LANDSCAPE OF ARTIFICIAL INTELLIGENCE / MACHINE LEARNING RESEARCH IN INDIA

A REPORT BY ITIHAASA RESEARCH AND DIGITAL
2018

itihaasa
Research and Digital

CONTENTS

Foreword 02

Executive Summary 04

1. Introduction 06

2. AI/ML research in India 07

2.1. What is AI/ML 08

2.2. A brief history of AI / ML research in India 09

2.3. Focus of research in India 12

2.4. Emerging research areas in AI/ML 14

2.5. Challenges faced by researchers in India 16

2.6. Funding for research 18

2.7. Measuring the impact of research 19

3. Research collaborations and learning from other countries 21

3.1. AI initiatives in other countries 22

3.2. Comparison of research publications and citations across countries 24

3.3. Research collaborations 24

3.4. Industry collaborations 26

4. Translating research into application 27

4.1. Applying traditional analytical and advanced AI techniques 29

4.2. Applying AI to Healthcare 30

4.3. AI and employment 31

5. Connect with Brain Research 32

5.1. A brief history of brain-inspired AI/ML 33

5.2. AI/ML and neuroscience – Perspectives from Indian researchers 33

6. Recommendations 36

6.1. Increasing the number of PhD students in India 37

6.2. Augment computing infrastructure for AI/ML research 38

6.3. Create India-specific AI challenges, tools and data-sets 39

6.4. Set up Centers of Excellence (CoEs) for AI/ML research 40

6.5. Adopt an AI Grand Challenges approach 40

6.6. Link institutional mechanisms to start-up ecosystem 41

Appendix: List of researchers interviewed for this study 43

About itihaasa Research and Digital 44

FOREWORD

We have come a long way since Homo Sapiens - our species - first made an appearance about two hundred thousand years ago. Stone tools were first used by Australopithecus Africanus about 3.4 million years ago, and fire was first controlled by Homo Erectus about six hundred thousand years ago. The start of the modern scientific revolution in the mid-16th century began with the publication of Nicolaus Copernicus' On the Revolutions of the Heavenly Spheres, the seminal work on the heliocentric theory. Ever since the advent of the steam engine in 1765, we have witnessed a continuous stream of technology inventions that has improved human productivity. According to a report by Barclays analysts, if human productivity was 100 units in 1765 it is increased to 3000 units today. In fact, the last five decades have seen an increase in human productivity from about 1500 units to 3000 units. And this steep increase coincides with the adoption of IT including personal computing, software, Internet, e-mail, mobile communications etc. The next phase, the fourth industrial revolution, is powered by advances in AI/ML and brain science and is blurring the lines between the physical, digital, and biological domains.

Let us look at the evolution of world from the economic lens. Robin Hanson, a researcher at the Future of Humanity Institute in Oxford University, modeled the step change in economic activity. If humans continued as a hunter – gatherer society, then the world's economy will double every 224,000 years. If we continued as a purely farming society, then the world economy doubles every 909 years. And if we continue as an industrial society - the first three industrial revolutions - then the world economy will double every 6.3 years. We can imagine the next economic shift powered by the science and technological advances of the fourth industrial revolution.

Where is India placed in shaping this next revolution? I am optimistic for India. It is heartening to note that India is continuing on the path of nurturing science and technology advancement, and using the resultant knowledge for the benefits of its citizens. However, there is more that should be done. India invests 0.8% of its GDP in research – 0.6% is by the Government and 0.2% is from private contribution. Countries like China and South Korea invest 4% of GDP in research and the USA 2.7%. We should increase research spending to 3% of GDP – 1.5% from Government and 1.5% from private sources. Private funding has played an important role in the transformation of emerging nations. For example, in the mid-2000s about 65% of Taiwan's R&D was from private sector funding up from about 30% in the past. India is at a point in history where private funding can supplement government funding in research, and this includes philanthropy. This may also be the time to tweak the current rules on corporate social responsibility that mandates that 2% of average net profit be spent on specified social activities.

What if instead, 1% of such funds can be used to fund science and technology research that shapes our future? Another current development that is heartening to note is the revival in philanthropic funding for Indian science. I believe that this will only grow in future.

The Government is seized of the importance of building national capability in AI. The Indian educational institutions and research community, industry and the start-up community are equally enthused about this task. India has a strong foundation in AI/ML and building one in brain science. If we treat this as a national priority, we will succeed in building world-class capabilities like we have successfully done in domains like IT services and space research. I hope that the landscape studies of AI/ML and brain research in India will identify important focus areas of research and translational research and help spark a dialog among researchers, policy makers, industry and philanthropists in India.

Let me end on a philosophical note. While artificial general intelligence - that is human-like intelligence - is still only an aspirational goal of AI, it may be prudent for us to deliberate on what it all means for furthering of human knowledge. I. J. Good, chief statistician of Turing's code breaking team wrote in 1965, "Let an ultra-intelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultra-intelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind. Thus, the first ultra-intelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control." Will we ever reach such a point of technological singularity? Will human ingenuity and spirit triumph? Only time will tell. We must nevertheless move forward on this journey enthusiastically and judiciously.

Kris Gopalakrishnan
Chairman itihaasa Research and Digital,
Co-founder Infosys

EXECUTIVE SUMMARY

This report provides insights into the current landscape of Artificial Intelligence (AI) and Machine Learning (ML) research in India. It is based on primary interviews of 25 AI/ML researchers. Our objective for this study is to provide a nucleus for starting a dialogue among stakeholders and catalyze action to significantly move the needle in AI/ML research in India.

AI is intelligence demonstrated by machines. ML is a subfield of AI and uses statistical techniques to give computer systems the ability to learn (e.g., progressively improve performance on a specific task) with data, without being explicitly programmed.

The first course in AI in India is believed to have been taught at IIT Kanpur in the 1960s. India initiated the Knowledge-Based Computing Systems (KBCS) project in 1986 as part of the Indian Fifth Generation Computer Systems (FGCS) research program. Today, the national institutes of importance like IISc, IIT Bombay, IIT Delhi, IIT Madras, IIIT Hyderabad, IIT Kanpur, IIT Kharagpur, etc. are the top universities for AI/ML in India. There are about 50 to 75 principal researchers in India in the domain.

The focus of Indian AI/ML research can be categorized under three buckets:
Sensing – document analysis, computer vision etc.
Comprehending – natural language processing, probabilistic decision making etc.
Responding – creating a complete situated learning agent, disease incidence prediction etc.

We observed some emerging themes of research like 1) unsupervised learning, 2) reinforcement learning, 3) explainable AI, 4) causal modeling of AI, and 5) resource efficient ML 6) AI and blockchain.

The researchers identified certain key challenges they have to overcome to achieve world-wide impact in AI/ML research – 1) quantity and quality of students entering AI/ML research in India, 2) computing infrastructure, 3) resources and administrative bottlenecks, 4) lack of good quality labeled data sets and 5) a siloed research approach within a university.

Indian AI/ML researchers in universities seem to receive adequate funding support from both Government agencies and industry for research apart from funding from the universities they are affiliated to.

More than 90% of the academic researchers in the study mentioned that good quality research publications and subsequent citations are the most important measures of impact of research. They also identified societal impact as another important measure of success.

About three fourths of the researchers we interviewed as part of the study had collaborations with researchers in universities and institutions other than their own. These research collaborations exist equally with researchers in other Indian universities as well as with researchers outside India.

About 90% of the researchers we spoke to had active collaborations with industry, or had collaborated with industry in the past, or in preliminary discussions to collaborate. Such a high percentage of collaboration is a reflection that the industry adoption of AI/ML to solve their challenges has reached an inflection point in terms of practical applications.

The researchers identified different domains where AI/ML research could be meaningfully applied – healthcare, financial services, hi-tech & communications, retail, education, agriculture, smart city and transportation, and India-specific challenges like monsoon prediction and Indian language processing. The researchers were predominantly of the belief that AI will not lead to a large scale loss of human jobs. While the researchers in general agree that neuroscience research will have a bearing on AI/ML development in the future, less than 10% of them currently have an active research project involving neuroscience / computational neuroscience. This contrasts with the study we did with neuroscience researchers in India, where we found at least 40% of them doing / collaborating on research that combined neuroscience & AI/ML.

The study makes six recommendations for furthering AI/ML research in India.

- Increasing the number of PhD students in India – by creating a special fund to attract world-class faculty to India, by instituting research fellowships for PhD and post-doc students and developing programs to inculcate an interest in AI/ML in the undergraduate students
- Augment computing infrastructure for AI/ML research by setting up a national high-performance computing infrastructure that is rich in Graphics Processing Units (GPUs) and specialized-hardware for AI research, while encouraging the capability to design and make such computing systems in India
- Create India-specific AI challenges, tools and data-sets – by focusing on India specific problems that affect large numbers of citizens and by creating resource repositories and data-sets for research in India
- Set up Centers of Excellence (CoEs) for AI/ML research which are multi-disciplinary including different engineering and humanities disciplines and which facilitates a close interaction with the industry
- Adopt an AI Grand Challenges approach so that the efforts of the different researchers from multiple disciplines are channelized towards a common purpose
- Link institutional mechanisms to start-up ecosystem and strengthen the academic incubators to help translate AI/ML research to market applications

1. INTRODUCTION

India has leapfrogged in building a world class capability in information technology (IT). The capability in IT has powered India's economic development in the past three decades. IT now contributes to 9.3% of India's GDP according to the Ministry of Electronics and Information Technology, Government of India[1]. While this is a significant achievement, India cannot afford to be complacent since IT is evolving rapidly. India has realized the impact of this change and is investing in the IT domains of modern artificial intelligence and machine learning (AI/ML). While the jury is still out on how well India has taken advantage of its dominant position in IT to acquire a capability in AI/ML, it is a good time to assess Indian AI/ML research landscape.

AI has captured India's public imagination. The Indian Prime Minister is reported to have discussed AI at least ten times in the recent past[2]. In the 2018 Budget of Government of India, the Finance Minister announced that NITI Aayog will establish a national program to conduct research and development in AI/ML[3]. These actions validate the importance India places on AI/ML. A study on research trends in AI in India based on analyzing data from SCOPUS journal and arXiv dataset reveals that India ranks fifth in the world in terms of number of citable documents in AI/ML between 2010 and 2016[4].

