इलेक्ट्रॉनिकी और सूचना प्रौद्योगिकी मंत्रालय
भारत सरकार

Ministry of Electronics & Information Technology
Government of India

REPORT OF COMMITTEE - B ON LEVERAGING A.I FOR IDENTIFYING NATIONAL MISSIONS IN KEY SECTORS

July 2019

Contents

1. Introduction

Artificial Intelligence (or AI, in short) is a technology which has suddenly come of age. A number of constitutive technologies and needed platforms are operational with some matun rity, and many enabling factors are in place. For example, the Internet, cloud infrastructure, cyber physical systems, Internet of things are in place or getting there at a fast pace. On the other side, availability of mobile phones as personal devices, all connected through a telecom network, provides a great source of data as well as the delivery mechanism of services to end users.

Data of all kinds is getting generated on a massive scale. However, sense needs to be made out of this data and privacy and security issues need to be handled with care. Unfortunately, even when there are no privacy or security issues or can be handled relatively easily, the data is not available to researchers and applications. It is in possession of companies, or locked in government databases. This is a major loss of opportunities for societal benefit.

2. Data, Knowledge, Algorithms and Systems

We present the architecture for AI systems in Figure 1. For ease of understanding, we split it into three major layers:

1. Data and knowledge (like life blood)

2. AI/machine learning algorithms (like organs)

3. AI applications and systems (like body systems) Each of the three layers is discussed next:

1. Data is necessary for AI applications to work. It is the life's blood that allows the systems to be developed and for them to work. However, the data has to be in wells defined formats, so that the relevant information can be appropriately retrieved and used.

At times, the data needs to be combined from diverse sources. This means that the inter-operability aspects have to be taken care. Thus, all such diverse pieces of data need to be put together much like the blood which contains many different elements which can be identified and made use by the body.

There are many instances when the data needs to be annotated with human help, so that machines can learn better. Such annotations are frequently not just data tags but knowledge, where the annotations provide knowledge. Such annotations or supervisory signals can be (i) collected from humans explicitly when it is required, (ii) gathered from the past historical records or (iii) harvested from the user interactions.

