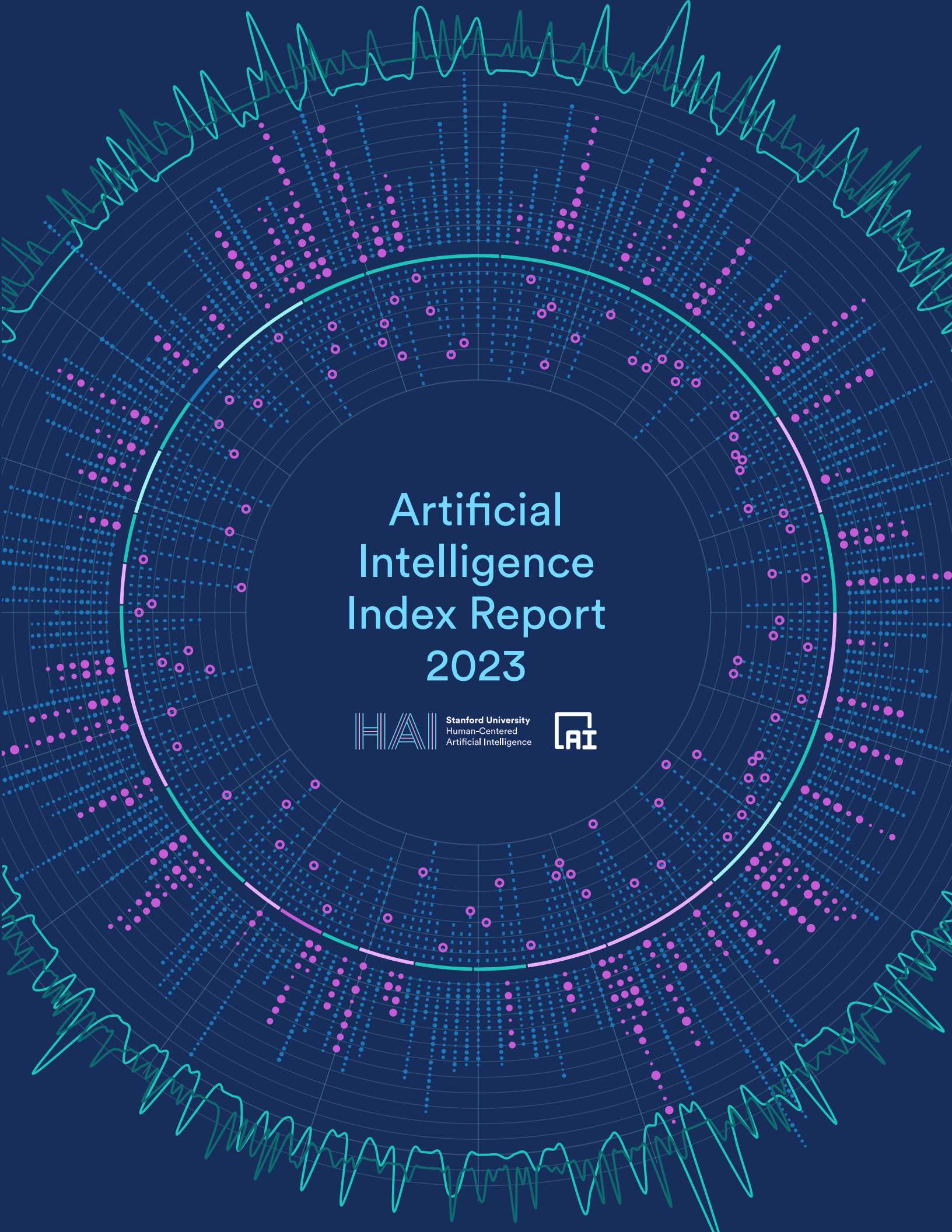
Artificial
Intelligence
Index Report
2023
HAI
Stanford University
Human-Centered
Artificial Intelligence
AI

Introduction to the AI Index Report 2023

Welcome to the sixth edition of the AI Index Report! This year, the report introduces more original data than any previous edition, including a new chapter on AI public opinion, a more thorough technical performance chapter, original analysis about large language and multimodal models, detailed trends in global AI legislation records, a study of the environmental impact of AI systems, and more.

The AI Index Report tracks, collates, distills, and visualizes data related to artificial intelligence. Our mission is to provide unbiased, rigorously vetted, broadly sourced data in order for policymakers, researchers, executives, journalists, and the general public to develop a more thorough and nuanced understanding of the complex field of AI. The report aims to be the world's most credible and authoritative source for data and insights about AI.

From the Co-Directors

AI has moved into its era of deployment; throughout 2022 and the beginning of 2023, new large-scale AI models have been released every month. These models, such as ChatGPT, Stable Diffusion, Whisper, and DALL-E 2, are capable of an increasingly broad range of tasks, from text manipulation and analysis, to image generation, to unprecedentedly good speech recognition. These systems demonstrate capabilities in question answering and the generation of text, image, and code unimagined a decade ago, and they outperform the state of the art on many benchmarks, old and new. However, they are prone to hallucination, routinely biased, and can be tricked into serving nefarious aims, highlighting the complicated ethical challenges associated with their deployment.

Although 2022 was the first year in a decade where private AI investment decreased, AI is still a topic of great interest to policymakers, industry leaders, researchers, and the public. Policymakers are talking about AI more than ever before. Industry leaders that have integrated AI into their businesses are seeing tangible cost and revenue benefits. The number of AI publications and collaborations continues to increase. And the public is forming sharper opinions about AI and which elements they like or dislike.

AI will continue to improve and, as such, become a greater part of all our lives. Given the increased presence of this technology and its potential for massive disruption, we should all begin thinking more critically about how exactly we want AI to be developed and deployed. We should also ask questions about who is deploying it—as our analysis shows, AI is increasingly defined by the actions of a small set of private sector actors, rather than a broader range of societal actors. This year's AI Index paints a picture of where we are so far with AI, in order to highlight what might await us in the future.

Jack Clark and Ray Perrault

Top Ten Takeaways

1 Industry races ahead of academia.

Until 2014, most significant machine learning models were released by academia. Since then, industry has taken over. In 2022, there were 32 significant industry-produced machine learning models compared to just three produced by academia. Building state-of-the-art AI systems increasingly requires large amounts of data, computer power, and money—resources that industry actors inherently possess in greater amounts compared to nonprofits and academia.

2 Performance saturation on traditional benchmarks.

AI continued to post state-of-the-art results, but year-over-year improvement on many benchmarks continues to be marginal. Moreover, the speed at which benchmark saturation is being reached is increasing. However, new, more comprehensive benchmarking suites such as BIG-bench and HELM are being released.

3 AI is both helping and harming the environment.

New research suggests that AI systems can have serious environmental impacts. According to Luccioni et al., 2022, BLOOM's training run emitted 25 times more carbon than a single air traveler on a one-way trip from New York to San Francisco. Still, new reinforcement learning models like BCOOLER show that AI systems can be used to optimize energy usage.

4 The world's best new scientist ... AI?

AI models are starting to rapidly accelerate scientific progress and in 2022 were used to aid hydrogen fusion, improve the efficiency of matrix manipulation, and generate new antibodies.

5 The number of incidents concerning the misuse of AI is rapidly rising.

According to the AIAAIC database, which tracks incidents related to the ethical misuse of AI, the number of AI incidents and controversies has increased 26 times since 2012. Some notable incidents in 2022 included a deepfake video of Ukrainian President Volodymyr Zelenskyy surrendering and U.S. prisons using call-monitoring technology on their inmates. This growth is evidence of both greater use of AI technologies and awareness of misuse possibilities.

6 The demand for AI-related professional skills is increasing across virtually every American industrial sector.

Across every sector in the United States for which there is data (with the exception of agriculture, forestry, fishing, and hunting), the number of AI-related job postings has increased on average from 1.7% in 2021 to 1.9% in 2022. Employers in the United States are increasingly looking for workers with AI-related skills.

Top Ten Takeaways (cont'd)

7 For the first time in the last decade, year-over-year private investment in AI decreased.

Global AI private investment was $91.9 billion in 2022, which represented a 26.7% decrease since 2021. The total number of AI-related funding events as well as the number of newly funded AI companies likewise decreased. Still, during the last decade as a whole, AI investment has significantly increased. In 2022 the amount of private investment in AI was 18 times greater than it was in 2013.

8 While the proportion of companies adopting AI has plateaued, the companies that have adopted AI continue to pull ahead.

The proportion of companies adopting AI in 2022 has more than doubled since 2017, though it has plateaued in recent years between 50% and 60%, according to the results of McKinsey's annual research survey. Organizations that have adopted AI report realizing meaningful cost decreases and revenue increases.

9 Policymaker interest in AI is on the rise.

An AI Index analysis of the legislative records of 127 countries shows that the number of bills containing "artificial intelligence" that were passed into law grew from just 1 in 2016 to 37 in 2022. An analysis of the parliamentary records on AI in 81 countries likewise shows that mentions of AI in global legislative proceedings have increased nearly 6.5 times since 2016.

10 Chinese citizens are among those who feel the most positively about AI products and services. Americans ... not so much.

In a 2022 IPSOS survey, 78% of Chinese respondents (the highest proportion of surveyed countries) agreed with the statement that products and services using AI have more benefits than drawbacks. After Chinese respondents, those from Saudi Arabia (76%) and India (71%) felt the most positive about AI products. Only 35% of sampled Americans (among the lowest of surveyed countries) agreed that products and services using AI had more benefits than drawbacks.

Steering Committee

Co-directors

Jack Clark
Anthropic, OECD

Raymond Perrault
SRI International

Members

Erik Brynjolfsson
Stanford University

John Etchemendy
Stanford University

Katrina Ligett
Hebrew University

Terah Lyons

James Manyika
Google,
University of Oxford

Juan Carlos Niebles
Stanford University,
Salesforce

Vanessa Parli
Stanford University

Yoav Shoham
(Founding Director)
Stanford University,
AI21 Labs

Russell Wald
Stanford University

Staff and Researchers

Research Manager and Editor in Chief

Nestor Maslej
Stanford University

Research Associate

Loredana Fattorini
Stanford University

Affiliated Researchers

Elif Kiesow Cortez
Stanford Law School
Research Fellow

Helen Ngo
Hugging Face

Robi Rahman
Data Scientist

Alexandra Rome
Freelance Researcher

Graduate Researcher

Han Bai
Stanford University

Undergraduate Researchers

Vania Chow
Stanford University

Siddhartha Javvaji
Stanford University

Mena Hassan
Stanford University

Naima Patel
Stanford University

Sukrut Oak
Stanford University

Stone Yang
Stanford University

Lucy Zimmerman
Stanford University

Elizabeth Zhu
Stanford University

How to Cite This Report

Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault, "The AI Index 2023 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023.

Public Data and Tools

The AI Index 2023 Report is supplemented by raw data and an interactive tool. We invite each reader to use the data and the tool in a way most relevant to their work and interests.

Raw data and charts: The public data and high-resolution images of all the charts in the report are available on Google Drive.

Global AI Vibrancy Tool: Compare up to 30 countries across 21 indicators. The Global AI Vibrancy tool will be updated in the latter half of 2023.

AI Index and Stanford HAI

The AI Index is an independent initiative at the Stanford Institute for Human-Centered Artificial Intelligence (HAI).

Artificial Intelligence Index

Stanford University Human-Centered Artificial Intelligence

The AI Index was conceived within the One Hundred Year Study on AI (AI100).

We welcome feedback and new ideas for next year.
Contact us at AI-Index-Report@stanford.edu.

Supporting Partners

Google

OpenAI

NSF

NETBASE QUID

Open Philanthropy

Analytics and Research Partners

CODE

CRA
Computing Research Association

CSET
CENTER for SECURITY and EMERGING TECHNOLOGY

GitHub

govini

Lightcast

LinkedIn

McKinsey & Company

OECD.AI
Policy Observatory

WiML
Women in Machine Learning

Contributors

We want to acknowledge the following individuals by chapter and section for their contributions of data, analysis, advice, and expert commentary included in the AI Index 2023 Report:

Research and Development

Sara Abdulla, Catherine Aiken, Luis Aranda, Peter Cihon, Jack Clark, Loredana Fattorini, Nestor Maslej, Besher Massri, Vanessa Parli, Naima Patel, Ray Perrault, Robi Rahman, Alexandra Rome, Kevin Xu

Technical Performance

Jack Clark, Loredana Fattorini, Siddhartha Javvaji, Katrina Ligett, Nestor Maslej, Juan Carlos Niebles, Sukrut Oak, Vanessa Parli, Ray Perrault, Robi Rahman, Alexandra Rome, Yoav Shoham, Elizabeth Zhu

Technical AI Ethics

Jack Clark, Loredana Fattorini, Katrina Ligett, Nestor Maslej, Helen Ngo, Sukrut Oak, Vanessa Parli, Ray Perrault, Alexandra Rome, Elizabeth Zhu, Lucy Zimmerman

Economy

Susanne Bieller, Erik Brynjolfsson, Vania Chow, Jack Clark, Natalia Dorogi, Murat Erer, Loredana Fattorini, Akash Kaura, James Manyika, Nestor Maslej, Layla O'Kane, Vanessa Parli, Ray Perrault, Brittany Presten, Alexandra Rome, Nicole Seredenko, Bledi Taska, Bill Valle, Casey Weston

Education

Han Bai, Betsy Bizot, Jack Clark, John Etchemendy, Loredana Fattorini, Katrina Ligett, Nestor Maslej, Vanessa Parli, Ray Perrault, Sean Roberts, Alexandra Rome

Policy and Governance

Meghan Anand, Han Bai, Vania Chow, Jack Clark, Elif Kiesow Cortez, Rebecca DeCrescenzo, Loredana Fattorini, Taehwa Hong, Joe Hsu, Kai Kato, Terah Lyons, Nestor Maslej, Alistair Murray, Vanessa Parli, Ray Perrault, Alexandra Rome, Sarah Smedley, Russell Wald, Brian Williams, Catherina Xu, Stone Yang, Katie Yoon, Daniel Zhang

Diversity

Han Bai, Betsy Bizot, Jack Clark, Loredana Fattorini, Nezihe Merve Gürel, Mena Hassan, Katrina Ligett, Nestor Maslej, Vanessa Parli, Ray Perrault, Sean Roberts, Alexandra Rome, Sarah Tan, Lucy Zimmerman

Public Opinion

Jack Clark, Loredana Fattorini, Mena Hassan, Nestor Maslej, Vanessa Parli, Ray Perrault, Alexandra Rome, Nicole Seredenko, Bill Valle, Lucy Zimmerman

Conference Attendance

Terri Auricchio (ICML), Lee Campbell (ICLR), Cassio de Campos (UAI), Meredith Ellison (AAAI), Nicole Finn (CVPR), Vasant Gajanan (AAAI), Katja Hofmann (ICLR), Gerhard Lakemeyer (KR), Seth Lazar (FAccT), Shugen Ma (IROS), Becky Obbema (NeurIPS), Vesna Sabljakovic-Fritz (IJCAI), Csaba Szepesvari (ICML), Matthew Taylor (AAMAS), Sylvie Thiebaux (ICAPS), Pradeep Varakantham (ICAPS)

We thank the following organizations and individuals who provided data for inclusion in the AI Index 2023 Report:

Organizations

Code.org
Sean Roberts

Center for Security and Emerging Technology, Georgetown University
Sara Abdulla, Catherine Aiken

Computing Research Association
Betsy Bizot

GitHub
Peter Cihon, Kevin Xu

Govini
Rebecca DeCrescenzo, Joe Hsu, Sarah Smedley

Lightcast
Layla O'Kane, Bledi Taska

LinkedIn
Murat Erer, Akash Kaura, Casey Weston

McKinsey & Company
Natalia Dorogi, Brittany Presten

NetBase Quid
Nicole Seredenko, Bill Valle

OECD.AI Policy Observatory
Luis Aranda, Besher Massri

Women in Machine Learning
Nezihe Merve Gürel, Sarah Tan

We also would like to thank Jeanina Casusi, Nancy King, Shana Lynch, Jonathan Mindes, Michi Turner, and Madeleine Wright for their help in preparing this report, and Joe Hinman and Santanu Mukherjee for their help in maintaining the AI Index website.

Table of Contents

ACCESS THE PUBLIC DATA

Report Highlights

Chapter 1: Research and Development

The United States and China had the greatest number of cross-country collaborations in AI publications from 2010 to 2021, although the pace of collaboration has slowed. The number of AI research collaborations between the United States and China increased roughly 4 times since 2010, and was 2.5 times greater than the collaboration totals of the next nearest country pair, the United Kingdom and China. However the total number of U.S.-China collaborations only increased by 2.1% from 2020 to 2021, the smallest year-over-year growth rate since 2010.

AI research is on the rise, across the board. The total number of AI publications has more than doubled since 2010. The specific AI topics that continue dominating research include pattern recognition, machine learning, and computer vision.

China continues to lead in total AI journal, conference, and repository publications. The United States is still ahead in terms of AI conference and repository citations, but those leads are slowly eroding. Still, the majority of the world's large language and multimodal models (54% in 2022) are produced by American institutions.

Industry races ahead of academia. Until 2014, most significant machine learning models were released by academia. Since then, industry has taken over. In 2022, there were 32 significant industry-produced machine learning models compared to just three produced by academia. Building state-of-the-art AI systems increasingly requires large amounts of data, computer power, and money—resources that industry actors inherently possess in greater amounts compared to nonprofits and academia.

Large language models are getting bigger and more expensive. GPT-2, released in 2019, considered by many to be the first large language model, had 1.5 billion parameters and cost an estimated $50,000 USD to train. PaLM, one of the flagship large language models launched in 2022, had 540 billion parameters and cost an estimated $8 million USD—PaLM was around 360 times larger than GPT-2 and cost 160 times more. It's not just PaLM: Across the board, large language and multimodal models are becoming larger and pricier.

Chapter 2: Technical Performance

Performance saturation on traditional benchmarks. AI continued to post state-of-the-art results, but year-over-year improvement on many benchmarks continues to be marginal. Moreover, the speed at which benchmark saturation is being reached is increasing. However, new, more comprehensive benchmarking suites such as BIG-bench and HELM are being released.

Generative AI breaks into the public consciousness. 2022 saw the release of text-to-image models like DALL-E 2 and Stable Diffusion, text-to-video systems like Make-A-Video, and chatbots like ChatGPT. Still, these systems can be prone to hallucination, confidently outputting incoherent or untrue responses, making it hard to rely on them for critical applications.

AI systems become more flexible. Traditionally AI systems have performed well on narrow tasks but have struggled across broader tasks. Recently released models challenge that trend; BEiT-3, PaLI, and Gato, among others, are single AI systems increasingly capable of navigating multiple tasks (for example, vision, language).

Capable language models still struggle with reasoning. Language models continued to improve their generative capabilities, but new research suggests that they still struggle with complex planning tasks.

AI is both helping and harming the environment. New research suggests that AI systems can have serious environmental impacts. According to Luccioni et al., 2022, BLOOM's training run emitted 25 times more carbon than a single air traveler on a one-way trip from New York to San Francisco. Still, new reinforcement learning models like BCOOLER show that AI systems can be used to optimize energy usage.

The world's best new scientist ... AI? AI models are starting to rapidly accelerate scientific progress and in 2022 were used to aid hydrogen fusion, improve the efficiency of matrix manipulation, and generate new antibodies.

AI starts to build better AI. Nvidia used an AI reinforcement learning agent to improve the design of the chips that power AI systems. Similarly, Google recently used one of its language models, PaLM, to suggest ways to improve the very same model. Self-improving AI learning will accelerate AI progress.

Chapter 3: Technical AI Ethics

The effects of model scale on bias and toxicity are confounded by training data and mitigation methods. In the past year, several institutions have built their own large models trained on proprietary data—and while large models are still toxic and biased, new evidence suggests that these issues can be somewhat mitigated after training larger models with instruction-tuning.

Generative models have arrived and so have their ethical problems. In 2022, generative models became part of the zeitgeist. These models are capable but also come with ethical challenges. Text-to-image generators are routinely biased along gender dimensions, and chatbots like ChatGPT can be tricked into serving nefarious aims.

The number of incidents concerning the misuse of AI is rapidly rising. According to the AIAAIC database, which tracks incidents related to the ethical misuse of AI, the number of AI incidents and controversies has increased 26 times since 2012. Some notable incidents in 2022 included a deepfake video of Ukrainian President Volodymyr Zelenskyy surrendering and U.S. prisons using call-monitoring technology on their inmates. This growth is evidence of both greater use of AI technologies and awareness of misuse possibilities.

Fairer models may not be less biased. Extensive analysis of language models suggests that while there is a clear correlation between performance and fairness, fairness and bias can be at odds: Language models which perform better on certain fairness benchmarks tend to have worse gender bias.

Interest in AI ethics continues to skyrocket. The number of accepted submissions to FAccT, a leading AI ethics conference, has more than doubled since 2021 and increased by a factor of 10 since 2018. 2022 also saw more submissions than ever from industry actors.

Automated fact-checking with natural language processing isn't so straightforward after all. While several benchmarks have been developed for automated fact-checking, researchers find that 11 of 16 of such datasets rely on evidence "leaked" from fact-checking reports which did not exist at the time of the claim surfacing.

Chapter 4: The Economy

The demand for AI-related professional skills is increasing across virtually every American industrial sector. Across every sector in the United States for which there is data (with the exception of agriculture, forestry, fishing, and hunting), the number of AI-related job postings has increased on average from 1.7% in 2021 to 1.9% in 2022. Employers in the United States are increasingly looking for workers with AI-related skills.

For the first time in the last decade, year-over-year private investment in AI decreased. Global AI private investment was $91.9 billion in 2022, which represented a 26.7% decrease since 2021. The total number of AI-related funding events as well as the number of newly funded AI companies likewise decreased. Still, during the last decade as a whole, AI investment has significantly increased. In 2022 the amount of private investment in AI was 18 times greater than it was in 2013.

Once again, the United States leads in investment in AI. The U.S. led the world in terms of total amount of AI private investment. In 2022, the $47.4 billion invested in the U.S. was roughly 3.5 times the amount invested in the next highest country, China ($13.4 billion). The U.S. also continues to lead in terms of total number of newly funded AI companies, seeing 1.9 times more than the European Union and the United Kingdom combined, and 3.4 times more than China.

In 2022, the AI focus area with the most investment was medical and healthcare ($6.1 billion); followed by data management, processing, and cloud ($5.9 billion); and Fintech ($5.5 billion). However, mirroring the broader trend in AI private investment, most AI focus areas saw less investment in 2022 than in 2021. In the last year, the three largest AI private investment events were: (1) a $2.5 billion funding event for GAC Aion New Energy Automobile, a Chinese manufacturer of electric vehicles; (2) a $1.5 billion Series E funding round for Anduril Industries, a U.S. defense products company that builds technology for military agencies and border surveillance; and (3) a $1.2 billion investment in Celonis, a business-data consulting company based in Germany.

While the proportion of companies adopting AI has plateaued, the companies that have adopted AI continue to pull ahead. The proportion of companies adopting AI in 2022 has more than doubled since 2017, though it has plateaued in recent years between 50% and 60%, according to the results of McKinsey's annual research survey. Organizations that have adopted AI report realizing meaningful cost decreases and revenue increases.

Chapter 4: The Economy (cont'd)

AI is being deployed by businesses in multifaceted ways. The AI capabilities most likely to have been embedded in businesses include robotic process automation (39%), computer vision (34%), NL text understanding (33%), and virtual agents (33%). Moreover, the most commonly adopted AI use case in 2022 was service operations optimization (24%), followed by the creation of new AI-based products (20%), customer segmentation (19%), customer service analytics (19%), and new AI-based enhancement of products (19%).