AI/ML is the immediate frontier in IT - a domain where India is a dominant global player. It is therefore imperative that India nurtures its capabilities in AI/ML to maintain its dominant position in IT. The fountainhead of building this capability is in nurturing research that in turn feeds into teaching and large scale capability building.

Our objective for this study is to provide a nucleus for starting a dialogue among stakeholders that results in concrete action plans to significantly move the needle in AI/ML research in India.

The report is organized as follows. Section 2 provides a brief history and the current state of AI/ML research in India; specific challenges researchers face and the state of funding. Section 3 analyzes the current collaborations of Indian researchers and provides an overview of national-level AI strategies of other countries. Section 4 looks at the state of translating Indian AI/ML research into applications, delves into application of AI/ML in healthcare and analyzes the impact AI will have on human employment. Section 5 analyzes the connections between AI/ML and neuroscience / computational neuroscience. Section 6 provides a set of recommendations for furthering AI/ML research in India.

1 http://meity.gov.in/content/fact-sheet-it-bpm-industry
2 https://www.analyticsindiamag.com/10-times-indian-prime-minister-narendra-modi-propagated-artificialintelligence/
3 https://economictimes.indiatimes.com/tech/software/budget-2018-government-to-push-research-effortsin-artificial-intelligence-says-arun-jaitley/articleshow/62738437.cms
4 https://neelshah18.github.io/Scopus-analysis.html

2. AI/ML RESEARCH IN INDIA

In this section we provide an overview of AI/ML, and its evolution in India. We also estimate the number of AI/ML researchers in India based on the expert estimates, and triangulate it with secondary data. A detailed description of the focus areas and projects for Indian researchers in AI/ML follows. And to complete this section, we look at how researchers measure impact of their research.

2.1. What is AI/ML

AI/ML is increasingly becoming all pervasive. The 337 million[5] Indians who use a smartphone are already experiencing the power of AI/ML even if they do not know what it is. Results for a search query in Google, feeds and facial recognition in Facebook, recommendations in Amazon or Netflix all use machine learning. Let us look at what AI and ML mean.

AI, sometimes called machine intelligence, is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals[6]. In computer science, AI research is defined as the study of intelligent agents which are devices that perceive their environment and take actions that maximize their chance of successfully achieving their goals. Colloquially, the term AI is applied when a machine mimics cognitive functions that humans associate with other humans, such as learning and problem solving. AI was a term coined by John McCarthy, an American computer scientist, in 1956 at the now famous Dartmouth Workshop[7].

ML is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" (e.g., progressively improve performance on a specific task) with data, without being explicitly programmed[8]. The term was first used by Arthur Samuel, an American computer scientist[9]. ML is considered as a subset of technologies that power AI. It provides AI the ability to automatically learn and improve from experience. ML is focused on the development of computer algorithms. These algorithms typically learn from existing data, and make predictions of new data.

AI/ML is not a new domain, and has been evolving from the 1950s. While AI/ML showed much promise, it had not lived up to its expectations in the past. For example, the mid 1970s and the period between the mid 1980s and mid 1990s is called the “AI winter”. The past few years has been marked by IT innovations that have advanced AI/ML to an extent that we are now in an “AI Spring”. The important drivers of this “AI Spring” are,

- Access to powerful and cheaper computing power and data storage via. the cloud
- Exponential increase in availability of (often labeled) digitized data

5 https://www.emarketer.com/content/more-than-a-quarter-of-india-s-population-will-be-smartphoneusers-this-year
6 https://en.wikipedia.org/wiki/Artificial_intelligence
7 https://en.wikipedia.org/wiki/Dartmouth_workshop
8 https://en.wikipedia.org/wiki/Machine_learning
9 Some Studies in Machine Learning Using the Game of Checkers by Arthur Samuel in IBM Journal of Research and Development. 3 (3): 210–229, 1959.

With an explosion of labeled digitized data in the past couple of decades, a new computing model is emerging – it is the learning machine paradigm. The core objective of a learner in the learning machine paradigm is to generalize from its experience. Generalization in this context is the ability of a learning machine to perform accurately on new, unseen case after having experienced a learning data set of many cases that have occurred before. The training examples come from some generally unknown probability distribution which is often considered a representative of all possible occurrences. The learning machine has to build a general model about this context that enables it to produce sufficiently accurate predictions in new cases. These models form the core of ML. Since training sets are finite and the future is uncertain, ML theory usually does not yield perfect guarantees of the performance of algorithms. Instead, probabilistic bounds on the performance are offered. ML is often at the heart of AI technologies like image recognition, autonomous driving, learning to play games like Go, etc.

AI/ML is on the world's center-stage today. Not surprisingly, three of the ten breakthrough technologies for 2018 identified by MIT Technology Review are AI/ML related[10]. First, Cloud Based AI/Ml which is democratizing this technology and making it accessible to one and all. Second, Sensing City which is a vision of a smart city powered by AI/ML and IoT technologies including autonomous vehicles, robots and self-regulating systems. Third, Dueling Neural Networks where two AI systems can spar with each other to create ultra-realistic original images or sounds, something machines have never been able to do before. This gives machines something akin to a sense of imagination, which may help them become less reliant on humans. While these are emerging applications, AI/ML is already having a profound impact today.

2.2. A brief history of AI / ML research in India

The first course on AI in an Indian educational institution is believed to have been introduced in IIT Kanpur in the late 1960s, when Prof. H. N. Mahabala returned from MIT after having spent a year and interacted with Prof. Marvin Minsky. Early research in AI was started by Prof. R. Narasimhan at TIFR. The Machine Translation for Indian languages project at IIT Kanpur and the Optical Character Recognition project at ISI Kolkata in the 1980s may have been some the earliest AI/ML research projects in India. The Department of Electronics, Government of India, with the assistance of the United Nations Development Program (UNDP) decided in 1986 to support a five-year project, the Knowledge-Based Computing Systems (KBCS) project as part of the Indian Fifth Generation Computer Systems (FGCS) research program[11].

The KBCS project focused on emerging developments in[12]:

10 https://www.technologyreview.com/lists/technologies/2018/
11 The Knowledge-Based Computer System Development Program of India: A Review, Patrick Saint-Dizier, AI Magazine Volume 12 Number 2 (1991),AAAI
12 Chapter 11, Computer Education in India: Past, Present and Future, Utpal Kumar Banerjee

- Intelligent man-machine interface
- Knowledge-based processing and management function
- Problem solving and inference making function
- Development of parallel processing platforms for KBCS

A key technical goal of the KBCS project was to develop state-of-the-art computer (AI) programming environment in which major R&D effort can be carried out.

The nodal centers under KBCS included Center for the Development of Advanced Computing (CDAC), the central government's Department of Electronics (DoE), Indian Institute of Science (IISc) for parallel processing, Indian Institute of Technology (IIT) Madras for expert systems for diagnosis, Indian Statistical Institute (ISI) for image processing, National Centre for Software Technology (NCST) for expert systems and natural language processing, and Tata Institute of Fundamental Research (TIFR) for speech processing[13]. During the period 1986-95 each nodal center received a total of INR 15 million. Approximately fifteen Ph.Ds are said to have been produced by the various nodal centers during this period. Each node had between 20 and 35 full-time researchers (not including students) involved in from 5 to 8 major projects.

The KBCS applications included expert systems for government administration, expert systems for engineering and medical applications, intelligent tutoring-authoring systems, computer vision system applications, and KBCS applications in and for ancient Indian sciences. They also included language-processing technologies, including natural language processing (access to databases and large corpora of texts), machine translation, speech-vision processing systems, and Indian language–processing environments. Some of the early research in AI was motivated by societal needs.

Early examples of AI research in applications in Indian context include:

- Eklavya, a knowledge-based program designed to support a community health worker in dealing with symptoms of illness in toddlers (developed by IIT Madras).
- Language teaching system Vidya (developed by NCST)
- Flight scheduling expert system Sarani (developed at CDAC, Mumbai),
- A speech synthesis system developed for the railways by TIFR
- IISc project group under sponsorship from ISRO, built a highly successful image processing facility using AI and computer vision techniques

The KBCS '89 conference was organized by the NCST. This was probably the first AI/ML conference in India. During the conference, 41 papers were presented and 6 invited lectures were given. Papers were presented by researchers from a large number of

13 A Detailed Report on R&D at Indian Computer-Science Establishments, Prof. Krithi Ramamritham, 1995

countries, including the United Kingdom, France, Germany, Japan, the United States, Canada, and Australia. The conference started with a one-day tutorial session on various aspects of AI: task-specific architectures for the construction of knowledge-based systems, knowledge engineering, logic programming systems, language processing, and computational vision

Two areas of AI research have thrived in India post the KBCS project[14]:

- Machine Translation – given the plethora of languages in India, research efforts have been undertaken in IIT Kanpur and later in IIIT Hyderabad and other institutions in machine translation and cross-language information retrieval
- Text to Speech – Research in text-to-speech synthesis systems especially for Indian languages have been undertaken in IIT Madras, IIT Kharagpur and CDAC and later in IIIT Hyderabad

The Centre for Artificial Intelligence and Robotics (CAIR) was established as a research centre in 1986 under the umbrella of the Defence Research and Development Organization (DRDO)[15]. In the early 1990s CAIR developed Nipuna, an expert system for monitoring the health of radars.

While the success of initiatives like KBCS and FGCs are mixed, it provided the nucleus for AI/ML research to nurture in India. Over the years, AI/ML research has grown in India. For instance, the 2001 paper, 'Improvements to Platt's SMO Algorithm for SVM Classifier Design' by S.S. Keerthi et al. is heavily cited (1767 citations) and is considered an important paper in the ML domain. Based on a study analyzing the number of research publications in the AI/ML domain between 2010 and 2016 and confirmed by our study, the national institutes of importance like IISc, IIT Bombay, IIT Delhi, IIT Madras, International Institute of Information Technology Hyderabad (IIIT Hyderabad), IIT Kanpur, IIT Kharagpur, Indian Statistical Institute, Kolkata etc. are the top universities for AI/ML in India[16].
This study also mentions the leading companies in India based on research publications in AI/ML during 2010-2016. They are multinational companies in India like Microsoft, IBM, Yahoo, Xerox, Texas Instruments, Google, Samsung, HP, etc. and Indian companies like TCS, HCL, etc.