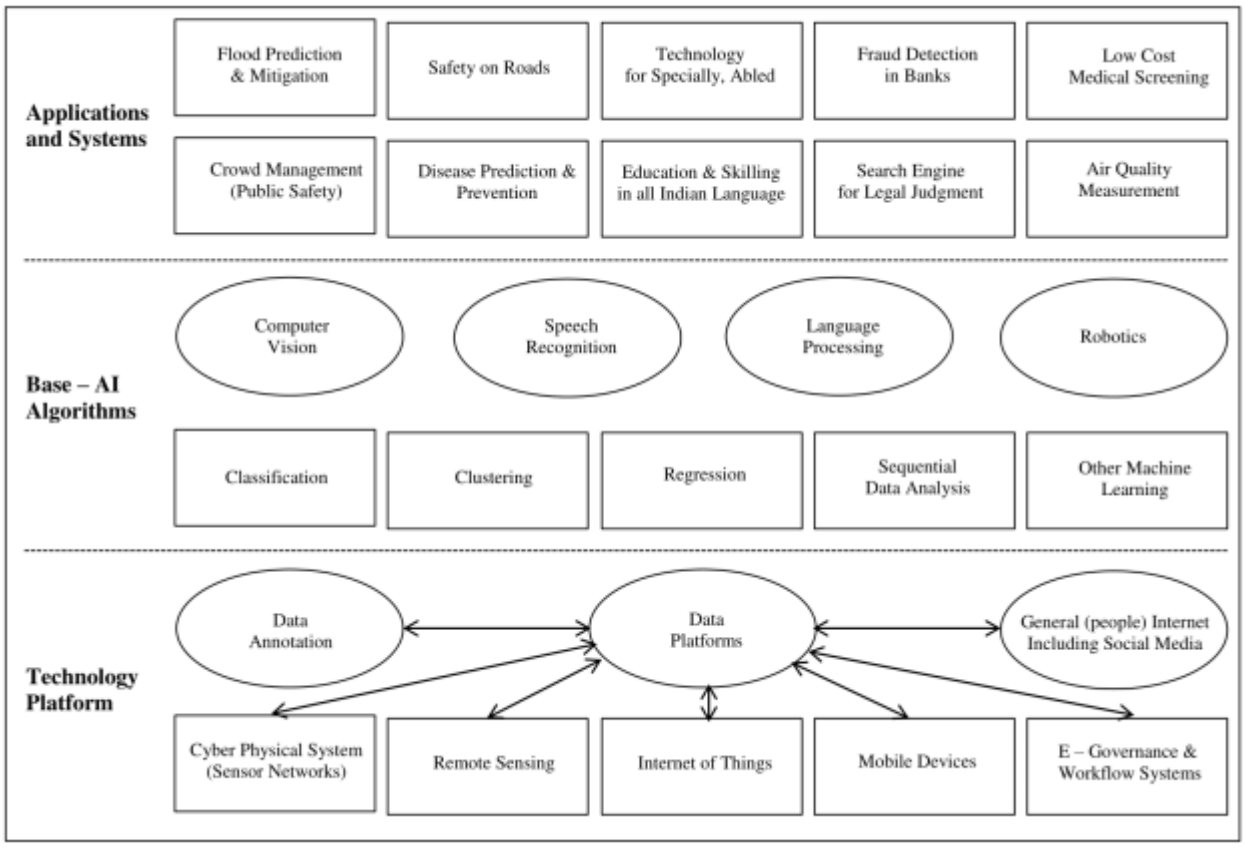

Figure 1: Three layered architecture of AI Systems

2. Algorithms are procedures or methods that operate to produce some useful results. In case of AI, they typically operate on the data or knowledge to do some tasks. Frec quently, these tasks relate to clustering, classification, pattern recognition, etc. These tasks are like the organs of the body through which blood (data and knowledge) passes, and allows useful tasks to be completed. For example, blood passes through lungs for oxygenation.

These algorithms discover structure from the data and very often extend the strucT ture and knowledge from the past to newer situations of the future leading to useful annotations, tags, alerts etc.

3. An AI application or AI system is one which combines many AI/machine learning algorithms with the right data and knowledge from diverse sources to accomplish useful work for end users. For example, pulmonary system in which lungs, chest muscles, blood, etc. all play a role.

Many real world applications require processing and understanding of data that is diverse and even multimodal. Each domain/application employ the power of the algorithms in different ways. Since the application systems could interact with humans, additional dimensions of usability and social acceptance come into the picture.

This is shown by a diagram (Figure 1) as a three layered structure, for the purpose of understanding. The bottom most layer relate to the data and knowledge available in the society, directly captured and often annotated and curated. The middle layer consists of a number of algorithms and modules that are popular in the AI space. The top layer consists of a number of applications which can be enabled with the help of the algorithms, given the appropriate data. Clearly, within a layer there might be further sub layers or structure, as shown for layer 1 and 2. While the three layered structure would help in understanding, planning and development, there might be cross connections for an application, at times or skipping a layer.

3. Indian Setting and Sectors

AI can play an impactful role in India if it is applied to:

• problems that are important, and

• which are amenable to AI technology

There is a lot of interest in AI in the industry. However, projections by the press either create a scary scenario where the robots will rule the human beings, or show trivial scenarios where AI will predict that you would buy yellow shampoo if you use blue soap. The former is not likely to happen, whereas the latter will not change the world even if it were to happen. While selecting problems to be tackled we should avoid both kinds of mistakes in assessment.

We have taken the approach of going sector by sector, identifying important problems as considered by people in the sector or general experts. Out of these problems, we have selected those that are likely to be amenable to AI technology in this report. Attention has been paid to problems which can be solved and deployed in the short term, medium term and long term. The time frames are roughly, 2 years, 7 years and 12 years, respectively. This means that some work would yield immediate benefits, but work would also be undertaken on deep technical tasks with large potential benefits in the medium and long term.

Some of the selected problems are such that the solution would be server based, and available to end users through mobile apps or dedicated applications. Whereas for others, it might require development of special purpose devices with embedded intelligence. For example, flood prediction or public safety information system would be server based with information dissemination through mobile apps, whereas for agriculture or health, there might be development of special devices that can sense and solve specific problems. In some cases, special coordinated efforts will be required to create the necessary data or annotation. In some other cases, users directly contribute the data. Creating special satellite imaging systems for accurately sensing the land use patterns may be a central focused effort, while capture of infected crop images on a mobile phone may be enabled with simple mobile applications.

Special attention is paid to develop an ecosystem involving the stakeholders. For example, it could mean promoting incubated companies as well as inviting established technology companies to play a role. Research and innovation ecosystem needs to be given a direction (rather than leaving it completely free for blue sky research). It is proposed to float a set of deep technological tasks whose solutions will connect with a large number of applications in specific sectors. These tasks would be worked upon by holding national contests which would catch the imagination of hundreds of thousands of Indian students. These would inspire Indian students to work on Indian problems and also motivate them to do self learning beyond the classroom.

The above measures would allow us to develop technology by inspiration and guided directions through students and research institutions. They can be delivered to end users through incubated as well as existing companies.