AI tools like Copilot are tangibly helping workers. Results of a GitHub survey on the use of Copilot, a text-to-code AI system, find that 88% of surveyed respondents feel more productive when using the system, 74% feel they are able to focus on more satisfying work, and 88% feel they are able to complete tasks more quickly.

China dominates industrial robot installations. In 2013, China overtook Japan as the nation installing the most industrial robots. Since then, the gap between the total number of industrial robots installed by China and the next-nearest nation has widened. In 2021, China installed more industrial robots than the rest of the world combined.

Chapter 5: Education

More and more AI specialization. The proportion of new computer science PhD graduates from U.S. universities who specialized in AI jumped to 19.1% in 2021, from 14.9% in 2020 and 10.2% in 2010.

New AI PhDs increasingly head to industry. In 2011, roughly the same proportion of new AI PhD graduates took jobs in industry (40.9%) as opposed to academia (41.6%). Since then, however, a majority of AI PhDs have headed to industry. In 2021, 65.4% of AI PhDs took jobs in industry, more than double the 28.2% who took jobs in academia.

New North American CS, CE, and information faculty hires stayed flat. In the last decade, the total number of new North American computer science (CS), computer engineering (CE), and information faculty hires has decreased: There were 710 total hires in 2021 compared to 733 in 2012. Similarly, the total number of tenure-track hires peaked in 2019 at 422 and then dropped to 324 in 2021.

The gap in external research funding for private versus public American CS departments continues to widen. In 2011, the median amount of total expenditure from external sources for computing research was roughly the same for private and public CS departments in the United States. Since then, the gap has widened, with private U.S. CS departments receiving millions more in additional funding than public universities. In 2021, the median expenditure for private universities was $9.7 million, compared to $5.7 million for public universities.

Interest in K–12 AI and computer science education grows in both the United States and the rest of the world. In 2021, a total of 181,040 AP computer science exams were taken by American students, a 1.0% increase from the previous year. Since 2007, the number of AP computer science exams has increased ninefold. As of 2021, 11 countries, including Belgium, China, and South Korea, have officially endorsed and implemented a K–12 AI curriculum.

Chapter 6: Policy and Governance

Policymaker interest in AI is on the rise. An AI Index analysis of the legislative records of 127 countries shows that the number of bills containing "artificial intelligence" that were passed into law grew from just 1 in 2016 to 37 in 2022. An analysis of the parliamentary records on AI in 81 countries likewise shows that mentions of AI in global legislative proceedings have increased nearly 6.5 times since 2016.

From talk to enactment—the U.S. passed more AI bills than ever before. In 2021, only 2% of all federal AI bills in the United States were passed into law. This number jumped to 10% in 2022. Similarly, last year 35% of all state-level AI bills were passed into law.

When it comes to AI, policymakers have a lot of thoughts. A qualitative analysis of the parliamentary proceedings of a diverse group of nations reveals that policymakers think about AI from a wide range of perspectives. For example, in 2022, legislators in the United Kingdom discussed the risks of AI-led automation; those in Japan considered the necessity of safeguarding human rights in the face of AI; and those in Zambia looked at the possibility of using AI for weather forecasting.

The U.S. government continues to increase spending on AI. Since 2017, the amount of U.S. government AI-related contract spending has increased roughly 2.5 times.

The legal world is waking up to AI. In 2022, there were 110 AI-related legal cases in United States state and federal courts, roughly seven times more than in 2016. The majority of these cases originated in California, New York, and Illinois, and concerned issues relating to civil, intellectual property, and contract law.

Chapter 7: Diversity

North American bachelor's, master's, and PhD-level computer science students are becoming more ethnically diverse. Although white students are still the most represented ethnicity among new resident bachelor's, master's, and PhD-level computer science graduates, students from other ethnic backgrounds (for example, Asian, Hispanic, and Black or African American) are becoming increasingly more represented. For example, in 2011, 71.9% of new resident CS bachelor's graduates were white. In 2021, that number dropped to 46.7%.

New AI PhDs are still overwhelmingly male. In 2021, 78.7% of new AI PhDs were male. Only 21.3% were female, a 3.2 percentage point increase from 2011. There continues to be a gender imbalance in higher-level AI education.

Women make up an increasingly greater share of CS, CE, and information faculty hires. Since 2017, the proportion of new female CS, CE, and information faculty hires has increased from 24.9% to 30.2%. Still, most CS, CE, and information faculty in North American universities are male (75.9%). As of 2021, only 0.1% of CS, CE, and information faculty identify as nonbinary.

American K–12 computer science education has become more diverse, in terms of both gender and ethnicity. The share of AP computer science exams taken by female students increased from 16.8% in 2007 to 30.6% in 2021. Year over year, the share of Asian, Hispanic/Latino/Latina, and Black/African American students taking AP computer science has likewise increased.

Chapter 8: Public Opinion

Chinese citizens are among those who feel the most positively about AI products and services. Americans ... not so much. In a 2022 IPSOS survey, 78% of Chinese respondents (the highest proportion of surveyed countries) agreed with the statement that products and services using AI have more benefits than drawbacks. After Chinese respondents, those from Saudi Arabia (76%) and India (71%) felt the most positive about AI products. Only 35% of sampled Americans (among the lowest of surveyed countries) agreed that products and services using AI had more benefits than drawbacks.

Men tend to feel more positively about AI products and services than women. Men are also more likely than women to believe that AI will mostly help rather than harm. According to the 2022 IPSOS survey, men are more likely than women to report that AI products and services make their lives easier, trust companies that use AI, and feel that AI products and services have more benefits than drawbacks. A 2021 survey by Gallup and Lloyd's Register Foundation likewise revealed that men are more likely than women to agree with the statement that AI will mostly help rather than harm their country in the next 20 years.

People across the world and especially America remain unconvinced by self-driving cars. In a global survey, only 27% of respondents reported feeling safe in a self-driving car. Similarly, Pew Research suggests that only 26% of Americans feel that driverless passenger vehicles are a good idea for society.

Different causes for excitement and concern. Among a sample of surveyed Americans, those who report feeling excited about AI are most excited about the potential to make life and society better (31%) and to save time and make things more efficient (13%). Those who report feeling more concerned worry about the loss of human jobs (19%); surveillance, hacking, and digital privacy (16%); and the lack of human connection (12%).

NLP researchers ... have some strong opinions as well. According to a survey widely distributed to NLP researchers, 77% either agreed or weakly agreed that private AI firms have too much influence, 41% said that NLP should be regulated, and 73% felt that AI could soon lead to revolutionary societal change. These were some of the many strong opinions held by the NLP research community.

Artificial Intelligence
Index Report 2023

CHAPTER 1:
Research and Development

CHAPTER 1 PREVIEW:

Research and Development

ACCESS THE PUBLIC DATA

Overview

This chapter captures trends in AI R&D. It begins by examining AI publications, including journal articles, conference papers, and repositories. Next it considers data on significant machine learning systems, including large language and multimodal models. Finally, the chapter concludes by looking at AI conference attendance and open-source AI research. Although the United States and China continue to dominate AI R&D, research efforts are becoming increasingly geographically dispersed.

Chapter Highlights

The United States and China had the greatest number of cross-country collaborations in AI publications from 2010 to 2021, although the pace of collaboration has since slowed.

The number of AI research collaborations between the United States and China increased roughly 4 times since 2010, and was 2.5 times greater than the collaboration totals of the next nearest country pair, the United Kingdom and China. However, the total number of U.S.-China collaborations only increased by 2.1% from 2020 to 2021, the smallest year-over-year growth rate since 2010.

Industry races ahead of academia.

Until 2014, most significant machine learning models were released by academia. Since then, industry has taken over. In 2022, there were 32 significant industry-produced machine learning models compared to just three produced by academia. Building state-of-the-art AI systems increasingly requires large amounts of data, computer power, and money—resources that industry actors inherently possess in greater amounts compared to nonprofits and academia.

AI research is on the rise, across the board.

The total number of AI publications has more than doubled since 2010. The specific AI topics that continue to dominate research include pattern recognition, machine learning, and computer vision.

China continues to lead in total AI journal, conference, and repository publications.

The United States is still ahead in terms of AI conference and repository citations, but those leads are slowly eroding. Still, the majority of the world's large language and multimodal models (54% in 2022) are produced by American institutions.

Large language models are getting bigger and more expensive.

GPT-2, released in 2019, considered by many to be the first large language model, had 1.5 billion parameters and cost an estimated $50,000 USD to train. PaLM, one of the flagship large language models launched in 2022, had 540 billion parameters and cost an estimated $8 million USD—PaLM was around 360 times larger than GPT-2 and cost 160 times more. It's not just PaLM: Across the board, large language and multimodal models are becoming larger and pricier.

This section draws on data from the Center for Security and Emerging Technology (CSET) at Georgetown University. CSET maintains a merged corpus of scholarly literature that includes Digital Science's Dimensions, Clarivate's Web of Science, Microsoft Academic Graph, China National Knowledge Infrastructure, arXiv, and Papers With Code. In that corpus, CSET applied a classifier to identify English-language publications related to the development or application of AI and ML since 2010. For this year's report, CSET also used select Chinese AI keywords to identify Chinese-language AI papers; CSET did not deploy this method for previous iterations of the AI Index report.[1]

In last year's edition of the report, publication trends were reported up to the year 2021. However, given that there is a significant lag in the collection of publication metadata, and that in some cases it takes until the middle of any given year to fully capture the previous year's publications, in this year's report, the AI Index team elected to examine publication trends only through 2021, which we, along with CSET, are confident yields a more fully representative report.

1.1 Publications

Overview

The figures below capture the total number of English-language and Chinese-language AI publications globally from 2010 to 2021—by type, affiliation, cross-country collaboration, and cross-industry collaboration. The section also breaks down publication and citation data by region for AI journal articles, conference papers, repositories, and patents.

Total Number of AI Publications

Figure 1.1.1 shows the number of AI publications in the world. From 2010 to 2021, the total number of AI publications more than doubled, growing from 200,000 in 2010 to almost 500,000 in 2021.

Number of AI Publications in the World, 2010–21

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

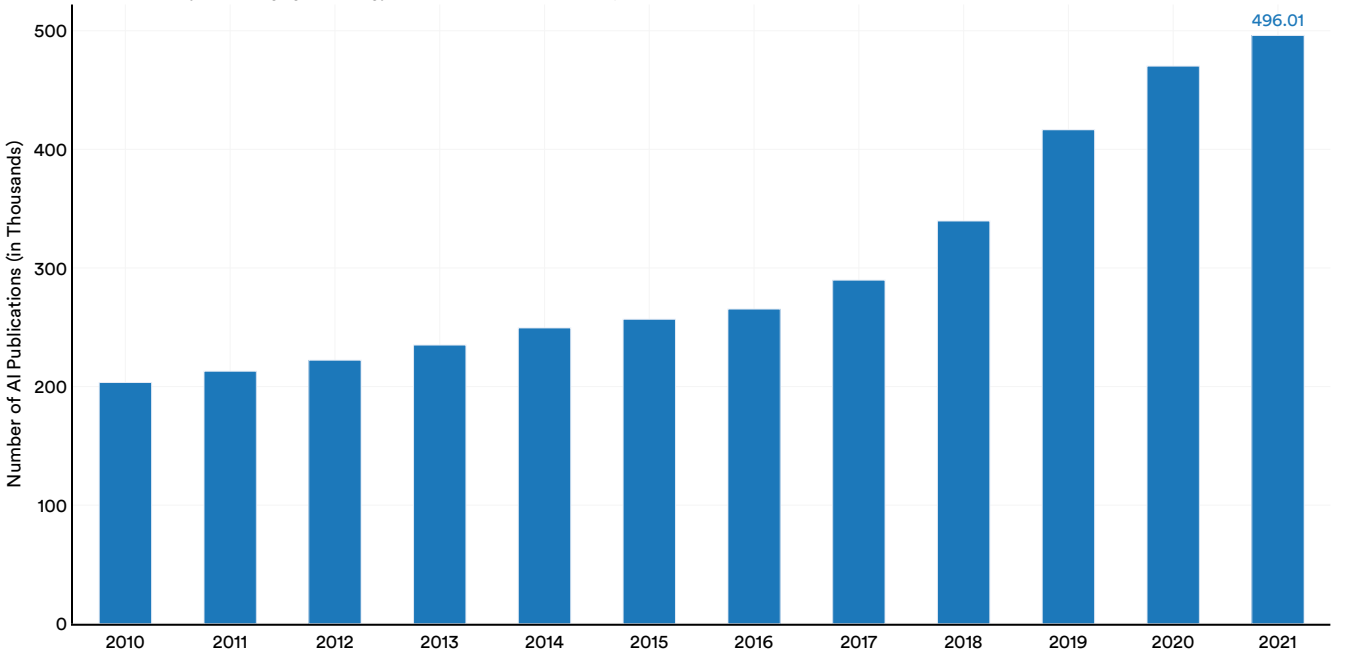

Figure 1.1.1

1 See the Appendix for more information on CSET's methodology. For more on the challenge of defining AI and correctly capturing relevant bibliometric data, see the AI Index team's discussion in the paper "Measurement in AI Policy: Opportunities and Challenges."

By Type of Publication

Figure 1.1.2 shows the types of AI publications released globally over time. In 2021, 60% of all published AI documents were journal articles, 17% were conference papers, and 13% were repository submissions. Books, book chapters, theses, and unknown document types made up the remaining 10% of publications. While journal and repository publications have grown 3 and 26.6 times, respectively, in the past 12 years, the number of conference papers has declined since 2019.

Number of AI Publications by Type, 2010–21

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

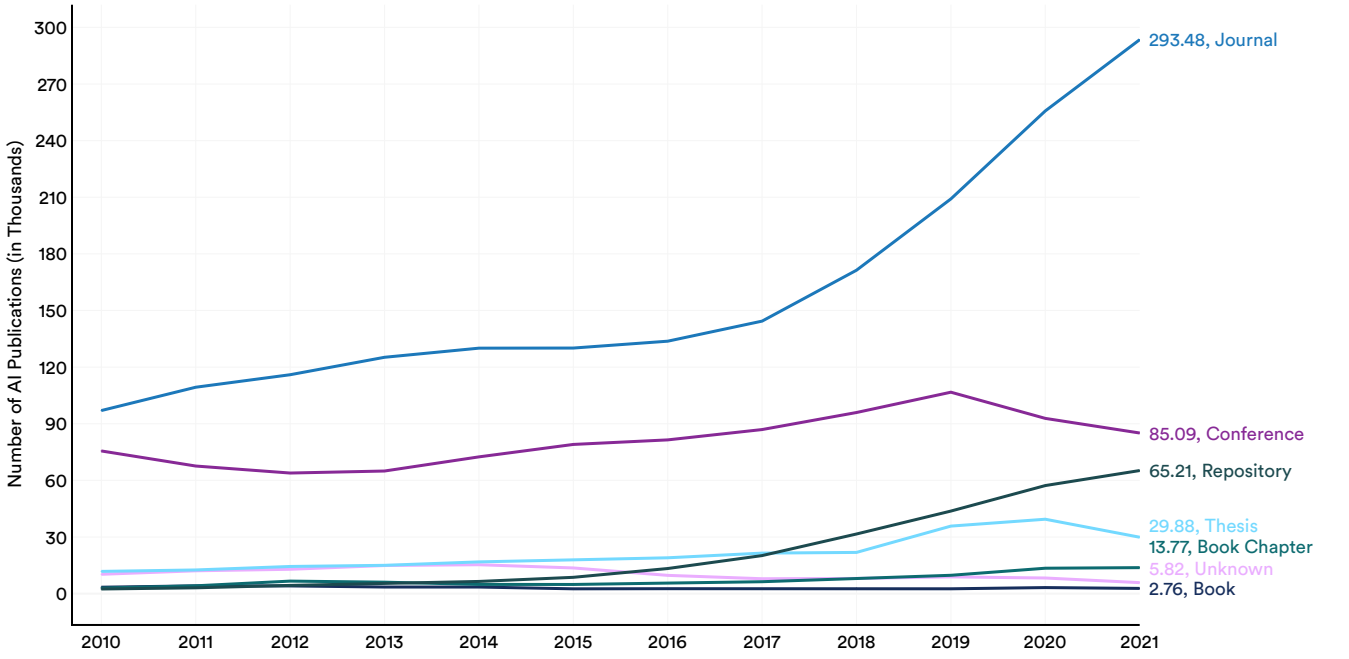

Figure 1.1.2

By Field of Study

Figure 1.1.3 shows that publications in pattern recognition and machine learning have experienced the sharpest growth in the last half decade. Since 2015, the number of pattern recognition papers has roughly doubled while the number of machine learning papers has roughly quadrupled. Following those two topic areas, in 2021, the next most published AI fields of study were computer vision (30,075), algorithm (21,527), and data mining (19,181).

Number of AI Publications by Field of Study (Excluding Other AI), 2010–21

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

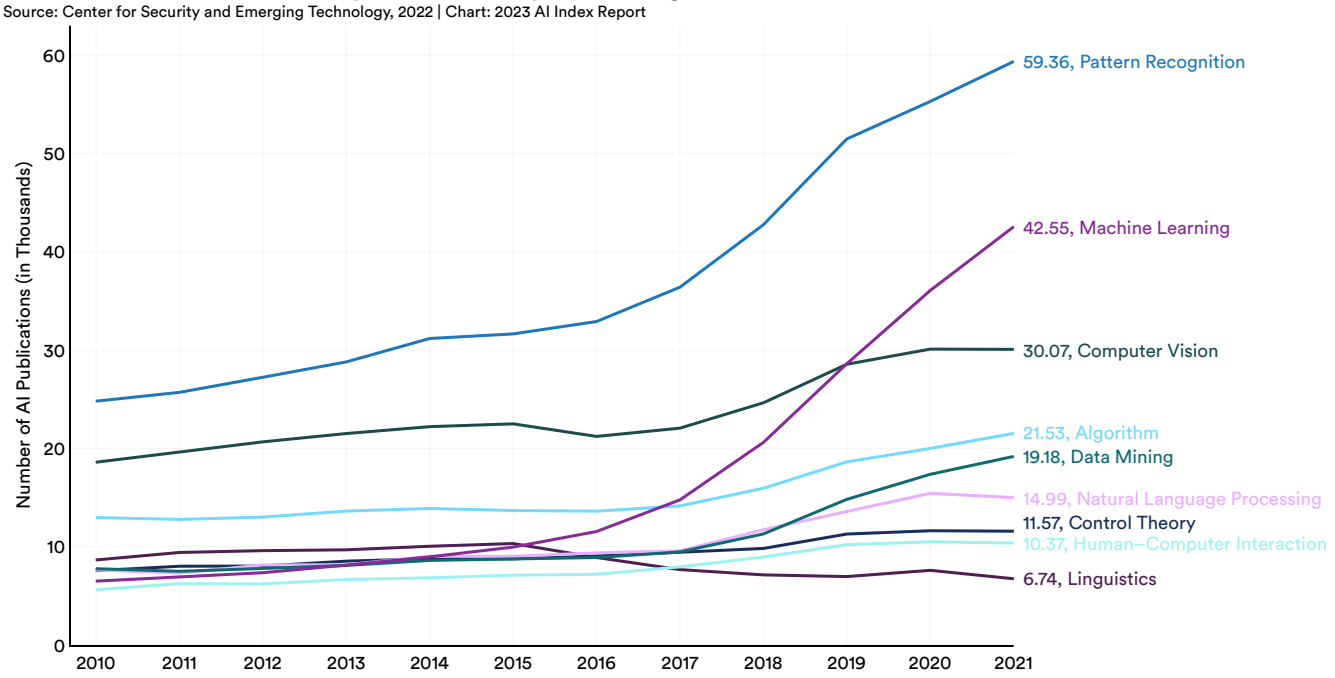

Figure 1.1.3

By Sector

This section shows the number of AI publications affiliated with education, government, industry, nonprofit, and other sectors—first globally (Figure 1.1.4), then looking at the United States, China, and the European Union plus the United Kingdom (Figure 1.1.5).[2] The education sector dominates in each region. The level of industry participation is highest in the United States, then in the European Union. Since 2010, the share of education AI publications has been dropping in each region.

AI Publications (% of Total) by Sector, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

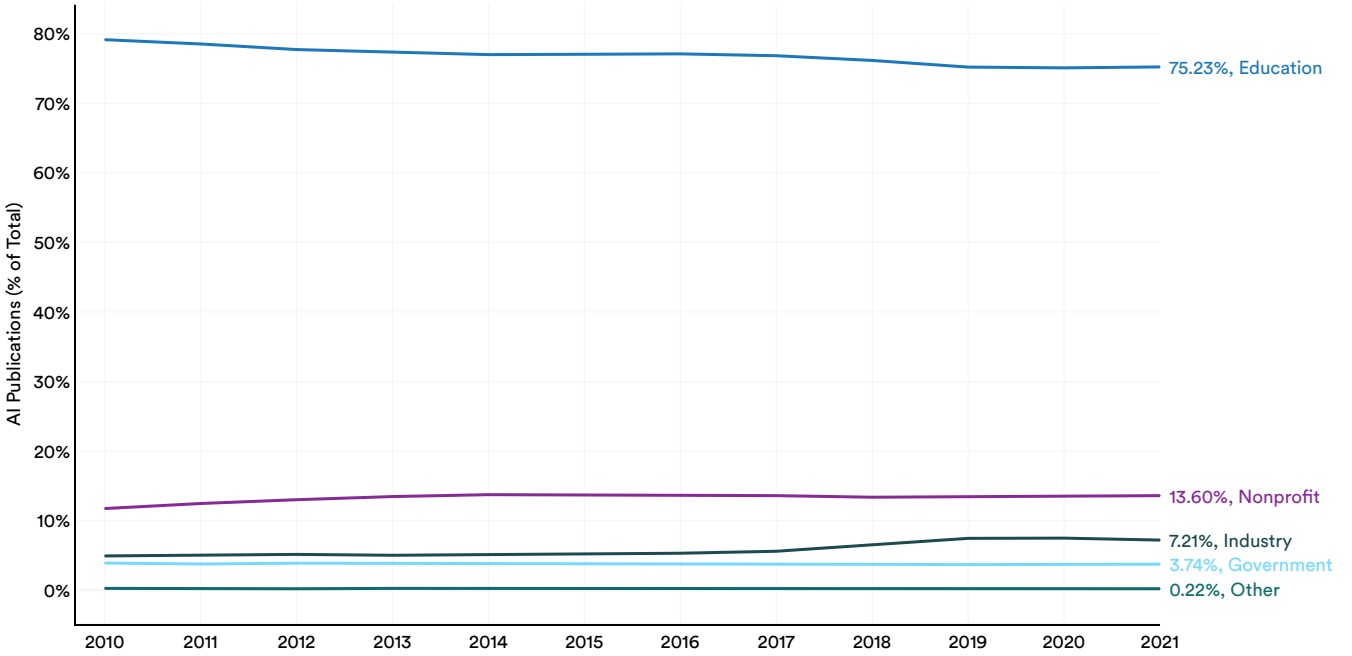

Figure 1.1.4

2 The categorization is adapted based on the Global Research Identifier Database (GRID). Healthcare, including hospitals and facilities, is included under nonprofit. Publications affiliated with state-sponsored universities are included in the education sector.

