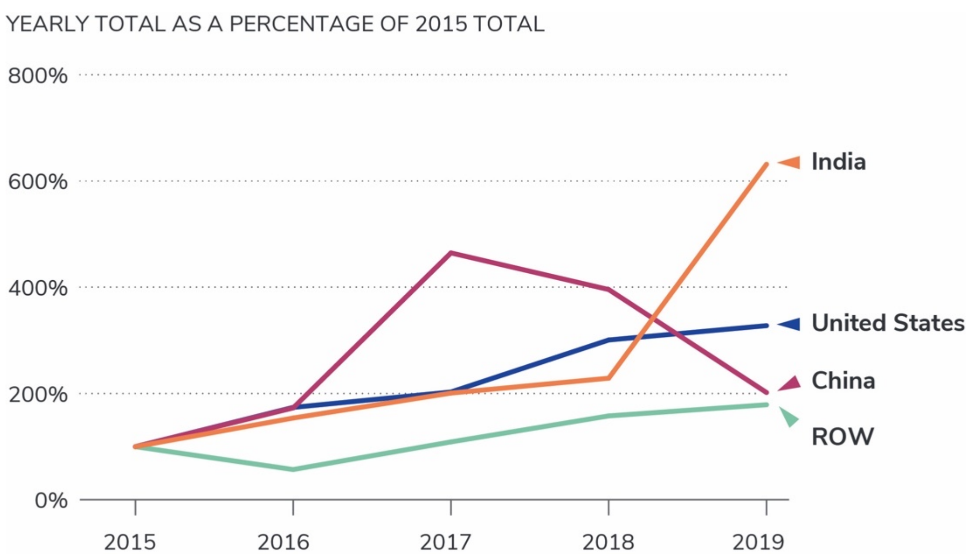


Source: CSET analysis of Crunchbase data, as of January 15, 2021.

In many AI investment transactions, the investment amount is not publicly revealed. Consequently, Figure 3 attempts to give a clearer view of the market as a whole by presenting the data as a) dollar values of investment disclosed publicly (disclosed investment), b) estimates based on the deals where the amounts are not revealed publicly (estimated investment), and c) number of AI investment transactions (transaction counts).[73] For comparison, corresponding AI investments in 2019 in the United States totaled $25 billion, disclosed, and $47 billion, estimated. In China, corresponding AI investments in the same year totaled $5 billion, disclosed, and $7 billion, estimated.

On all three fronts—disclosed investment, estimated investment, and transaction counts—India's AI investment figures appear small when compared to the U.S. and Chinese figures. However, investment activity in India has witnessed steady growth until 2018, after which it nearly tripled in dollar value. This jump in growth is depicted in Figure 4, which tracks estimated investment growth in dollars using 2015 numbers as a baseline.

Figure 4: Normalized growth in estimated AI investment transactions relative to 2015 baseline, by investment target.



Source: CSET analysis of Crunchbase data, as of January 15, 2021.

The figure illustrates India's steady growth in estimated investment in AI companies since 2015, and the steep rise of more than 400 percent from 2018 to 2019.[74] This was mostly driven by large mergers and acquisitions transactions in 2019 like Reliance Industries' acquisition of Haptik—a conversational AI platform focusing on customer engagement.[75]

Of all the investments Indian AI companies receive, U.S.-based and China-based investors are some of the most important players. To compare their activity, we calculated the number and value of AI company investments with at least one publicly disclosed Chinese, U.S., or Indian investor participating, results are presented in Table 8.[76]

Table 8: Investments into Indian AI companies, 2015–19.

		IN MILLIONS OF DOLLARS	
INVESTORS	TRANSACTIONS	Disclosed	Estimated
India	308	$647	$1,188
United States	140	$455	$858
China	20	$132	$159

Source: CSET analysis of Crunchbase data, as of January 15, 2021.

This analysis shows that from 2015 to 2019, U.S.-based investors were involved in seven times more AI investment transactions in India and invested roughly three to five times more disclosed and estimated investments when compared to their Chinese counterparts. Indian investors are evidently the biggest players in the market on all fronts, but the relatively small difference between U.S. and Indian investment figures compared with the stark gap between U.S. and Chinese investment figures, underlines the United States' presence as a major player in India.[77] Among other factors, this could be due to the larger network effects of Indian diaspora in Silicon Valley spurring private investment in the Indian market.

Indian investors were involved in more than double the transactions in comparison to their U.S. counterparts but the disclosed dollar amounts invested were relatively smaller. For estimated investments in Indian AI companies, the average size of a round involving Indian venture capitalists was about $3.8 million in comparison to $6.1 million for U.S. investors. This tendency to write smaller check sizes becomes slightly more clear when we look at the same data slicing transaction counts by investment stage (see Table B in the Appendix for details).

We learn that Indian investors tend to invest more in early-stage rounds (seed, pre-seed, angel) in comparison to U.S. or Chinese investors. This finding is not surprising given that seed investors tend to invest locally. In comparison, Chinese investors are more likely to participate in intermediate to late-stage venture financings (Series A, Series B, Series C) of Indian AI companies as compared to U.S. investors that engage comparatively lesser in intermediate rounds (Series B).

Finally, Table 9 gives an overview of the types of Indian AI-focused companies identified on Crunchbase, listing the counts of companies and shares of disclosed investment value under their primary application area.[78] It also presents the amount of disclosed investment value for each area as a share of the total dollar value invested.

Table 9: Distribution of AI companies in India by application area and investments.

	AI companies	Share of AI companies	Share of disclosed investment value
Business services and analytics	70	19.4%	25.9%
Sales, retail, and customer relations	64	17.7%	15.4%
General purpose	64	17.7%	4.2%
Finance	27	7.5%	12.5%
Medicine and life sciences	25	6.9%	8.1%
Arts, sports, leisure, travel, and lifestyle	24	6.7%	6.4%
Education	16	4.4%	8.8%
Process automation	12	3.3%	2.6%
Security and biometrics	11	3.1%	0.2%
Diversified/NOS/Unclear	11	3.1%	0.2%
Transportation	10	2.8%	8.1%
Agriculture	10	2.8%	3.0%
Utilities	5	1.4%	0.2%
Consumer goods	4	1.1%	0.2%
Broadcasting and media production	4	1.1%	0.8%
Military, public safety, and government	2	0.6%	2.6%
Construction and field services	2	0.6%	0.9%

Source: CSET analysis of Crunchbase data, as of January 15, 2021.