4. National Missions in Key Sectors

4.1 Agriculture

Prediction of Crop, Crop State and Yield Streamlining production and integration of demand and supply is imperative for Indian agricultural economy. Advancements in this area is possible with deeper understanding of high resolution satellite images. AI could augment remote sensing and overhead imagery data. Advances in fine grain image understanding capabilities and closely monitoring the crop type, state, and the geographical distribution. Also, the market for such yeild or the scarcity thereof for both short-term and long-term can be better analysed, predicted and planned for. Superimposition of climatic fluctuations on agro-ecology can be predicted and dynamic contingency operations initiated. Such advice can only be generated by developing dynamic systems to integrate macro, miso and micro level agro-meteorological data. Innovative crop planning strategies can neither be tracked nor disseminated without using AI applications, big data and advances in machine/deep learning.

Scalable and distributed health care for crops and animals Like in the medical space, agriculture experts are rare and also not instantly available in the remote areas in a reliable manner. A number of crop diseases are detectable through human naked eye, leading to a possible diagnostic solutions that can work remotely on images (say captured with simple mobile phone). Scaling such a system to the large potential cases requires image understanding capabilities. Here the objective could be to make the crop disease detection, diagnosis and treatments simpler, effective and low cost in rural areas. The point of care devices for identifying pest and diseases for plants, animals and humans require deep learning tools. Point of care diagnostics and treatment advisory for animals has only started in the recent time. Since livestock distribution is much less in equities than land distribution, applications of AI tools will have much higher social welfare impact.

Soil Mapping Characterizing the soil and environmental factors is important to recommend the right cultivation practices and also monitor the systematic and unwanted changes in the soil properties. From the agriculture point of view, one is interested in the presence and percentages of various chemicals in the soil. Low cost and scalable sensing has to be a thrust problem at this stage.

Automatic Weed Identification and Extraction Separation of weeds from the crop is still attempted manually. Advanced methods that can sense weeds from the signals can help in many agriculture sectors.

The majority of the farmer machinery whether manual, bullock drawn or motorized do not have feedback control system. Most mechanical machineries have practically no electronic applications. Dynamic feedback from individual farm does not get aggregated at eco-system level and therefore knowledge based efficiency inputs are not available to farmers. Given the agrological diversity response variabilities of inputs such as water, seed and other chemical and non chemical inputs is not analyzed dynamically to generate mid-course correction.

Live Stock Information System Live stock information system (i) in the dairies and

controlled environments, and (ii) in the wild.

There is enormous amount of data available on soil, plant, animal and human nutrition. However, we have not developed any analytical platform which can connect the variables for tracking eco-system health and ensuring preventive crop, animal and human nutrition.

Animal recognition schemes with advances in AI, one could aim at identification of individuals, species, and their health/status (eg. state of pregnancy) from images and other sensors. This can have impact in many sectors where the surveillance of a large farm/forest with cameras or other sensors is the only feasible solution to monitor the state.

Agriculture Microplanning Advice Farmers today rely on age old crop patterns or traditions, heuristics, hearsay or faulty advice to plan their crops. It will increasingly possible with IOT, weather and visual data, along with crop patterns, pricing data correlated with the vast body of agricultural research to offer personalized advice to farmers on what they should or should not produce, what methods should they deploy for production, help them in adverse circumstances such as pest attacks, where they would likely get the best prices for their output, etc. This could be manifest through natural language multi-modal personal assistant interfaces, with deep linkages to back end data bases and AI and data science models.

Farm Automation- creating a shared economy A very small percentage of farms currently deploy automation in their production, and most of the effort today is manual, which impacts productivity. The main reason is the capital intensive nature of automation equipment. By using a confluence of mobile, digital, voice based access, AI and data science, a shared economy of farm automation equipment for seeding, ploughing, harvesting, transportation and storage, can be created, which will be consumption based and time parcel based.

This will help farmers access modern automation equipment on lower rentals, and help them increase productivity of their farms. Higher farm productivity can lead to lower cost of production and transportation, as well as help farmers access far-away higher demand markets and can help in boosting farm incomes. In addition, shared economy around farm equipment can create a whole new range of jobs, similar to employment being created by mobile based shared taxi services in urban areas, boosting the rural economy.

Optimising Pesticide Consumption: It is very difficult for farmers to identify which pests are attacking their crop, and which corresponding pesticides should they use. As a result of this, pesticide consumption has sky-rocketed, potentially poisoning soil and increasing farmer's input costs, while a huge quantum of crops still gets damaged on account of crops. AI algorithms to classify image captures of different pests and then recommending the correct pesticides, could be a solution for farmers, so that they could access all the published research on the topic with just the power of a phone image.