AI Publications (% of Total) by Sector and Geographic Area, 2021

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

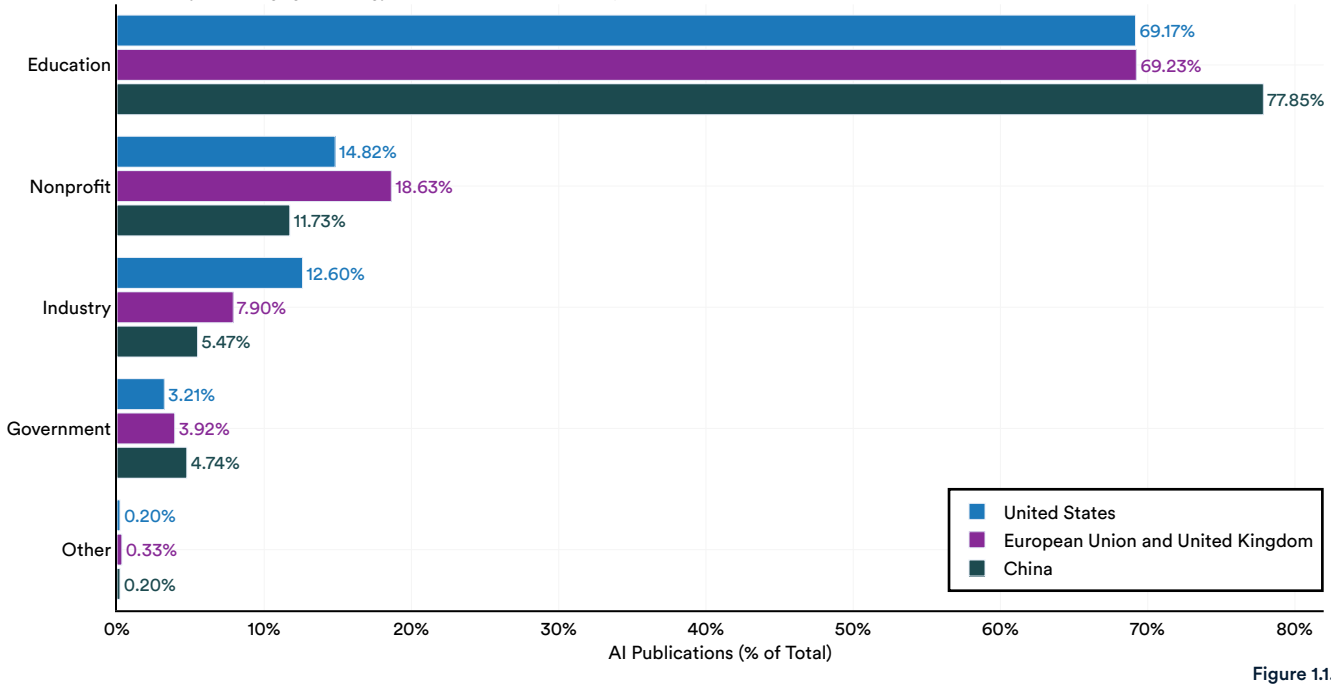

Figure 1.1.5

Cross-Country Collaboration

Cross-border collaborations between academics, researchers, industry experts, and others are a key component of modern STEM (science, technology, engineering, and mathematics) development that accelerate the dissemination of new ideas and the growth of research teams. Figures 1.1.6 and 1.1.7 depict the top cross-country AI collaborations from 2010 to 2021. CSET counted cross-country collaborations as distinct pairs of countries across authors for each publication (e.g., four U.S. and four Chinese-affiliated authors on a single publication are counted as one U.S.-China collaboration; two publications between the same authors count as two collaborations).

By far, the greatest number of collaborations in the past 12 years took place between the United States and China, increasing roughly four times since 2010. However the total number of U.S.-China collaborations only increased by 2.1% from 2020 to 2021, the smallest year-over-year growth rate since 2010.

The next largest set of collaborations was between the United Kingdom and both China and the United States. In 2021, the number of collaborations between the United States and China was 2.5 times greater than between the United Kingdom and China.

United States and China Collaborations in AI Publications, 2010–21

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

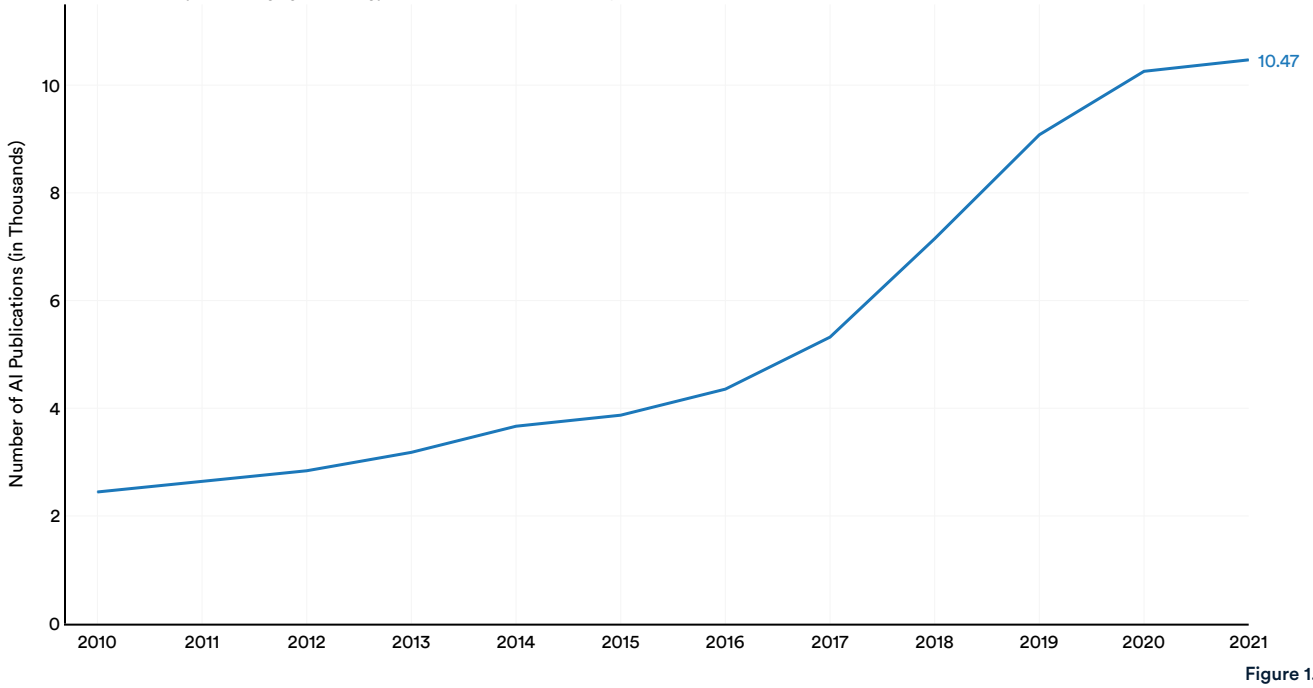

Figure 1.1.6

Cross-Country Collaborations in AI Publications (Excluding U.S. and China), 2010–21

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

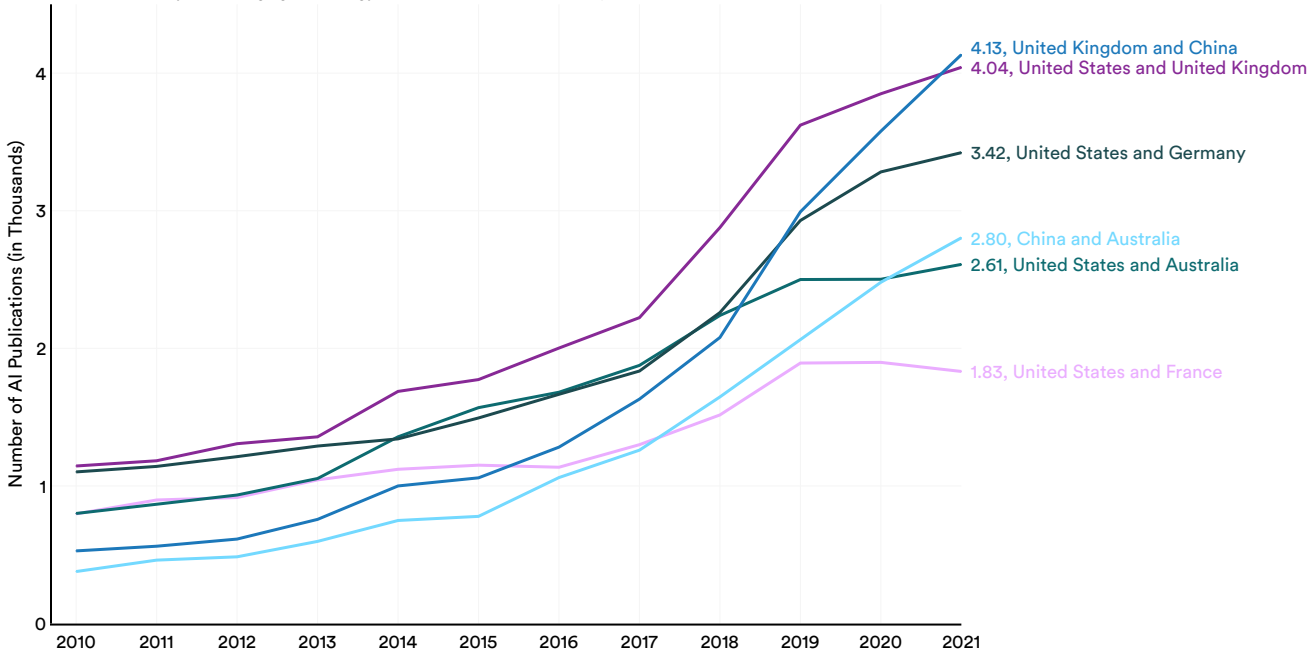

Figure 1.1.7

Cross-Sector Collaboration

The increase in AI research outside of academia has broadened and grown collaboration across sectors in general. Figure 1.1.8 shows that in 2021 educational institutions and nonprofits (32,551) had the greatest number of collaborations; followed by industry and educational institutions (12,856); and educational and government institutions (8,913). Collaborations between educational institutions and industry have been among the fastest growing, increasing 4.2 times since 2010.

Cross-Sector Collaborations in AI Publications, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

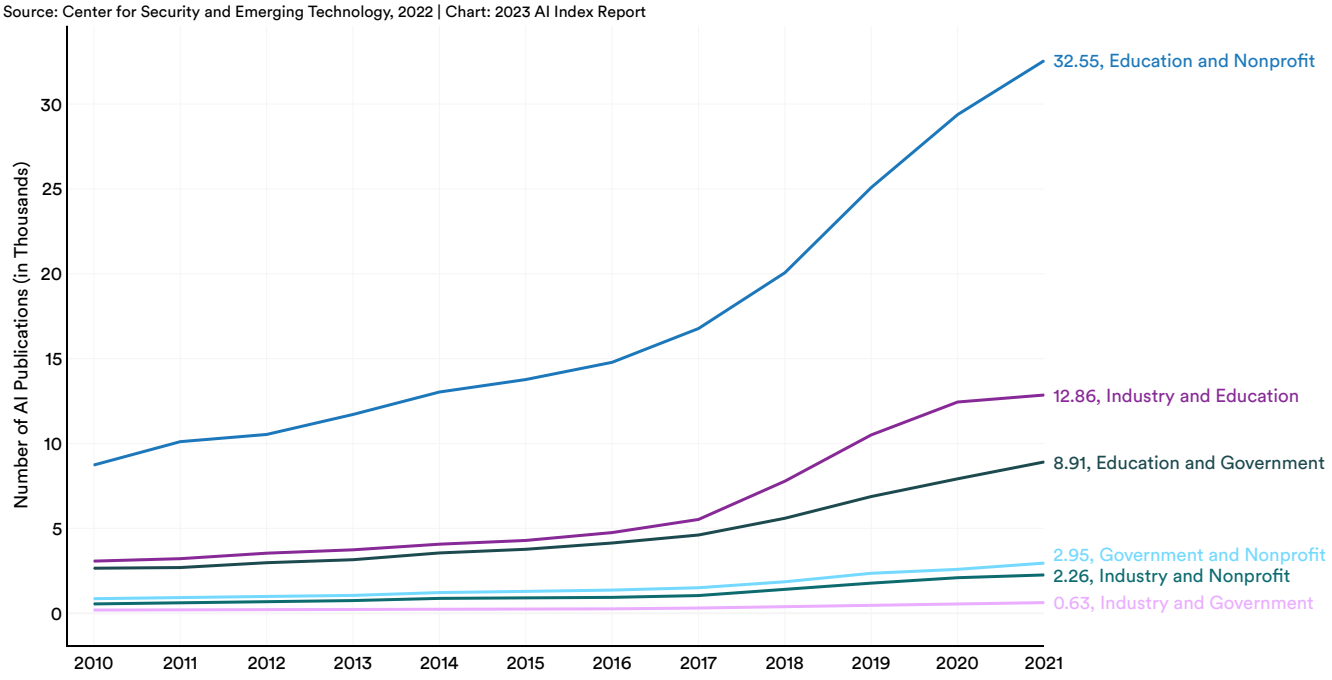

Figure 1.1.8

AI Journal Publications

Overview

After growing only slightly from 2010 to 2015, the number of AI journal publications grew around 2.3 times since 2015. From 2020 to 2021, they increased 14.8% (Figure 1.1.9).

Number of AI Journal Publications, 2010–21

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

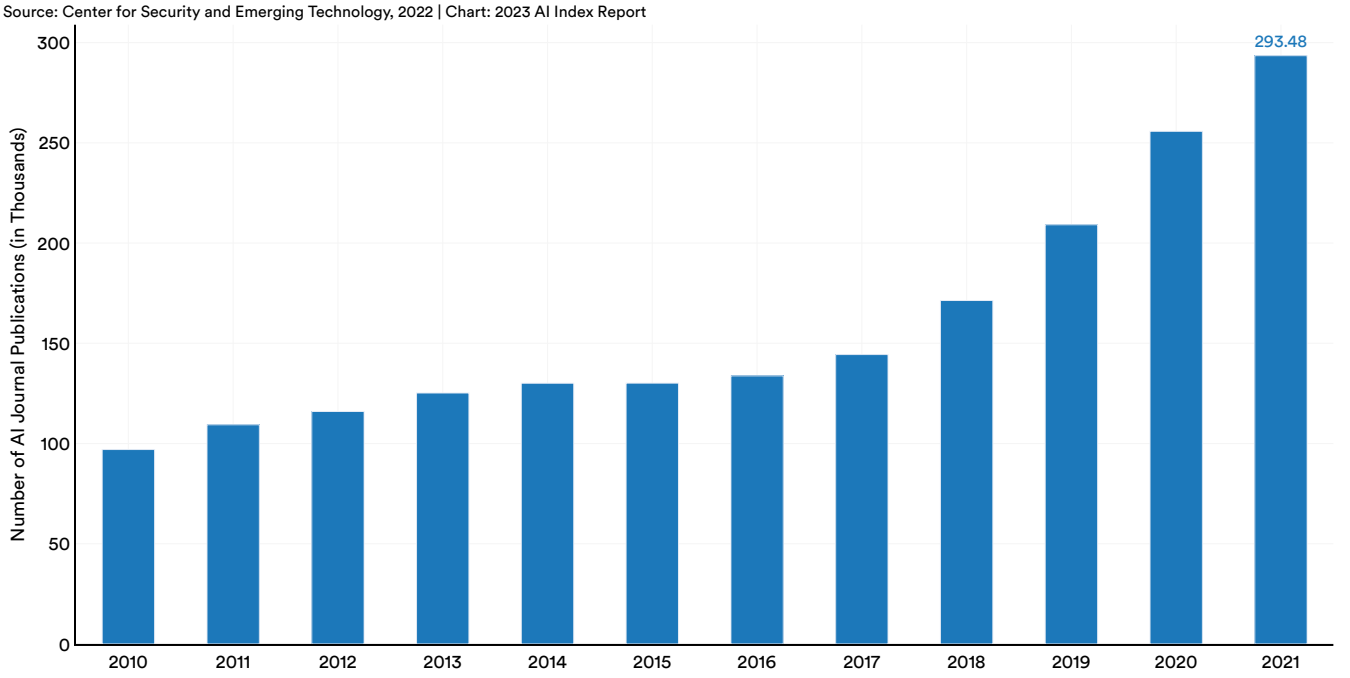

Figure 1.1.9

By Region[3]

Figure 1.1.10 shows the share of AI journal publications by region between 2010 and 2021. In 2021, East Asia and the Pacific led with 47.1%, followed by Europe and Central Asia (17.2%), and then North America (11.6%). Since 2019, the share of publications from East Asia and the Pacific; Europe and Central Asia; as well as North America have been declining. During that period, there has been an increase in publications from other regions such as South Asia; and the Middle East and North Africa.

AI Journal Publications (% of World Total) by Region, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

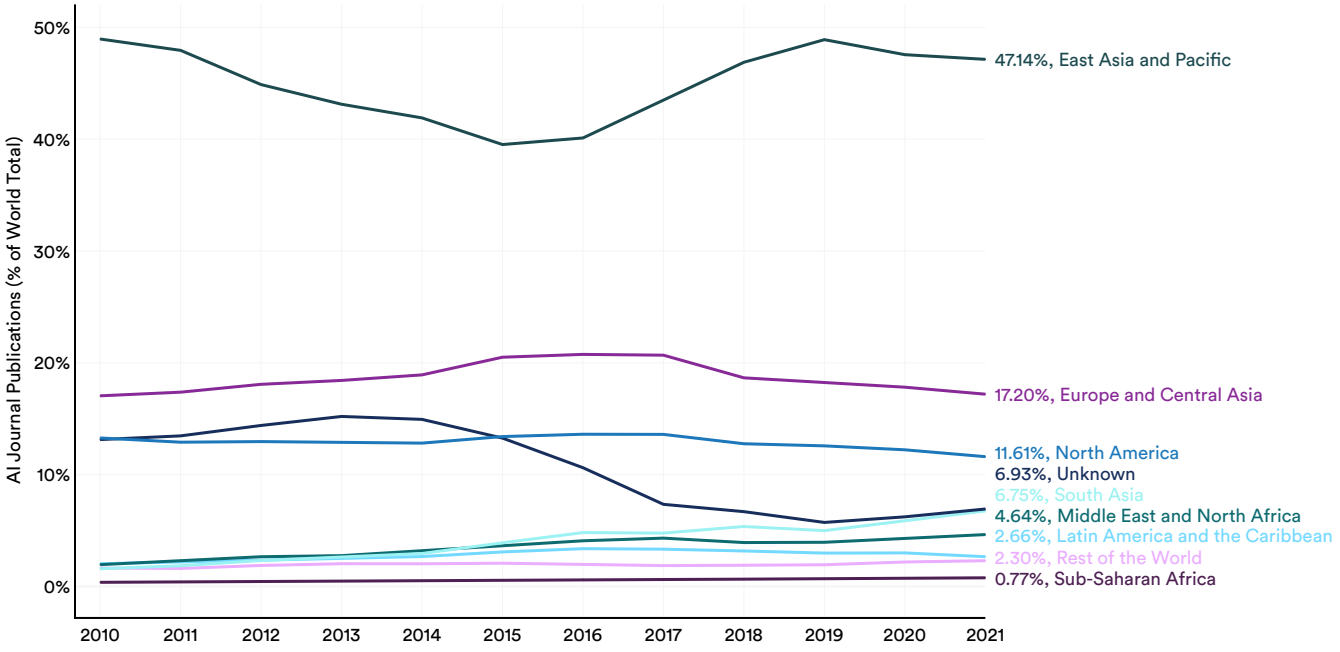

Figure 1.1.10

3 Regions in this chapter are classified according to the World Bank analytical grouping.

By Geographic Area[4]

Figure 1.1.11 breaks down the share of AI journal publications over the past 12 years by geographic area. This year's AI Index included India in recognition of the increasingly important role it plays in the AI ecosystem. China has remained the leader throughout, with 39.8% in 2021, followed by the European Union and the United Kingdom (15.1%), then the United States (10.0%). The share of Indian publications has been steadily increasing—from 1.3% in 2010 to 5.6% in 2021.

AI Journal Publications (% of World Total) by Geographic Area, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

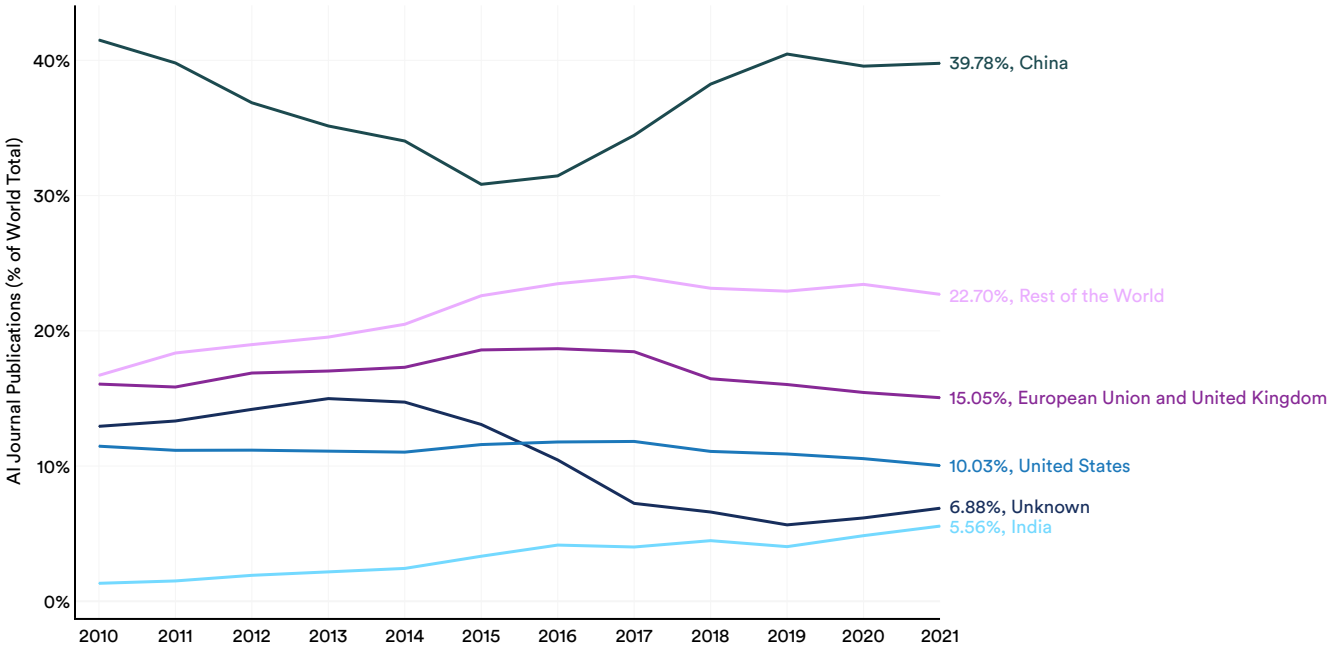

Figure 1.1.11

4 In this chapter we use "geographic area" based on CSET's classifications, which are disaggregated not only by country, but also by territory. Further, we count the European Union and the United Kingdom as a single geographic area to reflect the regions' strong history of research collaboration.

Citations

China's share of citations in AI journal publications has gradually increased since 2010, while those of the European Union and the United Kingdom, as well as those of the United States, have decreased (Figure 1.1.12). China, the European Union and the United Kingdom, and the United States accounted for 65.7% of the total citations in the world.