According to the Global AI Talent Report 2018[17] India ranks tenth globally in terms of number of PhDs in AI/ML and thirteenth in terms of presentations in top AI/ML research conferences. India has about 386 PhDs compared to about 9010 in the USA. India based researchers had about 44 papers in top conferences compared to 3017 by those in the USA. China has about 413 PhDs and about 206 papers in top conferences. This is just one estimate based on crawling LinkedIn, a popular professional social network.

14 A Perspective on AI Research in India, Deepak Khemani, AI Magazine, AAAI, 2012
15 https://www.drdo.gov.in/drdo/labs1/CAIR/English/indexnew.jsp?pg=about-lab.jsp
16 https://neelshah18.github.io/Scopus-analysis.html
17 http://www.jfgagne.ai/talent/

A popular peer-maintained online community Machine Learning India[18] has 41 ML faculty, 13 academic research groups in ML, 106 ML professionals spread over 90 companies. There is also a list of 748 research papers published by Indian ML researchers. Based on our interviews with AI/ML researchers in India, we estimate that there are about 50 to 75 principal investigators in the domain.

2.3. Focus of research in India

Our study reveals that Indian researchers are inspired to find answers to some challenging problems in AI/ML which can be categorized based on their similarity to what humans do[19].

- Sensing
- Comprehending
- Responding

The inspiration translates into focused research projects. Researchers in Indian universities are involved in different interesting projects covering different aspects of AI/ML including Supervised Learning, Unsupervised Learning and Reinforcement Learning (RL), and Deep Learning (DL) including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and other related research in India. A sample of these research projects that provide a flavor of the breadth of research include

- Sensing
  - Document analysis and recognition
  - Scene text understanding
  - AI/ML algorithms for face / iris recognition
  - Computer vision
  - Tools for building AI systems – neural network modeler
  - Multi-modal knowledge representation
  - Scaling / performance of DL algorithms

- Comprehending
  - Semantic and statistical approaches to text mining tasks, deep text understanding with applications in analyzing blogs, resume processing, comparing contracts, document set mining, question answering systems, novelty detection, summarization etc.
  - Advancing deep learning, reinforcement learning and weakly supervised learning models
  - Natural language processing (NLP), Cognitive NLP and domain specific knowledge models and graphs
  - Machine translation (MT), inter-lingual MT, sentiment analysis of languages –sarcasm, humble bragging
  - Statistical relational learning and lifted inference and learning models
  - Developing algorithms which have mathematical guarantee and involve large-scale data

18 http://ml-india.org/
19 http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion- Paper.pdf

- Non-convex optimization for machine learning
- Learning of visual routines that can tell an agent on what aspects of its visual input to focus
- Representing passages to compare them - how to embed entities, types and relations to infer new edges in knowledge graphs
- Building a search system for curating and searching the annotated web to extend the capabilities of current search engines
- Probabilistic decision making under uncertainty and Deep Reinforcement Learning models
- Detecting bias in and fairness of AI systems

Responding

- Create a complete situated learning agent
- Data mining of healthcare data to build better screening tools, disease incidence prediction, risk factor analysis, etc.
- Cognitively motivated representations of tasks and sub-tasks and implementing them in real robot platforms
- Applying algorithmic game theory in crowd-sourcing, auctions
- Large-scale information extraction and data integration– harnessing dark data for improved decision making
- Learning from limited amounts of information
- Using robust optimization and ML to reduce noise in data
- Automatically extracting information from natural language text in a domain-independent manner
- Code-switch algorithms for switching between languages

Applying machine learning models in

- Computational biology
- Detecting objects in images
- Video segmentation and summarization
- Detection of rare topics in text documents
- Statistical modeling of computer systems
- Biometrics – face / iris recognition. One unique Indian project is a Deep Learning algorithm for new-born face recognition by comparing photos at birth and at the time the baby is discharged from hospital

Robotics

- Multi robots: Effectively mapping an unknown environment using multiple robots
- Robot mechanism: Designing and controlling optimal robot mechanisms
- Mobile robots: Motion planning, Autonomous navigation
- Robotic vision: Spatially intelligent robotic vision system; Monocular reconstruction of dynamic images
- Aerial robots: Unmanned Aerial Vehicles (UAVs) or drones. Applications include visual surveying and parts inspection

The research subsidiaries of multinational companies are also working on AI/ML projects. They leverage the high-quality Indian talent in AI/ML. These applied research projects include

- Resource-efficient ML for Edge and Endpoint IoT Devices
- Machine translation
- Improving catalog quality and address quality using ML
- Improving efficiency of e-commerce using ML – recommendation engines, voice shopping, minimize fraud, etc.
- Medical image-recognition using ML

There are Indian startups working on applied research projects. These include

- Video indexing based on ML
- Analyzing dark data (unstructured data in the form of employee emails, customer communications, presentations, etc.) in enterprises
- Developing an autonomous cargo vehicle

2.4. Emerging research areas in AI/ML

We observed some emerging AI/ML research themes:

Unsupervised Learning

AI/ML research requires lots of data which is curated and labeled. AI/ML systems learn unlike humans who do not need labeled data. One direction of future research is to develop more data sets specific to India and learn to harness the power of the crowd (billion plus population) or leverage Generative Adversarial Networks (GANs) to create synthetic data sets. Another direction of research is to focus on unsupervised learning algorithms in situations where getting labeled data is challenging. There is also a large amount of unlabeled data on the World Wide Web and a never-ending learning agent continuously crawls the web to learn like a human.

Reinforcement Learning

One emerging area of research in ML is Reinforcement Learning (RL). RL models the trial-and-error learning process that humans typically adopt to learn new tasks. RL differs from standard supervised learning models and is needed in situations where explicit instructive signals are not available. RL uses rewards and punishment as signals for positive and negative behavior of the agent. Thus, the RL agent finds a balance between exploring its environment and maximizing its reward. Another area of research is imitation learning, with datasets or humans in the loop, which makes it possible to teach agents complex tasks with no need for explicit programming or task-specific reward function design.

Explainable AI

AI is intended for the benefit of humanity. Right now, this technology happens to be a

black-box i.e. we do not know exactly how AI systems are arriving at their decisions. And their algorithms may be inadvertently influenced by human-biases and may thus adversely impact people's lives. Hence it is very important to focus on 'explainability' or 'interpretability' of AI/ML. Researchers are thus studying ethics and fairness in AI. Researchers are also trying to determine more precise mathematical guarantees for AI/ML algorithms. While this does not unravel the black-box of AI, it helps users to understand how close the results are to the actual phenomenon.

Causal modeling of AI

AI was born 'Symbolic' and was characterized by its focus on expert systems and how to structurally code knowledge. In a couple of decades, the predominant model in AI became 'Connectionist'– where concepts are represented as a set of nodes, numbers, vectors, or matrixes. There is a need now to combine the two models, for instance, current statistical-mode machine learning systems with causal reasoning tools, in defining the next paradigm of AI.

Resource-efficient ML

Some of the researchers believe that India also provides interesting contexts for AI/ML research. The researchers also find the compute infrastructure and labeled-data constrained environment in India as contextual for developing resource-efficient ML algorithms.

Such algorithms help make edge devices smarter and address issues of bandwidth, latency, privacy, battery power etc. For example, the intersection of ML and IoT and an objective to make ML algorithms work in 2KB RAM resulted in a smart cane for the visually challenged in India.

AI and Blockchain

Two areas of research emerge –

- Blockchain as an enabler of AI especially with respect to provenance of data i.e. which data in the 'supply chain' that creates an AI model has come from where
- AI in the service of blockchain – various business objects collect onto a blockchain and data typically belongs to multiple owners. How does one do secure, confidentiality preserving AI?

We analyzed the research papers published in leading AI/ML conferences like Association for the Advancement of Artificial Intelligence (AAAI) and Neural Information Processing Systems (NIPS).

- The top categories of research (based on number of papers submitted in 2017 NIPS) include algorithms, deep learning, applications, probabilistic methods, optimization, theory, neuroscience and cognitive science, and reinforcement learning & planning[20]
- Analysis of AAAI papers database reveals the following top categories of research under

20 https://medium.com/syncedreview/a-statistical-tour-of-nips-2017-438201fb6c8a

AI – machine learning, representation and reasoning, robots, natural language; and within ML the top categories are – neural networks, statistical learning, learning graphical models and reinforcement learning[21]

The research in AI/ML in India is indeed varied and there are some pockets of world-class research excellence. While Indian AI/ML research has grown substantially over the past two decades, the researchers we interviewed pointed that a few fundamental challenges still remain.

2.5. Challenges faced by researchers in India

In order to truly achieve world-wide impact in AI / ML research, a few key challenges have to be overcome.

1.Quantity and quality of students entering AI / ML research in India

2.Computing infrastructure

3.Resources and administrative bottlenecks

4.Lack of good quality labeled data sets

5.Siloed research approach within a university

Quantity and quality of students entering AI / ML research in India

A consistent feedback across researchers is to increase the number of research students in AI / ML in India. India ranks tenth globally in terms of number of PhDs in AI/ML. One researcher estimated that the computer science departments of the top 20 Indian engineering universities would have about 500-600 PhD students enrolled. Around 100 students graduate every year in AI / ML in India, while in China, 2000 students graduate each year.

The number of PhD students is also proportional to the number of faculty. Increasing the number of faculty researchers will increase the number of PhD students enrolled. Post doctoral researchers are an integral part of academic research labs in universities around the world. India does not attract the best post-docs. Many Indian PhDs want to go to universities outside India for a post doc to gain an international experience. The quantum of post-doc fellowships (salary) is also low in India.

The primary pool for PhD students is Masters students. Many of the good Masters students do not seem to be inclined to pursue a PhD. And even if they do, they seem to prefer universities outside India. Thus, some of the Masters projects which have the potential to be expanded and pursued as a research topic often get abandoned. The researchers we spoke to felt that the tacit and background knowledge on the topic is often lost when the Masters student leaves the lab. Even the undergraduate students who are interested in pursuing research in AI/ML seem to prefer universities outside India.

