The Seybold REPORT

DEVELOPMENT OF ARTIFICIAL INTELLIGENCE (AI) IN DIFFERENT SECTORS

Author-1

Name -Ritika

Position- Phd Scholar (Economics)

E-mail- Ritscool61@gmail.com

Author-2

Name - Dr. Tareef Husain

E- mail- tareef.husain@galgotiasuniversity.edu.

ABSTRACT

According to many experts, artificial intelligence (AI) will dominate science and technology in the next century. As computing power has increased, big data quality and quantity have improved, and significant advances have been made in numerous research areas, including machine learning and speech recognition, in recent years, AI technology has advanced quickly and is now used extensively in all spheres of society. The use of AI technology in the financial sector for risk management, marketing, customer service, transactions, operations, and product optimization is getting more sophisticated, and some new business models have been developed. This study elaborates on the application, status quo, and development trend of AI in the financial business, starting with the application status and relevance of AI in the worldwide financial sphere. This article then outlines the steps to encourage the thorough, healthy, and sustainable growth of AI in the financial market in light of the dangers and practical problems that are present in the development process of AI based on the reality of worldwide financial development. The purpose of this article is to inform readers on the state of AI development in the financial sector and to serve as a theoretical resource for researchers in that area. As AI enters society and daily life, it will surely have a big impact on how India develops and advances. In addition to discovering methods to overcome more traditional hurdles like limited infrastructure and bureaucracy, AI has the potential to accelerate growth in India. It is crucial that the risks connected to investing in AI be evaluated at this early stage since they have long-term social ramifications. The opportunities and challenges of AI in India are covered in this article. We list potential solutions that are sector-specific (healthcare) and all-encompassing (bridging language barriers in India, mining public data). We draw attention to issues brought about by prevailing social standards (such caste and gender equations). Then, as India enters the era of artificial intelligence, we identify certain steps and safeguards that we believe are crucial for a robust and equitable development.

KEYWORDS

Artificial Intelligence, India, Financial sector, Economic sector.

INTRODUCTION

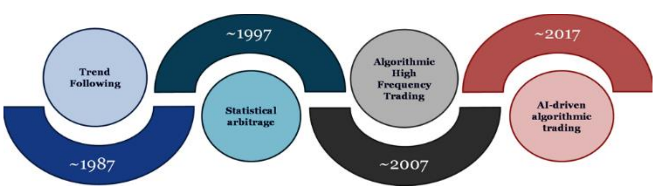
Studies of technology's societal impact are often organized "vertically" according to different factors, such as ethics [15], law [12], economic productivity and employment [11], and the consequences of various social markers, including gender [59] and race [16]. We use a "horizontal" approach in this study, viewing all of these problems from one central location: AI's function in India's current progress toward prosperity. Our decision to focus on one nation at a time is not novel. Little (1999) looks at how the global industrial system has affected numerous East Asian nations, while the recently conducted AI100 research [57] analyzes the junction of AI with a "typical North American city" in a number of different contexts. For this reason, it is important to address both technical and non-technical concerns in a coherent manner. The purpose of this study is to provide the groundwork for collaboration between engineers, social scientists, and policymakers. A sixth of the world's population lives in India, so it's important to keep an eye on the country's experimentation with artificial intelligence. The potential for AI's advantages and threats to be amplified in India's distinct social, cultural, economic, and political setting is of equal importance. India has a huge, youthful workforce [62], a rapidly expanding economy [41], and a strong, stable democracy [48], which all bode well for the potential size and scope of AI applications there. Public services may be improved by AI-driven interventions, such as by simplifying the public distribution system and lowering the cost of law enforcement. The usage of AI-enabled personalized healthcare or robots in manufacturing lines are only two examples of how AI might improve private services. However, India's vastness and diversity also amplify the country's problems, which range from economic disparity [1, 17] and caste-based discrimination [4] to linguistic diversity [42]. Other social issues, such as hunger [39] and the lack of female educators [43], may be better addressed by approaches that have nothing to do with AI. Even in areas where AI might make real improvements, its solutions will frequently have to compete with cultural norms that have evolved over the course of millennia.

Technologists and policymakers in India are looking to transfer concepts that have been effective in other settings since the country is so far behind others in terms of technical progress. A rising corpus of research [9, 37, 50] highlights the ineffectiveness and potential hazard of such a strategy. The central argument of this article is that, given India's unique social, political, cultural, and economic makeup, "AI for India" must be conceived from the ground up.

To introduce our thesis, we first describe the wide range of technological challenges that occur in the specific setting of India, including the country's linguistic diversity, its aging public archives, and its healthcare system. We provide this collection of exemplary situations in Section 2 in the hopes that it will motivate and equip technologists to address pressing societal issues. There may be significant gains. Similarly, the development of AI may pose new dangers. In Section 3, we emphasize how the progress led by AI might exacerbate existing inequities in Indiansociety (for example, depending on caste and gender). Section 4 summarizes our findings and makes recommendations for moving AI forward in India, including specific measures and safeguards. In Section 5 we summarize our findings.