Consistent with our findings for the patents data, roughly 19 percent of Indian AI companies are engaged in applications related to business services and analytics, and nearly 18 percent of them are active in sales, retail, and customer relations, making these two categories the primary application areas for AI companies in India. By disclosed value, companies active in business services attracted a greater proportion of equity investments in comparison to all other fields. Taken together, business services and analytics, finance, medicine, and sales, retail, and customer relations account for over 50 percent of all Indian AI companies and attract over 60 percent of all equity funding given to Indian AI companies. Other smaller sectors attracting a larger proportion of disclosed India-bound funding relative to their size include transportation, finance, and education, which could be indicative of their popularity among investors or higher capital requirements. Military, public safety, and government applications, along with security and biometrics account for a very small proportion of disclosed investments in India. Our findings corroborate the idea that profit is the primary driver of AI development in India, and a vast majority of equity investments going into privately held companies are not focused on government needs.

Looking at companies and investments, India as the third largest startup ecosystem globally attracts lesser investments when compared to the United States and China. However, there has clearly been a significant growth in both domestic Indian investment and foreign investment flows into privately-held Indian AI companies, particularly after 2018. Among foreign investors in the Indian market, U.S. investors have been in the lead, especially when compared with their Chinese counterparts. Among the Indian AI companies attracting most of this growing investment are those in sectors like business analytics, sales, retail, and customer relations, which reflect India's known comparative strengths.

Cloud Computing

Cross-country comparisons for computing capacity are hard to measure, especially for countries that do not manufacture semiconductors. However, market cloud computing spending figures offer useful proxies to estimate a country's capacity to support the computing needs of AI.[79] The data for this section draws from industry reports on cloud computing adoption in India as presented by the National Association of Software and Services Companies (NASSCOM) and Gartner. The figures below rely on information as presented by these two sources, and only provides limited insights into the methods used to collect and analyze the data.

In recent years, significant increases and growing availability of computing power have fueled breakthroughs in AI, especially ML-based systems. As AI systems become more sophisticated, they require specialized AI computer chips. These chips are specifically designed to accelerate training, reduce power consumption, and cost-effectively store and process petabytes of data. At present, India does not have the domestic capacity to produce these specialized chips and associated computing infrastructure. India does have a compute infrastructure enabled by high-performance computing, but these existing supercomputing facilities compare poorly to those in other countries: the November 2020 top 500 supercomputers list is dominated by China with 215 supercomputers, the United States with 113, and India with only 3 supercomputers, ranking 62nd, 77th, and 144th on the list of the top 500.[80] Moreover, these supercomputing facilities cater to small scale R&D work and are not designed for an AI workload especially for a larger ecosystem of startups and other institutions. The Indian government's existing cloud computing environment, MeghRaj, is also designed for cloud services with a central processing unit based underlying architecture that cannot be upgraded to add graphics processing unit nodes, keeping its capabilities modest for cutting-edge AI solutions.[81]

As a result, many companies in India rely on public cloud computing, like computing power available on-demand from vendors in the private sector like Amazon Web Services (AWS), Google Cloud, and Microsoft Azure.[82] Table 10 presents the current state of cloud penetration across different countries.[83]

Table 10: Global IT spending and public cloud spending, 2018.[84]

	SPENDING, IN BILLIONS OF USD				
	GDP	Total IT	Public cloud	Total IT spending as % of GDP spending	Cloud spending as % of total IT spending
United Kingdom	$2,622	$137	$15.6	5.2%	11.4%
United States	$19,391	$911	$103.9	4.7%	11.4%
Canada	$1,653	$63	$7.1	3.8%	11.3%
Australia	$1,323	$48	$3.7	3.6%	7.7%
World	**$80,000**	**$2,362**	**$186.6**	**3.0%**	**7.9%**
Brazil	$2,056	$40	$3.2	1.9%	7.9%
India	**$2,597**	**$42**	**$2.5**	**1.6%**	**6.0%**
China	$12,238	$172	$4.6	1.4%	2.7%
Russia	$1,578	$15	$0.4	1.0%	2.9%

Source: NASSCOM cloud penetration across countries.

India's cloud adoption is still at a nascent stage in comparison to the world average. Data from NASSCOM tells us that in 2018, business enterprises in India spent $42 billion, or 1.6 percent of the GDP, on their IT budget.[85] This is roughly half that of the world average of 3 percent. Of the total amount spent on IT, public cloud spending stood at 6 percent, again lagging behind the global average of 7.9 percent. India's cloud spending totaled $2.5 billion in 2018. This amount is estimated to be more than Russia's expenditures, but lower in comparison to many other technologically advanced countries. It is roughly half of China's cloud spending estimate, one-fifth of the United Kingdom's figure, and only 2.4 percent of the United States, which spent $103.8 billion on cloud.

While the cloud computing market in India is at its early stages of adoption, it is growing very fast. Between 2016 and 2018, the growth rate for cloud spending in India was at 40.2 percent on average, making it the second fastest growing cloud computing market at the time (with China being the first).[86] The cloud computing adoption rate is expected to continue to accelerate further due to a projected increase in demand and increased capacity.[87]

As per forecasts presented by NASSCOM in 2019, India's cloud market was projected to grow at a 30 percent compound annual growth rate, reaching three times its 2018 figures by 2022.[88] NASSCOM projected the country's cloud spending to grow from $2.5 billion to $3.9 billion in 2020 per its conservative estimates, $4.4 per base estimates, and $5.1 per aggressive estimates. Indian cloud spending was roughly $3.16 billion in 2020, missing even the conservative estimates.[89] By 2022, these numbers were forecasted to be around $5.6 billion conservatively, $7.1 billion per base estimates, and $9.3 in an aggressive case. With the 2020 slowdown, the $7 billion market forecast for India is likely to take longer than 2022, which keeps it from matching the world average in the short term, let alone emulating developing economies in cloud computing.[90]

Among the AI development levers we reviewed in this paper, India's capabilities in cloud computing are the least advanced. So far, cloud adoption in India is at a nascent stage with its figures on cloud spending as a share of IT spending falling below the world average. In recent years, India's cloud computing demands have grown significantly. It fell short of meeting conservative forecasts for 2020—indicating factors that could be slowing it down.[91] Nevertheless, based on figures of existing forecasts, we concur that India's cloud computing demands will continue to grow in the double-digits in the near future. In order to make progress in cloud computing, the Indian government should emphasize the various infrastructural and operational factors that could drive cloud growth further in the country, including reliable power, sustainable land regulations, high-speed connectivity, among others.[92] Additionally, a pragmatic and cost-effective path forward for India is building a domestic centralized AI-specific cloud infrastructure facility to facilitate the work of a broader spectrum of stakeholders in the Indian AI ecosystem. Notably the Indian AI strategy is moving in this direction.[93]

Conclusion

India has the potential to position itself as an important player and a valuable partner in the AI ecosystem, but serious obstacles remain.