21 https://aitopics.org/search Accessed on 10 Oct 2018

Computing infrastructure

The researchers were very clear that computing infrastructure poses a challenge with respect to conducting cutting-edge AI / ML research. Researchers recount anecdotal evidence suggesting that researchers in multinational information technology companies have close to 800 TPUs (Tensor Processing Units) for crunching results for a research paper. In contrast, the Indian researchers stated that the size of computing infrastructure in top universities in India is nowhere close. We estimated that IIIT Hyderabad has a cluster of over 100 GPUs while each of the older IITs have clusters that have between 70 and 100 GPUs. To compare, Canada's National Infrastructure for AI has more than 3000 GPUs. The researchers need access to more GPUs and specialized hardware for their research purposes. There are certain categories of research problems that are empirical and require intensive calculations. Non-availability of high-performance computing infrastructure will limit the researchers in such fields and also limit their ability to publish papers in Tier 1 conferences and journals in AI/ML.

Resources and administrative bottlenecks

Besides computing infrastructure, the researchers reported challenges in terms of other resources and administrative processes. Most researchers indicated that travel grant for attending leading conferences and paper-submissions is limited or comes with many procedural requirements. One researcher indicated that international travel money needs enhancement from the existing limit of INR 300,000 over 3 years for a faculty and INR 100,000 over 5 years for a student. Another researcher indicated that Government funding for travel poses significant procedural challenges because request for funding has to be put up several months in advance while travel for conferences happens at shorter notice.

Another researcher highlighted the challenge of poor infrastructure support and maintenance in Indian universities. For instance, the air-conditioning infrastructure for the computer centers that house the clusters of GPUs is critical. Air-conditioning should work 24X7 at optimum installed capacity. Sometimes even this is a challenge in India with frequent disruptions to the air-conditioning service. The underlying support infrastructure for high performance computing installations is a soft underbelly in India.

Lack of good quality labelled data sets

The availability of good quality, domain specific public data sets in India, is fairly limited. Labelled, curated data is required to train AI systems. Open Government Data (OGD) Platform India - data.gov.in - is a platform for supporting Open Data initiative of Government of India[22]. The Ministry of Electronics and IT conducted a 24hours hackathon themed on drinking water & sanitation, transport, education, crime & health, based on data sets in OGD platform[23]. There are other data sources in India, like the RBI database of Indian economy, Ministry of Statistics and Programme Implementation dataset etc.[24].

22 https://data.gov.in/about-us
23 http://pib.nic.in/newsite/PrintRelease.aspx?relid=170690
24 https://www.analyticsindiamag.com/top-10-indian-government-datasets-that-you-can-use-for-analyticsprojects/

However, the quality of such data sets needs to be significantly improved. The challenges include data unreliability, insufficient standardization, lack of system (e.g. software systems) and semantic (different departments gathering different information under the same heading, or the same information under different headings) inter-operability, limited data anonymization etc.[25].

Siloed research approach within a university

Some researchers opined that the Indian research system has certain characteristics that come in the way of research, especially AI/ML. The Indian research system tends to be more theoretical rather than empirical. Indian researchers are not 'hands-on' and not 'builders'. AI/ML requires a balance of both theoretical and empirical research. This requires teams of researchers with complementary expertise in theoretical and empirical research to coalesce as a team. Also, the Indian research system adopts a more siloed approach. Typically, we do not find more than a couple of faculty researchers from the same institution working together on the same problem area. There are very few multi-disciplinary AI/ML research centers across Indian universities. In contrast, in the German research model, the chaired professors recruit a few assistant professors to work with them in the research projects in their labs. This provides more researchers and teams with complementary skills to work on difficult research problems. Some of the challenges like improving the quality of high-performance computing infrastructure and funding for research can be mitigated by improving funding. And there are already green shoots on this front.

2.6. Funding for research

Indian AI/ML researchers in universities seem to receive adequate funding support from both Government agencies and industry for research apart from funding from the universities they are affiliated to. Over 90% of the academic researchers in our study confirmed that they have received funding / done consulting engagements for the private industry. This is in sharp contrast to the percentage of Neuroscience / Computational Neuroscience researchers in India who collaborated with industry which stood at around 25%.

- Public sources of funding for research in AI/ML comes from Government organizations like Department of Science and Technology (DST) and its IMPRINT program, Ministry of Electronics and Information Technology (MeitY), Defence Research and Development Organization (DRDO), Bhabha Atomic Research Centre (BARC), Department of Biotechnology etc.

  Multinational subsidiaries like Google India, IBM Research Lab India, Microsoft Research India etc. fund research projects and provides travel grants for Indian researchers to travel to conferences outside India.

- Some illustrative industry-funded research projects for AI in the last couple of year include:
  - IIT Madras has partnered with Bosch and set up the Robert Bosch Centre for Data

25 https://cis-india.org/openness/publications/ogd-report

Science and Artificial Intelligence with a proposed funding of about INR 40 million per year for five years[26]

- At IIT Madras, the Pratiksha Trust has provided INR 300 million to fund three chairs in the Center for Computational Brain Research
- At IIIT Hyderabad, Intel funded a project that led to development of public dataset of Indian driving conditions[27]
- IIT Bombay will collaborate with IBM in its AI Horizons Network as part of a multi-year collaboration to advance AI research[28]
- IIT Kharagpur will set up a Center for AI, with seed funding from Capillary Technologies[29]
- At IIIT Delhi, the Infosys Centre for AI was set up with a grant of INR 240 million from Infosys Foundation in 2016[30]

- The National Supercomputing Mission (NSM) was approved by the Government in 2015 at a total cost of INR 45 billion a seven year period of implementationaccording to the Annual Report 2017-18 of the DST[31]. One of the application areas for using these super computers is AI. 1 petaflop supercomputer is likely to be installed in IIT Kanpur and IIT Kharagpur as part of NSM initiatives
- In mid-2018, IIT Bombay installed a supercomputer that can reach a maximum speed of 1 petaflop[32]. IISc also has a cluster computing infrastructure that can process at speeds up to 1 petaflop[33]
- The report of the Task Force on AI set up by the Department of Industrial Policy and Promotion, Ministry of Commerce and Industry recommends budgetary support for setting up of an inter-ministerial National Artificial Intelligence Mission (N-AIM) with a budget of INR 2.4 billion per year for five years[34]. Out of this budget the Task Force proposes setting aside INR 500 million per year for core activities in AI research.

2.7. Measuring the impact of research

More than 90% of the academic researchers we spoke to as part of the study mentioned that good quality research publications and subsequent citations are the most important measures of impact of research. For the industrial researchers, the predominant measure of their research impact was financial i.e. in terms of monetary benefits either in terms of reduced costs or time, or increased revenues or enhanced product features for the company.

26 https://www.analyticsindiamag.com/bosch-to-invest-%E2%82%B917-billion-in-india-will-focus-onexpanding-bengaluru-smart-campus/
27 https://factordaily.com/india-driving-dataset-iiit-hyderabad/
28 http://www.iitb.ac.in/en/breaking-news/iit-bombay-and-ibm-team-to-accelerate-ai-research-india
29 https://www.indiatoday.in/pti-feed/story/iit-kharagpur-partners-capillary-technologies-for-ai-centre-1180197-2018-03-01
30 https://indianexpress.com/article/education/delhi-starting-july-college-course-on-artificial-intelligence-5023560/
31 https://drive.google.com/file/d/1IPKUdbSx0Da2Zi_ufzC4u-T3jCFzPred/view
32 https://globenewswire.com/news-release/2018/07/03/1532942/0/en/Indian-Institute-of-Technology-Bombay-deploys-Cray-to-Power-Research-and-Education.html
33 http://www.serc.iisc.in/facilities/nvidia-dgx-1-cluster/
34 http://dipp.nic.in/sites/default/files/Report_of_Task_Force_on_ArtificialIntelligence_20March2018_2.pdf

The researchers identified another important aspect of research which is societal impact. This is typically achieved by translating the research to products or services that serve Indian citizens. Some examples include AI/ML based solutions to enhance access to healthcare services, leveraging AI/ML to measuring agricultural yields, predicting weather patterns and determine appropriate sowing cycles, automatic translation to and from Indian text and speech, assessing credit worthiness for microfinance loans, etc.

One note of caution that emerged from the researchers with respect to applying AI/ML to societal challenges is to ensure that the solutions be fail-safe. AI/ML solutions are for augmenting humans and they should not harm the humans directly or indirectly due to inherent biases in the data that was used to train the AI systems. In cases of financial services and healthcare, researchers emphasized the need for explainable or interpretable AI in order to understand precisely how the AI system came to a particular recommendation or prediction.

While India is nurturing a world class research capability in AI/ML, it is imperative to understand and adopt relevant best practices from across the world to accelerate building this capability.

3. RESEARCH COLLABORATIONS AND LEARNING FROM OTHER COUNTRIES

In this section we shall highlight some of the AI/ML research initiatives in other countries and detail the collaborations researchers have with other universities in India and outside India as well as with the industry.

3.1. AI initiatives in other countries

In the past fifteen months, besides India, Canada, China, Denmark, the EU Commission, Finland, France, Italy, Japan, Mexico, the Nordic-Baltic region, Singapore, South Korea, Sweden, Taiwan, the UAE, and the UK have all released strategies to promote the use and development of AI[35]. They share five key themes: ethics; using AI in government and public services; research & development; capacity, skills & education; data & digital infrastructure[36]. Let us look at the national AI strategies of a few countries.

