Bringing India at the Forefront of AI Research

Vijay Daultani
vijay.daultani@gmail.com

Abstract

Recent advancements in Artificial Intelligence (AI) are driving a transformative era in human history. Recognizing AI's crucial role, numerous countries have significantly invested in advancing the field over the past decade. Surprisingly, despite being the world's largest exporter of engineers, India lags behind in developing the advanced AI capabilities needed to compete with leading nations. We argue that India's relatively low contribution to AI research is the primary reason for this shortfall. This article emphasizes the urgent need for India to boost its initiatives on AI research. We advocate for a fundamental shift from an over-reliance on premier institutions to fostering an environment where young engineers from diverse backgrounds can engage in research. We propose establishing an AI research lab focused on publishing cutting-edge research with an India-first approach. Introducing the concept of Mass Research, we emphasize the need for large-scale research efforts. Finally, we outline specific steps to position India at the forefront of AI research and innovation.

1 Introduction

Recent advancements in artificial intelligence, particularly in computer vision (CV), and natural language processing (NLP) have seen remarkable growth. Many contend that this represents the most significant revolution in human history, surpassing even the Industrial Revolution of the mid-18th century. The impact of these technologies on daily human life continues to expand, fueled by substantial investments from various countries and major corporations seeking to capitalize on these innovations.

Countries like the USA and China have made substantial investments in developing advanced AI systems and creating environments that foster continuous innovation. In contrast, despite being one of the largest exporters of engineers and IT services globally, India significantly lags behind in AI research. One effective way to gauge a country's influence in a field is by examining the proportion of research articles published in top venues. Our analysis indicates that India is significantly lagging in AI research (§3). If this trend persists over the next decade, India risks falling so far behind that it may no longer remain self-reliant in accessing advanced technologies, thereby raising national security concerns.

Publishing research is crucial as it often necessitates the development of systems that demonstrate the proposed ideas' effectiveness. These systems, developed through research, play a key role in defining access to state-of-the-art technology over time. Field of AI is no different, therefore, advancing in the field of AI requires a concerted effort to enhance research capabilities within the field. In other words, we cannot simply have access to state-of-the-art AI products/technologies without doing state-of-the-art AI research.

This article explores the critical importance of advancing AI research for India and argues why it should be one of the national priorities (§2). We assess India's current contributions to AI research (§3) and outline a goal for future progress (§4). Despite India's strong education system, we investigate why India's leading institutions are not producing research on a scale competitive to nations at the

forefront of AI innovation (§5). Subsequently, we present a comprehensive blueprint for achieving this goal (§6). Finally, we conclude with specific action items aimed at bridging the current gaps (§7).

2 Why Does India Need to Ramp up AI Research

Investing in AI technology is crucial for several reasons, particularly for a developing nation like India with its vast and diverse population. Firstly, achieving self-sufficiency in advanced technology access is vital for our national interests. While some AI systems are open source, many of the most sophisticated, like ChatGPT and Gemini, remain proprietary behind closed doors. At present, only a handful of foreign organizations like Google, OpenAI, and Microsoft dictate the future of AI for everyone in the world. The lack of transparency and inaccessibility to these closed AI systems presents significant challenges. For e.g., if harmful effects are identified in specific use cases within India, we must depend on other organizations to address these issues, often at their discretion. These reasons underscore the necessity for India to conduct its own research and develop AI systems that are entirely domestically owned.

Secondly, India faces unique challenges that necessitate solutions tailored to its societal context. One significant issue with AI systems is bias, which can lead to discrimination based on race, gender, age, and other factors [6, 7, 8]. Such biases can disproportionately affect specific groups, a concern particularly relevant in a country as diverse as India. Therefore, it is crucial to conduct AI research that focuses on developing systems capable of mitigating these risks in a nation with such a varied populace.

Thirdly, another major challenge is the limited availability of AI systems in various languages. Most AI research and development are conducted in English, leading to AI systems that do not perform equally well across different languages. This is a significant issue for India, which has 22 official languages and numerous other languages and dialects spoken across the subcontinent. Ensuring that AI systems are accessible and effective in all these languages is essential for their widespread and equitable adoption.

3 Where Does India Stand at Present

A common method to gauge a country's research impact in a particular field involves analyzing its authors' publications in leading journals and conferences. In this study, we examine data in Table 1 from two renowned conferences: ACL, a prominent venue for Natural Language Processing (NLP), and CVPR, a major conference in Computer Vision (CV). The dataset is sourced from publicly available information from ACL [1, 2, 3] and CVPR [4, 5]. The data, as presented in Table 1, indicates that the USA and China are leading contributors to AI research. Consequently, it is unsurprising that these countries have developed some of the world's most advanced AI systems.

For clarity and focus, we specifically compare India's contributions to these conferences with those of the two nations with the highest publication rates. The analysis indicates that India's participation in these forums has either consistently declined or remained static over the years. Notably, India's share of publications has not surpassed 5% of the total articles presented at these conferences. This is in sharp contrast to the USA and China, whose contributions average around 29% each.

4 What Should India's AI Research Goal

Having evaluated India's standing in AI fields of natural language processing and computer vision, it's apparent that a significant gap exists between India (averaging around 5%) and leading contributors like the USA and China (averaging around 29% each). Recognizing this gap prompts the question: what should be our objective to establish a competitive edge in AI research?

We advocate benchmarking against the current frontrunners in the field. Considering India's population and strong emphasis on the STEM (Science, Technology, Engineering, and Mathematics) compared to these nations, we suggest setting a target of 25% to ensure our research endeavors are conducted at a level that maintains competitiveness on a global scale.