- India has a large pool of potential AI talent but weaknesses in its higher education sector limit its ability to train a highly skilled AI workforce at scale. India produces a very small number of doctoral students with most of them going overseas, especially to the United States.

- It is the fourth largest producer of AI research but has relatively fewer international collaborations, which may have implications for research quality and impact. Established AI research fields in India like speech and pattern recognition offer opportunities for Indian researchers to increase their outreach, especially with their counterparts in the United States with whom they have the strongest research ties internationally.

- India has witnessed a high growth in AI patent applications since 2012, but without administrative improvements in granting patents and increased IP protections, it will not be able to ride the AI patent growth wave.

- Investments feeding the vast network of India's privately held AI companies is small in comparison to China and the U.S., but has grown significantly and the fastest globally between 2018–19, with the United States as the primary foreign investor. Sectors like business analytics, life sciences, sales, retail and customer relations reflect India's known comparative strengths among its AI companies.

- India lags in the capacity to support its cloud computing needs for AI, and has a rapidly growing domestic cloud market with rising demands for computing capacity.

India-U.S. Cooperation on AI and the Way Forward

Alignment between the United States and India on AI is clear and the potential to build on an already strong relationship is enormous. An assessment of the United States' and India's national AI strategies reveals a shared commitment for a safe and reliable deployment of AI. There is a growing consensus in the two countries to showcase a democratic way of emerging technologies—one that promotes developing AI tools ethically and responsibly. This will lay ground for the two like-minded partners to defend values like liberty, equality, and justice by providing competitive alternatives to counter the export of censorship and surveillance technologies attempting to develop a global authoritarian AI ecosystem. The AI sector in India and the United States is interlinked and congruous and the two countries should work with each other given their talent interlinkages and well-meshed IT sector, high academic AI research collaborations, and existing investor ties.

While countries across the globe race to become AI leaders competing to control the best expertise, flow of talent and information is uniquely beneficial to both India and the United States and is more an avenue of cooperation than conflict. The inability to nurture India's large talent pool by providing it with the right environment will be a loss to India, the United States, and the international AI community. Talent and compute issues faced by India offer opportunities for the United States to strengthen India as an AI partner, by exploring ways to help Indian innovators acquire greater access to computing capacity and vitalize its higher education sector.

A successful partnership between India and the United States will rely on their ability to further bolster international research collaborations across the various AI fields. Progress is already underway on this front through efforts of organizations like the IUSSTF that recently launched a U.S.-India Artificial Intelligence (USIAI) initiative to promote bilateral AI R&D collaboration through workshops and networking, among others.[94] The initiative could serve as a good platform to address the major issues faced by scholars from both countries seeking international collaboration in order to further strengthen partnerships.[95]

Additionally, IP is critical for sustained economic growth and further advancement of AI. India should continue to build upon existing legislative progress in IP protection and take steps to further improve its infrastructure to address judicial delays. This will enable AI innovators to reap the financial rewards resulting from their creativity and also create a welcoming environment for businesses and potential investors. Additionally, India needs to also work towards improving its regulatory apparatus and issues pertaining enforcement of contracts to build more confidence amongst foreign investors who continue to be major players in its private market.

Any successful AI partnership between India and the United States will also require them to overcome several other challenges, including differing viewpoints on issues covering data privacy and storage rules.[96] The process will be more productive if a diverse range of stakeholders that seek to balance profit, innovation, and social welfare, are involved. Bilateral India-U.S. dialogues on AI can lay the foundation for future discussions in multilateral forums and can potentially build a global consensus by bridging the divide in thinking across the developing and developed economies.

Authors

Husanjot Chahal is a research analyst with CSET, where Sara Abdulla is a data research analyst, Jonathan Murdick is a research intern, and Ilya Rahkovsky is a data scientist.

Acknowledgments

The authors are grateful to Dewey Murdick, Igor Mikolic-Torreira, and Margarita Konaev for critical support throughout the research and writing process. Thanks to C. Raymond Perrault at SRI International, Trisha Ray at Observer Research Foundation (ORF), New Delhi, Andreas Kuehn and Michael Depp at ORF America, and Saurabh Mishra at Stanford University's Institute for Human-Centered Artificial Intelligence, for reviewing the report and providing feedback and suggestions. We thank Zachary Arnold and Patrick Thomas for their guidance, and Andrew Imbrie, Catherine Aiken, Melissa Flagg, Remco Zwetsloot, and Saif Khan for their valuable insights and comments. Many thanks to Farhana Hossain for help with data visualization, and Shelton Fitch, Alex Friedland, Lynne Weil, Matt Mahoney, and Melissa Deng for their editorial support.

Document Identifier: doi: 10.51593/20200096

Appendix

This appendix presents additional charts and tables of relevance to this topic.

In reference to the research section, Table A presents the top academic entities in the United States and India collaborating on AI research. The Indian Institute of Technology in India (23 institutes grouped together) and Carnegie Mellon University emerge as the top collaborating pair.

Table A: Top 10 academic India-U.S. collaborating entities on AI research

	Indian Entity	Papers	U.S. Entity
1.	Indian Institute of Technology	74	Carnegie
2.	TCS Innovation Labs	69	Harvard
3.	Indian Institute of Technology	66	University of Southern California
4.	Indian Institute of Technology	60	University of California System
5.	Indian Institute of Technology	58	Harvard
6.	National Institute of Science and Technology	49	National Institute of Standards and Technology
7.	Tata Consultancy Services	49	Harvard
8.	Indian Institute of Technology	42	University of Texas System
9.	Indian Institute of Technology	39	Johns Hopkins University
10.	Indian Institute of Technology	39	MIT

Source: CSET merged corpus of scholarly literature, as of January 15, 2021.