Canada[37]

- Canada was the first country to release a national AI strategy
- The Pan-Canadian Artificial Intelligence Strategy of 2017 is a five-year, C$125 million plan
- It primarily focuses on enhancing research and talent availability in Canada
- The Canadian Institute for Advanced Research leads the Government strategy and three new AI Institutes have been set up: the Alberta Machine Intelligence Institute (AMII) in Edmonton, the Vector Institute in Toronto, and MILA in Montreal

China

- China released its first AI plan in 2017 – 'A Next Generation Artificial Intelligence Development Plan'; subsequently China released 'Three-Year Action Plan to Promote the Development of New-Generation Artificial Intelligence Industry'
- The 3-year plan highlights four key tasks: (1) develop intelligent and networked products such as vehicles, service robots (2) develop intelligent sensors and neural network chips, (3) intelligent manufacturing, and (4) invest in industry training resources, standard testing, and cybersecurity
- China recruits Baidu, Alibaba, Tencent and iFlyTek to the AI 'national team'
- China will build a USD2.1 billion technology park for AI research in Beijing

France[38]

- France has announced a Euro 1.5 billion plan, 'For a Meaningful Artificial Intelligence: Towards a French and European Strategy' – Euro 700 million towards research, Euro 100 million towards startups, and Euro 400 million to industrial projects
- Plans to create a network of four or five research institutes across France
- Create "innovation sandboxes" which would provide an open platform for innovation and offer resources (like super-computers designed for AI) for use in field-testing

35 https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd
36 https://www.oxfordinsights.com/insights/2018/1/23/aistrategies
37 https://www.cifar.ca/ai/pan-canadian-artificial-intelligence-strategy
38 https://www.gouvernement.fr/en/artificial-intelligence-making-france-a-leader

Japan[39]

- Japan released 'Artificial Intelligence Technology Strategy' in 2017. It plans increase its science and innovation budget by JPY 900 billion by 2020 for AI
- It envisages 3 centers of research - RIKEN Center for Advanced Intelligence Project, National Institute of Advanced Industrial Science and Technology (AIST), National Institute of Information and Communications Technology (NICT)
- Areas of interest include industrial productivity improvement, healthcare, medical care and welfare, mobility and information security

Singapore[40]

- Singapore announced 'AI Singapore' in 2017 – a five-year, SGD 150 million national program
- Key initiatives included fundamental research in AI technologies of the future, grand challenges for society and industry, conduct 100 AI experiments, and Makerspace of shared computing
- The AI Apprenticeship Program (AIAP) is a 9-month full-time structured training program comprising 2 months of AI coursework consisting of classrooms, online, mini-projects, and 7 months of on-the-job training on a real-world AI problem.

United Kingdom[41]

- The British government released the AI Sector Deal in April 2018. It plans to increase its R&D spend to 2.7% of its GDP by 2027, created a GBP 725 million industrial strategy challenge fund, and a GBP 1.7 billion for transforming cities fund
- UK plans to build over 1,000 government supported PhD institutions by 2025 and set up a Turing fellowship to support an initial cohort of AI fellows and launch of the Centre for Data Ethics and Innovation
- Has envisaged 4 AI Grand Challenges –AI and data, ageing society, clean growth, and future of mobility

United States[42]

- In May 2018, the US government outlined its approach to AI
- The Federal Government's investment in unclassified R&D for AI and related technologies has grown by over 40% since 2015. The Department of Defense alone plans to spend USD 2.4 billion on AI. Other large spenders include departments of agriculture, veteran affairs and homeland security
- FY2019 Budget Request designates AI and autonomous and unmanned systems as Administration R&D priorities
- Committed USD 200 million in grant funds that were matched by a private industry commitment of USD 300 million towards STEM education with a particular focus on computer science

39 http://www.nedo.go.jp/content/100865202.pdf
40 https://www.aisingapore.org/wp-content/uploads/2017/09/201705031442082191-Press-Release- AI.SG-FINAL-web.pdf
41 https://www.gov.uk/government/publications/artificial-intelligence-sector-deal
42 https://www.whitehouse.gov/wp-content/uploads/2018/05/Summary-Report-of-White-House-AISummit.pdf

3.2. Comparison of research publications and citations across countries

We also analyzed the data from Scimago Journal and Country Rank (SJR) for the above countries over the period 2013 to 2017 and compared it to India[43]. Our analysis shows that:

Country	Citable Documents	Citations	Citable Docs ratio Country:India	Citation ratio Country:India
Canada	5293	19186	0.4	1.03
China	37918	117927	3.1	6.3
France	6815	16781	0.6	0.9
India	12135	18700	1	1
Japan	10653	13863	0.9	0.7
Singapore	2497	12081	0.2	0.6
United Kingdom	9958	34836	0.8	1.9
United States	32421	92612	2.7	5.0

Source: Scimago Journal and Country Rank data for years 2013-2017

- Over the five years, India ranked third in the world in terms of number of citable documents in 'Artificial Intelligence', and ranked fifth in terms of citations
- USA produced 2.7 times the number of citable documents and they were cited 5 times more often
- China produced 3.1 times the number of citable documents and they were cited 6.3 times more often
- UK produced lesser number of citable documents than India (0.8 times) but they were cited more often (1.9 times)
- Canada produced much lesser number of citable documents (0.4) but they were cited slightly more (1.03)
- India compared favourably as compared to France, Japan, and Singapore in terms of number of both citable documents and also the citations

3.3. Research collaborations

About three fourths of the researchers we interviewed as part of the study had collaborations with researchers in universities and institutions other than their own. These research collaborations exist equally with researchers in other Indian universities as well as with researchers outside India.

Based on our interviews with AI/ML researchers in India, there are only about 50 to 75 principal investigators in the domain. Consequently, they tend to collaborate with one another for further research.

43 https://www.scimagojr.com/countryrank.php?category=1702

Some of the Indian universities that have active collaborations include IISc, IIT Bombay, IIT Delhi, IIT Madras, IIT Kanpur, IIT Guwahati, and IIT Gandhinagar. The International Institute of Information Technology (IIIT) are also active in collaborations especially IIIT Hyderabad, IIIT Delhi and IIIT Bangalore.

Research collaborations also take place through research consortiums. For example, SANDHAN was a mission-mode cross-lingual search project involving a consortium of 120 researchers of 12 institutions over a period of 6 years and was funded by Technology Development for Indian Languages (TDIL).

AI/ML research in India typically involves multiple disciplines within computer science (CS), besides mathematics, statistics and electrical engineering. These CS disciplines include fields like core systems, graphics, videos, speech, natural language processing, web language processing, computer vision, and robotics. Only in a few instances in our study did the researchers indicate a multi-disciplinary approach to AI/ML research, spanning other disciplines like biology, materials science, and mechanical engineering. Most Indian researchers maintain the umbilical cord with the university where they completed their stint as a graduate student, post-doctoral researcher, or visiting researcher. Researchers who hold a dual appointment including in an Indian university are helping Indian researchers build a strong collaboration with their parent universities and international professional networks.

Some of the international collaborations by Indian researchers are leading to cutting-edge research projects being funded by international funding agencies. For example, researchers from IIT Delhi and UT Dallas collaborated on a Defense Advanced Research Projects Agency (DARPA) funded project on explainable AI. Similarly, researchers from IISc collaborated in a big data research project funded by DST India- French Institute for Research in Computer Science and Automation (INRIA) France.

Some of the international universities or research institutions with which Indian researchers collaborate include (in alphabetical order): Allen Institute, Carnegie Mellon University, Chalmers University, Duke University, INRIA, France, NUS Singapore, Ohio State University, Princeton University, Purdue, University College London, University of Illinois at Urbana Champaign, University of Maryland, University of Norte Dame, University of Rome, University of Southern California, University of Texas Austin, University of Texas Dallas, University of Washington etc.
(Refer Appendix 2)

The quality of the universities and research institutions that Indian researchers are collaborating signals the quality of AI/ML research in India. There is also an increased pull from AI/ML researchers from other countries wanting to work with Indian researchers based on the quality of Indian research.

3.4. Industry collaborations

About 90% of the researchers we spoke to had active collaborations with industry, or had collaborated with industry in the past, or in preliminary discussions to collaborate. Such a high percentage of collaboration is a reflection that the industry adoption of AI/ML to solve their challenges has reached an inflection point in terms of practical applications.

An illustrative list of companies mentioned as collaborators by the researchers include (in alphabetical order): 1mg, 99 Acres, Accenture, Amazon, Bloomberg, Bosch, Ericsson, Flipkart, GE, Google, Honeywell, IBM, Intel, MathWorks, Microsoft Research, Mindtree, NetApp, Nvidia, Qualcomm, Siemens, TCS, Wipro.

Some of the salient points with respect to industry collaboration include:

- The industry collaborators are from a spectrum of industries – traditional technology majors, industrial majors, IT services and consulting majors, product companies, and startups – both Indian companies and multinationals
- The nature of collaboration ranges from consulting engagements for specific problem statements to research collaboration leading to journal / conference papers
- The extent and specialized capabilities of computing infrastructure in the industry far exceeds what is available in academic institutions in India. In certain types of research problems, this provides a significant leg-up to industrial researchers who dominate research papers published in leading journals / conferences in those problem areas. In some cases, the industry is also challenged often with respect to sharing with outsiders their customer data required for AI/ML research. Keeping these factors in mind, we notice a trend of AI/ML faculty researchers from academia taking up sabbaticals in companies for a period of 1 to 2 years
- Industry is also sponsoring travel grants to academic researchers for travel to good conferences when they have paper presentations
- In the healthcare industry context, the collaborations between technology and medical devices companies and hospitals existed for purpose of obtaining patient data to train the AI/ML models. Privacy considerations guide such data transfers

Academia – industry collaborations work best when there is balance between the academic goals of publications and industry's goals of a solution to a problem. The age old tensions in academia – industry collaboration seem to exist in the AI/ML domain as well. In some cases the industry partner sponsoring a research project in a university has a tendency to equate researchers on a project as employees. And in some cases, the researchers working in an industry-sponsored project are not proactive in showing the results to the industry partners sponsoring the research project. It also appears that industry seems to prefer projects that last between 6 months to a year, while academic researchers tend to favor projects that are a few years in duration. There needs to be some mediating mechanism of extending one-year industry collaboration projects to multi-year projects.

The nature of current collaborations especially with industry leads us to an associated aspect, that of translating research into applications.

4. TRANSLATING RESEARCH INTO APPLICATIONS

AI/ML has become one of the most transformative technologies that businesses have encountered, and its application is being sought across a multitude of industries.

McKinsey has analyzed more than 400 use cases of AI/ML across 19 industries and 9 business functions and found industries like retail, automotive, banking, insurance, logistics and travel to be ones with the greatest potential to derive value from AI/ML[44].

In the Indian context, translational research is considered as a proverbial Achilles' heel. This context is amplified in the AI/ML domain. AI/ML is becoming an engine for the fourth industrial revolution. The fourth industrial revolution is the emerging era powered by the current advances in science and technology especially the intersection of cyber, physical, and biological systems. It is characterized by a fusion of technologies that is blurring the lines between the physical, digital, and biological domains. AI/ML and brain science are all technologies that will power the fourth industrial revolution. AI/ML is now a new factor of production along with labor, capital, and technology.