Farm economy modernization Farmers are not able to access formal financial institutions on account that banks, microfinance institutions just don't have the wherewithal to reach out and build successful credit models or detect fraud. The high level of manual intervention in the interactions as well lending decision making leads to faulty lending, high level of NPAs, and inadequate delivery of financial instruments to farmers. Using AI and data science, credit models could be developed that could lead to smarter lending and record ery decisions. Modernisation of the farm economy could help reduce hardships to farmers, reduce the extent of loan waivers, etc.

4.2 Food

Inspection of Food Quality of food directly correlate to the health of the country. In the past, there have been dedicated labs and testing systems in place. It is now possible to develop scalable and affordable techniques, kits, solutions for variety of food items with advances in sensors, and sensor understanding. AI can contribute to the development of techniques for inspection of food and allied materials, including fruits, vegetables, grains, milk etc. Food inspection for presence of pesticides, method of ripening, residuals of chemical agents etc, can directly help in controlling the penetration of unhealthy practices in fruits and vegetables.

Minimizing Post Harvest Losses Indian agriculture reports a significant loss after harvesting. Techniques that help in minimizing post harvest losses are needed in many sectors. Advances in storage, logistics and financing infrastructure can reduce the post harvest losses. Distributed and coordinated storage, better technologies for packing and transportation, etc. could include the directions.

Sensors for baskets In a large storage center, often only one isolated pack/bag gets deteriorated in quality and there is no systematic process to keep track the health/state of the individual bags/packs. Identification of packs/baskets from a large collection with food items that have started to degrade/rot with intelligent sensing is required.

4.3 Health

Capture and Analytics of Health Data from Hospitals Leading public hospitals such as AIIMS and PGI get as many as 10,000 patients a day. A large number of patients also visit the public hospital network in the country. Intelligent tools are required to analyze the interaction between treatment, disease variability, gender, age and other variables. While intuitive understanding just grows among experts. In the absence of analytical tools, patterns cannot be identified easily.

Low cost Medical Screening Solutions In many small places, advance diagnostic facilities such as MRI and basic diagnostic tool such as X-Ray are not available. India has reported less than 10,000 radiologists. Low cost medical screening solutions are required to reach better health care to rural areas. This includes low cost imaging devices, portal powerful screening equipment's, Internet screening support with AI algorithms running in the back end etc.

Prediction of disease outbreaks : Early warning of outbreaks of vector-borne and infectious diseases like dengue, chikungunya, flu, and meningitis can help contain the scale of outbreaks, enable better preparedness by the healthcare system, and reduce case severity. Machine-learning systems on signals helpline call analysis, symptomatic case reports, weather and environmental conditions, and other correlates hold promise.

Clinical decision support systems: ML/AI-based decision support systems can aid doctors and healthcare workers in resource constrained settings, improving quality of screening, pathology, diagnosis, therapeutic pathways, prevention of drug interactions and contradictions, and triaging of referrals.

Microplanning for village healthcare workers: Healthcare workers regularly have to decide how to prioritize their time across following up on high risk households, health surveillance across regions, prevention drives like screening and vaccination, and administrative tasks like reporting, and training. Data-driven probabilistic decision systems promise improved effectiveness of their efforts.

Targeting awareness and education campaigns: Behavior-change campaigns are a significant component of public health efforts. Data-driven methods that allow targeting and sustaining campaigns based on need and effectiveness can reduce cost and improve effectiveness.

AI for supply chains: Use of tracking and prediction can reduce stock-outs in dispensaries due to seasonal and local variation of demand, wastage due to expiration and pilferage, and spoilage during supply.

4.4 Water Resources

Prediction of floods Prediction of floods and water levels in river basins is an important problem for a large part of the country, specially on rivers like Ganga and Kaveri. Make accurate predictions based on:

- the remote sensing and meteorological data.
- river basin models, topology, water usage patterns etc.

Ground water aquifer monitoring Modeling and prediction of the ground water level is very important. Activities life construction, large scale use of ground water lead to unexi pected changes in the ground water levels at other parts. Modeling of this and predicting with systematically capturing the data and modeling them is important for the sustainability of the planet.

4.5 Environment and Pollution

Portable Water Quality Testing Accurate and low cost water quality testing needs to be developed and distributed throughout the country. Regular monitoring of the water quality in public spaces, tanks needs to be displayed. Availability of the data over time can create alerts on sudden or systematic changes in the quality with the help of appropriate algorithms.

Air Quality Measurement Regular monitoring of air quality is becoming critically important, specially in cities. Sensors for air quality measurement deployed in various parts of the city could produce accurate data of the state of the quality of air. Logging of such data over time can also lead to understanding of trends, prediction of the air quality, warning system for citizens etc.