There does not seem to be any prior work on AI in India that matches the breadth of our article. When compared to other summaries of the status of AI research in various nations, such as Israel [21], Singapore [60], and India [28], our focus is wider. Vempati's (2016) study, which serves as a "wake-up call" to Indian policymakers, comes closest to ours in terms of overall tone. As Vempati urges the rapid adoption of AI, he cites geopolitical implications, including analogies with China. We discuss a contrasting perspective that focuses on events occurring inside India, describing the advantages and disadvantages of this approach. The use of AI methods like evolutionary computation, deep learning, and probabilistic logic for the discovery of trading strategies and their automated implementation without human involvement has the greatest disruptive potential in the trading industry. Despite the fact that algorithmic trading has been around for some time (see Figure 2.4), AI-powered algorithms bring a new level of complexity and development to the practise. These algorithms eventually become fully automated, computer-programmed systems that depend less on human intervention and learn from the data input used. While classic systematic tactics would take longer to alter parameters owing to the significant human involvement, reinforcement learning enables the model to adapt to changing market circumstances.

Figure 2.3. Historical evolution of trading and AI



Additionally, the application of ML models pushes the analysis away from traditional back-testing procedures based on historical data and toward prediction and real-time trend analysis, for instance by using "walk forward" tests3 instead of back testing. 4 In order to prevent over-fitting in back tests based on past data and trends (Liew, 2020[10]), these tests forecast and react to trends in real time. They also get around the drawback of forecasts based on historical data when previously detected trends fail.

Advanced AI-based algorithms are now being used to find signals from 'low informational value' events in flow-based trading, while traditional algorithms have been employed to detect 'high informative events' and give speed of execution (for example, in high frequency trading or HFT).

These events are made up of harder-to-identify occurrences that are challenging to get value from. As a result, AI is now being used to extract signal from noise in data and turn this information into trading choices rather than providing speed of execution to front-run transactions. However, it is anticipated that as AI approaches advance, these algorithms will enable the augmentation of "conventional" algorithm capabilities, especially during the execution stage. With ramifications for financial markets, AI might support the full chain of events around a transaction, from signal detection through strategy development and automated execution without human participation.

1 OPPORTUNITIES FOR AI-DRIVENDEVELOPMENT

For India, AI holds promise as a catalyst to accelerate progress and to leapfrog traditional hurdles such as poor infrastructure and bureaucracy. In nearly every sector—finance, healthcare, law enforcement, transportation, agriculture, environmental conservation—one finds applications in which AI can be effective. In a timely move, the Indian government has recently constituted a task force precisely to identify openings for AI across sectors and guide policy [40]. In this section we describe some uniquely (at any rate, typically) Indian problems and their amenability to AI. Rather than enumerate a long list, we restrict our focus to three illustrative problems, which we present in some detail. The first two (in sections 2.1 and 2.2) cutacross different sectors; the third (Section 2.3) relates to a particularsector.

1.1 Scaling up NLP/ASR for Indian Languages

India is one of the nations with the highest number of languages spoken, with over 700. Indo-Aryan, Dravidian, Austroasiatic, Sino-Tibetan, Tai-Kadai, and Great Andamanese are only few of the language families represented here. Over a million people speak one of at least 20 languages as their native tongue. Language is a barrier to communication and information sharing since a big percentage of the population is either monolingual or bilingual. Because of the language barrier, only 13% of Indians are able to read this material.

Reality is different for the world's 422 million Hindi speakers and 60 million Tamil speakers. A quick online search will give you an idea of the quality of digital services available in various languages, while we are not aware of any comprehensive research on the topic. As an example, the authors played around with Google Translate, which is generally regarded as the best machine translation service out there. Simple English lines were translated into Hindi and Tamil and shown in Figure 1. Several "obvious" errors exist.

The availability of huge training corpora [5, 63] is largely responsible for the success of current NLP systems like Google Translate. Unfortunately, most Indian languages have far smaller data sets available compared to major Western languages. An open parallel text corpus in 11 European languages, for instance, has several millions of words in each language [29]. There is nearly one hundred times less information available to researchers working with Indian languages [31]. The lack of digitized data therefore limits the effectiveness of NLP, despite its popularity in the Indian AI community [28].