More importantly, AI/ML has the power to improve the lives of Indians. According to a report by the National Institution for Transforming India (NITI Aayog) the Government of India's apex policy organization, AI/ML has the potential to transform India's nucleus – healthcare, agriculture, education, smart cities and infrastructure, and smart mobility and transportation[45]. According to an estimate by Accenture, AI/ML can add about USD 957 billion to India's GDP by 2035[46].

The Artificial Intelligence Task Force set up by the Ministry of Commerce and Industry, Govt. of India has identified 10 important domains of relevance to India including manufacturing, healthcare, fin-tech, agriculture, education, retail, accessibility, national security and public utility services[47].

In India, AI/ML should enable enhancing access to services and taking services to the marginalized sections of the society especially in sectors like agriculture, healthcare etc.

The researchers we spoke to as part of this study have identified the following application areas of their research.

- Healthcare – hospital productivity systems, diagnosis systems
- Financial services – risk modeling, fraud detection, automated report generation from unstructured text
- Hitech and communications – better search outcomes, information security and biometrics, crowd-sourcing platforms, text analytics, language processing, dialog systems
- Retail – eCommerce applications like improving correctness of an address, recommendation systems

44 https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontierapplications-and-value-of-deep-learning
45 http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf pg 24 - 46
46 https://www.accenture.com/t20171220T030619Z__w__/in-en/_acnmedia/PDF-68/Accenture-ReWire-For-Growth-POV-19-12-Final.pdf%20-%20zoom=50
47 http://dipp.nic.in/sites/default/files/Report_of_Task_Force_on_ArtificialIntelligence_20March2018_2.pdf pg 9-30

- India specific challenges – monsoon prediction, flood modeling, Indian language processing systems, transportation modeling
- Other domains like agriculture, biomedical, climate modeling, education, manufacturing, food, materials industry, mobility, smart city, urbanization and transportation, process industry like petrochemicals, and software engineering

Some of the researchers we spoke to are of the opinion that one of the ways to jumpstart translation is to motivate undergraduate and Masters students working on AI/ML projects to leverage the university incubators to startup.

4.1. Applying traditional analytical and advanced AI techniques

Industry is applying advanced AI / ML techniques like deep learning (DL) neural networks (feed forward networks, convolutional neural network (CNNs), Recurrent neural network (RNNs), Generative adversarial network (GANs)), transfer learning and reinforcement learning alongside more traditional analytics like decision tree learning, clustering and regression analysis.

For instance, the financial services industry has been using linear regression techniques for a long time in areas like risk modeling. But with the advent of more sophisticated computing environments and availability of more consumer data, advanced AI/ML techniques are helping areas like retail credit risk assessment.

Take the familiar example of online search. Query processing on text and web has traditionally leveraged string-matching algorithms to deliver results. More advanced search solutions focus on entities, types, and relations rather than merely string-matching.

Here are a few examples of seemingly simple searches but which require more sophisticated algorithms working at the background – 'Which scientist played violin?' There may be artifacts in the world wide web that mentions Einstein playing the violin but never mentioning that Einstein is a scientist. Other artifacts in the world wide web may detail Einstein as a scientist. The search algorithm needs to connect both these categories of artifacts to conclude that Einstein is a scientist who plays the violin.

In some industrial applications, AI/ML models have worked well where traditional analytical models were not effective. One example is predictive maintenance of machines in a manufacturing-IoT context. Another example is from healthcare - disease diagnosis relying on rich data sets incorporating image and video inputs, including from MRIs (Magnetic Resonance Imaging).

4.2. Applying AI to Healthcare

We will detail out one of these application areas – applying AI to healthcare.

Healthcare systems have proven to be a challenge all over the world, with many developed countries struggling with spiralling costs of healthcare. In India, besides the cost of healthcare, access to healthcare is also a significant challenge – shortage of medical facilities and qualified healthcare professionals. India has just 0.7 doctors and 1.1 beds for every 1,000 of its citizens[48]. Hence there is a great demand in India for leveraging digital technologies to create affordable healthcare solutions which can scale rapidly.

AI can help in creation of electronic health data repositories, clinical decision support systems, self-learning systems in fields like radiology, pathology and genomics[49]. Other top applications of AI in healthcare include robot assisted surgery, virtual nursing assistants, administrative workflow assistance, dosage error reduction etc[50].

Global history of AI/ML in medical imaging

- In early 90s, computer-vision assisted expert systems introduced – for identifying image of the heart and correlating to coronary disease
- Rule-based expert systems followed. But AI was not scalable then
- Medical Visualization became important – to extract meaningful objects from volumes of data
  - Algorithms like the Marching Cubes developed and the field moved away from AI
- From 2004 onwards, combining visualization with computer vision to handle multi-modal data
  - As computational capabilities improved, ML algorithms helped with better assessment of the boundary of images
- In late 2000s, ML became more important. Initially to identify "features" and extracting from images
  - DL has changed that - "features" are now learnt by the neural network

The researchers in the study identified some examples of AI applications in India:

- AI is improving productivity of the medical practitioner and operator – For example, there are solutions developed for an ASHA (Accredited Social Health Activist) worker to communicate with a doctor
- AI is used for disease diagnosis and model how the diseases will progress
  - For example, AI solutions can provide solutions for diagnosing diseases like cancer, like the case of Tata Memorial Hospital & Navya (a machine-based diagnosis) for cancer. Manipal Group has tied up with IBM Watson for oncology
  - Aravind Eye Care works with Google on retinal screening of diabetes-related eye diseases, and LV Prasad Eye Institute and Microsoft collaborate to predict regression rates for eye operations
  - There are over a dozen startups in India like Niramai, Sigtuple, Forus Health and others that leverage AI to detect in early stage various medical conditions like breast cancer, cervical cancer, cardiovascular and eye diseases and analyse medical tests
  - 99DOTS is a low-cost approach for monitoring and improving TB medication adherence
- AI used to increase access of medical care – For example, in India even if there is enough capital to put capable scanners in Tier 2 and Tier 3 towns as well, there are not enough medical

48 https://yourstory.com/2018/01/healthtech-startups-india/
49 http://dipp.nic.in/sites/default/files/Report_of_Task_Force_on_ArtificialIntelligence_20March2018_2.pdf, pg 15
50 https://www.accenture.com/us-en/insight-artificial-intelligence-healthcare

professionals to read the scans. AI plays a role in tele-radiology solutions. AI solutions help in providing healthcare away from hospitals and at homes or work

- AI techniques like Generative adversarial networks (GANs) are used to synthesize data & realistic images for cancer, blockage of artery etc. and thus help increase availability of labelled data

4.3. AI and employment

There have been dooms-day predictions in the popular press about AI taking over the world or displacing humans from work. In this study, we asked researchers whether AI-led automation would lead to a significant loss of human jobs. The researchers were predominantly of the belief that AI will not lead to a large-scale loss of human jobs. Jobs in 2030 have not been conceptualized now. New jobs will get created even as some old ones get eliminated. This would be similar to trends we experienced during the 'computerization' era in the 80s in India.

- We see instances of 'narrow AI' in action – AI applied for a specific task and in most cases, these AI systems are not continuously learning as well. Such 'narrow AI' systems cannot cause the destruction of human race
- While displacement of human work is a possibility, it may significantly impact only one generation of humans
- The pace of change of AI revolution is significantly higher compared to changes caused by earlier technological disruptions. While earlier society experienced change management over a few generations, today it is intra-generational
- A significant effort in re-skilling people will be required to be relevant in the AI-first world
- Jobs that are routine and generating labeled data would likely be eliminated by AI
- Jobs that are considered 'risky' for humans such as extractives or mining, working in radioactive environments etc. could be automated by AI
- In medical diagnosis, the prognosis was that AI will not replace the radiologist, but it will replace the radiologist not using AI
- One view that emerged from our study is that Governments may accept robots as the army of the future. Drones used as unmanned combat aerial vehicles are examples of AI systems that once activated, can find and destroy the target on their own

The sections so far focused on understating the research landscape of AI / ML research in India. The next section describes Indian researchers' perspective on how neuroscience and computational neuroscience will make an impact on AI/ML.

5. CONNECT WITH BRAIN RESEARCH

5.1. A brief history of brain-inspired AI/ML

In 1956, some of the founders of AI including Marvin Minsky, John McCarthy, Claude Shannon, and Nathan Rocheste sent a proposal for the Dartmouth summer research project on AI[51]. They proposed a 2 month, 10 person study to "*find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer*."

AI originally envisaged creating an entity that possessed human like intelligence. While related academic fields such as operations research, statistics, pattern recognition, information theory and control theory already existed, and were often inspired by human and animal intelligence, these fields were arguably focused on "low-level" signals and decisions. "AI" was meant to focus on something different—the "high-level" or "cognitive" capability of humans to "reason" and to "think."[52]

Thus, began the brain-inspired journey of AI. In the initial years, neuroscientists and AI scientists had similar backgrounds and were well connected. But by the 1980s, AI diverged from a neural-type system to a logic driven system, one that focused on big stores of heuristics and rules to make decisions - more databases than neuroscience.

Recently, we have had multiple AI/ML researchers highlight problems that are keeping them from creating human-like artificial general intelligence, such as data hungriness, lack of transparency, and limited capacity for transfer[53]. Some have recommended that AI/ML should look into human cognition and deeply study natural language understanding and commonsense reasoning, some have argued the need for new research into unsupervised learning and understanding how the brain works. There is a view that human-level AI cannot emerge from current statistical mode machine learning systems, and they should be equipped with causal reasoning tools[54]. Some researchers recommend creating a virtuous circle that advances both AI/ML and neuroscience and learning higher level concepts like memory and imagination from the brain[55]. AI/ML and neuroscience are coming closer again.

5.2. AI/ML and neuroscience – Perspectives from Indian researchers

In our study, while all agreed that neuroscience research will have a bearing on AI/ML development in the future, less than 10% of them had any active work currently involving neuroscience / computational neuroscience. This contrasts with the study we did with neuroscience researchers in India, where we found at least 40% of them doing / collaborating on research that combined neuroscience and AI/ML.