Forest Fire monitoring and mitigation Prediction and detection of forest fire with remote sensing, special sensors and associated technologies has become a necessity. Sensors and images could create immediate alarms in the event of forest fire. Creating warning system for forest fire with inputs from various sensors can save life and resources.

Distribution and movement of animal species Animal species, their density, movement patterns, distribution at any point needs to be monitored. Animal population dynamics using satellite telemetry may be a possible direction. This information also needs to connect

to the water and other data which is locally available. Monitoring animal/bird movements also helps in identifying the endangered species.

4.6 Education

Intelligent Tutoring Systems With web based technologies and MOOCs, there has been widespread use of electronic content in the higher educational space. However, there needs significant additional efforts to make them accessible and effective in Indian setting. This include ((i) Student centric personalization and pacing, and (ii) content accessible in Indian languages.

Educational Content in Indian Languages Availability of the educational content and supplementary information on Internet such as TED lectures and MOOCs has been a boon to many of the students to overcome the limited teaching resources or infrastructures that schools are able to provide. However, most of these contents are in English. This demands the development of technologies that can make the content available in Indian languages.

Scalable and Reliable Evaluation Schemes India needs scalable, distributed and reifiable evaluation models. Distribution of the importance of single high pressure exam to multiple smaller but reliable evaluation can make big impact in the learning process. This require technologies such as (i) automatic question generation (ii) student centric evaluation schemes (iii) automatic and assistive evaluation schemes etc.

Language Learning and Using Language is fundamental to the effective communication. Students need to develop the capabilities to read, comprehend and speak Indian languages as well as English. This demands development of schemes that measures the quality (eg. reading/comprehending, speaking/pronunciation quality). AI can help in designing tools and environments that can make this happen in schools and houses.

Multilingual Learning With availability of language access tools, creative solutions are possible so that a person can read any text in any Indian language or English, in their own language. This could take us beyond the 3-language formula.

4.7 Culture

Virtual and Digital History Capturing and presentation of the history and culture of the country can be enabled with a set of AI and related technologies. Richer digital displays, museums, online streaming of the content, etc. can help in accumulating and presenting the rich knowledge, culture and history.

Richer Experience to the Culture Richer experience of culture can be enabled through technologies. Experiencing dance, music, monuments etc. can be improved well beyond the existing light and sound shows, tour guides etc. Technology can also help in the training/education of classical art forms.

Archival, Restoration, Augmentation Archival of art forms ranging from manuscripts, painting to music and dance can be enabled through technology. The archived data can be restored, augmented (say with text in an alternate language) automatically to make the user experience more detailed and effective.

4.8 Specially Abled

Technology for Perception and Communication India has 2.7 crores of people with some form of disability. This include people who are visually challenged, cerebral palsy, hearing impaired, Dyslexia, etc. We need to develop technology for them to see, hear, speak and communicate. AI enabled solutions could have immediate impact in their life with (i) technology support for access to information and communication (ii) technology support for mobility (iii) Intelligent interface technologies that makes the social interactions of specially abled easier and effective (iv) technology for the education and training for the specially abled. This requires various vernacular speech and language interfaces for the people. This will bring the challenged to the mainstream.

Security of the Disabled Solutions are required to improve the security of the disabled when they are alone at home, at work place and while traveling. This includes detection of potential threats (from stray dogs to humans, from electrical hazards etc.) and preventing the attack by taking helps, creating alerts etc.

4.9 Transportation

Smart Public Transport Large percentage of the population depends on public transport system for day to day activities. However, their timely availability, promptness and the schedule create practical discomforts to the people. This can be addressed to a large extent with coordination, spacing and synchronization of bus, multimodal and other public transport services. This needs aggregation of data across agencies and coordination consist ing the demands, availability, density on roads. At the end-user level, this should lead to mobile applications and public displays (say at bus stops) for the information dissemination. Technologies for this are quite mature and can lead to immediate visible results.

Sensing the Traffic State of the traffic, specially on roads, change over time. The number of vehicles and pedestrians on the road may vary due to the day, time and many special events. The type and amount of traffic participants needs to be sensed on a regular basis and used for dynamically changing the solutions (eg. traffic lights, routing). Sensing the

traffic situation with a wide a set of distributed sensors is an immediate step that can enable a number of applications.

Traffic control and monitoring Understanding the traffic needs deeper intelligence. Each traffic participant may have its own behaviors and preferences. Characterizing these from the individual actions and history/prior knowledge helps in the planning. Modeling of traffic flow on the road networks and usage patterns is fundamental to scientifically model the problem and simulate the situations while planning and scheduling.