Use of resource-rich languages as "pivots" to develop software for languages with less available resources is a common practice in the natural language processing (NLP) field at present [45]. This strategy may have its own value in languages and has immediate practical applications. However, we foresee that developing systems that collect and distribute data in Indian languages would be necessary in the long run. We urge the AI community to take this up as a serious endeavor. Some potential avenues for the digitization of linguistic data include automatic speech recognition (ASR) and the resourceful use of crowdsourcing. Curiously, these subfields lament a dearth of information and support for their own work. Similar to natural language processing (NLP), automatic speech recognition (ASR) has designated under-resourced languages as a distinct area of study [8]. A new analysis by Pavlick et al. (2014) identifies languages that linguists should prioritize studying because they have a high likelihood of finding translators on Amazon Mechanical Turk.

Instead of focusing on practical solutions, our plan takes the need to bridge India's linguistic divides as a given. Complex activities, such as translation, should not be prioritized in the beginning phases. Increasing the amount of information available in each language online, as well as enabling services like search and voice interfaces and same-language subtitling in movies to increase functional literacy, might pave the way to thriving digital local language ecosystems.

1.2 Structuring and Mining Public Data

Every department of the government generates records that are available in the public domain. In 2005, the Indian government passed the “Right to Information” act, which enables individuals to query governmental organisations for particular types of information. This facility—a positive step towards bringing transparency—has already been used to good effect by individuals and civil society.Yet, to make accountability and efficiency *intrinsic* to public-related offices, it is necessary to build pipelines that deliver *structured* data. In this regard, it is instructive to consider Berners-Lee's 5-starcategorisation of open data [7], which is reproduced in Table 1.

No matter the file type (scan, picturetable) or encoding, data that is freely accessible online falls under the lowest category. When data is organized (in a table, rather than as free-form text) and connected to other sources, it is much easier to use. Dependable, well-organized data is the backbone upon which useful programs and services may be constructed. In the future, it will be imperative that data system design strive for perfection. Looking backward also reveals an interesting possibility for AI. However, there is a significant quantity of data, particularly within the past several decades, that does fit at least the 1-star condition (being available in digital form, and accessible over the Internet), even if the vast majority of historical data does not. Unfortunately, until this data is put into an organized form, the useful information it may contain may stay buried.

Table 1: 5-star categorisation of open data, reproduced fromthe web page maintained by Berners-Lee (2010).

⋆	Available on the web (whatever format) but with an open licence, to be Open Data
⋆⋆	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
⋆⋆⋆	as (2) plus non-proprietary format (e.g. CSV instead of excel)
⋆⋆⋆⋆	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
⋆⋆⋆⋆⋆	All the above, plus: Link your data to other people's data to provide context

the data makes it impractical for human annotators to undertake the structuring exercise—but this is certainly something within the reach of modern AI techniques (such as computer vision and NLP).We delve into the details of a specific case study to illustrate that

(1) even "macro" patterns in various data sets are often not known,and (2) gleaning them can provide invaluable inputs for course correction and policy making.

It is well-known that legal matters in India may become bogged down in court for years on end at different appeals stages [34]. We zero concentrate on income tax matters, which face court delays despite having designated appellate authorities (the Assessing officer, the Commissioner of Income Tax Appeals (CIT(A), and the Income Tax Appellate Tribunal [18]). The High Court (HC) and thereafter the Supreme Court (SC) hear appeals from the ITAT. As of March 2015, the number of appeals and the disputed sums are shown in Table 2 at various stages of litigation.

One may assume that situations with bigger sums in dispute would be appealed to higher levels, however the last column of the chart reveals otherwise. If you compare the average amount in dispute at ITAT to that at CIT(A), you'll see that it lowers by around a third as you go from ITAT to HC and another third as you move from HC to SC. Obviously, knowing the cause of this tendency is essential for developing plans to shorten the time it takes for cases to go through the different stages of appeal. Two possible causes include (1) the high number of cases filed over the last decade, many of which are still ongoing at lower levels and impacting the averages, and (2) the fact that the government is the more frequent appellant at levels beyond ITAT, for purposes of creating precedent. It seems like it would be easy to check whether any of these hypotheses is right, but unexpectedly, we don't know the answers to the questions we've posed below.

How many lawsuits at each level of appeal are started by taxpayers against the government?

How often do the government and taxpayers win at each stage of appeal?

How much do taxpayers and governments often disagree on, on average, in appeals?

How long does the typical case take to get from first evaluation to a final ruling?

For starters, the ITAT and SC both maintain their own website with its own unique database structure; until almost a decade ago, neither website was even searchable. Legal researchers in India can't do their jobs without Indian Kanoon (https://indiankanoon.org/), a website that provides free search services and custom scrapers.

Appellate authority	Number of appeals	Amount in dispute (Rs)	Average per case (Rs)
CIT(A)	2.32 L	3.84 LC	1.6 C
ITAT	37,506	1.45 LC	3.9 C
HC	34,281	37,684 C	1.09 C
SC	5,661	4,654 C	82 L

Table 2: Appeals at different levels of litigation [20] (L = lakh = 10^5; C = crore = 10^7).