With regard to the analysis under the Companies and Investment section, Table B lays out the transaction counts for India-bound investments into privately held AI companies by Indian investors, U.S. investors, and Chinese investors, presented by investment stage. We learn that Chinese investors are more likely to participate in intermediate-stage venture financings (Series A, Series B) as compared to its Indian and American counterparts that invest more in early rounds (seed, pre-seed, angel).

Table B: Percent count of India-bound investment by investment stage (2015–19).

	Indian investors	U.S. investors	Chinese investors
Seed, pre-seed, angel	64%	55%	53%
Series A	14%	20%	21%
Series B	4%	5%	21%
Series C	2%	2%	5%
Series D+	0%	1%	0%
Other/Unknown	16%	18%	0%

Source: CSET analysis of Crunchbase data, as of January 15, 2021.

Table C illustrates the top five countries of origin for inbound international students in the United States in 2016. The share of Chinese students is the highest followed by India, both of which respectively comprised 34.5 percent and 14.1 percent of the international student body in the United States for the period. Together Chinese and Indian students represented nearly half of all international students in the United States in 2016.

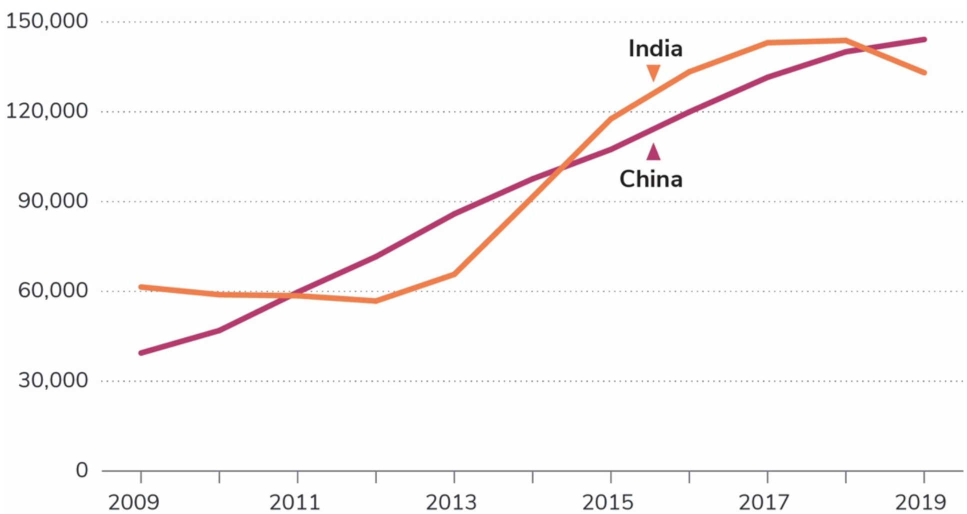
Table C: Top 5 countries of origin for international students in the United States, 2016.

		SHARE OF INTERNATIONAL STUDENTS
1.	China	34.5%
2.	**India**	**14.1%**
3.	South Korea	5.4%
4.	Saudi Arabia	4.9%
5.	Canada	2.7%

Source: UNESCO UIS, Global flow of tertiary-level students.

Figure D presents the total number of math, computer science, and engineering students in the United States coming from India and China for bachelor's, master's, and doctoral degrees between 2009 and 2019. Chinese students in this field have witnessed roughly a steady growth in comparison to Indian students that dipped in numbers between 2010 and 2012, then witnessed a high growth until 2017, after which the numbers dipped again.

Figure D: Math/Computer Science & Engineering students in the United States from India and China, 2009–19.



Source: Institute of International Education, Open Doors Report on International Educational Exchange.

As Figure D depicts, there is a greater proportion of Indian students studying in the United States are in STEM fields in comparison to their Chinese counterparts. This can be seen when we compare Figure C and Figure D. Despite a difference of almost 20 percentage points between shares of Indian and Chinese students in the United States, when we look at their breakdown under math, computer science, and engineering in Figure D, we see their numbers are nearly tied with surges in Indian STEM students outpacing Chinese students in some years. Indian math, computer science, and engineering students in the United States were almost twice the amount of their Chinese counterparts in 2009. In fact, 2009 onwards, Indian STEM students surpassed Chinese STEM students for six years out of 10, specifically 2009–10 and 2015–18.

The total number of international Chinese students has always been higher because in comparison to Indian students, there are many more Chinese students in the United States pursuing degrees in physical and life sciences, and slightly more non-STEM fields like business.

Endnotes

[1] Along with an advantage of a demographic dividend, India also has a youth bulge in comparison to several advanced countries where the working age population is declining.

[2] "National Strategy for Artificial Intelligence #AIForAll" (NITI Aayog, June 2018), https://niti.gov.in/sites/default/files/2019-01/NationalStrategy-for-AI-Discussion-Paper.pdf.

[3] This metric is taken from HAI Stanford's AI Index 2019 report. HAI Stanford laid out its assessment after evaluating skill data drawn from member profiles of professionals on LinkedIn. One reason why India ranks so high on this metric is maybe because India just lacks people with high-tech skills that are not AI. Raymond Perrault, Yoav Shoham, Erik Brynjolfsson, Jack Clark, John Etchemendy, Barbara Grosz, Terah Lyons, James Manyika, Saurabh Mishra, and Juan Carlos Niebles, "The AI Index 2019 Annual Report" (Human-Centered AI Institute, Stanford University, December 2019), 78, https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf.

[4] Raymond Perrault et al., "The AI Index 2019 Annual Report," 73.

[5] This statistic is confirmed by the authors' assessment of Crunchbase data. Department for Promotion of Industry and Internal Trade, "Indian Startup Ecosystem," Ministry of Commerce and Industry, https://www.startupindia.gov.in/content/sih/en/international/go-to-market-guide/indian-startup-ecosystem.html ; "Start-up India - Momentous Rise of the Indian Start-up Ecosystem," (NASSCOM, 2015), https://nasscom.in/knowledge-center/publications/start-report-momentous-rise-indian-start-ecosystem; Gerard J. Tellis, "2016 Startup Index of Nations, Cities: (Startups worth $1 billion or more: "Unicorns")" (University of Southern California, Marshall School of Business), https://www.marshall.usc.edu/sites/default/files/Unicorn-Index-Report-GT17.pdf.

[6] "National Strategy for Artificial Intelligence #AIForAll."