51 A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, www.aaai.org
52 Artificial Intelligence — The Revolution Hasn't Happened Yet, Prof. Michael I. Jordan, UC Berkeley
53 Deep Learning: A Critical Appraisal, Gary Marcus
54 Theoretical Impediments to Machine Learning With Seven Sparks from the Causal Revolution, Prof.Judea Pearl, University of California Los Angeles
55 Neuroscience-Inspired Artificial Intelligence, Hassabis, Demis et al., Neuron , Volume 95 Issue 2, 2017, pp. 45 - 258

Another point of contrast in the two studies is the background of researchers. The researchers in AI / ML are predominantly from computer science backgrounds while neuroscience research is multi-disciplinary. A large proportion of researchers in neuroscience, other than the ones focused on clinical neuroscience aspects, have backgrounds in computer science or actively collaborate with such researchers.

A few salient themes that emerged from our discussions with researchers on AI/ML and neuroscience include

- AI/ML can learn from the higher-level cognitive abilities of the brain
- AI/ML can also learn from the workings of the brains of other smaller species
- Need new AI/ML specific computer architectures

AI/ML can learn from the higher-level cognitive abilities of the brain

Deep learning AI/ML systems rely heavily on labeled data. In contrast, the human brain has the ability to learn from very little information. Unlike current AI/ML systems, the brain does not need to be pre-trained with labelled data.

AI/ML algorithms work very well for specific areas. For example, some image processing AI/ML algorithms perform even better than the human brain in recognizing images. But that same image processing AI/ML algorithm cannot perform other functions well, like catching a ball or processing sounds instead of images. At some point in the future, AI/ML systems may mimic multi-tasking like humans.

Especially in areas of learning, AI systems have a lot to learn from how the human brain works. There is research work happening in language-independent brain processing of words – for example, studying how brain processes the words ‘glass’ and ‘ainak’ (a Hindi word for glass).

AI/ML can also learn from the workings of the brains of other smaller species

Just like airplanes do not mimic birds for flight, it is not necessary for AI/ML to work like the human brain. To develop a small drone or a micro air vehicle that goes to a disaster zone and takes pictures, researchers can draw inspiration to develop an AI/ML system to navigate a drone based on a dragonfly’s brain. It’s conceivable that we develop full brain models of smaller animals first.

In robotic mapping and navigation, simultaneous localization and mapping (SLAM) is the computational problem of constructing or updating a map of an unknown environment while simultaneously keeping track of an agent's location within it[56]. New bio-inspired robotic algorithms are studying the model of the hippocampus of a rat.

Need new AI/ML specific computer architectures

Some researchers are learning from the architecture of the brain and developing new AI/ML specific computer architectures for improvements in processing speeds and energy consumed.

56 https://en.wikipedia.org/wiki/Simultaneous_localization_and_mapping

The new computer architecture may be characterized as fine-grained parallelism, of being data driven, and event driven in their manner of operating the neurons and where at any given time only 1% or so of these neurons are working.

This kind of model is not suited to existing hardware. Researchers will have to build a specific hardware or processor to suit this kind of computational pattern. It is not just newer hardware, but newer algorithms are also being developed using event-based processing. In the case of speech synthesis and recognition, these newer algorithms use 10X less data for training.

At least one of the researchers in our study is researching on augmenting neural networks with memory. Some cutting-edge research in this space include work on Capsule net (CapsNet), a structure called capsules are added to a CNN and output from several of these capsules are reused to form more stable representation; and LSTM (Long Short-term memory) architectures, an RNN architecture that remembers values over arbitrary intervals[57].

57 https://en.wikipedia.org/wiki/Long_short-term_memory

6. RECOMMENDATIONS

In the previous sections, we have focused on the current landscape of AI/ML research in India. In this section, we will present a few salient recommendations for strengthening the AI/ML research ecosystem in India and addressing the key challenges identified in earlier sections.

6.1. Increasing the number of PhD students in India

AI / ML researchers are unanimous in their call for increasing the number of research students available in India and also for making it attractive for PhDs to stay or return to India. Researchers we spoke to as part of this study mentioned that each AI/ML lab requires about 5 to 6 PhD students to provide a critical mass for research projects. Top universities in the world in the AI/ML domain have tens of PhD students in an AI/ML lab. This provides them with the foundation to do impactful research. Research students follow research faculty. India needs to create a special fund to attract faculty who want to come (back) to India from the USA or other parts of the world. The fund will provide fellowships to such faculty for a period of at least 3 years.

India needs to create fellowships for PhDs and post-doc students. The Prime Minister's Research Fellowship (PMRF) allows students to be awarded INR 70,000 on a monthly basis along with an annual research grant of INR 200,000. Students from all Indian universities will now be eligible for PMRF for admission to full-time PhD[58]. Some researchers also feel that special fellowships in "hot areas" like AI/ML is required to attract and retain the best students in research. It is also important to encourage the post-doc model in India so that not only do Indian candidates but also those from outside India spend quality time in Indian AI/ML labs. An important aspect that was emphasized was the need for continuity of researchers – principal investigators, post-docs, doctoral students, and research associates in the lab for high impact research programs.

The researchers also suggested programs for introducing undergraduate students to AI/ML. This can motivate those interested to pursue a Masters or PhD in AI/ML. Some of the programs suggested include

- Summer schools on AI/ML for research-oriented students
- Competitions / hackathons on AI/ML - provide India-specific problems that include need for data acquisition as well
- Undergraduate 4th year project - students should be encouraged to participate in a national-level AI/ML solution development challenge. India should create a platform that offers problems and data and the winners should be promoted on social media so that there is a pull created to participate in such projects
- Government-academia-industry may offer an internship program (online classroom and training on a real-world India-specific AI problem) for students who pass out from recognized universities

The plans of China in this regard are a reference point for policy makers. China has launched a five-year university program to train at least 500 teachers and 5,000 students working on AI/ML technologies. It plans to develop 50 world-class teaching institutes and research institutions,

58 https://www.hindustantimes.com/education/students-from-all-universities-will-be-allowed-to-apply-for-pm-fellowship/story-tp6B9wT4OXrw2JUI44rwAM.html

50 national-level high quality online open courses and 50 AI/ML facilities by 2020 as part of the "AI+X" program[59]

6.2. Augment computing infrastructure for AI/ML research

The top universities in the world in AI/ML have computing infrastructure that has thousands of GPUs, while the next tier universities in the world have an infrastructure that has hundreds of GPUs. Compared to this Indian universities typically have computing infrastructure that has tens of GPUs. This is expected to change with new investments that the top Indian universities are making in upgrading their computing infrastructure for AI/ML research.

Setting up world-class computing infrastructure for AI/ML research is expensive. The NITI Aayog report mentions a proposal to have a national computing infrastructure, AIRAWAT, a 100-petaflop supercomputing system for AI/ML applications. Extrapolating based on the costs of commercially available similar smaller systems, the cost of this computing infrastructure is likely to be in the order of tens of millions of USD[60].

A researcher provided some back-of-the-envelope calculations for setting up specialized computing infrastructure in a typical lab in the CS department of a university.

- It would cost approximately INR 250 thousand per GPU or INR 1.8 million for an 8-GPU cluster with a system made up of cheapest GPUs
- A lab typically needs 50-60 GPUs. Thus, the overall cost for a lab is approximately between INR 10.8 million to INR 14.4 million

In parallel to procuring and setting up such national computing infrastructure, the Government should augment capability to design and make such computing systems in India. For example, Shakti is an open-source hardware design & processor development initiative by the RISE group at IIT Madras (and with funding from MeitY) and aimed at developing industrial-grade processors[61]. It has released India's first indigenously-built microprocessor.

- The processors will help in aiding research related to architecture, where one can run simulations on the actual hardware and obtain much accurate results, rather than settling with a lesser accurate software simulation[62]
- They can be adopted by the industry and different categories of System-on-chip (SoCs) are envisaged – for micro-edge devices (like IoT sensors), medium-edge devices (used in industrial control, vehicular electronics), large-edge devices (like mobile / desktops, routers) with AI/ML and neuromorphic accelerators and decision making and monitoring (cloud, supercomputing) with optional AI/ML accelerators
- A budget of INR 250 million is required for the prototype development of reference SoCs. The Indian industry should take responsibility for translating from prototype to real-world implementations

59 http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf, pg 103
60 http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf, pg 58-59
61 http://shakti.org.in/e-class.html
62 https://dl.acm.org/citation.cfm?id=2925106

6.3. Create India-specific AI challenges, tools and data-sets

Another aspect that came up in our study is that Indian public funded AI/ML research will have an obligation to focus on India specific problems that affects large numbers of citizens. There is a high possibility that the solutions to these India specific problems may have a use in other parts of the word. For example, our interviews revealed some interesting India specific use cases that Indian AI/ML researchers are working on.

- A computer vision algorithm in autonomous vehicle context working at about 80% efficiency in Western conditions, may work only at about 40% efficiency on Indian roads. Hence the need for India road conditions data to be captured and used for training the autonomous vehicles.
- Image recognition solutions that identify the condition (potholes, water stagnation, etc.) of roads.
- Character / text recognition with a focus on Indian languages.
- Speech recognition when it is multilingual in the same sentence. Like in the case when Hindi and English are used in the same sentence.
- In the Indian agricultural context of small land-holdings of 0.5 to 1 acre per farmer, applying sensors to fields becomes prohibitive. AI solutions combining weather data and earth-sensing satellite data (considered IoT in the sky) are producing fine-grained insights for the farmer including crop yield forecast, and pest and disease prediction

One such initiative is NITI Aayog's project on Indian languages. The initiative aims to create a repository of basic tools and resources for Indian languages. Any industry or startup wanting to develop an Indian language application can access this repository.

The researchers identified some interesting ideas for creating data-sets for research in India.

- Every researcher should take a pledge to make available the data sets they have created and allow use by other Indian researchers for free. Some leading international journals / conferences are insisting on this
- User Generated Content is a big opportunity in India - even if a fraction of the population can help in generating and labeling the content, it will greatly benefit
- In the context of Indian language speech recognition research, the scripts used for movies or television serials in Bollywood, Indian television production houses and OTT (Over the top) media companies could provide the voice and script
- Synthetic data (which is computer generated data that mimics real data) provides a powerful approach to dealing with the challenges of creating large-scale training data for AI systems
- Bring in legislation to manage data availability for AI researchers
  - Make AI solution providers to open up data, especially the training data which was used to train the AI model in order to understand the inherent biases in the system
  - The European Commission published a communication on "Artificial Intelligence for Europe" which identifies initiatives to make data available to researchers, especially data resulting from

publicly funded research[63]

- Japan introduced the idea of an AI specific Regulatory Sandbox which is to develop an environment in which businesses are able to conduct demonstration tests and pilot projects for new technologies like AI[64]
- NITI Aayog has suggested the idea of a National AI Marketplace that includes a data marketplace[65]

6.4. Set up Centers of Excellence (CoEs) for AI/ML research

The researchers recommended setting up Centers of Excellence (CoEs) for AI research in educational institutions.