Typical ITAT judgements have multiple mentions of Rupee amounts; the only way to extract the dispute amount is from a description in natural language. For exam- ple, in *M/S Jain Furnishing vs. ACIT* (accessed November 10, 2017, https://indiankanoon.org/doc/78538052/), the dispute amount is thesum of the two amounts mentioned in the following sentence:

> "The assessee in this appeal challenged the addition of Rs. 15,609/- on account of municipal taxes and ad-dition of Rs. 4,80,000/- disallowing part of the rent."

While it would not be trivial, it certainly appears feasible to train an NLP method to extract relevant fields such as the dispute amount from tax judgements, especially if domain knowledge can also be exploited. This simple technical intervention could eventually helpidentify blockages in the tax appeal hierarchy, and save precious time and resources. Similar opportunities abound in other areas of India's legacy data. For example, several opportunities in the political sphere are explored at the Trivedi Centre for Political Data (accessed November 11, 2017, https://tcpd.ashoka.edu.in/new-about-us/).

1.3 HEALTHCARE

Artificial intelligence (AI) technology have the potential to dramatically improve access to excellent health care in underdeveloped nations. A lack of trained medical professionals who are willing to work outside of major cities is a major issue in this field [51]. India has an estimated 15.2 trained health professionals per 10,000 people, but the WHO criteria is between 22.8 and 59.4 [23].

The advent of contemporary AI has made it possible to compensate for a lack of consistent laboratory infrastructure and supplement the skills of limited human resources. For instance, Gann et al. (2017) showed that traits that human pathologists aren't normally trained to look for may predict the recurrence of prostate malignancies. Beck et al. (2011) do the same thing for breast cancer prediction, using computational approaches to find and use new, improved characteristics. Another example of computational pathology's effectiveness is a software-controlled microscope that can reliably diagnose malaria in the field with the precision of an expert [19]. Sepsis in newborns is a major cause of death [53]. Time series data from common non-invasive metrics, such as heart rate and respiration in the first few hours of a preterm baby's life, have been shown to predict morbidity with the same level of accuracy as invasive (and sometimes costly and unavailable) laboratory testing [54].

Epidemiological analysis, which seeks to comprehend disease load and response, may also be aided by data-driven algorithms. The POSEIDON research [52] was a well-planned operation that collected information from medical facilities in 880 different Indian cities and towns on the same day. Preliminary study of this data, which includes information from over 200,000 individuals, suggests trends in the number of times people of all ages and genders, with different types of ailments, visit medical institutions. Compared to data sets collected in other countries like Sri Lanka and Singapore, there are also notable discrepancies. Thus, it follows that analysis of massive data sets may give substantial contributions to healthcare policy. Using automated capturing techniques, such as Internet of Things (IoT) enabled medical equipment and app-based forms with location and image-based inputs, AI may provide the technology to digitize health data. The goal would be to build pipelines that automatically provide high-quality data with little to no human interaction.

Computer vision-based inventory and employee monitoring solutions might help reduce such delivery losses.

Despite our focus on healthcare as an illustration "vertical" in this section, it is important to recognize that challenges and good solutions often cut across industry lines. Issues including restricted access to information and education, low-quality service delivery because of a lack of infrastructure and corruption, and the debt trap caused by high out-of-pocket costs all contribute to the overall bad health of a community. In general, it's best to take stock of the interconnected nature of issues rather than diving headfirst into fixing them.

2 RISKS OF AN AI-CENTRIC APPROACH

The fact that AI has so many potential (and sometimes surprising) methods to aid in progress is a source of optimism. It would be foolish, though, not to plan for and prepare for the possible downsides of AI-driven development. We highlight the most pressing issues that arise in India's socioeconomic setting below.

Worker replacement. The tsunami of artificial intelligence that is starting to displace employees throughout the world is starting to hit India as well [11]. McKinsey and Company (2014) conducted a research in which they concluded that "routine clerical, customer service, and sales jobs could be affected by advancements in machine learning and natural language interfaces (speech recognition)," employing a total of 6-8 million people." For a middle-income nation working to lift a big portion of its population out of poverty, a loss of employment on this size might have devastating effects on the economic well-being of many individuals who rely on these wage-earners. There are signs that automation is having an impact on India's much-touted IT sector [58], which might lead to a catastrophe in the coming years as people lose their jobs. Some of AI's potential drawbacks may not become apparent until far later.