[7] "AIRAWAT – Establishing an AI-specific cloud computing infrastructure for India" (NITI Aayog, January 2020), https://niti.gov.in/sites/default/files/2020-01/AIRAWAT_Approach_Paper.pdf.

[8] This strategy document was launched by the National Institution for Transforming India, also called NITI Aayog—the Indian government's think tank providing both directional and policy inputs. "National Strategy for Artificial Intelligence #AIForAll."

Recent examples of the Indian government's partnership with tech players to implement AI project include the launching of an AI-driven virtual assistant across online platforms to provide customized and accurate information on the COVID-19 pandemic by India's Ministry of Electronics and Information Technology in partnership with Accenture and Microsoft. In 2018, the government of Karnataka collaborated with Intel to use AI-based technology solutions for dealing with traffic issues and preventing road accidents.

[9] Ministry of Electronics and Information Technology, *Office Memorandum* (New Delhi: Government of India, February 7, 2018), https://www.meity.gov.in/writereaddata/files/constitution_of_four_committees_on_artificial_intelligence.pdf.

[10] Ministry of Electronics and Information Technology, *Artificial Intelligence Committee Reports* (New Delhi: Government of India) https://www.meity.gov.in/artificial-intelligence-committees-reports.

[11] "AI in India," OECD AI Policy Observatory, https://oecd.ai/dashboards/countries/India.

[12] *Interim Report and Third Quarter Recommendations* (Washington, D.C.: National Security Commission on Artificial Intelligence, October 2020), https://www.nscai.gov/wp-content/uploads/2021/01/NSCAI-Interim-Report-and-Third-Quarter-Recommendations.pdf.

[13] Jeffrey Ding, "Deciphering China's AI Dream" (Centre for the Governance of AI, March 2018), http://www.fhi.ox.ac.uk/wp-content/uploads/Deciphering_Chinas_AI-Dream.pdf; Andrew Imbrie, Elsa B. Kania, and Lorand Laskai, "The Question of Comparative Advantage in Artificial Intelligence: Enduring Strengths and Emerging Challenges for the United States" (Center for Security and Emerging Technology, January 2020), https://cset.georgetown.edu/wp-content/uploads/CSET-The-Question-of-Comparative-Advantage-in-Artificial-Intelligence-1.pdf; Daniel Castro, Michael McLaughlin, and Eline Chivot, "Who is Winning the AI Race: China, the EU or the United States?" (Center for Data Innovation, August 19, 2019), https://datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/.

[14] Husanjot Chahal, Ryan Fedasiuk, and Carrick Flynn, "Messier than Oil: Assessing Data Advantage in Military AI" (Center for Security and Emerging Technology, July 2020), https://cset.georgetown.edu/research/messier-than-oil-assessing-data-advantage-in-military-ai/.

[15] Patents are a noisy measure of innovation, as some companies do not file patents believing that secrecy offers a better protection of intellectual property, while others file numerous weak or even troll patents with questionable product potential.

[16] We do not present data for bachelor's students because of measurement challenges we faced with Chinese data. China reports its master's and doctoral graduates by academic field where the discipline is listed as "Engineering." Its numbers on undergraduates are presented by the "type of courses," where engineering as a category is missing and the only relevant category is lumped together as "Natural Sciences & Tech." Consequently we decided to not present data on bachelor's students in our main findings.

[17] Remco Zwetsloot, Roxanne Heston, and Zachary Arnold, "Strengthening the U.S. AI Workforce" (Center for Security and Emerging Technology, September 2019), https://cset.georgetown.edu/wp-content/uploads/CSET_U.S._AI_Workforce.pdf.

[18] For instance, China's Ministry of Education does not publicly report graduates in computer science or does not subcategorize engineering degrees. In India, apart from having "computer engineering" as a sub-field of engineering, "computer science" as a discipline exists separately under the field of "IT & computer." Even though the United States provides a breakdown of engineering degrees, unlike India, it does not have a separate category of "computer engineering" as reported by the National Center for Education Statistics or the National Science Foundation.

[19] The percentages do not add up to 100, because this table does not list the 51 MPhil graduates under computer engineering, and 10288 MPhil graduates under IT & Computer.

[20] Only 34.9 percent of all universities run graduate programs and just 2.5 percent run PhD programs. Yojana Sharma, "India in 'initial stages' of higher education massification – Report," *University World News*, November 28, 2019, https://www.universityworldnews.com/post.php?story=20191128104421724; Philip G. Altbach, "India's higher education challenges," *Asia Pacific Educ. Review* 15 (December 2014): 503-510, https://link.springer.com/article/10.1007/s12564-014-9335-8; Shamika Ravi, Neelanjana Gupta, and Puneeth Nagaraj, "Reviving Higher Education in India" (Brookings India, November 2019), https://www.brookings.edu/wp-content/uploads/2019/11/Reviving-Higher-Education-in-India-email.pdf.

[21] Ravi, Gupta, and Nagaraj, "Reviving Higher Education in India," 16.

[22] "Students from a given country studying abroad (outbound mobile students)," United Nations Data, http://data.un.org/Data.aspx?d=UNESCO&f=series%3AED_FSOABS.

[23] Figures as estimated and presented in Ravi, Gupta, and Nagaraj, "Reviving Higher Education in India," 16.

[24] In 2017, international students in Canada grew at nearly 20 percent in comparison to the United States where growth rates halved to 3.5 percent from previous numbers. MMA Services, "Indian Students Mobility Report" (M.M. Advisory, 2018), http://mdotm.in/wp-content/uploads/2019/03/Indian-Students-Mobility-Report-2018.pdf.

[25] Remco Zwetsloot, Jacob Feldgoise, and James Dunham, "Trends in U.S. Intention-to-Stay Rates of International Ph.D. Graduates Across Nationality and STEM Fields" (Center for Security and Emerging Technology, April 2020), https://cset.georgetown.edu/research/trends-in-u-s-intention-to-stay-rates-of-international-ph-d-graduates-across-nationality-and-stem-fields/.

[26] The total number of engineering PhDs in India were 2785 in 2015, 3366 in 2016, and 4907 in 2017, which averages to 3686 across 2015-17 (AISHE reports).

[27] U.S. Citizenship and Immigration Services, *Characteristics of H-1B Specialty Occupation Workers* (Washington, DC: Department of Homeland Security, March 5, 2020), https://www.uscis.gov/sites/default/files/document/reports/Characteristics_of_Specialty_Occupation_Workers_H-1B_Fiscal_Year_2019.pdf.