Year	ACL			CVPR		
	USA	China	India	USA	China	India
2023	22	31	2	-	-	-
2022	-	-	-	21	45	2
2020	29	17	12	40	23	2
Avg.	26	24	7	31	34	2

Table 1: The distribution of authors by country in ACL and CVPR conferences reveals that, on average, 29% of authors are from the USA, 29% are from China, and 5% are from India over the last three years. The notation "-" indicates that information is "Not Publicly Available" for the respective cells.

5 Why Premium Institutes Are Unable to Produce Mass Research

India has made significant investments in its education system, boasting one of the largest pools of engineers and doctors globally, underscoring the importance of higher education in the lives of its citizens. Despite hosting numerous world-renowned institutions and brilliant minds, a pertinent question arises: Why does India still face challenges in generating competitive AI research on a global level?

We believe the primary reason for this is a misalignment of priorities for students. Often, students from these prestigious institutes prioritize securing a promising future, aiming for well-paying jobs at reputable companies. Consequently, their focus tends to be on excelling in coursework to attain high grades, as this is typically the primary criterion for evaluating prospective candidates during placements. Henceforth, the absence of any incentives for students to pursue or contribute in research hampers the cultivation of a vibrant research culture.

It's important to recognize that the bulk of research originating from India stems from these elite institutes. However, this output falls considerably below the level required for India to establish itself as a frontrunner in AI technology. Thus, it's evident that while the talent needed for this endeavor exists, there's a disconnect between students' immediate goals of entering the workforce post-graduation and the imperative for India to assert itself as a prominent player in the AI domain.

6 How Can India Ramp up AI Research

Jeff Bezos once stated, "If you double the number of experiments you do per year, you're going to double your inventiveness." This quote resonates deeply with our understanding of research dynamics. Increasing experimentation frequency can significantly enhance creativity, particularly in the realm of research. To elevate the quality and quantity of publications in esteemed academic venues, embracing a culture of experimentation is paramount.

A widespread misconception about research is that focusing solely on grand ideas and significant problems is the key to producing quality work. However, this notion is flawed. We believe quality is often a product of quantity. In simpler terms, achieving high-quality research begins with conducting a substantial volume of research. Considering India's demographic advantage with its vast pool of engineering students, leveraging this abundance should be a priority.

While elite institutions undoubtedly play a pivotal role in advancing research, relying solely on them for scalability is limiting. To effectively address the gap of over 18%, it is imperative to promote research on a broader national scale. We advocate for tapping into the vast potential of non-elite academic engineering students, thereby fostering what we term "Mass Research".

Mass Research involves harnessing the intellect of engineering students on a large scale, not limited to elite institutes. It's acknowledged that students across all institutions prioritize coursework and grades. To effectively utilize this broad talent pool, two significant shifts are imperative. Firstly, there should be a concerted effort to instill in future engineers the perspective of research as a promising career path. Secondly, strategic incentives must be implemented to support this shift in mindset.

7 What Are the Action Items to Achieve AI Self-reliance for India

We propose the establishment of a AI research lab targeting third and fourth-year engineering undergraduates and master's students enrolled in computer science engineering. This initiative offers students the opportunity to engage in research alongside their university studies, providing them with hands-on research experience without the need for full-time commitment. Our focus on these students stems from their foundational understanding of research methodologies acquired during their early years in the engineering program. Additionally, this approach aligns with their priorities, as many students in this demographic are inclined towards joining internship programs. By joining our institute, students can view it as akin to participating in an internship, albeit with a distinct emphasis on honing research skills and contributing to publications in esteemed AI journals and conferences.

Our lab will be dedicated to a singular goal: publishing cutting-edge research in premier AI forums. We prioritize tackling real-world challenges over theoretical pursuits. Through our research efforts, we aim to develop tailored systems geared towards addressing issues in India. We perceive this as a means to counterbalance the dominance of a few organizations from the USA and China, ultimately enhancing India's presence in AI research.

We believe it is crucial to update India's education system to incorporate AI and other emerging technologies from an early stage. However, such systemic changes are long-term solutions that may take decades to yield significant results. In contrast, our proposed research lab aims to address the current gap in research publication within the next one to two decades. An institution like this will provide immediate opportunities for students to contribute to cutting-edge research and elevate India's presence in the global AI research community.

8 How to Measure Success

We measure the success of our proposed lab by a singular metric: the number of articles published in esteemed AI journals and conferences. While we anticipate additional outcomes like the creation of AI systems stemming from our research, our paramount objective, by which we gauge our achievement, lies in our publication output.

9 Conclusion

We foresee a future for India where it stands as a global leader in AI technology. Currently, there is a significant disparity in the number of articles published by Indian researchers in top AI conferences. To address this, we propose an ambitious target: 25% of papers at leading AI conferences and journals should come from India. We also examine why India's elite institutions struggle to produce research output that rivals other nations, despite having one of the largest pools of engineering students worldwide. Our proposed solution involves a large-scale initiative focused on specific demographics of engineering students, aimed at equipping them with the necessary skills for AI research and publication in top-tier venues. Despite the challenges, we firmly believe that with a clear understanding of our current position, a well-defined goal, and a strategic, results-oriented approach, we can reach this target and position India at the forefront of AI research.

References

[1] ACL. acl 2020 proceedings of the conference.

[2] ACL. acl 2022 proceedings of the conference.

[3] ACL. acl 2023 proceedings of the conference.

[4] CVPR. cvpr 2022 chairs opening message.

[5] CVPR. cvpr 2023 chairs opening message.

[6] Harvard. Racial discrimination in face recognition technology.

[7] Reuters. Insight - amazon scraps secret ai recruiting tool that showed bias against women.

[8] MIT Technology Review. Predictive policing algorithms are racist. they need to be dismantled.