Adding fuel to the fire of prejudice. India's caste system is a social stratification that goes back centuries. As a result, prejudice persists in subtle and covert ways, influencing things like pay [4, employment [36], incarceration [46], and bank credit [30]. Concerns about the possibility of bias being picked up by data-driven algorithms have grown in tandem with the development of AI. In the United States, for instance, recidivism rate assessment algorithms [2] are accused of displaying racial biases [16]. Data-driven algorithms used to evaluate employment, loan, or bail applications are vulnerable to bias due to the presence of caste and religious identity markers in names and addresses [33]. Banerjee et al. (2009) ran an experiment that showed indications of caste-based discrimination in applicants to call-center jobs. Even if we assume that human assessors made the judgments in this situation, it is still concerning that those decisions may be used to teach an algorithm to filter applications.

Intensifying the gap between the sexes. Both desktop and mobile Internet users in India are projected to increase, reaching 420 million and 300 million by 2017 [26]. In rural areas of India, where 60% of people have access to the Internet, mobile phones are the major point of entry. While widespread smartphone use appears promising for AI, it may also serve to further marginalize women. There is a 38% gender gap in mobile phone ownership in South Asia, with the gap likely widening due to patriarchal and sexist societal norms [24]. As a result, gender divides (along with other differences deriving from economic and geographical restrictions) may become a dividing line in AI's reach.

The second cause for concern is the persistent gender gap that exists across India's software sector [32]. As a result, there is a serious possibility that the general public's use of AI will be developed with an extreme sexism. The inequity may have devastating long-term effects [59].

Targeted exclusion of the vulnerable. Due of the substantial initial investment required, it is possible that private firms may provide the first impetus. Companies have no special responsibility to solve socially important concerns like equal access, thus it is only natural for them to seek income in areas with substantial profit pools. As a result, the demands of the less lucrative may be overlooked.

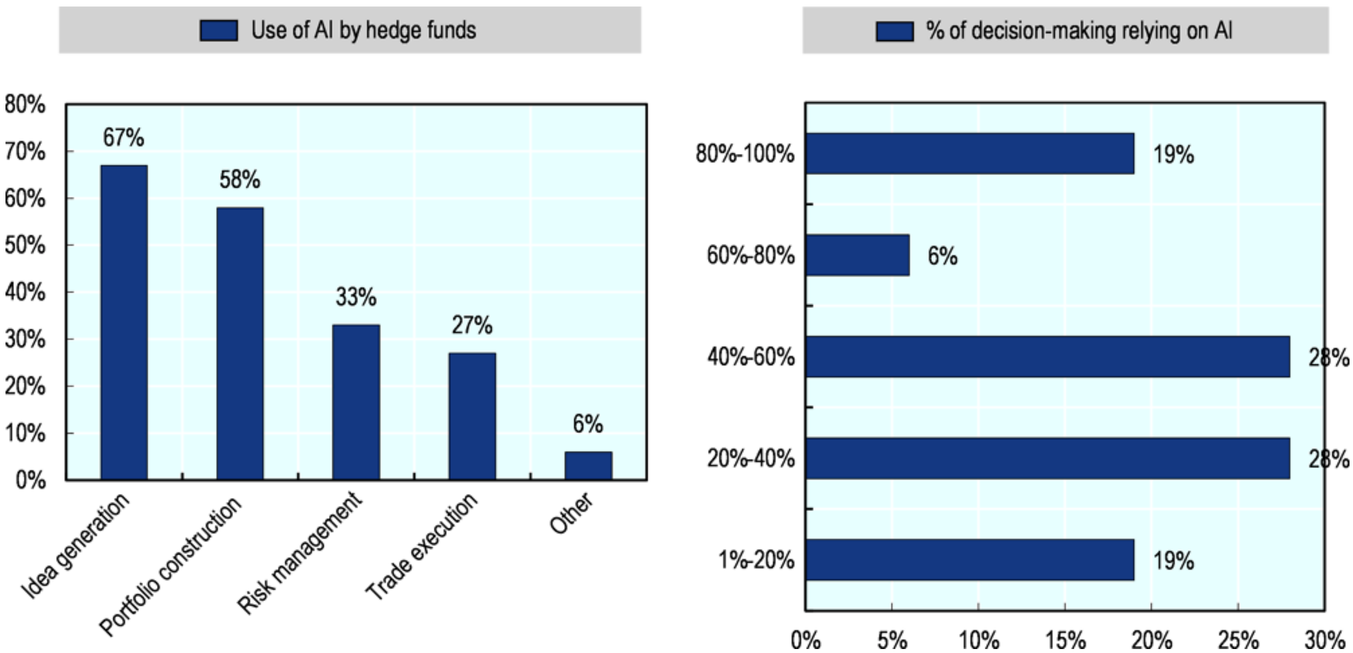
If we take Google's translation efforts between Tamil and Hindi as an example (Figure 1), we can safely assume that the company will not give the latter as much emphasis as it does the English-Mandarin engine. There is a danger that the poor may be more marginalized when commercial interests are superimposed over AI-based markets. A persuasive explanation of this frightening scenario may be found in a recent piece by Calo and Rosenblat (2017).