[28] Department of Science and Technology, *S&T Indicators Tables* (New Delhi: Ministry of Science & Technology, March 2020), https://dst.gov.in/sites/default/files/S%26T%20Indicators%20Tables%202019-20.pdf.

[29] Zwetsloot, Feldgoise, and Dunham, "Trends in U.S. Intention-to-Stay Rates"; Remco Zwetsloot, James Dunham, Zachary Arnold, and Tina Huang, "Keeping Top AI Talent in the United States" (Center for Security and Emerging Technology, December 2019), https://cset.georgetown.edu/wp-content/uploads/Keeping-Top-AI-Talent-in-the-United-States.pdf.

[30] Tina Huang, Zachary Arnold, and Remco Zwetsloot, "Most of America's "Most Promising" AI Startups Have Immigrant Founders" (Center for Security and Emerging Technology, October 2020), https://cset.georgetown.edu/wp-content/uploads/CSET-Most-of-Americas-Most-Promising-AI-Startups-Have-Immigrant-Founders.pdf.

[31] As per reports this was said by Abid Hussain, former Planning Commission member and Indian Ambassador to the United States, but has been reused several times after that. S. Murlidharan, "Brain drain better than brain in the drain," *The Hindu Business Line*, October 23, 2012, https://www.thehindubusinessline.com/opinion/columns/s-murlidharan/brain-drain-better-than-brain-in-the-drain/article22985300.ece.

[32] Rupa Chanda and Niranjana Sreenivasan, "India's experience with skilled migration," in *Competing for Global Talent*, editors Christine Kuptsch and Pang Eng Fong, *International Institute for Labour Studies*, 2006, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.466.5123&rep=rep1&type=pdf#page=224.

[33] "US visa policy that supports Indian talent is win-win for both: Jaishankar," *Business Standard*, September 1, 2020, https://www.business-standard.com/article/current-affairs/us-visa-policy-that-supports-indian-talent-is-win-win-for-both-jaishankar-120090100109_1.html.

[34] For more information on how an AI-relevant paper is defined, please see James Dunham, Jennifer Melot, and Dewey Murdick, "Identifying the development and application of Artificial Intelligence in scientific text," *arXiv [cs.DL]*, arXiv, May 28, 2020, https://arxiv.org/pdf/2002.07143.pdf.

[35] We attribute papers to a country if one or more of the authors' organizational (universities, companies, government, etc.) affiliation is in that country. This means that a paper with multiple country affiliations will figure in the counts of multiple countries. Therefore, all of the research papers linked with one country should not be seen as exclusive for that country.

[36] An author's affiliation to a country is seen through the author's organizational affiliation as listed on the paper published.

[37] The figures represent total spending on R&D in PPP$. Mark Boroush, "Research and Development: U.S. Trends and International Comparisons," National Science Board, January 15, 2020, https://ncses.nsf.gov/pubs/nsb20203/cross-national-comparisons-of-r-d-performance.

[38] "How much does your country invest in R&D?," UNESCO Institute for Statistics, http://uis.unesco.org/apps/visualisations/research-and-development-spending/.

[39] A counterargument to this could question the possibility of India focusing more of its R&D investments in AI. However, this is unlikely for several reasons. To begin with, most Indian government reports on AI recently have focused their recommendations on increasing research mechanisms and associated funding. NITI Aayog's "National Strategy for AI" report, for instance, specifically proposed setting up research centers to boost both core and applied research in AI. Such a focus is not indicative of a well-resourced research environment. Secondly, India's ranking in scientific publication output in general is third globally as per the National Science Foundation in United States, which makes it unlikely that the AI sector is given a special R&D boost to give it a higher rank versus other sectors.

[40] This is based on CSET's merged corpus of scholarly literature, as of January 21, 2021.

[41] Kirstin R.W. Matthews, Erin Yang, Steven W. Lewis, Brandon R. Vaidyanathan, and Monica Gorman, "International scientific collaborative activities and barriers to them in eight societies," *Accountability in Research* 27, no. 8 (June 2020): 477-495, https://www.tandfonline.com/doi/full/10.1080/08989621.2020.1774373.

[42] K.A. Khor, and L.G. Yu, "Influence of international co-authorship on the research citation impact of young universities," *Scientometrics* 107 (March 15, 2016): 1095-1110, https://link.springer.com/article/10.1007/s11192-016-1905-6.

[43] This is based on CSET's merged corpus of scholarly literature, as of January 21, 2021. The total citation count for all AI research with at least one author affiliated to India is 18,188,313.

[44] India's paper collaboration counts with one country are not exclusive. For instance, India-U.S. collaborations represent the counts of papers with at least one Indian and one American author. This does not rule out the possibility of having a third author from another country.

[45] Researchers in the United States coauthor papers the most with their Chinese counterparts with a total of 38,502 papers, 13,698 papers with the United Kingdom, 10,434 with Germany, 8,763 with Canada, and 6,703 with Italy (as of January 22, 2021).

[46] Coauthorship is a tangible form of research collaboration, and can occur due to several factors including personal and professional connections. Additional research on this topic includes, B.D. Fonseca, R.B. Sampaio, and M.V.D Fonseca, "Co-authorship network analysis in health research: method and potential use," *Health Research Policy and Systems* 14, article no. 34 (2016), https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-016-0104-5.

[47] In our calculations, each paper is only counted in one category based on their association with various AI fields. Papers are allocated into their respective topics based on MAG fields. MAG provides multiple topic scores for each paper indicating each topic's relevance to the paper; we choose the highest scoring topic when available. Additionally, when a paper lacks MAG classification but is closely linked to other papers' "citation networks," we impute its MAG fields based on the papers it is closely related to. Papers without MAG fields and whose MAG fields could not be imputed are allocated in the "Unclassified" category.

[48] The fields of study are derived from MAG L1 classification, which provides scores for papers indicating their association with various fields. In calculating

the distribution of papers across fields, each paper is counted only in one category based on the highest field score.
The MAG L1 classification also lists artificial intelligence as a general sub-field within all AI papers. This likely captures AI papers that generally belong to the field and do not specialize in one area.

[49] To calculate the growth rate, we took the mean of the 2010 values subtracted from the 2019 values, divided by the 2010 values and multiplied the factor by 100 to get percentages. This is a standard growth rate calculation (e.g. "Planning Analysis: Calculating Growth Rates," University of Oregon, https://pages.uoregon.edu/rgp/PPPM613/class8a.htm).