Indian Open Maps have become the basic backbone infrastructure for many digital applications in the transportation sector. India needs its own maps. A number of semantics that are used in the Indian setting for navigation, localization (addresses) and planning are missing in the existing map systems. Also this needs to be fully open and free. This demands building interactive and intelligent maps with open access. This will enable large scale applications on top of it.

Traffic Simulator with Indian Flavor A simulation system that can simulate a specific traffic situation is important. Such a simulator needs to model and use the volume and variety of traffic participants, behavioral models of the participants, resource constraints (such as road width and varying widths), and dynamically changing artifacts (such as potholes and water logs).

Minimzing Traffic Jams and Improving Road Efficiency With the limited infrasW tructure, one needs to obtain larger throughput. This is particularly for highways passing through small towns, facing jams. Planning and routing schemes that can use the dynamic traffic conditions can help in this.

Reduction in Road Fatalities Road fatalities on Indian roads are far higher compared to that in the west. A set of steps that lead to reduction in road fatalities needs to be undert taken. Technology can help in this. This include identification of accident prone zones, speed control, road user education, developing simulators etc. Simple low cost driver assistance systems are critically needed in the Indian setting. This may be dedicated hardware with minimal calibration/customization or even implemented on a mobile phone itself.

4.10 Highways and Waterways

Monitoring of Road and Associated Fixtures Regular monitoring of the state of the roads can help in streamlining traffic and planning the regular maintenances. An estimate of potholes, waterlogs, encroachments and other damages can be estimated with automated monitoring systems mounted on vehicles. They can also help in rapidly acting after a natural disaster. In addition to the road surface quality, one can also log the road accessories such as traffic signs, traffic lights, street lights, etc.

Construction Quality Estimation and Certification An automated scanning and quality estimation system will provide objective profiles, analyze and correlate with any potential damages in the later stage. Results and observations from the novel sensor based scanning could also correlate with the classical metrics employed in the domain.

4.11 Railways

Fog Vision A number of trains get delayed in the winter due to the adversarial weather. The weather conditions may be bad in some parts, however, the effects gets percolated to other parts of the country resulting in massive disruptions in schedule. Intelligent and computational vision systems need to be developed and made available to the locomotive drivers.

Track Inspection Railway tracks get inspected by humans. This is effective in many situations but challenging in practice. A complementary direction is to do automated inspection with sensors (such as cameras) that can image tracks on a regular basis from all the moving trains. Potential faults could be tagged early and human intervention is then possible on time.

Undercarriage Monitoring Safety of the rail carriages is very important. One should regularly monitor the state of the carriages and tag the potential issues specially due to wear and tear. This may also help against the unwanted sabotage attempts. Undercarriage monitoring technologies at stations (or places where trains often slow down) can be a regular security measure in this direction.

Freight planning and wagon utilization Freight transportation is an important activity of railways. However, demands are often asymmetric and skewed. This leads to an

optimization problem with immediate applications.

4.12 Energy

Sustainable Energy and Smarter Grids A significant percentage of the population still has no access to reliable energy sources. This affects the local economy, life and even agriculture. There has been initiatives in expanding the renewable energy sources. The generation of such power is uneven. This demands smarter ways to distribute and consume the energy in the society. This leads to a complex control and optimization problem with multiple constraints and behaviors. Managing incentive and pricing models can also lead to systematic changes in the usage patterns leading to better efficiency.

4.13 Habitat

Better use of Resources India is still in the infancy of construction of buildings. We expect to see lot more buildings getting constructed in the next two decades. Given that our resources are constrianed, we need these buildings to be smart. This requires the tighter integration of AI, IoT and Information technology. Develop technologies that lead to intellii gent operation of buildings by efficient and effective use of resources. Resources could include natural resources, energy etc. We need to meet the potentially contradicting demands of the users, economies and sustainability. AI can help in making the right and optimal decisions automatically and provide intelligent automation by combining signals/observations from different sources and reasoning by meeting the complex optimization requirement.

Buildings interfacing with smart grids Buildings require external inputs (such as power, water etc.) on a regular basis. They could be rare resources or could be costly. Smart grids allow the participants to contribute to the demand supply equilibrium. Each player has different objectives to optimize (eg. comfort vs cost). Buildings could be buying or selling resources to the grids, and have complex objective to optimize and the collective optimization of the performance across a set of buildings (or a city or even country) is far more involved. AI, Multi Agent Systems, Game Theory, Optimization etc. are critical technologies that will have impact in this space. Many of the decisions to be made here are dynamic and time varying. This naturally forces us to use machines and technology.