Asset management2 and the buy-side

Since a few years ago, asset managers and the buy-side of the market have employed AI, mostly to increase risk management and back-office operations while also allocating portfolios. By decreasing investment managers' back-office costs, automating reconciliations, and speeding up operations, the use of AI techniques has the potential to improve operational workflow efficiency. This would reduce friction (costs associated with direct and indirect transactions) and improve performance overall by reducing noise (irrelevant features and information) in decision-making (Blackrock, 2019[3]). (Deloitte, 2019[4]). Asset managers and other institutional investors also use AI to improve risk management since it enables the daily, cost-effective monitoring of hundreds of risk factors as well as the modelling of portfolio performance across a vast array of market and economic scenarios.

The primary use of AI in asset management is the creation of strategies that affect decisions about portfolio allocation. This application depends on the utilisation of large data and ML models trained on such datasets. The asset management sector and the investing community as a whole have traditionally been built on information, and many investment techniques before the arrival of AI relied heavily on data (e.g. fundamental analysis, quantitative strategies or sentiment analysis). The presence of enormous volumes of raw or unstructured data combined with the predictive ability of ML models gives investors who use AI to digest such enormous datasets and reveal insights that subsequently guide their strategies at very short timescales a new informational advantage.

Figure 2.2. **Use of AI techniques by hedge funds (H1 2018)**



Note: based on Industrial research by Barclays, as of July 2018.

Source: (BarclayHedge, 2018[5]).

There is a possible danger of concentration in a limited number of major financial services organisations given the investment necessary for the deployment of AI techniques, since larger and more powerful players may outrun some of their smaller competitors (Financial Times, 2020[6]). Such an investment is not only related to talent and staff capabilities, but also to the financial resources needed to invest in AI technology. Utilizing third-party suppliers helps to some extent reduce this concentration risk, but it also creates new issues with governance, responsibility, and reliance on others (including concentration risk when outsourcing is involved) (see Section 2.3.5).

3 AI FOR DEVELOPMENT: STEPS ANDSAFEGUARDS

In this part, we lay out some guidelines for developing India's artificial intelligence industry. The simple thermometer was essential in the development of both the automotive engine and the air conditioner. India's "vital statistics" can't be improved unless the tools to measure them are put in place, therefore now is the time to go to work. Artificial intelligence relies on having ready access to relevant data in the digital realm. We urge that the development of 5-star data pipelines be prioritized, as described in Section 2.2. As such, it is encouraging to see the government take efforts like "Digital India" (accessed October 6, 2018, https://www.digitizeindia.gov.in/) and "Open Government Data" (accessed October 6, 2018, https://data.gov.in/). The creation of regionally relevant public open sets concerning language, health, crops, markets, and so on would complement government-issued public data. Artificial intelligence (AI) tools like computer vision and crowdsourcing might be used to kickstart the process of building such databases. The development of AI-based solutions must be an isolated activity, or it will not be successful or sustainable. It is crucial that more people, particularly those from underrepresented groups like women, language minorities, and rural regions, get training in AI system development and upkeep. Figure 1 still serves as an illustration of the illuminating power of Google's erroneous translations. Evidently, local speakers who are aware of existing gender biases in their native languages may correct such errors more promptly and efficiently. To create and operate their own translation engines, however, they will want access to data as well as expertise in the relevant technical areas. In this regard, Jain's (2002) suggestion that the Nehruvian, top-down model of information-generation and distribution be actively complemented by the Gandhian, bottom-up model is of particular significance to the development of the AI knowledge network. India's open source community has had some success, and it might be extended to aid in the creation of artificial intelligence (AI) library, standard, and API resources.

India benefits from a well-developed educational infrastructure and a skilled labor force. However, a vast, diversified, and growing nation creates a need for knowledge and expertise that is not met by the current supply. Flagship demonstrations, like Deepmind's AlphaGo program [56], may inspire young minds to seek careers in AI, particularly if they can be set in the Indian context. The release of compelling data sets and the hosting of contests would also help. Research hubs on the home front might lead the way not only in traditional AI but also in related fields. If AI is the new electricity, then society will require both electricians and electrical engineers. It would be helpful if measures were taken to prepare a big number of people to work on application development employing vision, speech, and so on. This might help cushion the blow of employment losses.

The private sector, and particularly startups, will be important in determining and realizing AI's potential advantages across a wide range of industries. In India, IT startups may find a supportive environment with easy access to skilled workers, funding, and extensive consumer bases. As of May, 2017 [44], there were over 300 AI-centric businesses in India, and over USD 100 million had been invested in them since 2014 [55]. This sum, however, pales in contrast to the United States (where investments total over USD 4 billion) and China (where investments total over USD 3 billion). Startups will have difficulties in acquiring data sets and skilled workers; but, if they work more closely with educational institutions, they may be able to overcome the latter problem. on order to minimize losses, young companies may choose to specialize on high-volume, low-margin markets. Savings of even 5 percent in areas like as raw-material waste, energy usage, or rejection rate may add up to significant sums in the manufacturing sector.