[50] The Indian government has been worried about security risks related to research collaborations with China and in recent years has publicly raised concerns and taken steps to regulate the same. Sanjay Kumar, "Indian government tightens rules on academic collaboration with China," *Science*, October 17, 2019 https://www.sciencemag.org/news/2019/10/indian-government-tightens-rules-academic-collaboration-china; Remco Zwetsloot, "China's Approach to Tech Talent Competition: Policies, Results, and the Developing Global Response" (Center for Security and Emerging Technology, April 2020), https://cset.georgetown.edu/research/chinas-approach-to-tech-talent-competition-policies-results-and-the-developing-global-response/.

[51] For more details on the data and the methodological approach for identifying AI patents and classifying them into AI fields, please see Patrick Thomas and Dewey Murdick, "Patents and Artificial Intelligence: A Primer" (Center for Security and Emerging Technology, September 2020), https://cset.georgetown.edu/wp-content/uploads/CSET-Patents-and-Artificial-Intelligence.pdf.

[52] For more details on 1790 Analytics, see "Data resources," 1790 Analytics, https://1790analytics.com/#data.

[53] A patent family contains all patent documents associated with the same invention across multiple countries. This includes patent applications or requests pending at a patent office for the grant of a patent and patent grants approved requests awarding a property right for that invention.

[54] Both patent applications and patent grant numbers are deduplicated. If a family contains a granted patent then it is classified as a granted patent. If it does not, and the "highest" level document the family contains is a patent application, then it is classified as a patent application.

[55] India filed only 2053 patent applications in 2019 under WIPO's Patent Cooperation Treaty system for all fields. This was roughly half of the total patent applications by Huawei alone that filed 4411 applications. Dipti Jain, "India's patent problems," *Livemint*, November 24, 2014,

https://www.livemint.com/Politics/LkKhP62yJrhSRJZDoqDIiN/Indias-patent-problems.html; Banikinkar Pattanayak, "Global patents applications: India puts up a poor show, even Huawei has more filings," *Financial Express*, August 24, 2020, https://www.financialexpress.com/industry/global-patent-filings-huawei-alone-has-more-claims-than-indias-dismal-share-at-below-1/2063285/.

[56] Department of Science and Technology, "Research & Development Statistics at a Glance 2019-20," Ministry of Science & Technology, March 2020, https://dst.gov.in/sites/default/files/R%26D%20Statistics%20at%20a%20Glance%202019-20.pdf.

[57] Patent families are by organized by 'priority country,' whuch is where the original patent application was filed. For more details on this, see page 20 in Thomas and Murdick, "Patents and Artificial Intelligence: A Primer."

[58] Between 1983 and 2002, China had 149 patent families, United States 3618, Japan 2747, South Korea 353, Germany 540, United Kingdom 221, Taiwan 60, Russia 30, and France 131. India had none.

[59] World Intellectual Property Organization, *World Intellectual Property Indicators 2018* (Geneva: WIPO, 2018), https://www.wipo.int/edocs/pubdocs/en/wipo_pub_941_2018-chapter2.pdf; Department of Science & Technology, "Research & Development Statistics at a Glance 2019-20."

[60] Thomas and Murdick, "Patents and Artificial Intelligence: A Primer."

[61] World Intellectual Property Organization, "The Story of AI in Patents," WIPO, https://www.wipo.int/tech_trends/en/artificial_intelligence/story.html.

[62] In comparison to Indian patent system, applications are examined automatically in the U.S. system.

[63] Lack of sufficient patent examiners is one of the factors contributing to the slow patent application processing pace in India. The Indian government is taking steps to address this issue. Chadha & Chadha Intellectual Property Law Firm, "India: Patent Office Welcomes Hundreds of New Examiners, Hoping to put Major Dent in Backlog," Lexology, https://www.lexology.com/library/detail.aspx?g=adef0e21-d394-4684-bd4b-d6f360728832.

[64] See for example, Manisha Singh and Varun Sharma, "Patent litigations in India – Emerging Trends," The Legal 500, https://www.legal500.com/guides/hot-topic/patent-litigations-in-india-emerging-trends/.

[65] Compared to past years, India's ranking in the global IP Index released by the U.S. Chamber of Commerce has witnessed successive improvements. Global

Innovation Policy Center, *Art of the possible: U.S. Chamber International IP Index* (Washington, DC: U.S. Chamber of Commerce, February 5, 2020), https://www.uschamber.com/report/2020-international-ip-index.

[66] These categories were defined by CSET. A detailed description of this classification along with an index for patent documents can be found on GitHub: https://github.com/georgetown-cset/1790-ai-patent-data. More details on the approach are available here: Thomas and Murdick, "Patents and Artificial Intelligence: A Primer."

Patent counts for application fields are not exclusive, such that one patent can be associated with multiple application fields. For the purposes of the figure below, we undertook fractional counting for each application field to get estimates for exclusive counts. The total count for fractionally counted personal devices and computing patents gives us a total of 396.

[67] For an overview of the subcategories under these application fields, as well as a context summary on how can AI technologies be applied to these fields, see World Intellectual Property Organization, *WIPO Technology Trends 2019: Artificial Intelligence* (Geneva: WIPO, 2019), https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf.

[68] The category of business was associated with 248 patent families, telecommunications with 185, and life sciences with 142.

[69] World Intellectual Property Organization, *WIPO Technology Trends 2019: Artificial Intelligence*.

[70] AI companies were identified by Arnold et al., who searched across company descriptions for keywords or keyword combinations consistent with CSET's definition of "AI companies." For more details on the methodology and other assumptions, please refer to: Zachary Arnold, Ilya Rahkovsky, and Tina Huang, "Tracking AI Investment" (Center for Security and Emerging Technology, September 2020), https://cset.georgetown.edu/wp-content/uploads/CSET-Tracking-AI-Investment.pdf.

[71] As per reports, India is estimated to have about 50,000 startups overall (in 2018). Of these 8,900-9,300 were technology startups, and 1300 tech startups were born in 2019 alone. Department for Promotion of Industry and Internal Trade, "Indian Startup Ecosystem."

In our methodology, we identified all Indian companies listed on Crunchbase. As of February 2, 2021, these were 51,131 in total. Out of the pool, 361 were identified as AI-focused, privately held companies based on our methodology. Therefore, what this paper has analyzed is all Indian AI companies listed on

Crunchbase. Any comparisons of numbers presented in this report with estimates in other reports should be undertaken keeping this in mind.