- The CoEs facilitate collaboration among stakeholders in industry, research, education, government & policy on a specific or a set of AI/ML themes
- They ensure an interdisciplinary approach (e.g. financial services or healthcare or manufacturing in combination with AI/ML)
- The CoE will help provide faculty travel grants and required computing infrastructure. It may also offer incentives for publishing papers in top journals
- The CoEs are practical environments or field validation labs in which institutions develop, test and implement effective AI/ML solutions

One such CoE is the Robert Bosch Centre for Data Science and Artificial Intelligence (RBC-DSAI), at IIT Madras[66]. It is a multi-disciplinary AI research center involving 24 faculty from 8 different departments. It aims to carry out

- Fundamental research in the areas of deep learning, network analytics, learning with limited and partial data, causal modelling, data science specific system architecture and reinforcement learning.
- Applied research in four verticals - manufacturing analytics, financial analytics, smart cities, systems biology & health care

Another important aspect that is emerging is to include members with humanities background into larger centers focusing on AI/ML. This is important since the ethics of AI/ML models is becoming important. There are cases where the AI/ML system is biased since the data used to train the system was biased.

6.5. Adopt an AI Grand Challenges approach

In addition to setting up Centers of Excellence (CoEs) for AI research, it is also important to seed each CoE with 1 or 2 Grand Challenges so that the efforts of the different researchers from multiple disciplines are channelized towards a common purpose.

63 https://fpf.org/2018/07/19/policy-brief-european-commissions-strategy-for-ai-explained/
64 http://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/miraitousi2018_en.pdf
65 http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf, pg 77-84
66 http://rbc-dsai.iitm.ac.in/about

An excellent example of AI grand-challenge approach led success is that of the autonomous vehicle challenge sponsored by DARPA and which ran in 2004, 2005 and 2007[67]. In the Urban Challenge 2007, the Tartan Racing team, comprising members from Carnegie Mellon University, won the first prize. The second prize went to the Stanford Racing team from Stanford University. Post this grand challenge, roboticists from Stanford and CMU helped form the core team of Google's self-driving car project (now Waymo), which is one the most successful autonomous vehicle project today. What was a grand challenge a decade back is on the cusp of disrupting urban transport today.

The four grand challenges for AI in India identified by The AI Task Force are 1) improving manufacturing, especially in the SME (small and medium-sized enterprises) sector, 2) improving healthcare quality, reach and cost, 3) improving agriculture yields and profitability, and 4) improving delivery of public services[68]

6.6. Link institutional mechanisms to start-up ecosystem

Given the importance of AI/ML to India's progress, it is imperative to ensure that there are adequate institutional support and freedom for translational research. The Indian Government's Department of Science and Technology DST has recently launched a new program - Interdisciplinary Cyber Physical Systems (ICPS) - to foster and promote research and development in cyber physical systems that are mechanism controlled or monitored by computer-based algorithms, tightly integrated with internet and its users. AI/ML is an integral part of cyber physical systems.

ICPS has proposed a four-layer institutional structure to support research and its translation[69]

- ICON (International Centres of New Knowledge): focusing on creation of new knowledge through basic research
- CROSS (Centre for Research on Sub-Systems): focusing on developing and integrating core technologies developed at ICON layer and any other sources
- CASTLE (Centre for Advanced Studies, Translational research and Leadership): focusing on development and deployment of application-based research and
- CETIT (Centre of Excellence in Technology Innovation and Transfer): focusing on commercialization of technologies developed

NITI Aayog in its report on AI has simplified this four-layer structure and proposed a two-tier structure[70].

- The functions of ICON and CROSS are combined into COREs (Centres of Research Excellence in Artificial Intelligence): COREs will specialize in creating new knowledge through basic research and will source for fundamental knowledge / technologies that will be needed to keep India prepared for

67 http://www.foundingfuel.com/article/a-mission-to-solve-grand-challenges/
68 http://dipp.nic.in/sites/default/files/Report_of_Task_Force_on_ArtificialIntelligence_20March2018_2.pdf, pg 10
69 http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf, pg 54
70 http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf, pg54-57

the next generation of technologies.

- The functions of CASTLE and CETIT are combined into ICTAI (International Centre for Transformational Artificial Intelligence): This will provide the ecosystem for application-based technology development and deployment. ICTAIs will also be responsible for delivering commercial technology and taking ideas / concepts or prototypes and turning them into marketable products by way of proactive coordination, communication and interfacing for technology transfer to the industry. For example, NITI Aayog, Intel and TIFR are collaborating to set up an ICTAI for application-based AI research in healthcare, agriculture and smart mobility[71].

We have proposed the multi-disciplinary CoEs and Grand Challenges model as institutional mechanisms.

Based on our study, we believe that while institutional support for translating basic research in AI/ML to applications is important, it may not be sufficient. Our interviews with researchers reveal that about 40% of them are working with startups or have startups getting incubated based on the work done in their labs. The links between AI/ML labs and startup ecosystem needs to be encouraged and strengthened. This may be the quickest and most natural means of translating research. This may also motivate the undergraduate, Masters, and PhD students to stay back in India for longer once they complete the requirements for their educational programs. Many of the principal researchers we spoke to as part of this study felt that both continuity and pace of progress is lost when students move out of the labs and projects.

This is where academic incubators can play an important role of helping labs translate their research into applications. Many of the IITs, IISc, and IIITs have strong academic incubator programs. The Atal Innovation Mission (AIM) under NIT Aayog has an ongoing scheme to support academic incubators[72]. There are green shoots visible already. Take the example of the IIT Madras Incubation Cell – about 7% of all companies incubated here have AI/ML as a core technology[73]. The mortality rate of start-ups in this incubation cell is around 15 per cent which is 2 to 3 times lower than other incubators. About 35% of the total portfolio of 170 deep-tech startups established at this incubator have IITM faculty members as founders or minority shareholders. The technology business incubator at IIT Bombay, Society for Innovation and Entrepreneurship (SINE) has partnered with DST and Intel to set up Plugin, a collaborative incubator for hardware & systems startups especially with a focus on AI and IoT[74].

Apart from academic incubators, we also believe that there is an opportunity for Indian angel and venture capital to invest in startups coming from Indian universities. This not only helps scale up solutions that can benefit India and the world, but also build high quality Indian intellectual property.

71 http://pib.nic.in/newsite/PrintRelease.aspx?relid=183377
72 http://niti.gov.in/content/atal-incubation-centres-aics
73 http://www.incubation.iitm.ac.in/home
74 https://plugin.org.in/

Appendix: List of researchers interviewed for this study

We thank the researchers who graciously gave us time to understand different aspects of AI/ML research in India.

1. Anand Rangarajan, Ashwani Sharma, *Google India*
2. Balaraman Ravindran, *IIT Madras*
3. Chiranjib Bhattacharyya, *IISc*
4. C. V. Jawahar, *IIIT Hyderabad*
5. Kajoli Krishnan, Independent Researcher – *Medical Devices*
6. Madhava Krishna, *IIIT Hyderabad*
7. Manish Gupta, *Videoken*
8. Mausam, *IIT Delhi*
9. Mayank Vatsa, *IIIT Delhi*
10. Parag Singla, *IIT Delhi*
11. Partha Pratim Talukdar, *IISc*
12. Piyush Rai, *IIT Kanpur*
13. Prateek Jain, *Microsoft Research India*
14. Preethi Jyothi, *IIT Bombay*
15. Pushpak Bhattacharya, *IIT Patna & IIT Bombay*
16. Rakesh Mullick, *GE John F. Welch Technology Center*
17. Rajeev Rastogi, *Amazon*
18. Shourya Roy, *American Express*
19. Shudeshna Sarkar, *IIT Kharagpur*
20. Soumen Chakrabarti, *IIT Bombay*
21. Sriram Raghavan, *IBM Research*
22. Subrat Panda, *Capillary Technologies*
23. V. Kamakoti, *IIT Madras*
24. V. Vinay, *Ati Motors*
25. Y. Narahari, *IISc*

Appendix 2: Illustrative list of researcher-researcher collaboration

The following is an illustrative list of research collaborations between researchers from Indian and other Indian / foreign universities as mentioned by participants in our study (in alphabetical order).

1. IIIT Delhi and University of Notre Dame (USA)
2. IIIT Hyderabad and Universitat Autonoma de Barcelona
3. IIIT Hyderabad and University of Maryland (USA)
4. IISc and Carnegie Mellon University (USA)
5. IISc and Chalmers University of Technology (Sweden)
6. IISc and INRIA (France)
7. IIT Bombay and IIT Patna and Kyoto University
8. IIT Bombay and University of Illinois Urbana-Champaign (USA)
9. IIT Bombay and University of Rome (Italy)
10. IIT Delhi and University of Texas at Dallas (USA)
11. IIT Delhi and University of Washington (USA)
12. IIT Kanpur and Duke University
13. IIT Kharagpur and Jadavpur University
14. IIT Madras and University of Birmingham (UK)

About itihaasa Research and Digital

itihaasa Research and Digital (www.itihaasa.com) is a non-profit Section 8 company that aims to understand and chronicle the history and evolution of technologies and businesses in India. Kris Gopalakrishnan, co-founder Infosys, is the founder and Chairman of itihaasa Research and Digital.

Our flagship project is itihaasa history of Indian IT, a first-of-its-kind free digital museum app that recounts the history of Indian IT since the 1950s. This app makes the incredible history of Indian IT accessible to an audience across the world and is available on App Store (iOS) and Play Store (Android). It is also available through a chatbot on our website. itihaasa history of Indian IT is featured in the IEEE Annals of the History of Computing.

You can reach out to **N. Dayasindhu, PhD** (dayasindhu@itihaasa.com) and **Krishnan Narayanan** (krishnan@itihihaasa.com) for any queries / feedback on the study.

itihaasa
Research and Digital