India will need to develop safety and quality standards; legislative frameworks addressing data security, privacy, and responsibility; and ethical review committees as artificial intelligence develops.

4 CONCLUSION

The purpose of this study is to begin sketching out the boundaries of artificial intelligence in India. First and foremost, we have seen that AI presents both immediate and alluring benefits and hidden dangers that may not become evident for some time. We think that AI may not only have a net beneficial influence on India's development, but also assist the country skip over conventional growth barriers with careful planning and management.

With so much at risk, it is crucial that the development of AI in India be investigated thoroughly in the classroom. With this publication, we want to plant the seed for future research along these lines. We hope the offered template, as well as the many references, will serve as a jumping off point for further study. We've just done a quick scan of one industry (healthcare), so we know we've missed some ground. Other previous publications [40, 47, 57] have devoted whole sections to agriculture, transportation, education, urban planning, security, employment, entertainment, manufacturing, robotic automation, and the environment, so we direct you there for a more complete listing of verticals. More detailed and quantitative investigation is necessary for the threats we identified in Section 3.

REFERENCES

1. Nisha Agrawal. 2016. Inequality in India: what’s the real story? *World Economic Forum* (2016). Accessed October 26, 2017,www.weforum.org/agenda/2016/10/inequality-in-india-oxfam-explainer.

2. Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2017. Machine Bias. *Pro Publica* (2017). Accessed October 10, 2017, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

3. Abhijit Banerjee, Marianne Bertrand, Saugato Datta, and Sendhil Mullainathan. 2009. Labor market discrimination in Delhi: Evidence from a field experiment. *J.Comparative Economics* 37, 1 (2009), 14–27.

4. Biswajeet Banerjee and J. B. Knight. 1985. Caste Discrimination in the Indian Urban Labour Market. *Journal of Development Economics* 17, 3 (1985), 277–307.

5. Michele Banko and Eric Brill. 2001. Scaling to Very Very Large Corpora for Natural Language Disambiguation. In

Proceedings of the 39th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 26–33.

6. Andrew H Beck, Ankur R Sangoi, Samuel Leung, Robert J Marinelli, Torsten O Nielsen, Marc J van de Vijver, Robert B West, Matt van de Rijn, and Daphne Koller. 2011. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Science Translational Medicine* 3, 108 (2011), 108ra113.

7. Tim Berners-Lee. 2010. Linked Data. (2010). Accessed October 28, 2017, www.w3.org/DesignIssues/LinkedData.html.

8. Laurent Besacier, Etienne Barnard, Alexey Karpov, and Tanja Schultz. 2014. Au-tomatic Speech Recognition for Under-resourced Languages: A Survey. *Speech Communication* 56 (2014), 85–100.

9. Nicola J. Bidwell. 2016. Moving the centre to design social media in rural Africa.

10. *AI & Society* 31, 1 (February 2016), 51–77.

11. Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam Kalai. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *Advances in Neural Information Processing Systems 29*. Curran Associates, 4349–4357.

12. Erik Brynjolfsson and Andrew McAfee. 2016. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company.

13. Ryan Calo. 2016. Robots as Legal Metaphors. *Harvard Journal of Law and Technology* 30, 1 (2016), 209–237.

14. Ryan Calo and Alex Rosenblat. 2017. The Taking Economy: Uber, Information, and Power. *Columbia Law Review* 117 (2017), 1623–1690.

15. Subrata Chattopadhyay. 2015. Corruption in healthcare and medicine. *Indian Journal of Medical Ethics* 10, 3 (2015), 153–189.

16. Mike Cooley. 1995. The myth of the moral neutrality of technology. *AI & Society*

17. 9, 1 (March 1995), 10–17.

18. Kate Crawford. 2016. Artificial Intelligence's White Guy Prob- lem. *The New York Times* (2016). Accessed October 4, 2017, https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences- white-guy-problem.html.

19. Credit Suisse. 2014. Global Wealth Databook 2014. (2014). Accessed Oc- tober 27, 2017, http://publications.credit-suisse.com/index.cfm/publikationen- shop/research-institute/global-wealth-databook-2014/.

20. Pratik Datta, B. S. Surya Prakash, and Renuka Sane. 2017. Understanding Judicial Delay at the Income Tax Appellate Tribunal in India. *National In- stitute for Public Finance and Policy* (2017). Accessed October 28, 2017, nipfp.org.in//media/medialibrary/2017/10/WP_2017_208.pdf.