AI Journal Citations (% of World Total) by Geographic Area, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

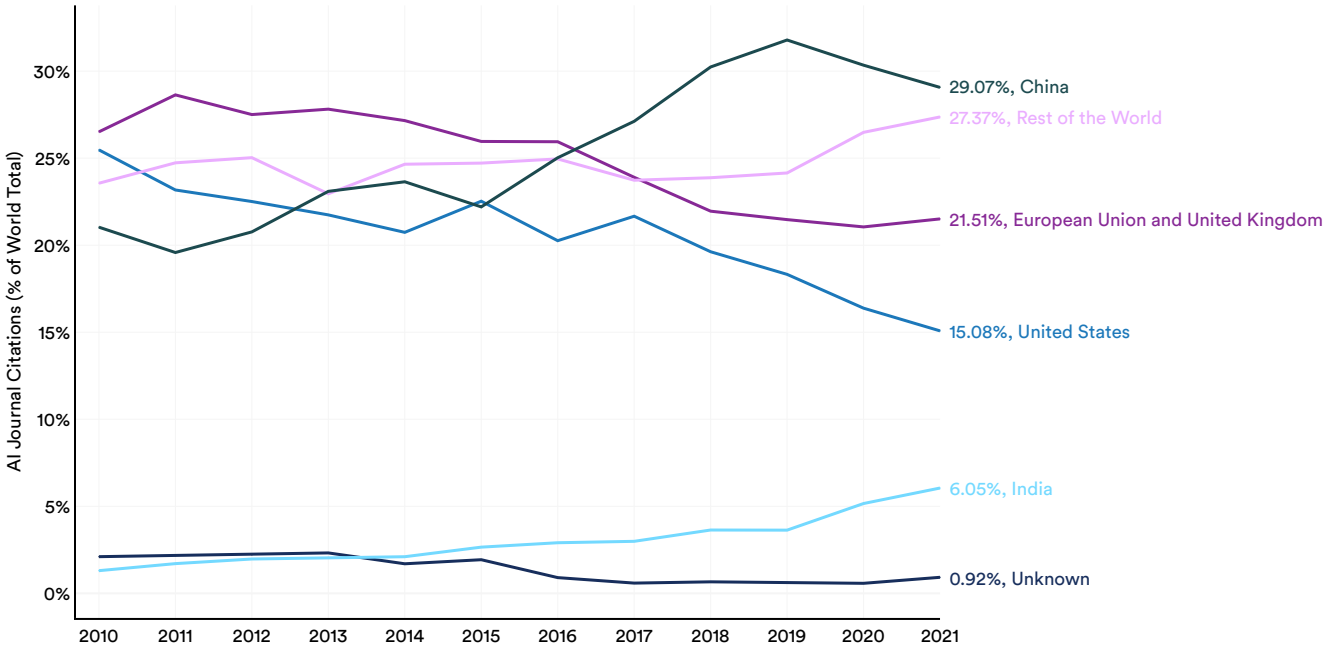

Figure 1.1.12

AI Conference Publications

Overview

The number of AI conference publications peaked in 2019, and fell 20.4% below the peak in 2021 (Figure 1.1.13). The total number of 2021 AI conference publications, 85,094, was marginally greater than the 2010 total of 75,592.

Number of AI Conference Publications, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

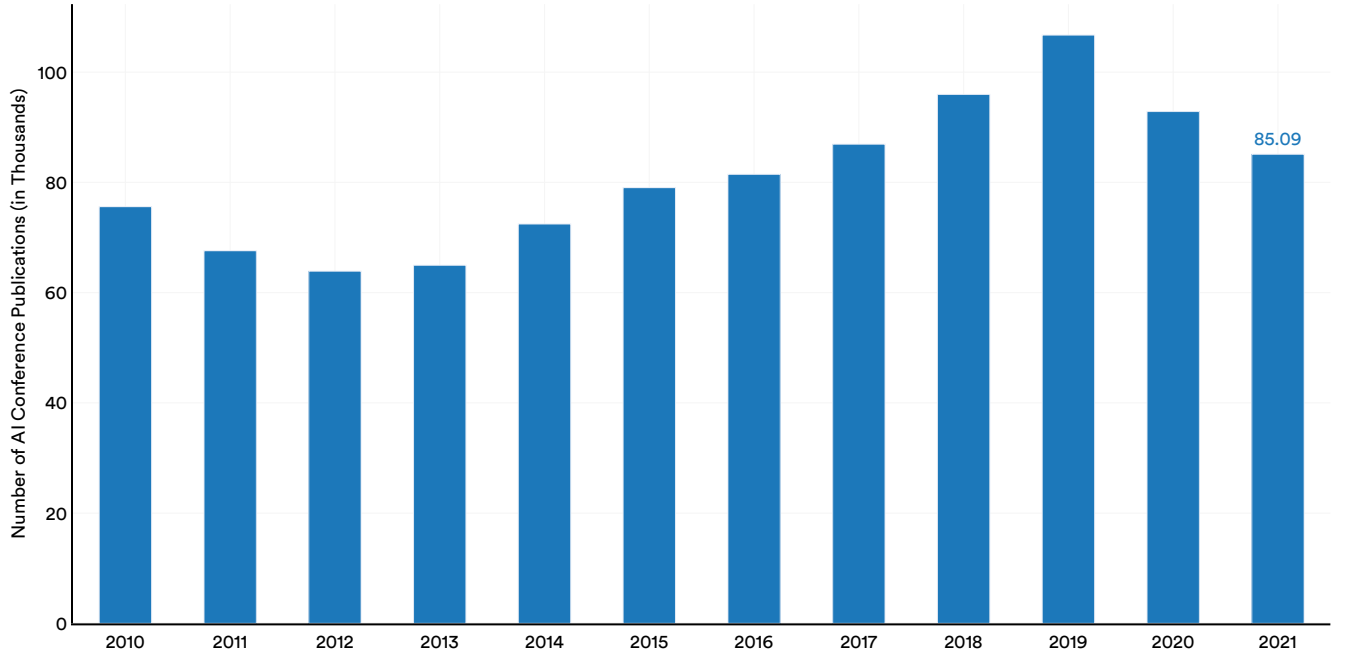

Figure 1.1.13

By Region

Figure 1.1.14 shows the number of AI conference publications by region. As with the trend in journal publications, East Asia and the Pacific; Europe and Central Asia; and North America account for the world's highest numbers of AI conference publications. Specifically, the share represented by East Asia and the Pacific continues to rise, accounting for 36.7% in 2021, followed by Europe and Central Asia (22.7%), and then North America (19.6%). The percentage of AI conference publications in South Asia saw a noticeable rise in the past 12 years, growing from 3.6% in 2010 to 8.5% in 2021.

AI Conference Publications (% of World Total) by Region, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

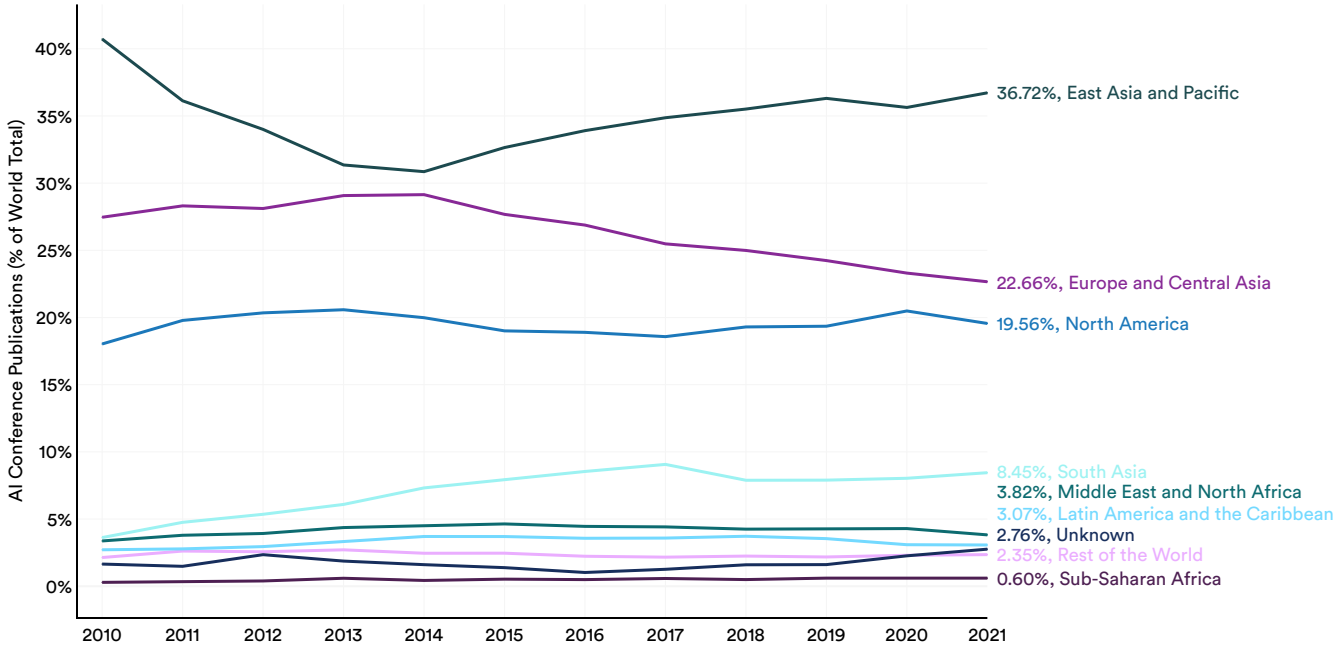

Figure 1.1.14

By Geographic Area

In 2021, China produced the greatest share of the world's AI conference publications at 26.2%, having overtaken the European Union and the United Kingdom in 2017. The European Union plus the United Kingdom followed at 20.3%, and the United States came in third at 17.2% (Figure 1.1.15). Mirroring trends seen in other parts of the research and development section, India's share of AI conference publications is also increasing.

AI Conference Publications (% of World Total) by Geographic Area, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

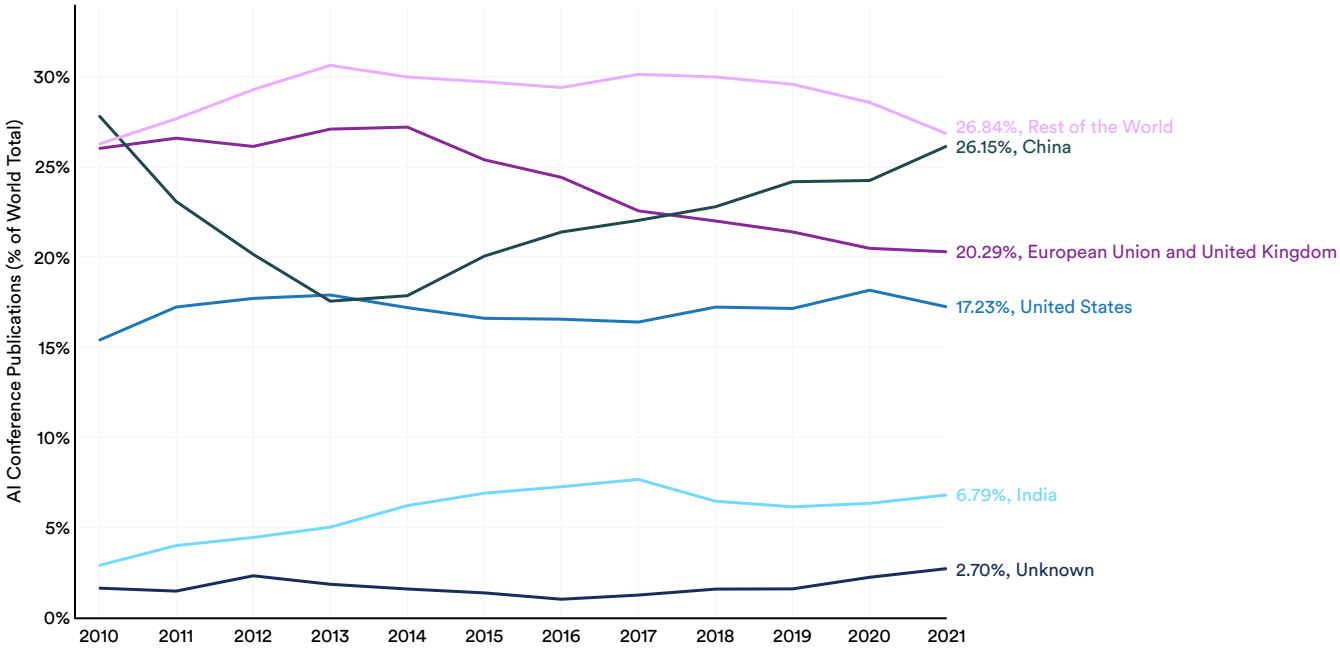

Figure 1.1.15

Citations

Despite China producing the most AI conference publications in 2021, Figure 1.1.16 shows that the United States had the greatest share of AI conference citations, with 23.9%, followed by China's 22.0%. However, the gap between American and Chinese AI conference citations is narrowing.

AI Conference Citations (% of World Total) by Geographic Area, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

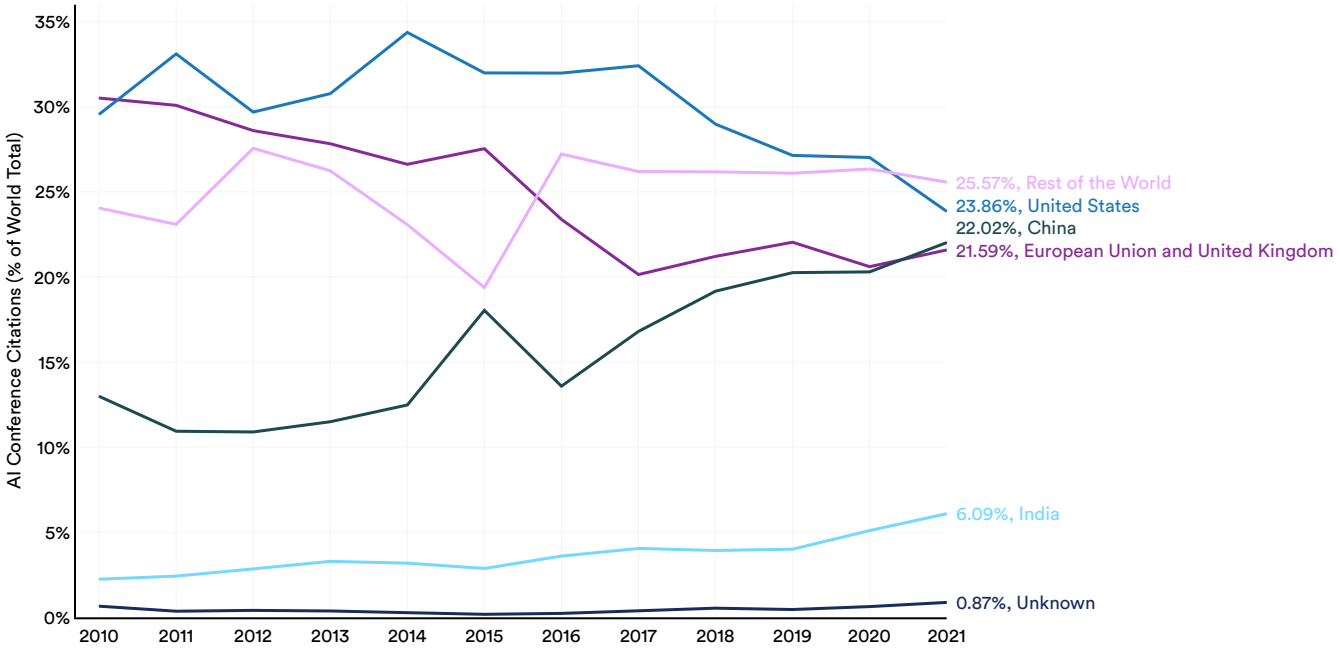

Figure 1.1.16

AI Repositories

Overview

Publishing pre-peer-reviewed papers on repositories of electronic preprints (such as arXiv and SSRN) has become a popular way for AI researchers to disseminate their work outside traditional avenues for publication. These repositories allow researchers to share their findings before submitting them to journals and conferences, thereby accelerating the cycle of information discovery. The number of AI repository publications grew almost 27 times in the past 12 years (Figure 1.1.17).

Number of AI Repository Publications, 2010–21

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

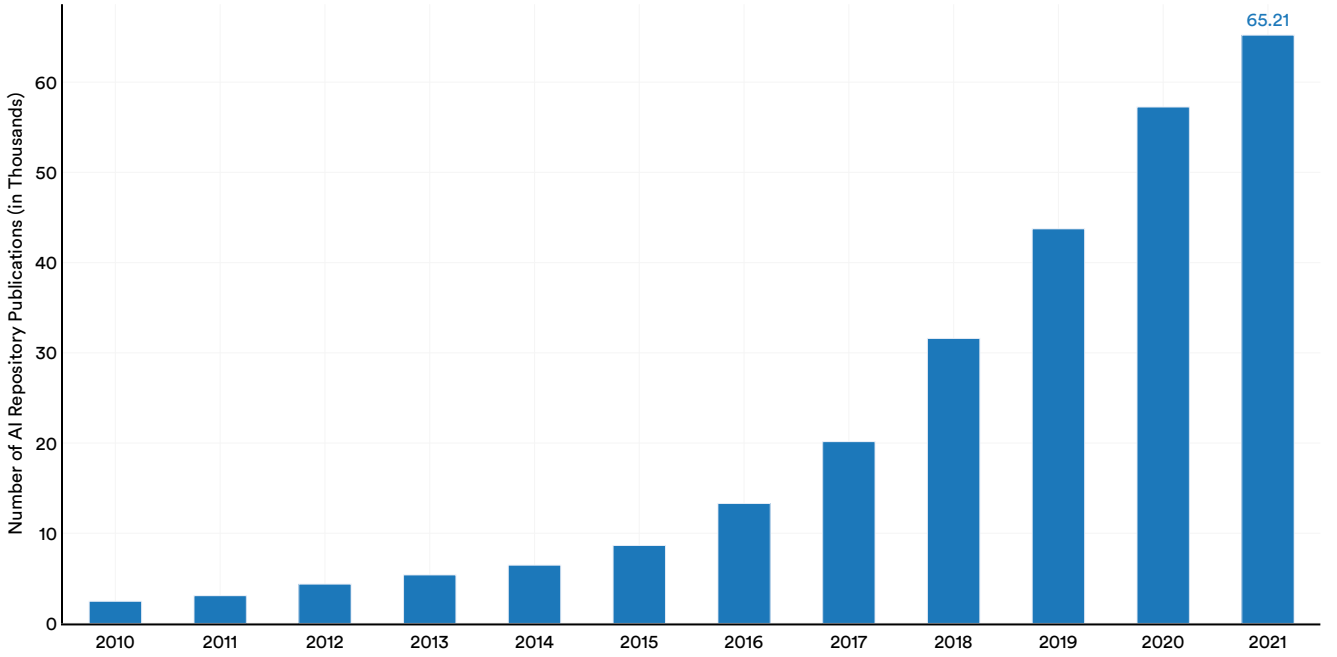

Figure 1.1.17

By Region

Figure 1.1.18 shows that North America has maintained a steady lead in the world share of AI repository publications since 2016. Since 2011, the share of repository publications from Europe and Central Asia has declined. The share represented by East Asia and the Pacific has grown significantly since 2010 and continued growing from 2020 to 2021, a period in which the year-over-year share of North American as well European and Central Asian repository publications declined.

AI Repository Publications (% of World Total) by Region, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

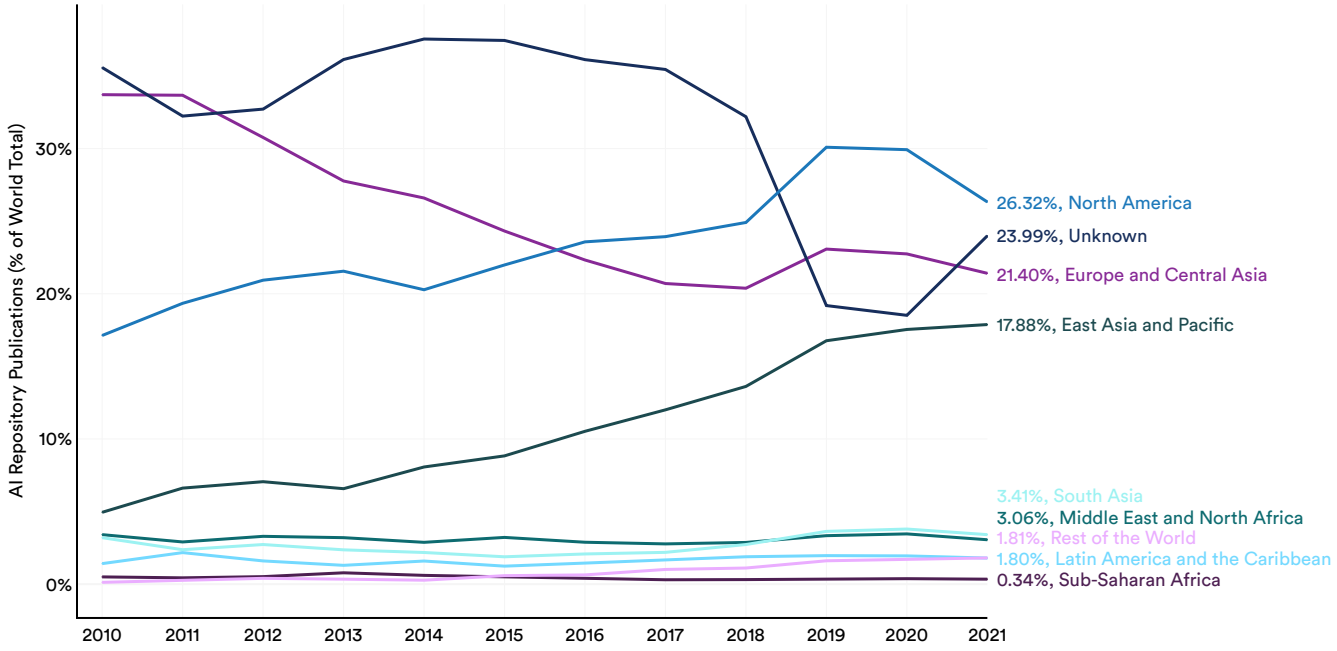

Figure 1.1.18

By Geographic Area

While the United States has held the lead in the percentage of global AI repository publications since 2016, China is catching up, while the European Union plus the United Kingdom's share continues to drop (Figure 1.1.19). In 2021, the United States accounted for 23.5% of the world's AI repository publications, followed by the European Union plus the United Kingdom (20.5%), and then China (11.9%).

AI Repository Publications (% of World Total) by Geographic Area, 2010–21

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

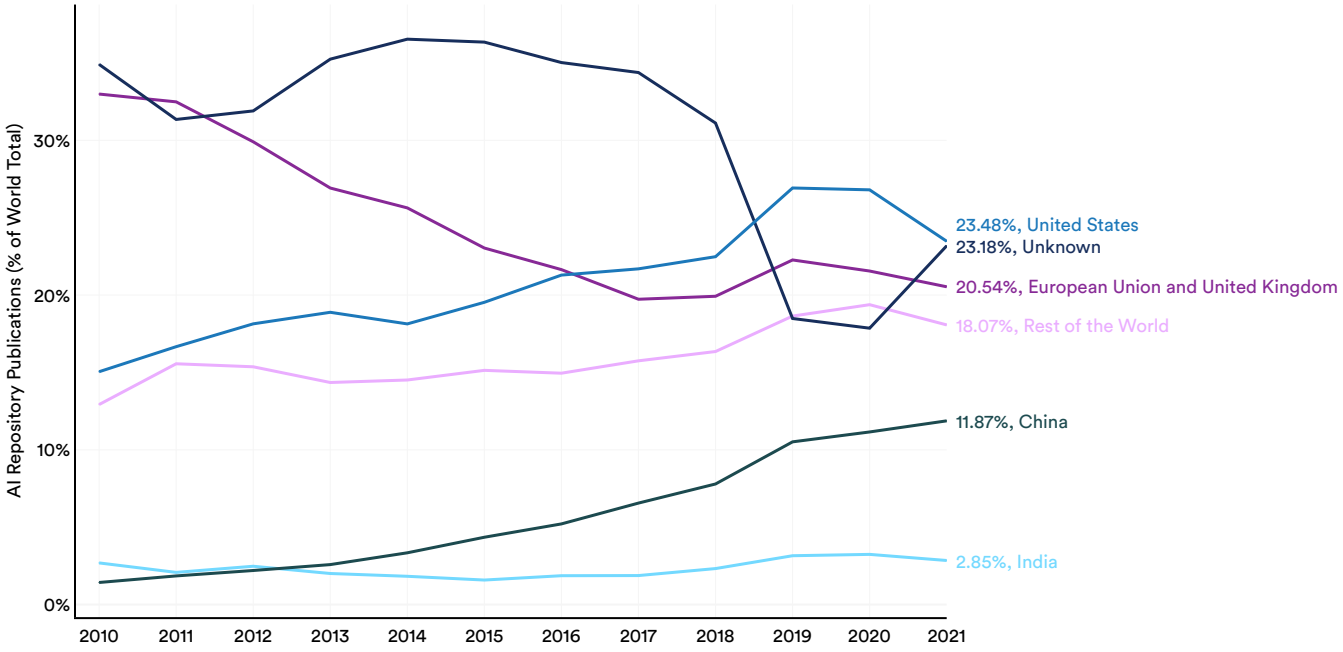

Figure 1.1.19

Citations

In the citations of AI repository publications, Figure 1.1.20 shows that in 2021 the United States topped the list with 29.2% of overall citations, maintaining a dominant lead over the European Union plus the United Kingdom (21.5%), as well as China (21.0%).