4.14 Public Safety

Crowd Management during Large Gathering It is common to large festive and religious gatherings with many people located in smaller area with limited exit points. In case of any accident or even in presence of hoax, a stampede could happen. It is important to develop technologies that can be ported and deployed at short notice in different places. This could involve crowd measurement, crowd movement pattern, prediction of imminent accidents or events, prediction on unwanted incidents (warnings). Such portable and easy to deploy crowd management and monitoring system could use drones, surveillance cameras and other signals, together with AI vision systems that can count or estimate crowd in delimited areas and Coming with evacuation plans/strategies and developing technologies that can deliver the recommendations to the individual users and contribute towards avoiding dangerous high build ups of crowds.

4.15 Disaster Management

Floods see under Water Resources

Forest Fires see under Environment and Pollution

Early warning Natural disasters leave strong negative effects on the society, economy and country. Advance prediction of these will have major impact on saving life and economy. Many recent advances in prediction techniques including complex modeling and sequence prediction can enhance the performance of the early warning systems.

Also the advances in perception and processing of different sensor outputs will have major impact. Prediction of Flood (see also under water resources), Earthquake, Typhoon, forest fire etc. will benefit by this.

Rescue, Recovery and Rehabilitation Damages like earthquake, flood leave a city or a larger area isolated from the rest of the world in communication and reachability. A set of technology including (i) sensing life, humans in the debris (ii) drones and similar technology for surveillance, estimation of the damage and providing immediate help (iii) alternate communication channels for maintaining connectivity etc. are required.

4.16 Legal

Enhancing Effectiveness of the Legal System With a number of years of processing experience, Indian judicial system has accumulated a large number of records of judgments. Some of them are in English and some of them are in different Indian languages. Making them accessible across languages through various language processing related technologies (OCR, Search, Machine Translation etc.) is required.

Reducing Pendency of judicial cases There are approximately 27 million cases pending in Indian courts. Speeding up the judicial process is a critical need. Data analysis on flow of cases (with appeals) type of sections, can help in speeding up this significantly.

Search engine with information extraction Effectiveness of the information access can improve the legal process since the participants of the system at any point will have access to all the related information in a compact form. This requires domain specific information extraction and retrieval techniques.

Judicial awareness among people Increasing awareness of judiciary including the process, facts and interpretation, can help in enhancing confidence of the people in the system. Also this can lead to reducing the number of potential disputes that could emerge due to the lack of awareness. AI and information technology can help in scaling such an effective awareness program to the wide spectrum of the public.

4.17 Finance

Detection of Financial Frauds Financial frauds are increasing. This is primarily due to the fact that new digital channels, payment devices etc. leave a lot of loopholes to be exploited by fraudsters. Consequently, cyber security incidents in India have grown significantly very often resulting in huge financial loss. While there are some commercial tools available in the market, there is an urgent need to develop more sophisticated tools incorporating cutting-edge algorithms that can perform incremental learning, which is required for near-real time predictions.

Prediction of bad loans/non-performing assets Prediction of potential bad loans ahead of time can lead to early actions in a Banking system. This requires AI algorithms and predictive analytics.

5. Grand Challenges

Health To have an AI primary-healthcare center, with healthcare workers providing patient interfacing, AI support for screening, pathology, diagnosis, therapeutic recommendations, and referral triaging, with doctors handling invasive interventions like childbirth, surgery, etc. with machine learning.

Flood Prediction To predict floods and water levels in river basins, specially Ganga and Kaveri, based on remote sensing and meteorological data using models for river basin topology, water flow and aquifer exchange.

Forest Fire Mitigation To map forest fires, predict their spread based on remote sensing, meteorological data, special sensors (drones), forest models, and suggest strategies for their mitigation.

Universal Language Access To translate spoken lectures into all Indian languages automatically and instantly, thus opening up access to knowledge for all.

Culture and History To create virtual tours around monuments recreating the bygone era including music and dance.

Transportation To reduce road fatalities by identification of accident prone zones, speed controls and enforcement (using road signs, speed breakers, driver education, monitoring through AI patrols), and fast response in case of accidents.

Fog Vision for Railways To provide visibility to loco drivers during fog etc using multimodal cameras and other sensors on locomotives, vision on bends in track, etc.

Public Safety To monitor and manage crowds during large gatherings in small delimited areas using cameras on drones. Dissemination of information to general public as well as safety services and police.

Legal To develop search engine and information extraction for legal domain to assist courts, lawyers, police, and general public.

Finance To detect fraud in banking transactions.

6. Discussions

Section 5 lists problems in each sector. We have selected those problems whose solution is important for the sector, as well as which are more amenable to AI.

On the other hand, the problem areas are large and the research and development manO power in India is limited. Our education lacks problem solving approach and is rote learning oriented. It has much less project orientation. There are also international studies that indicate that Indian students are much poorer in their comprehension of subject compared to Russian or Chinese students. Indian society on the other hand is highly creative (jugaad), except that the creativity seldom finds organized expression.