[72] Arnold, Rahkovsky, and Huang, "Tracking AI Investment."

[73] For details on the methodology for estimating investments, please see Arnold, Rahkovsky, and Huang, "Tracking AI Investment."

[74] The figures for disclosed investments and transaction counts present a similar jump for India from 2018 to 2019.

[75] Jon Russell, "Reliance Jio's latest acquisition is a $100M bet on the future of internet users in India," *Techcrunch*, April 4, 2019, https://techcrunch.com/2019/04/04/reliance-jio-buys-haptik-for-100-million/.

[76] For more detailed explanation and limitations of this method, please see Appendix 1 in Arnold, Rahkovsky, and Huang, "Tracking AI Investment."

[77] As indicated above, these figures are based on investment transactions that had at least one publicly disclosed Chinese, American, or Indian investor participating. We compared our results by looking at investment rounds with only Indian investors, only U.S. investors, and only Chinese investors separately. The results were consistent with our assessment about the United States as a major player in the Indian market. The number of India-bound China only investments between 2015-19 are about $9 million (disclosed), $34 million (estimated) with 8 transaction counts. U.S. only investments are $159 million (disclosed), $539 million (estimated), and 68 transaction counts. India only investments are $293 million (disclosed), $793 million (estimated), and 191 transaction counts.

[78] The identified AI companies were divided into these application areas by CSET researchers after examining every company's Crunchbase profile and website. For the methodology behind these tables, please see Appendix 1 in Arnold, Rahkovsky, and Huang, "Tracking AI Investment."

[79] Cloud compute spending is only a rough proxy of the availability of compute and hopefully better indicator data will become available in the future. This proxy has several limitations: 1) it doesn't reflect the compute capacity of "in-house" private data centers, 2) it makes it unclear how much a country's investments in cloud computing are useful for the purposes of AI, among others.

[80] "TOP500 Release," TOP500, November 2020, https://www.top500.org/lists/top500/2020/11/.

[81] "AIRAWAT – Establishing an AI-specific cloud computing infrastructure for India."

[82] There are three types of cloud deployments —private cloud, public cloud, and hybrid cloud. Private cloud is a computing model dedicated to the needs of a single business entity, and whose infrastructure may be managed by the organization itself or a third party, and may exist on premise or off premise. Public cloud is a model available to the general public or enterprises while its infrastructure is owned by a third-party cloud service provider. Hybrid cloud is a composition of two or more clouds (private on premise/off premise, or public) that are unique to an organization but bound by proprietary technology that enables data and application portability.

[83] Cloud penetration refers to cloud spending as a percent of IT spending. Cloud spending includes domestic spend on SaaS, IaaS (only public cloud), PaaS, Cloud business process services, and Cloud management and security services.

[84] Please note that the figures for public cloud here do not refer to an exclusive category of total spending only on public cloud. India's $2.5 billion public cloud market here could also include entities that may also have a private cloud infrastructure for core assets, or a hybrid cloud.

[85] End user IT spending includes spending by "enterprises on Data center systems, Devices (excluding Mobile phones), Software, IT Services, and Communication Services (excluding consumer communication)." "Cloud – Next Wave of Growth in India" (NASSCOM, 2019), https://nasscom.in/knowledge-center/publications/nasscom-cloud-next-wave-growth-india-2019.

[86] "Cloud – Next Wave of Growth in India," 30.

[87] Over 13 percent of India's computational workload is expected to move to Cloud, in addition to recent investments by hyperscalers like AWS and Google cloud. "Cloud – Next Wave of Growth in India," 30.

[88] "Cloud – Next Wave of Growth in India."

[89] Manasi Sakpal, "Gartner forecasts end-user spending on public cloud in India to grow 29% in 2021," Gartner, November 24, 2010, https://www.gartner.com/en/newsroom/2020-11-24-gartner-forecasts-end-user-spending-on-public-cloud-in-india; Shivani Anand, "IDC Forecast the India Public Cloud Services Market to Grow at CAGR 22.2% for 2020-24," IDC, December 11, 2020, https://www.idc.com/getdoc.jsp?containerId=prAP47129220.

[90] The growth rates for 2021 are expected to stay at 29.4 percent bringing 2021 figures to $4.1 billion. Even if the market grows at the same rate for the following year, it will equal $5.3 billion at most, again missing the conservative $5.6 billion forecasts for 2022.

[91] Possible factors include economic slowdown, high costs, and a downturn in the market globally. Natalya Yezhkova, Lidice Fernandez, Kuba Stolarski, and Michael Shirer, "Cloud IT Infrastructure Revenues Decline in Q2 2019 Amid a Slow Down in Overall Spending, According to IDC," IDC, September 25, 2019, https://www.idc.com/getdoc.jsp?containerId=prUS45552219; Therese Poletti, "The biggest question looming over tech: Is the cloud boom over?," *MarketWatch*, January 21, 2020, https://www.marketwatch.com/story/the-biggest-question-looming-over-tech-is-the-cloud-boom-over-2020-01-21.

[92] All of these factors, and others are discussed in this report: "Cloud: Next Wave of Growth in India."

[93] As mentioned above in the introduction, for details on the proposed AI-specific cloud infrastructure by the Indian government - AIRAWAT - please refer to "AIRAWAT – Establishing an AI-specific cloud computing infrastructure for India."

[94] "U.S.-India Artificial Intelligence (USIAI) Initiative," Indo-U.S. Science and Technology Forum, March 10, 2021, https://www.iusstf.org/assets/pdf/usiai-broucher.jpg.

[95] Among others this includes lack of funding access for Indian scholars, controls over material and sharing of data, samples, and equipment for American scholars, along with issues of 'reported or perceived bias against or toward working with a specific region/country' indicated by both. Kirstin R.W. Matthews, Erin Yang, Steven W. Lewis, Brandon R. Vaidyanathan, and Monica Gorman, "International scientific collaborative activities and barriers to them in eight societies."

[96] U.S. India Business Council, "USIBC Statement on Data Privacy Bill," https://www.usibc.com/press-release/usibc-statement-on-data-privacy-bill/; See '"forced" localization' in "U.S.-India Trade Relations" (Congressional Research Service, 23 December 2020), https://crsreports.congress.gov/product/pdf/IF/IF10384.