AI Repository Citations (% of World Total) by Geographic Area, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

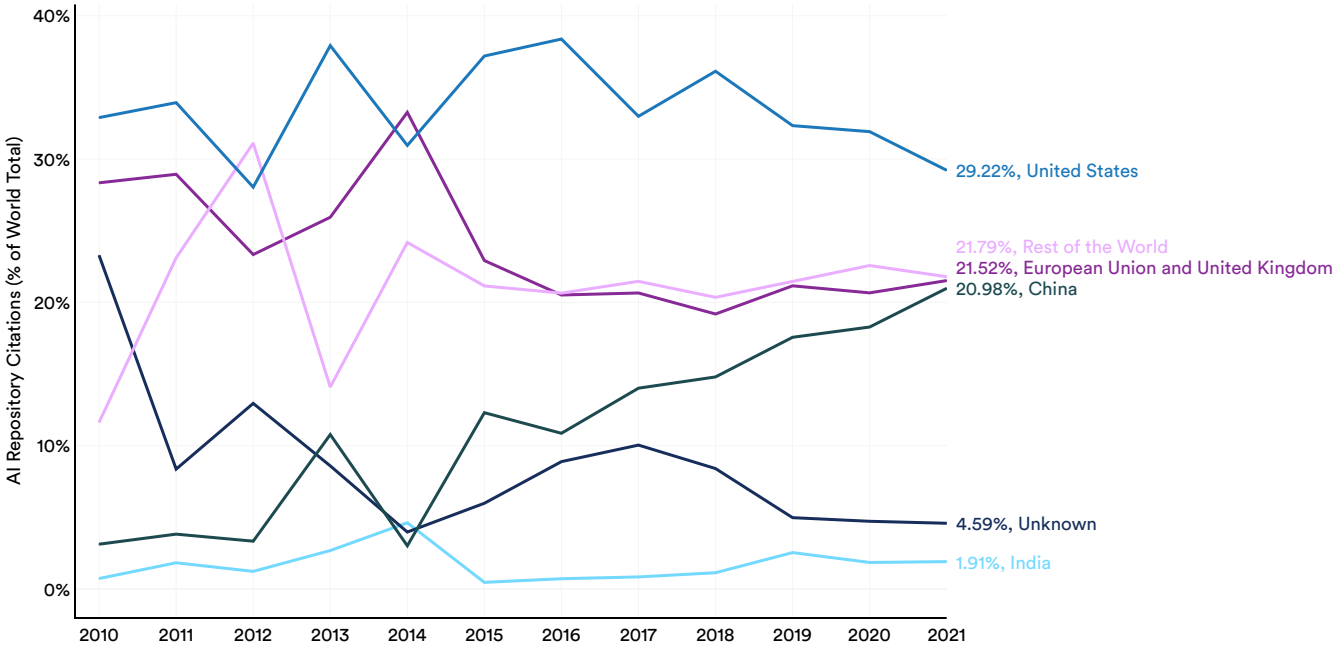

Figure 1.1.20

Narrative Highlight:
Top Publishing Institutions

All Fields

Since 2010, the institution producing the greatest number of total AI papers has been the Chinese Academy of Sciences (Figure 1.1.21). The next top four are all Chinese universities: Tsinghua University, the University of the Chinese Academy of Sciences, Shanghai Jiao Tong University, and Zhejiang University.[5] The total number of publications released by each of these institutions in 2021 is displayed in Figure 1.1.22.

Top Ten Institutions in the World in 2021 Ranked by Number of AI Publications in All Fields, 2010–21
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

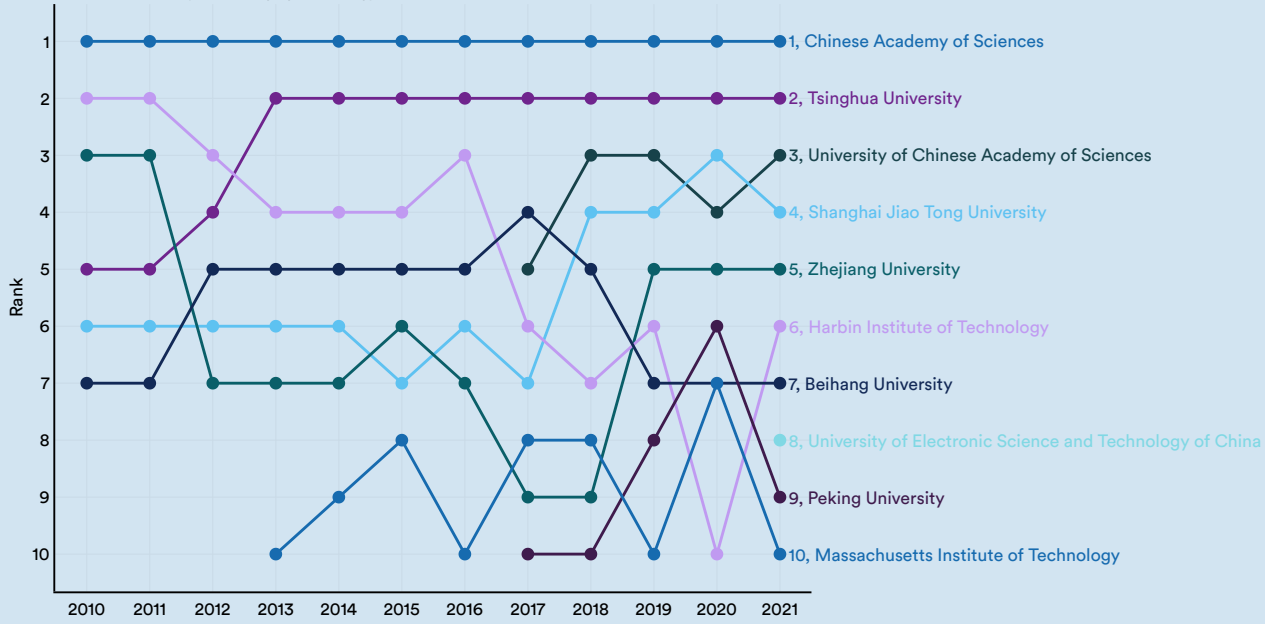

Figure 1.1.21

5 It is important to note that many Chinese research institutions are large, centralized organizations with thousands of researchers. It is therefore not entirely surprising that, purely by the metric of publication count, they outpublish most non-Chinese institutions.

Narrative Highlight:
Top Publishing Institutions (cont'd)

Top Ten Institutions in the World by Number of AI Publications in All Fields, 2021
Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

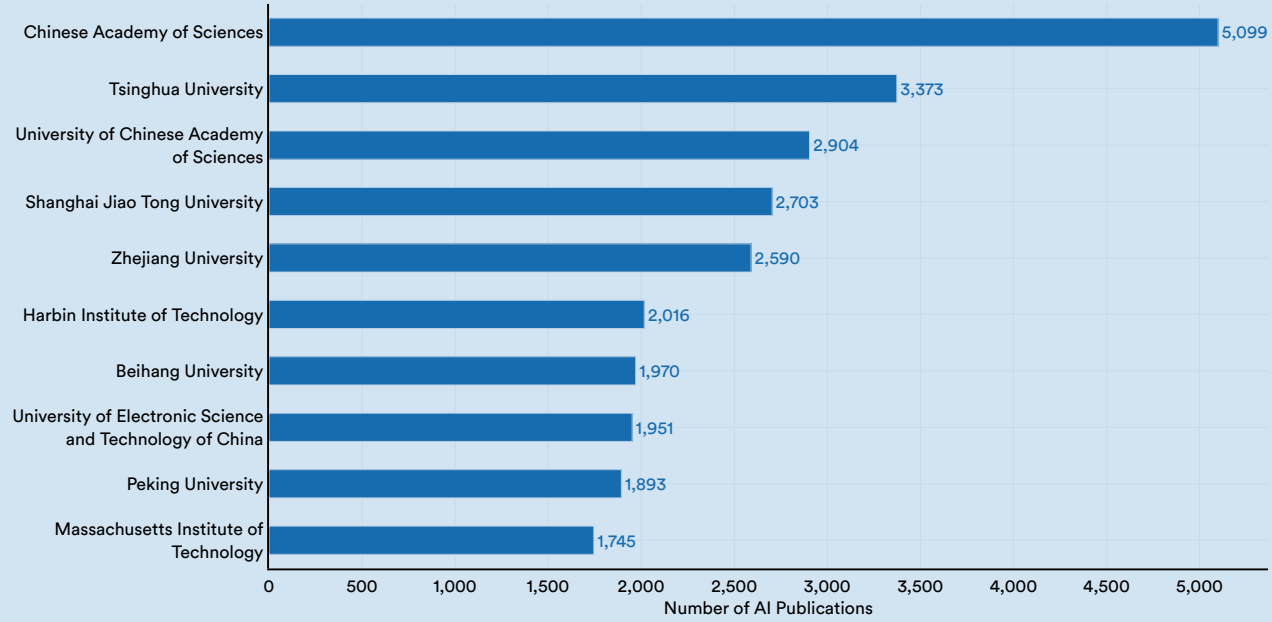

Figure 1.1.22

Narrative Highlight:

Top Publishing Institutions (cont'd)

Computer Vision

In 2021, the top 10 institutions publishing the greatest number of AI computer vision publications were all Chinese (Figure 1.1.23). The Chinese Academy of Sciences published the largest number of such publications, with a total of 562.

Top Ten Institutions in the World by Number of AI Publications in Computer Vision, 2021

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

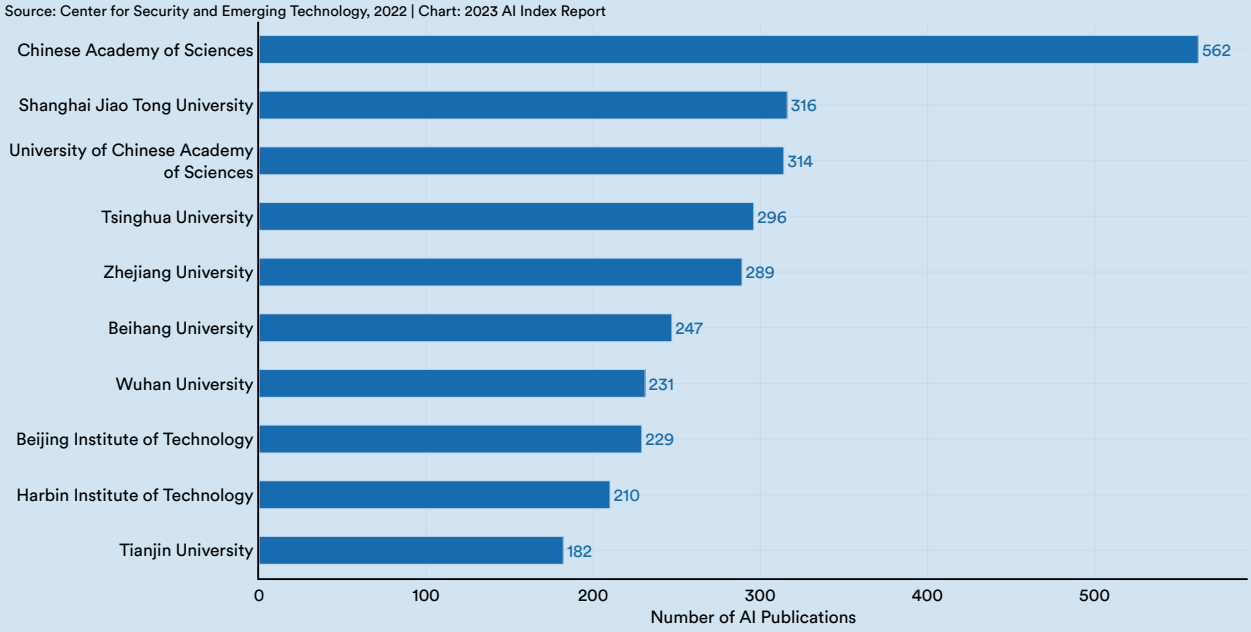

Figure 1.1.23

Narrative Highlight:

Top Publishing Institutions (cont'd)

Natural Language Processing

American institutions are represented to a greater degree in the share of top NLP publishers (Figure 1.1.24). Although the Chinese Academy of Sciences was again the world's leading institution in 2021 (182 publications), Carnegie Mellon took second place (140 publications), followed by Microsoft (134). In addition, 2021 was the first year Amazon and Alibaba were represented among the top-ten largest publishing NLP institutions.

Top Ten Institutions in the World by Number of AI Publications in Natural Language Processing, 2021

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

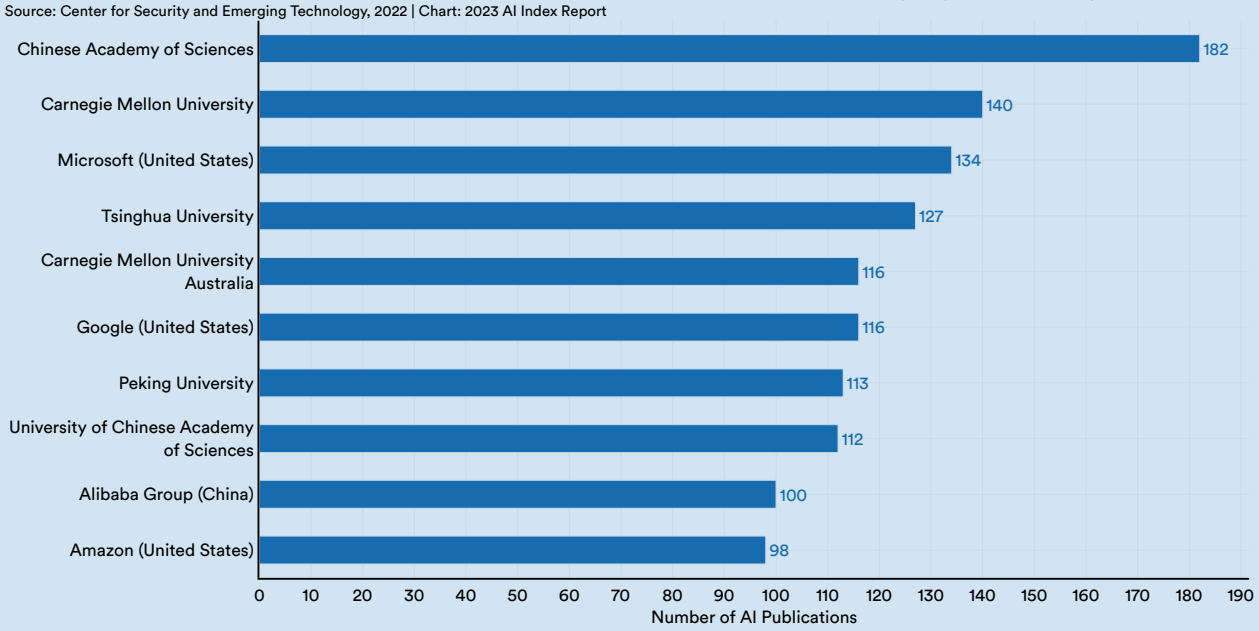

Figure 1.1.24

Narrative Highlight:

Top Publishing Institutions (cont'd)

Speech Recognition

In 2021, the greatest number of speech recognition papers came from the Chinese Academy of Sciences (107), followed by Microsoft (98) and Google (75) (Figure 1.1.25). The Chinese Academy of Sciences reclaimed the top spot in 2021 from Microsoft, which held first position in 2020.

Top Ten Institutions in the World by Number of AI Publications in Speech Recognition, 2021

Source: Center for Security and Emerging Technology, 2022 | Chart: 2023 AI Index Report

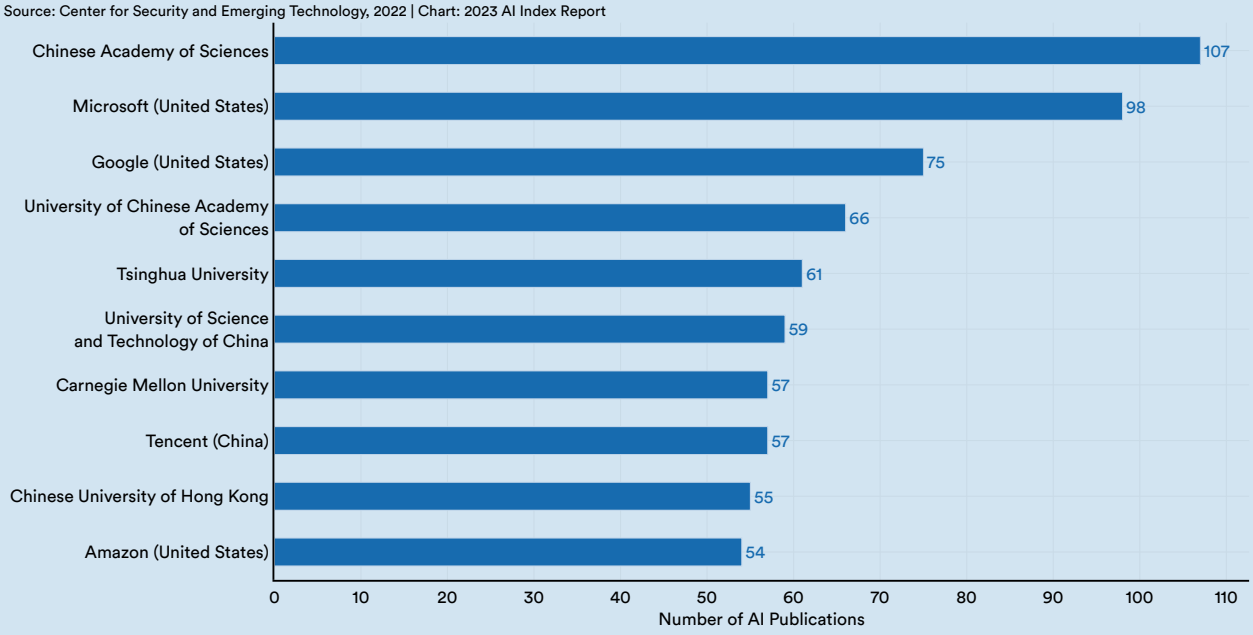

Figure 1.1.25

Epoch AI is a collective of researchers investigating and forecasting the development of advanced AI. Epoch curates a database of significant AI and machine learning systems that have been released since the 1950s. There are different criteria under which the Epoch team decides to include particular AI systems in their database; for example, the system may have registered a state-of-the-art improvement, been deemed to have been historically significant, or been highly cited.

This subsection uses the Epoch database to track trends in significant AI and machine learning systems. The latter half of the chapter includes research done by the AI Index team that reports trends in large language and multimodal models, which are models trained on large amounts of data and adaptable to a variety of downstream applications.

1.2 Trends in Significant Machine Learning Systems

General Machine Learning Systems

The figures below report trends among all machine learning systems included in the Epoch dataset. For reference, these systems are referred to as *significant machine learning systems* throughout the subsection.

System Types

Among the significant AI machine learning systems released in 2022, the most common class of system was language (Figure 1.2.1). There were 23 significant AI language systems released in 2022, roughly six times the number of the next most common system type, multimodal systems.

Number of Significant Machine Learning Systems by Domain, 2022

Source: Epoch, 2022 | Chart: 2023 AI Index Report

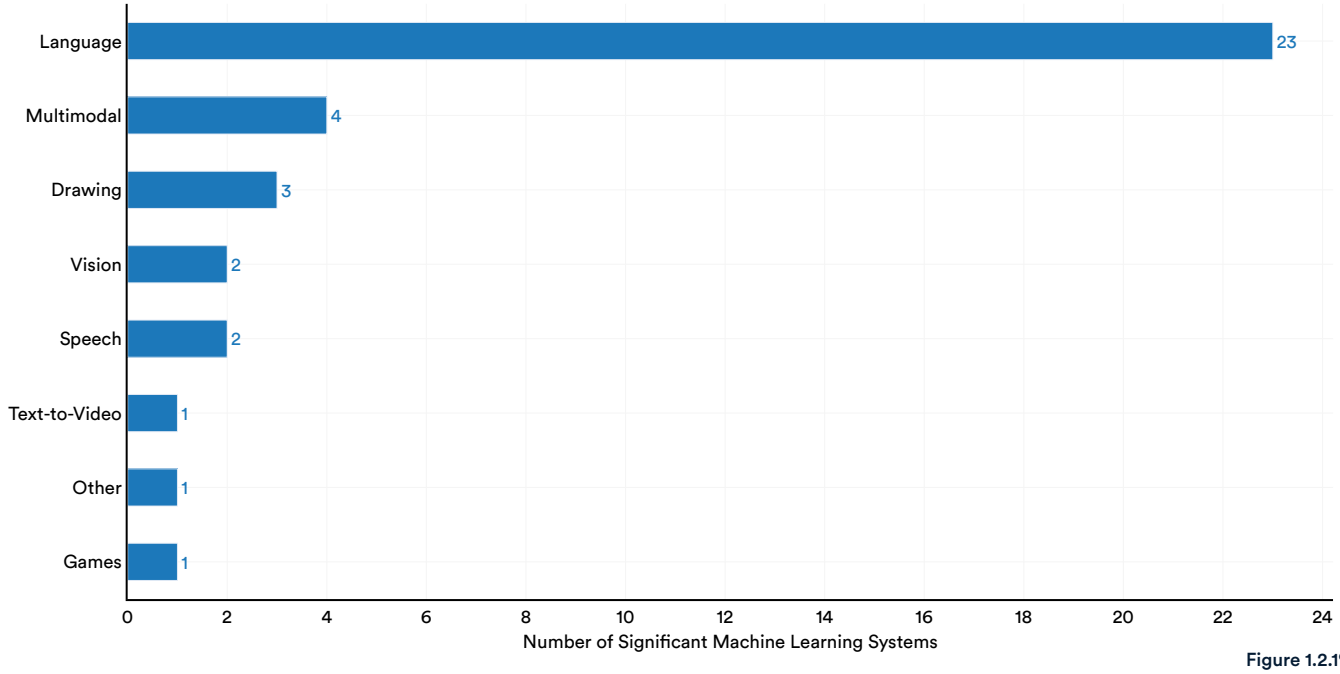

Figure 1.2.1[6]

6 There were 38 total significant AI machine learning systems released in 2022, according to Epoch; however, one of the systems, BaGuaLu, did not have a domain classification and is therefore omitted from Figure 1.2.1.