21. Charles B Delahunt, Courosh Mehanian, Liming Hu, Shawn K McGuire, Cary R Champlin, Matthew P Horning, Benjamin K Wilson, and Clay M Thompon. 2015. Automated microscopy and machine learning for expert-level malaria field diagnosis. In *Global Humanitarian Technology Conference (GHTC), 2015 IEEE*. IEEE, 393–399.

22. Department of Revenue. 2016. Report of the Comptroller and Auditor General of India. (2016). Accessed Novem- ber 13, 2017, www.cag.gov.in/sites/default/files/audit_report_files/ Union_Direct_Tax_Compliance_Revenue_Report_3_2016_ Depart-ment_Revenue.pdf.

23. Ariel Felner. 2016. The Israeli AI Community. *AI Magazine* 37, 3 (2016), 118–122.

24. Peter H Gann, Lingdao Sha, Virgilia Macias, Neeraj Kumar, and Amit Sethi. 2017. Computer Vision Detects Subtle Histological Effects of Dutasteride on Benign Prostate Tissue. *The FASEB Journal* 31, 1 Supplement (2017), lb520.

25. Global Health Workforce Alliance and World Health Organization. 2013. A Uni-versal Truth: No Health Without a Workforce. *WHO Report* (2013). Accessed November 8, 2017, www.who.int/workforcealliance/knowledge/resources/hrhre-port2013/en/.

26. GSMA. 2015. Bridging the gender gap: Mobile access and us-age in low- and middle-income countries. (2015). Accessed October 10, 2017, www.gsma.com/mobilefordevelopment/wp-content/uploads/2016/02/GSM0001_03232015_GSMAReport_ NEWGRAYS-Web.pdf.

27. Julia Hirschberg and Christopher D. Manning. 2015. Advances in natural language processing. *Science* 349, 6245 (2015), 261–266.

28. IAMAI and Kantar IMRB. 2017. Internet in India – 2016. (2017). Accessed October 26, 2017, bestmediainfo.com/wp-content/uploads/2017/03/Internet-in- India-2016.pdf.

29. Ashok Jain. 2002. Networks of Science and Technology in India: The Elite and the Subaltern Streams. *AI & Society* 16, 1–2 (May 2002), 4–20.

30. Deepak Khemani. 2012. A Perspective on AI Research in India. *AI Magazine* 33, 1 (2012), 96–98.

31. Philipp Koehn. 2005. Europarl: A Parallel Corpus for Statisti-cal Machine Translation. (2005). Accessed November 8, 2017, https://homepages.inf.ed.ac.uk/pkoehn/publications/europarl-mtsummit05.pdf

32. Sunil Mitra Kumar. 2013. Does Access to Formal Agricultural Credit Depend on Caste? *World Development* 43, C (2013), 315–328

33. Anoop Kunchukuttan, Abhijit Mishra, Rajen Chatterjee, Ritesh M. Shah, and Pushpak Bhattacharyya. 2014. Śhata-Anuvādak: Tackling Multiway Translation of Indian Languages. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014)*. European Language Resources Association, 1781–1787.

34. Jadine Lannon. 2013. Research Initiative: Women in India's IT Industry. *The Centre For Internet & Society* (February 2013). Accessed November 13, 2017, cis-india.org/internet-governance/blog/women-in-indias-it-industry.

35. Issie Lapowsky. 2017. One State's Bail Reform Exposes the Promises and Pitfalls of Tech-Driven ustice. *Wired* (2017). Accessed October 10, 2017, www.wired.com/story/bail-reform-tech-justice.

36. Law Commission of India. 2014. Report No. 245: Arrears and Backlog: Cre- ating Additional Judicial (wo)manpower. (2014). Accessed October 27, 2017, lawcommissionofindia.nic.in/reports/Report_No.245.pdf.

37. Stephen E. Little. 1999. Science, technology and society in East Asia: Frameworks for the challenges of the next century. *AI & Society* 13, 3 (September 1999), 247–262.

38. S Madheswaran and Paul Attewell. 2007. Caste Discrimination in the Indian Urban Labour Market: Evidence from the National Sample Survey. *Economic and Political Weekly* 42, 41 (2007), 4146–4153.

39. Miwao Matsumoto. 1999. The .Japan problem. in science and technology and basic research as a culture. *AI & Society* 13, 1–2 (March 1999), 4–21.

40. McKinsey and Company. 2014. India's technology opportunity: Transforming work, empowering people. (2014). Accessed Oc-tober 10, 2017, www.mckinsey.com//˜media/McKinsey/Industries/High Tech/Our Insights/Indias tech opportunity Transforming

41. work empowering people/MGI India tech_Full report_ December 2014.ashx.

42. Anthony R. Measham and Meera Chatterjee. 1999. *Wasting Away: The Crisis of Mulnutrition in India*. IBRD/World Bank.