Sector Analysis

Which sector among industry, academia, or nonprofit has released the greatest number of significant machine learning systems? Until 2014, most machine learning systems were released by academia. Since then, industry has taken over (Figure 1.2.2). In 2022, there were 32 significant industry-produced machine learning systems compared to just three produced by academia. Producing state-of-the-art AI systems increasingly requires large amounts of data, computing power, and money; resources that industry actors possess in greater amounts compared to nonprofits and academia.

Number of Significant Machine Learning Systems by Sector, 2002–22
Source: Epoch, 2022 | Chart: 2023 AI Index Report

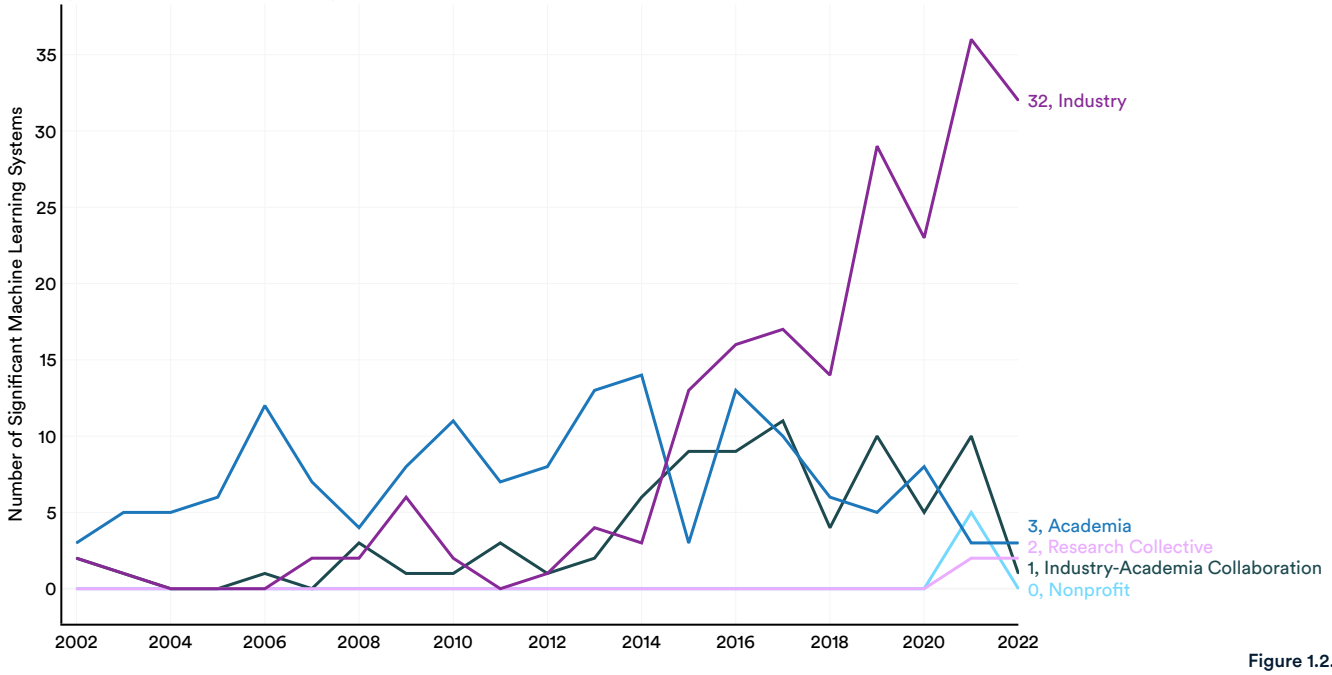

Figure 1.2.2

National Affiliation

In order to paint a picture of AI's evolving geopolitical landscape, the AI Index research team identified the nationality of the authors who contributed to the development of each significant machine learning system in the Epoch dataset.[7]

Systems

Figure 1.2.3 showcases the total number of significant machine learning systems attributed to researchers from particular countries.[8] A researcher is considered to have belonged to the country in which their institution, for example a university or AI-research firm, was headquartered. In 2022, the United States produced the greatest number of significant machine learning systems with 16, followed by the United Kingdom (8) and China (3). Moreover, since 2002 the United States has outpaced the United Kingdom and the European Union, as well as China, in terms of the total number of significant machine learning systems produced (Figure 1.2.4). Figure 1.2.5 displays the total number of significant machine learning systems produced by country since 2002 for the entire world.

Number of Significant Machine Learning Systems by Country, 2022
Source: Epoch and AI Index, 2022 | Chart: 2023 AI Index Report

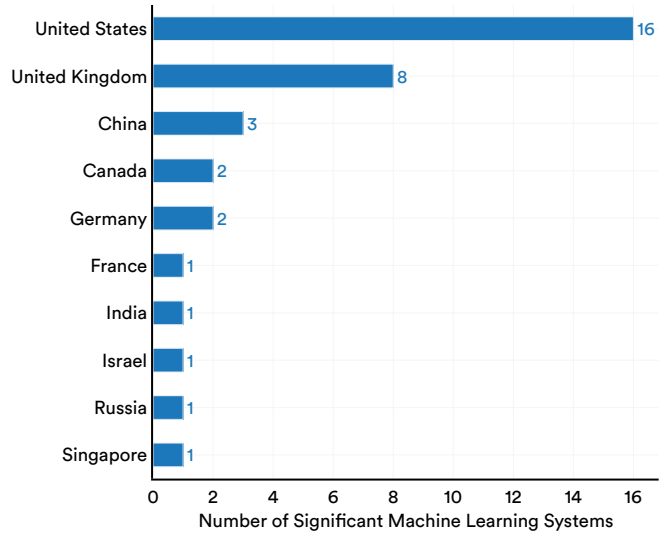

Figure 1.2.3

Number of Significant Machine Learning Systems by Select Geographic Area, 2002–22
Source: Epoch and AI Index, 2022 | Chart: 2023 AI Index Report

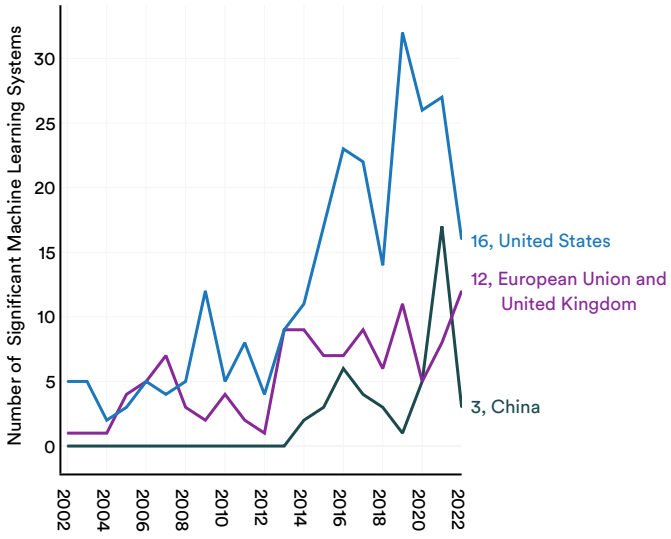

Figure 1.2.4

7 The methodology by which the AI Index identified authors' nationality is outlined in greater detail in the Appendix.
8 A machine learning system is considered to be affiliated with a particular country if at least one author involved in creating the model was affiliated with that country. Consequently, in cases where a system has authors from multiple countries, double counting may occur.

Number of Significant Machine Learning Systems by Country, 2002–22 (Sum)

Source: AI Index, 2022 | Chart: 2023 AI Index Report

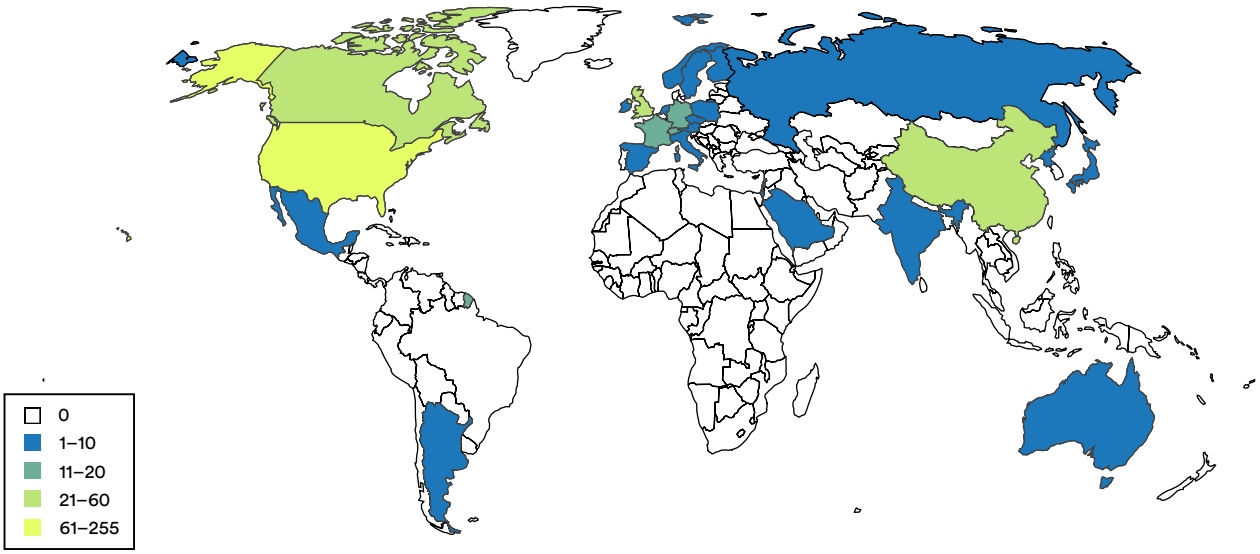

Figure 1.2.5

Authorship

Figures 1.2.6 to 1.2.8 look at the total number of authors, disaggregated by national affiliation, that contributed to the launch of significant machine learning systems. As was the case with total systems, in 2022 the United States had the greatest number of authors producing significant machine learning systems, with 285, more than double that of the United Kingdom and nearly six times that of China (Figure 1.2.6).

Number of Authors of Significant Machine Learning Systems by Country, 2022
Source: Epoch and AI Index, 2022 | Chart: 2023 AI Index Report

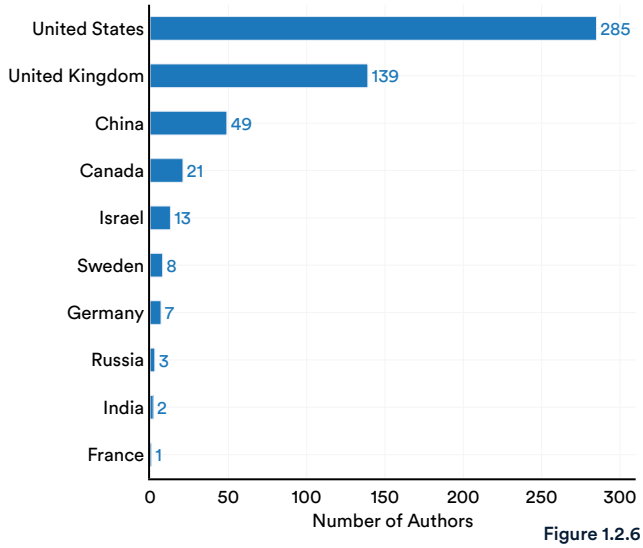

Figure 1.2.6

Number of Authors of Significant Machine Learning Systems by Select Geographic Area, 2002–22
Source: Epoch and AI Index, 2022 | Chart: 2023 AI Index Report

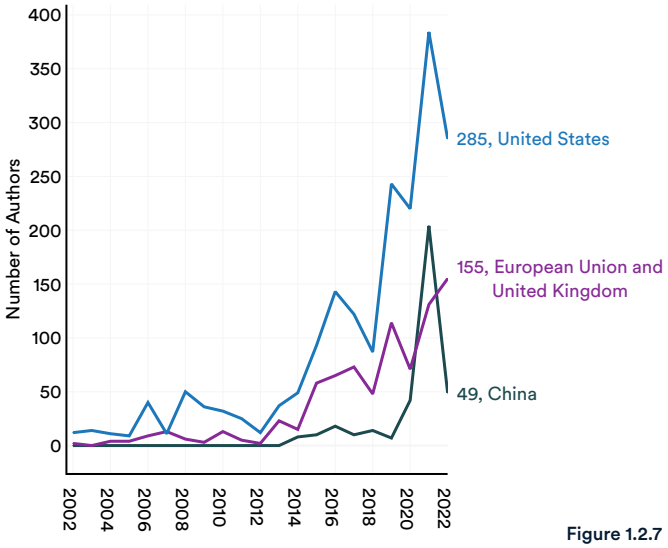

Figure 1.2.7

Number of Authors of Significant Machine Learning Systems by Country, 2002–22 (Sum)
Source: AI Index, 2022 | Chart: 2023 AI Index Report

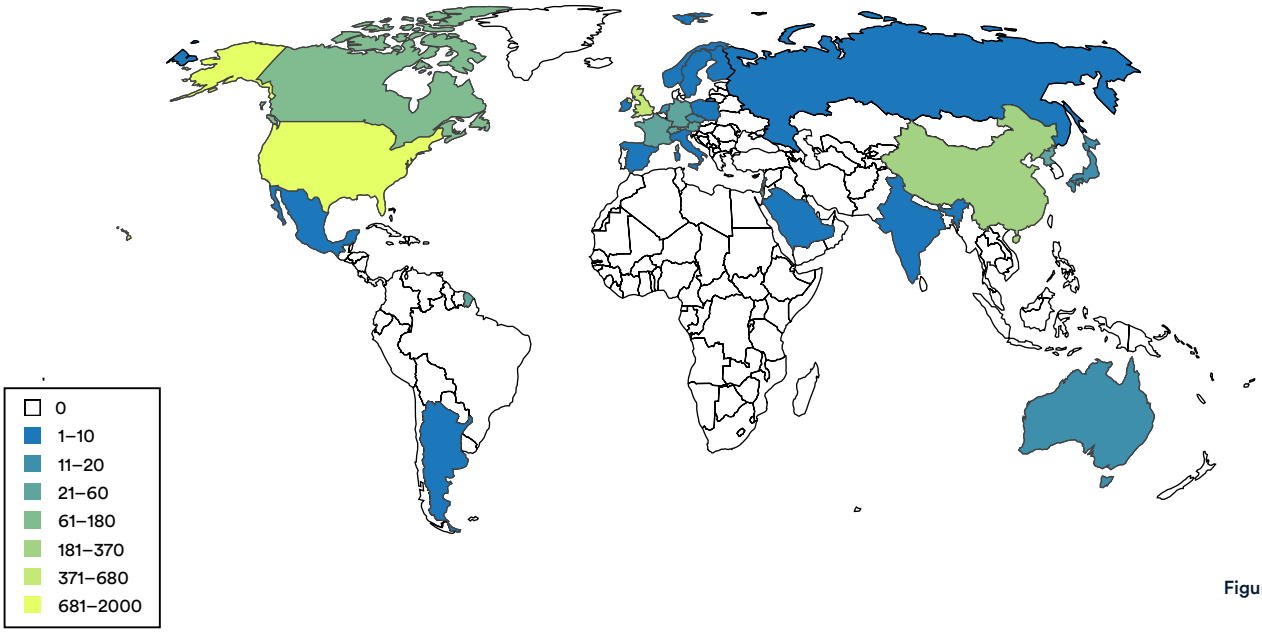

Figure 1.2.8

Parameter Trends

Parameters are numerical values that are learned by machine learning models during training. The value of parameters in machine learning models determines how a model might interpret input data and make predictions. Adjusting parameters is an essential step in ensuring that the performance of a machine learning system is optimized.

Figure 1.2.9 highlights the number of parameters of the machine learning systems included in the Epoch dataset by sector. Over time, there has been a steady increase in the number of parameters, an increase that has become particularly sharp since the early 2010s. The fact that AI systems are rapidly increasing their parameters is reflective of the increased complexity of the tasks they are being asked to perform, the greater availability of data, advancements in underlying hardware, and most importantly, the demonstrated performance of larger models.

Number of Parameters of Significant Machine Learning Systems by Sector, 1950–2022

Source: Epoch, 2022 | Chart: 2023 AI Index Report

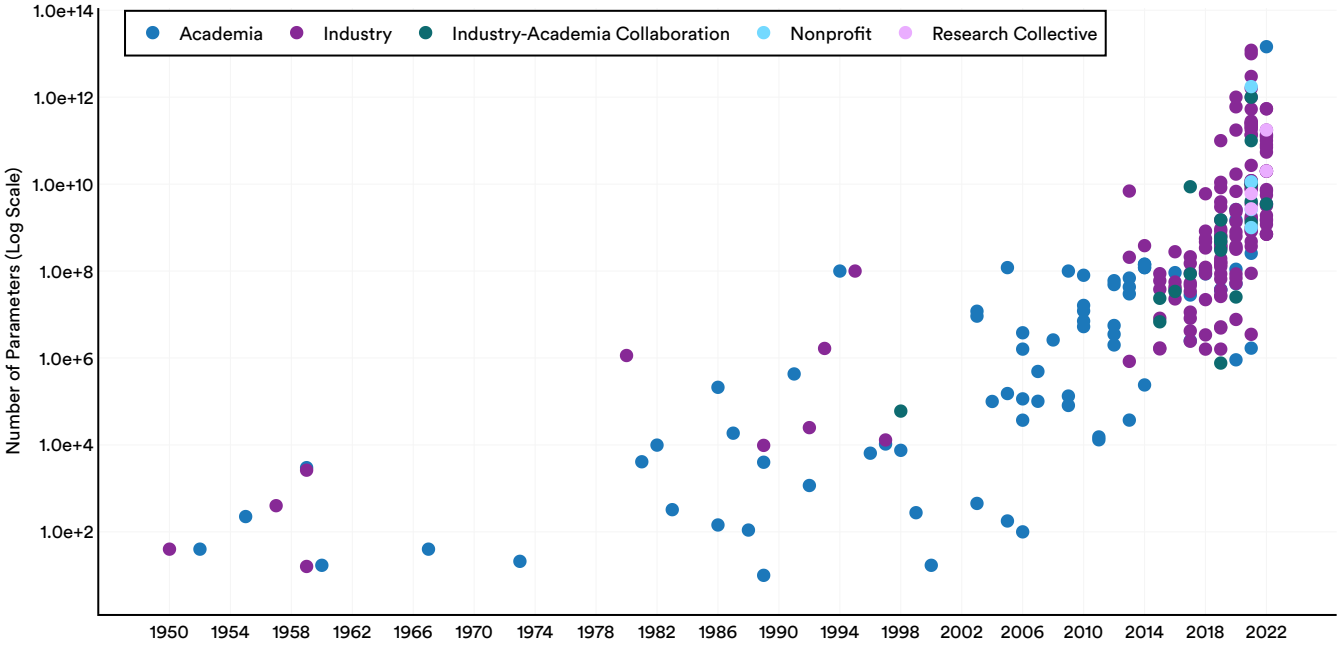

Figure 1.2.9

Figure 1.2.10 demonstrates the parameters of machine learning systems by domain. In recent years, there has been a rise in parameter-rich systems.

Number of Parameters of Significant Machine Learning Systems by Domain, 1950–2022

Source: Epoch, 2022 | Chart: 2023 AI Index Report

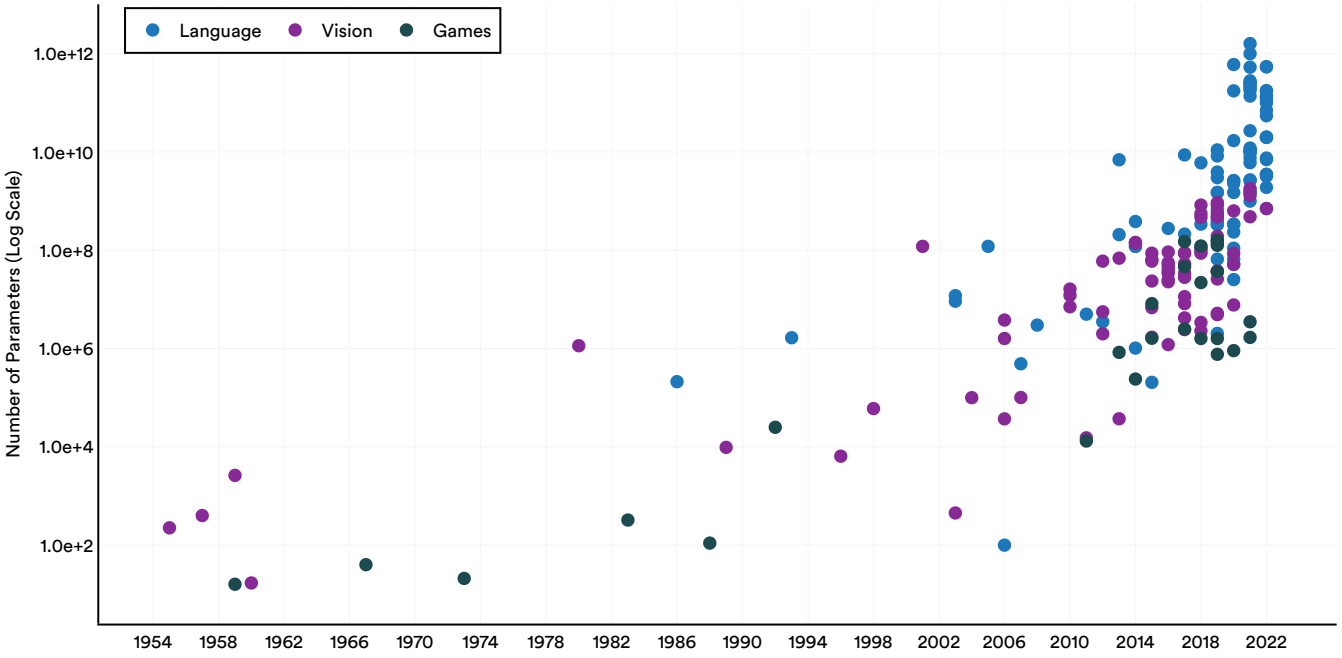

Figure 1.2.10

Compute Trends

The computational power, or "compute," of AI systems refers to the amount of computational resources needed to train and run a machine learning system. Typically, the more complex a system is, and the larger the dataset on which it is trained, the greater the amount of compute required.

The amount of compute used by significant AI machine learning systems has increased exponentially in the last half-decade (Figure 1.2.11).[9] The growing demand for compute in AI carries several important implications. For example, more compute-intensive models tend to have greater environmental impacts, and industrial players tend to have easier access to computational resources than others, such as universities.

Training Compute (FLOP) of Significant Machine Learning Systems by Sector, 1950–2022
Source: Epoch, 2022 | Chart: 2023 AI Index Report

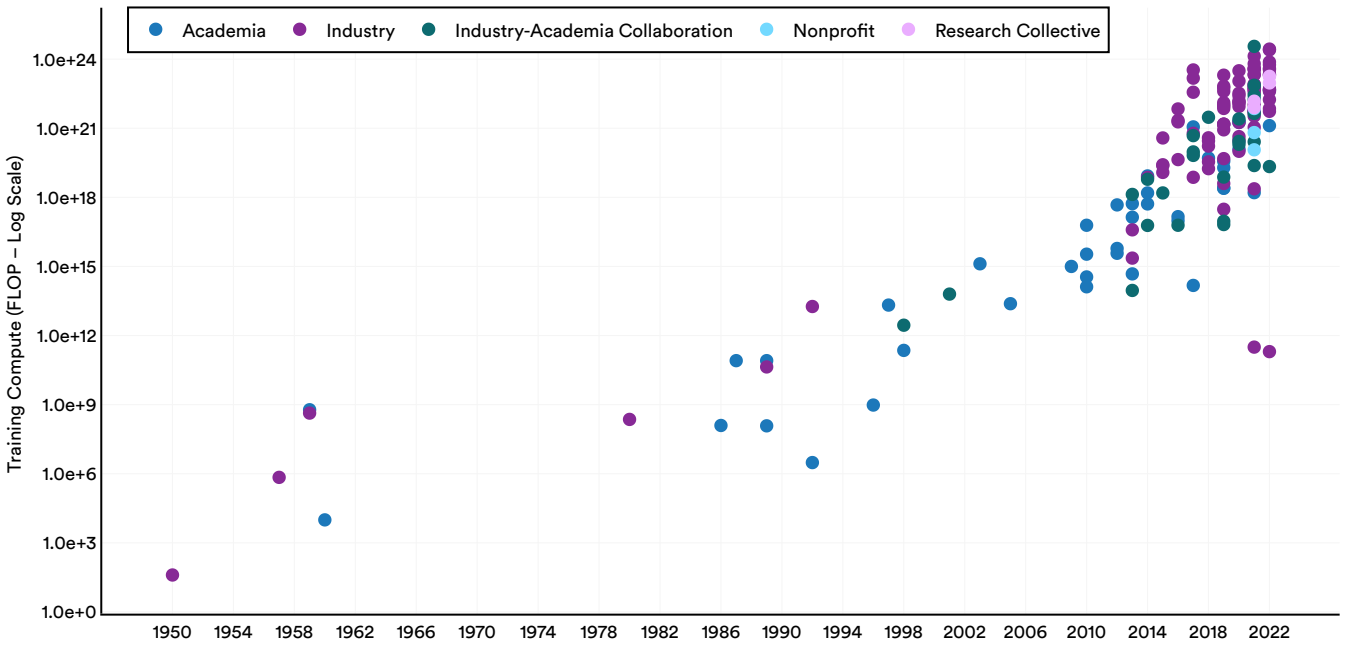

Figure 1.2.11

9 FLOP stands for "Floating Point Operations" and is a measure of the performance of a computational device.

Since 2010, it has increasingly been the case that of all machine learning systems, language models are demanding the most computational resources.

Training Compute (FLOP) of Significant Machine Learning Systems by Domain, 1950–2022
Source: Epoch, 2022 | Chart: 2023 AI Index Report

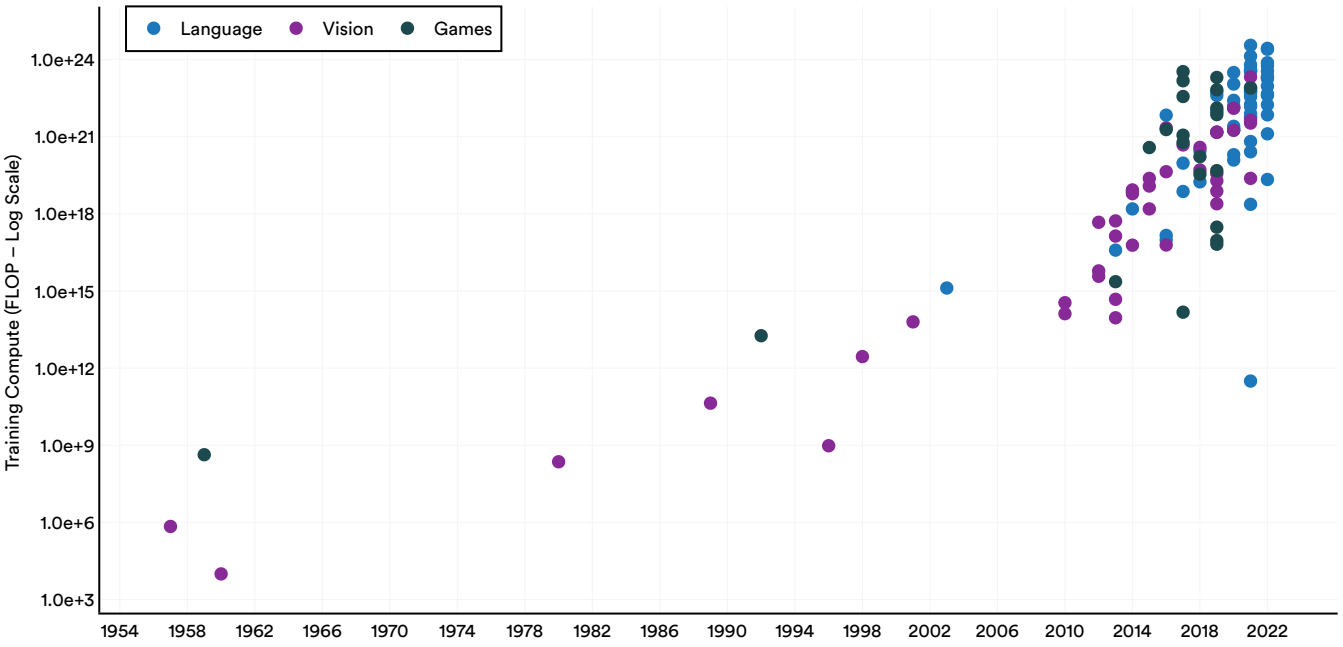

Figure 1.2.12

Large Language and Multimodal Models

Large language and multimodal models, sometimes called foundation models, are an emerging and increasingly popular type of AI model that is trained on huge amounts of data and adaptable to a variety of downstream applications. Large language and multimodal models like ChatGPT, DALL-E 2, and Make-A-Video have demonstrated impressive capabilities and are starting to be widely deployed in the real world.

National Affiliation

This year the AI Index conducted an analysis of the national affiliation of the authors responsible for releasing new large language and multimodal models.[10] The majority of these researchers were from American institutions (54.2%) (Figure 1.2.13). In 2022, for the first time, researchers from Canada, Germany, and India contributed to the development of large language and multimodal models.

Authors of Select Large Language and Multimodal Models (% of Total) by Country, 2019–22
Source: Epoch and AI Index, 2022 | Chart: 2023 AI Index Report

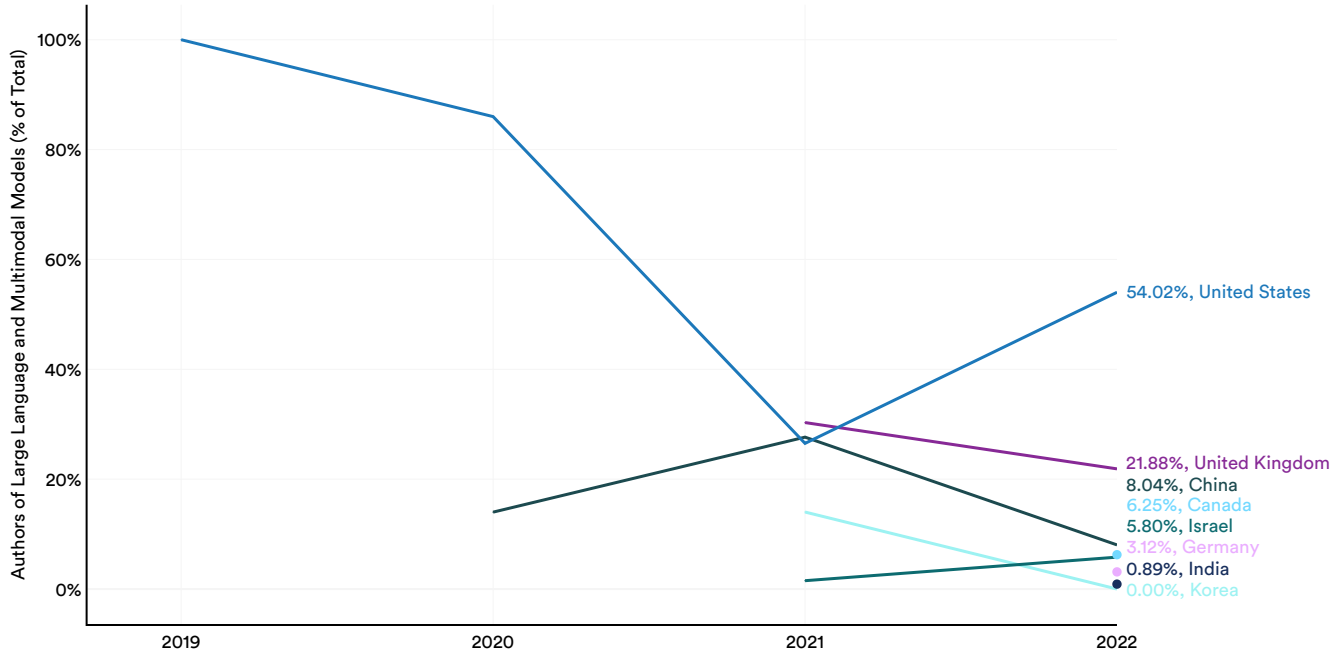

Figure 1.2.13

Figure 1.2.14 offers a timeline view of the large language and multimodal models that have been released since GPT-2, along with the national affiliations of the researchers who produced the models. Some of the notable American large language and multimodal models released in 2022 included OpenAI's DALL-E 2 and Google's PaLM (540B). The only Chinese large language and multimodal model released in 2022 was GLM-130B, an impressive bilingual (English and Chinese) model created by researchers at Tsinghua University. BLOOM, also launched in late 2022, was listed as indeterminate given that it was the result of a collaboration of more than 1,000 international researchers.

10 The AI models that were considered to be large language and multimodal models were hand-selected by the AI Index steering committee. It is possible that this selection may have omitted certain models.

Timeline and National Affiliation of Select Large Language and Multimodal Model Releases

Source: AI Index, 2022 | Chart: 2023 AI Index Report

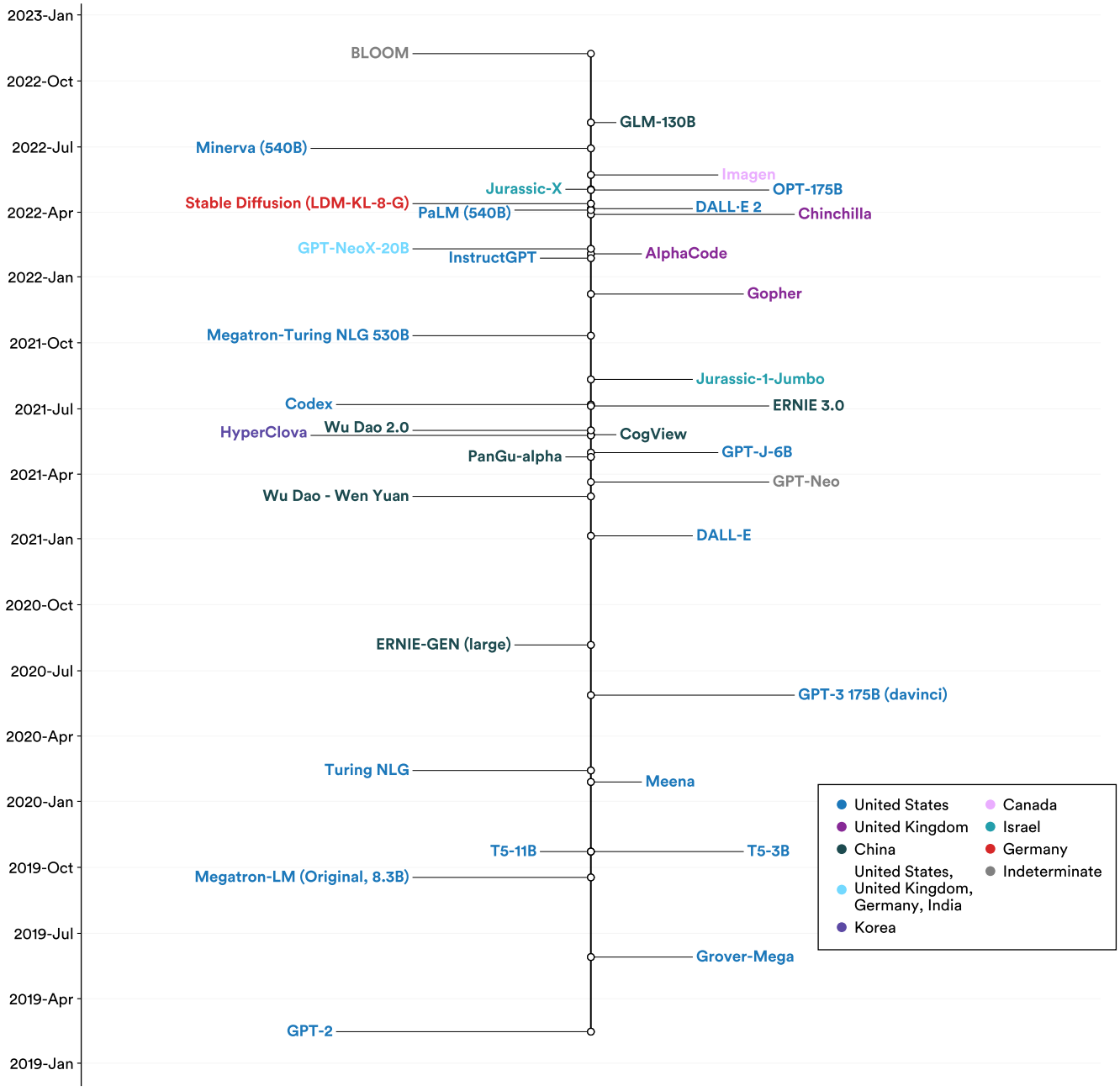

Figure 1.2.14[11]

11 While we were conducting the analysis to produce Figure 1.2.14, Irene Solaiman published a paper that has a similar analysis. We were not aware of the paper at the time of our research.

Parameter Count

Over time, the number of parameters of newly released large language and multimodal models has massively increased. For example, GPT-2, which was the first large language and multimodal model released in 2019, only had 1.5 billion parameters. PaLM, launched by Google in 2022, had 540 billion, nearly 360 times more than GPT-2. The median number of parameters in large language and multimodal models is increasing exponentially over time (Figure 1.2.15).

Number of Parameters of Select Large Language and Multimodal Models, 2019–22

Source: Epoch, 2022 | Chart: 2023 AI Index Report

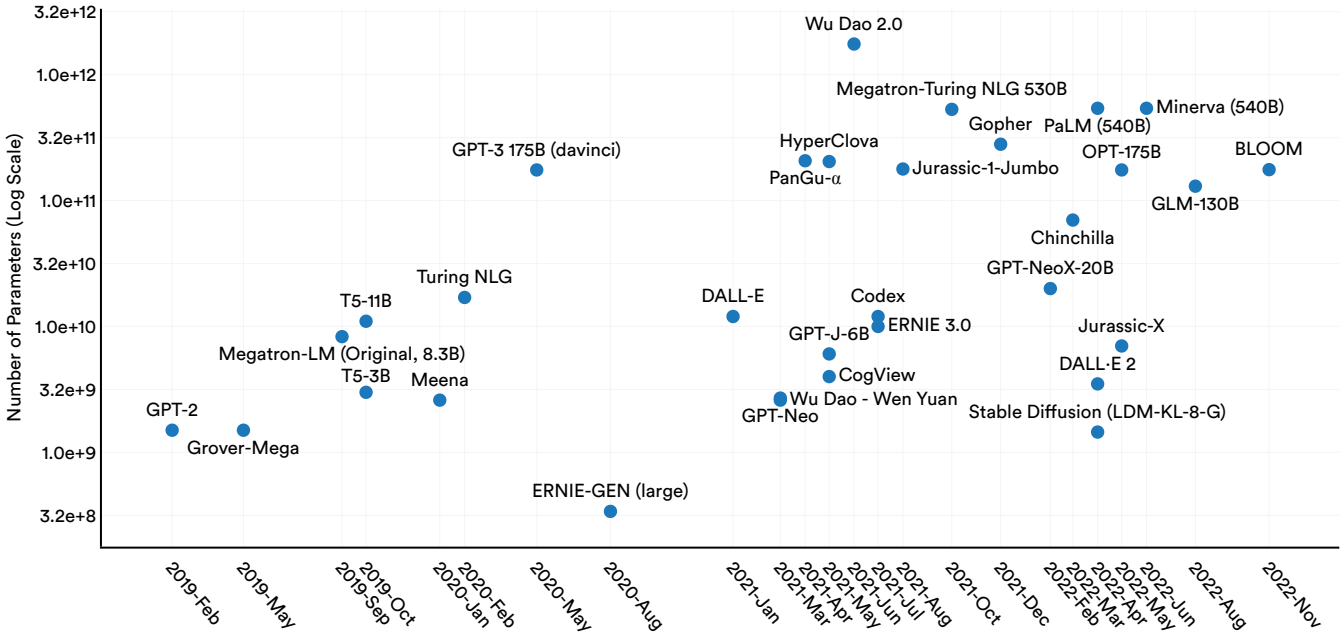

Figure 1.2.15

Training Compute

The training compute of large language and multimodal models has also steadily increased (Figure 1.2.16). The compute used to train Minerva (540B), a large language and multimodal model released by Google in June 2022 that displayed impressive abilities on quantitative reasoning problems, was roughly nine times greater than that used for OpenAI's GPT-3, which was released in June 2022, and roughly 1839 times greater than that used for GPT-2 (released February 2019).

Training Compute (FLOP) of Select Large Language and Multimodal Models, 2019–22
Source: Epoch, 2022 | Chart: 2023 AI Index Report

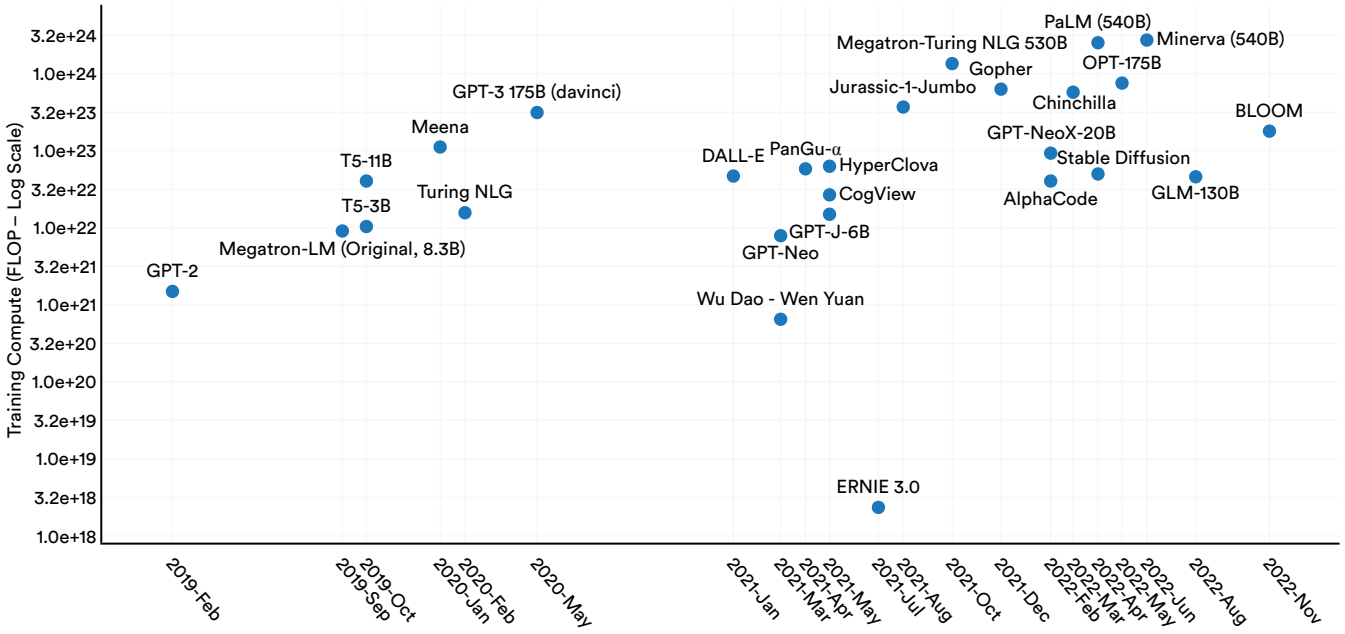

Figure 1.2.16

Training Cost

A particular theme of the discourse around large language and multimodal models has to do with their hypothesized costs. Although AI companies rarely speak openly about training costs, it is widely speculated that these models cost millions of dollars to train and will become increasingly expensive with scale.

This subsection presents novel analysis in which the AI Index research team generated estimates for the training costs of various large language and multimodal models (Figure 1.2.17). These estimates are based on the hardware and training time disclosed by the models' authors. In cases where training time was not disclosed, we calculated from hardware speed, training compute, and hardware utilization efficiency. Given the possible variability of the estimates, we have qualified each estimate with the tag of mid, high, or low: mid where the estimate is thought to be a mid-level estimate, high where it is thought to be an overestimate, and low where it is thought to be an underestimate. In certain cases, there was not enough data to estimate the training cost of particular large language and multimodal models, therefore these models were omitted from our analysis.

The AI Index estimates validate popular claims that large language and multimodal models are increasingly costing millions of dollars to train. For example, Chinchilla, a large language model launched by DeepMind in May 2022, is estimated to have cost $2.1 million, while BLOOM's training is thought to have cost $2.3 million.

Estimated Training Cost of Select Large Language and Multimodal Models
Source: AI Index, 2022 | Chart: 2023 AI Index Report

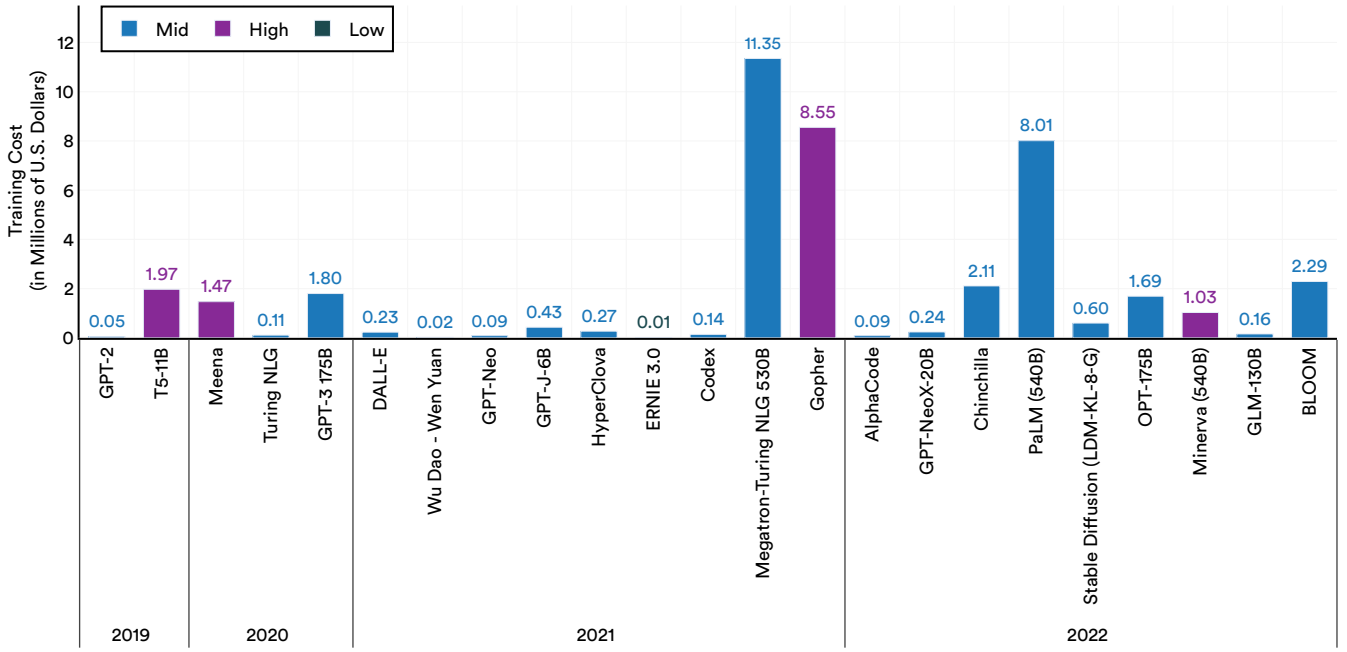

Figure 1.2.17

12 See Appendix for the complete methodology behind the cost estimates.

There is also a clear relationship between the cost of large language and multimodal models and their size. As evidenced in Figures 1.2.18 and 1.2.19, the large language and multimodal models with a greater number of parameters and that train using larger amounts of compute tend to be more expensive.

Estimated Training Cost of Select Large Language and Multimodal Models and Number of Parameters
Source: AI Index, 2022 | Chart: 2023 AI Index Report

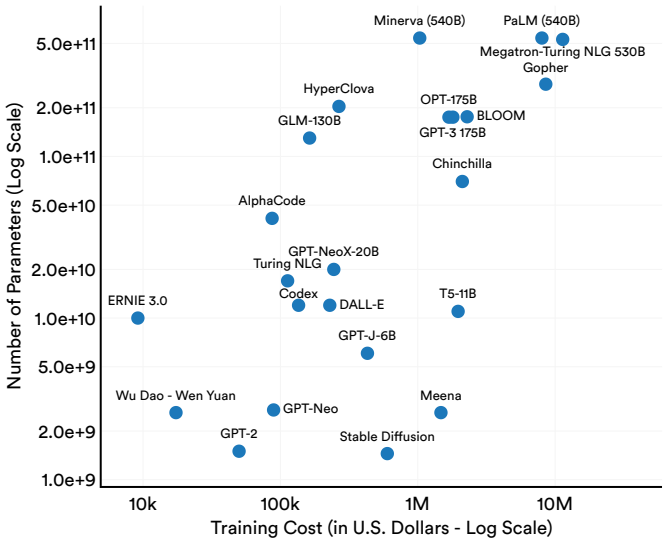

Figure 1.2.18

Estimated Training Cost of Select Large Language and Multimodal Models and Training Compute (FLOP)
Source: AI Index, 2022 | Chart: 2023 AI Index Report

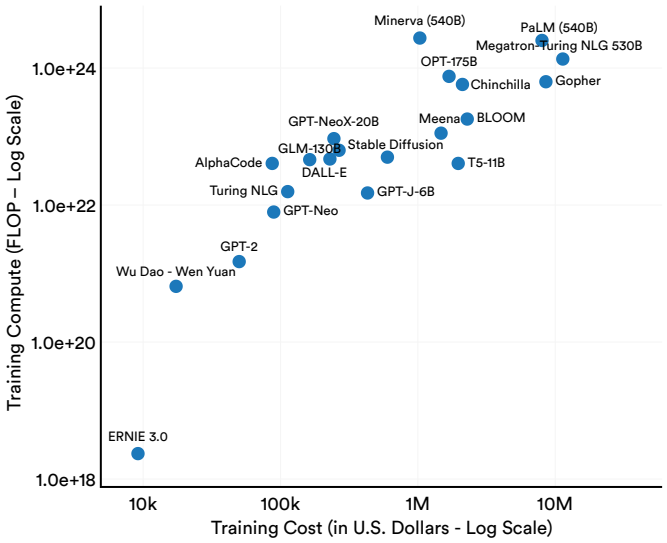

Figure 1.2.19

AI conferences are key venues for researchers to share their work and connect with peers and collaborators. Conference attendance is an indication of broader industrial and academic interest in a scientific field. In the past 20 years, AI conferences have grown in size, number, and prestige. This section presents data on the trends in attendance at major AI conferences.

1.3 AI Conferences

Conference Attendance

After a period of increasing attendance, the total attendance at the conferences for which the AI Index collected data dipped in 2021 and again in 2022 (Figure 1.3.1).[13] This decline may be attributed to the fact that many conferences returned to hybrid or in-person formats after being fully virtual in 2020 and 2021. For example, the International Joint Conference on Artificial Intelligence (IJCAI) and the International Conference on Principles of Knowledge Representation and Reasoning (KR) were both held strictly in-person.

Neural Information Processing Systems (NeurIPS) continued to be one of the most attended conferences, with around 15,530 attendees (Figure 1.3.2).[14] The conference with the greatest one-year increase in attendance was the International Conference on Robotics and Automation (ICRA), from 1,000 in 2021 to 8,008 in 2022.

Number of Attendees at Select AI Conferences, 2010–22

Source: AI Index, 2022 | Chart: 2023 AI Index Report

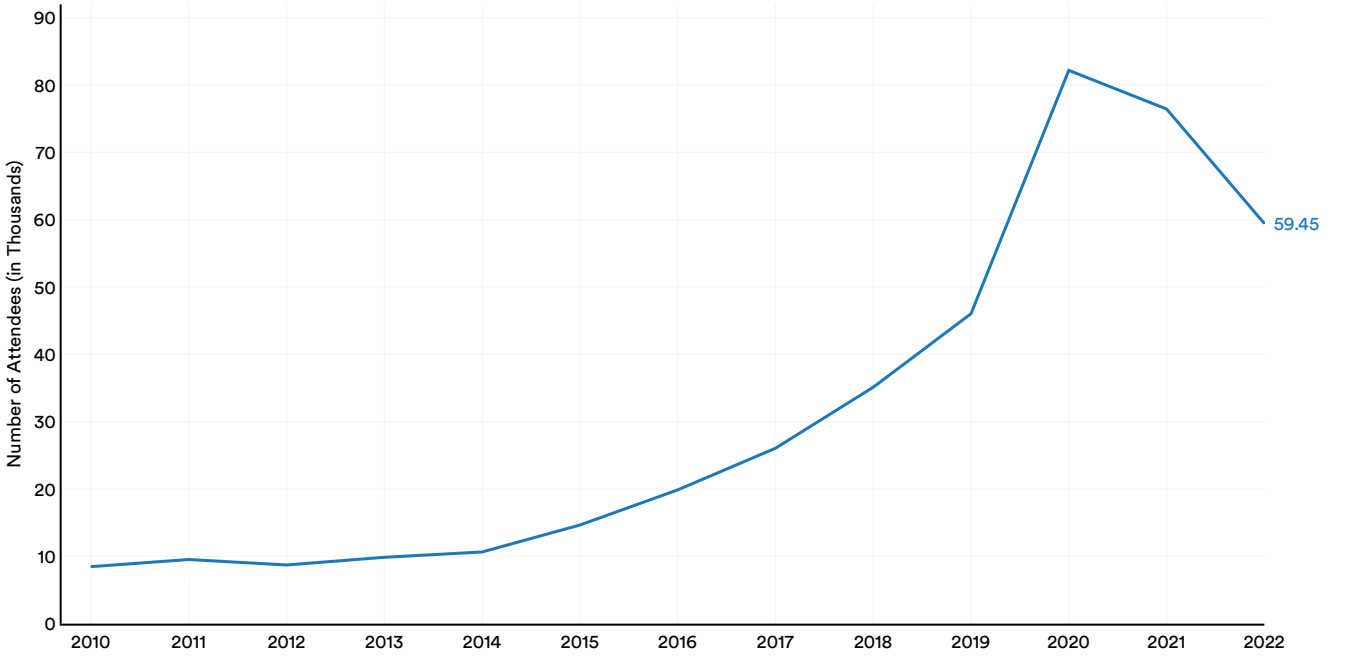

Figure 1.3.1

13 This data should be interpreted with caution given that many conferences in the last few years have had virtual or hybrid formats. Conference organizers report that measuring the exact attendance numbers at virtual conferences is difficult, as virtual conferences allow for higher attendance of researchers from around the world.
14 In 2021, 9,560 of the attendees attended NeurIPS in-person and 5,970 remotely.

Attendance at Large Conferences, 2010–22

Source: AI Index, 2022 | Chart: 2023 AI Index Report

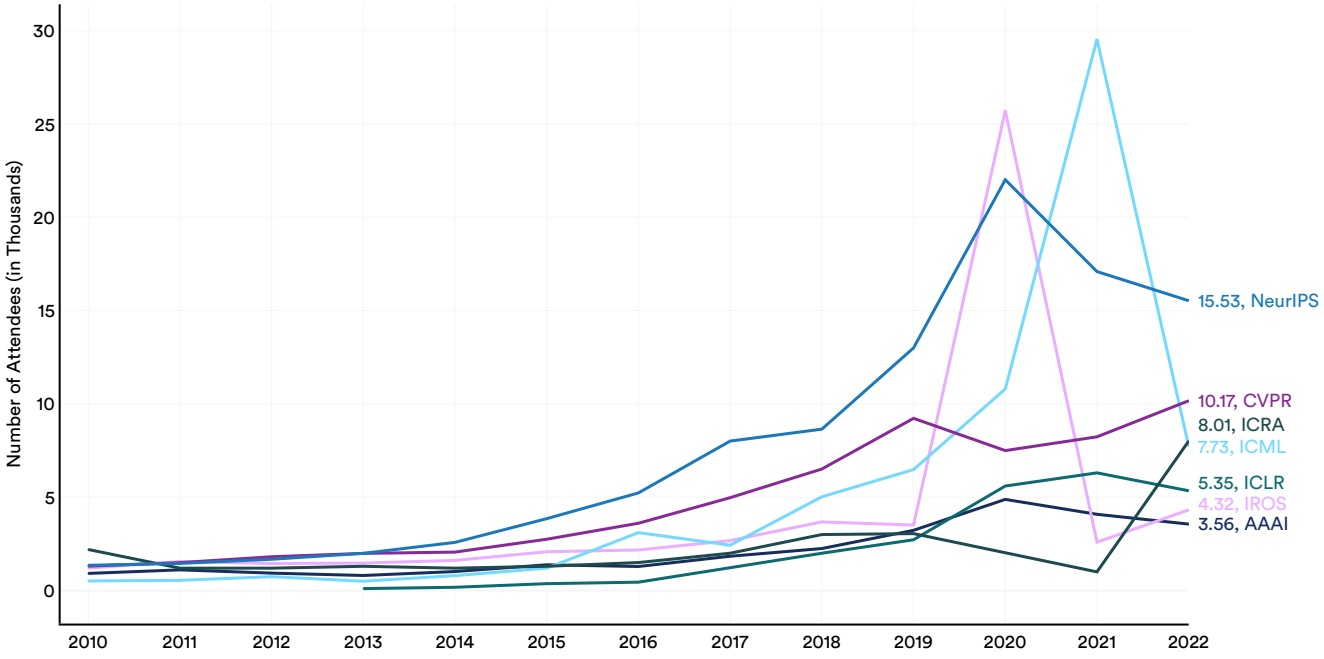

Figure 1.3.2

Attendance at Small Conferences, 2010–22

Source: AI Index, 2022 | Chart: 2023 AI Index Report

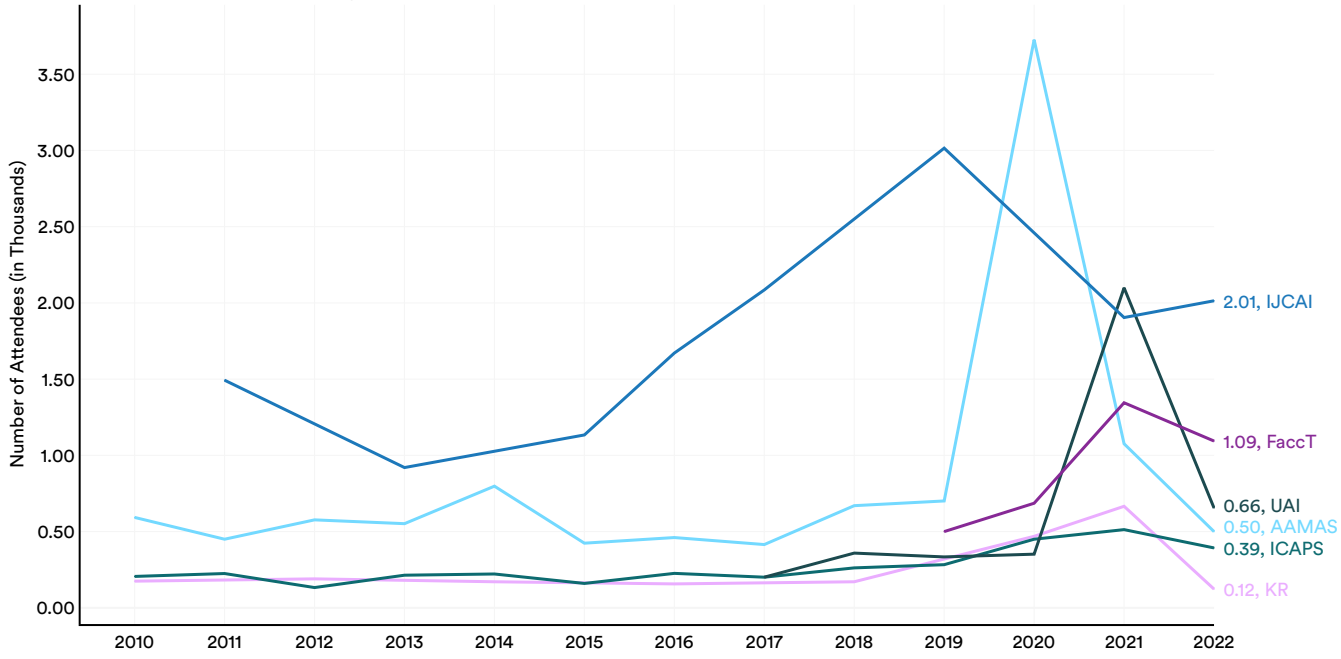

Figure 1.3.3

GitHub is a web-based platform where individuals and coding teams can host, review, and collaborate on various code repositories. GitHub is used extensively by software developers to manage and share code, collaborate on various projects, and support open-source software. This subsection uses data provided by GitHub and the OECD.AI policy observatory. These trends can serve as a proxy for some of the broader trends occuring in the world of open-source AI software not captured by academic publication data.

1.4 Open-Source AI Software

Projects

A GitHub project is a collection of files that can include the source code, documentation, configuration files, and images that constitute a software project. Since 2011, the total number of AI-related GitHub projects has steadily increased, growing from 1,536 in 2011 to 347,934 in 2022.

Number of GitHub AI Projects, 2011–22
Source: GitHub, 2022; OECD.AI, 2022 | Chart: 2023 AI Index Report

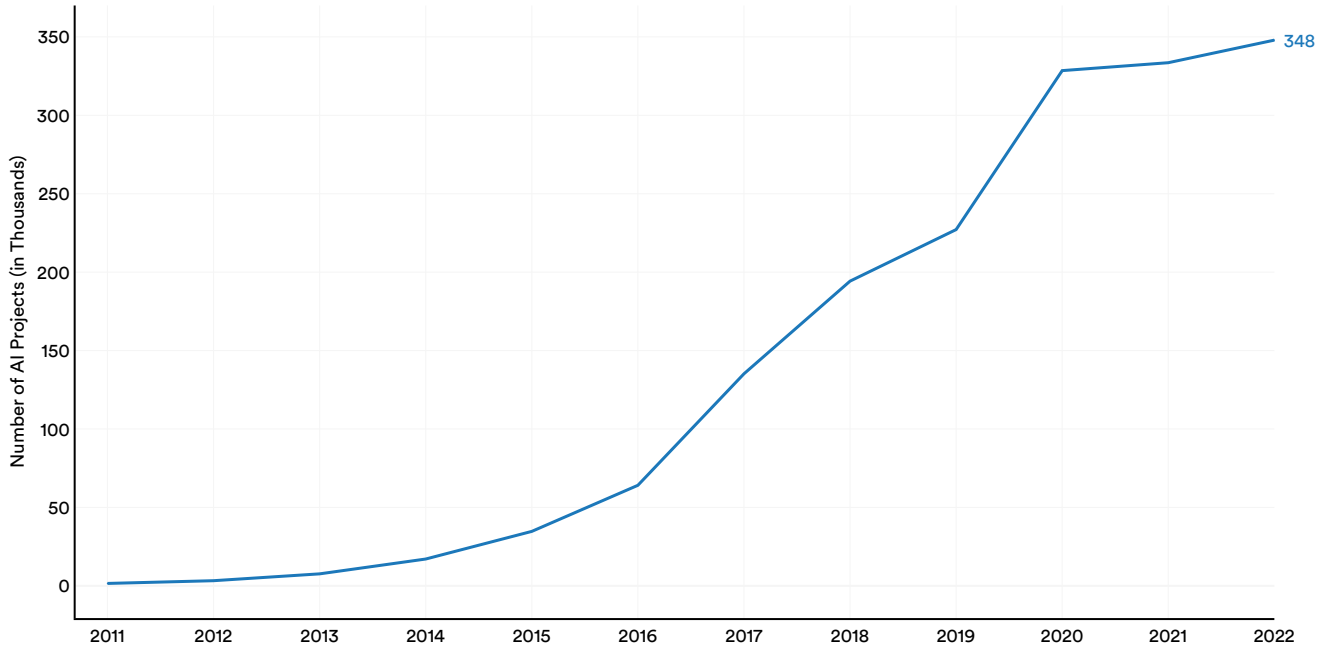

Figure 1.4.1

As of 2022, a large proportion of GitHub AI projects were contributed by software developers in India (24.2%) (Figure 1.4.2). The next most represented geographic area was the European Union and the United Kingdom (17.3%), and then the United States (14.0%). The share of American GitHub AI projects has been declining steadily since 2016.

GitHub AI Projects (% Total) by Geographic Area, 2011–22
Source: GitHub, 2022; OECD.AI, 2022 | Chart: 2023 AI Index Report

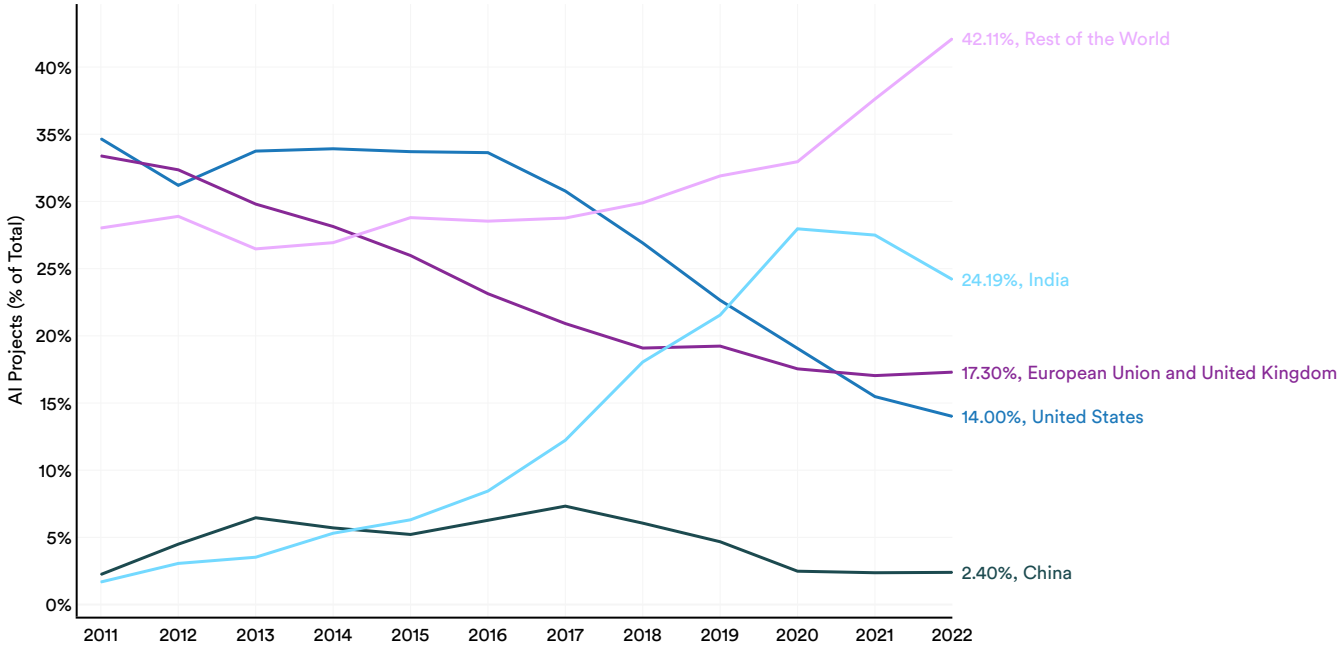

Figure 1.4.2

Stars

GitHub users can bookmark or save a repository of interest by "starring" it. A GitHub star is similar to a "like" on a social media platform and indicates support for a particular open-source project. Some of the most starred GitHub repositories include libraries like TensorFlow, OpenCV, Keras, and PyTorch, which are widely used by software developers in the AI coding community.

Figure 1.4.3 shows the cumulative number of stars attributed to projects belonging to owners of various geographic areas. As of 2022, GitHub AI projects from the United States received the most stars, followed by the European Union and the United Kingdom, and then China. In many geographic areas, the total number of new GitHub stars has leveled off in the last few years.

Number of GitHub Stars by Geographic Area, 2011–22
Source: GitHub, 2022; OECD.AI, 2022 | Chart: 2023 AI Index Report

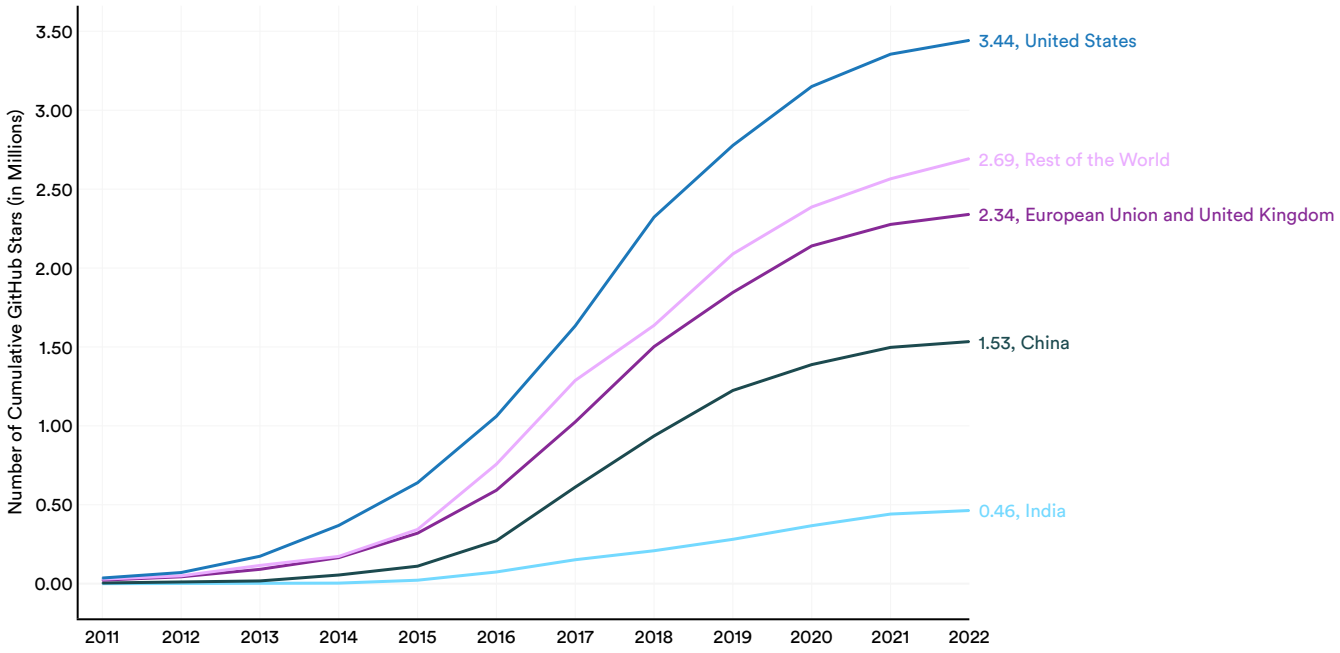

Figure 1.4.3