6.1 Methodology of Development of Solutions

Development of solutions to problems requires the work by knowledgeable, creative and motivated people. Standard methods of first identifying research areas, then asking researchers to submit proposals (Request for Proposals or RPFs), their evaluation and award of projects can be used. Major changes are needed, however, Instead of granting projects to individual researchers in academic & research institutions, it should become more team oriented and goal directed approach. In addition, we propose a different method as well.

Main idea is to inspire a large number of students in engineering colleges to work on technical problems in AI that are strongly related to national missions. This can be done by organizing national contests { say, a contest or two in each sector. Teams of students would be invited to work on the contest problems. The total duration of the contest would be 6 to 9 months, with pre-announced dates and timelines.

Each contest itself would have a well defined deep technical task, relevant training data, together with some test data, all released at the same time. The enrolled student teams would have 3 to 4 months to develop their algorithms to carry out the tasks. The first round of the contest could be carried out at the end of the 3 to 4 months period on unknown data at their own college. The teams would have to run their solution on the unknown data set and upload the result. Evaluation would be done by automatic evaluation instantly. Based on the evaluation scores, 20 to 25 winning teams would be shortlisted for the 2nd round of the contest. The 2nd round would take place at a selected centre to which the shortlisted teams would travel to. In the 2nd round, besides running their solution on the new unknown data, each team would also present their approach in front of all the other teams (through a research and innovation presentation). The winners would be decided based on the performance of the algorithm as well as innovativeness of the approach.

The presentation serves to reveal the different approaches to everyone which lead to much better performance in the next year's contest. Within a couple of years, strong approaches and teams emerge. The solutions would far surpass any known algorithm for the task. This has been the experience wherever such contests have been organized. They are a fast means to move the research frontier. This methodology has the potential to arouse the imagination of tens of thousands of students in the country. If this is done over 10-15 sectors in each of which some 250 teams participate, consisting of, say, four students each, it would directly engage 10,000 to 15,000 students. They would be motivated to learn on their own, wherever required. For the first time, perhaps, Indian students would be inspired to work on Indian problems . For the above to work, it requires careful selection of the technical task for the contest. For example, under public safety mission, images of crowds from drones could be used to estimate crowd density in an identified geographical area. A technical AI task for this mission would be to count people (or their heads) from the aerial images. As the accuracy improves for the technical tasks, the AI application for public safety would be able to perform better. The selection of the technical task should be done by the experts in AI as well as experts in domain areas. For example, to predict flooding in river basins, it would require an expert team consisting of AI experts dealing with image processing, remote sensing experts and river modeling experts. The technical tasks could be river flow estimation for a given segment of the river using recent rainfall data and river models.

6.2 Methodology of Delivery of Solutions

For sectors and missions where sufficient advances have been made in AI algorithms (for major tasks), one should promote development and delivery of solutions to end users. The end users might be common people or service delivery organizations. Several methods are available to deliver solutions:

1. Incubating companies

2. Technology transfer to established companies,

3. Public organizations with open software models

A proper combination is needed of all the three models in the sectors. The first one is most agile, the second one is less agile but has a larger reach, and the third one works better in those areas where direct revenue generation and market development is some distance away.

With a large number of students inspired to work on national problems and develop solution for related technical tasks, it is hoped that some of them would become entrepreneurs. As mentioned, if a few hundred teams work in each sector, some 10 of them might get inspired to also incubate companies in each sector

6.3 Availability of Data

AI and Machine learning applications thrive on data. If data is not available to researchers and later for deploying the developed solutions, we would have largely missed the AI revolution. The applications would not get developed in the first place, and even if they are developed, they would languish without getting deployed.

A large amount of data exists today in a plethora of sectors. However, it mostly resides in stand alone mode without it being used effectively. Even when it is used, it gets used only in silos. Many times, data integration across sectors produces spectacular benefits, which are missed.

For these benefits to be realized fully, data should be hosted on cloud platforms in well defined data formats. Much of the data can be made open, taking care of the privacy and security concerns. In any case, the data must be opened up to researchers and then later for deployed solutions.

Special effort with support of the highest level is needed to put the data on cloud platforms in well defined formats. Ministries have to come forward and take a lead in this direction.

7. Conclusions

AI as a technology holds a lot of promise. It uses data platforms, AI algorithms and AI applications. It can be applied in many sectors, particularly where a large amount of data is available or likely to be available. About 20 sectors are identified and national missions are worked out in each sectors vary from agriculture and food to water resources, environment, education, transportation, legal, finance, governance etc. Finally, a set of grand challenges are proposed. Grand challenges have the power to catch the imagination of the young. A major new method proposed to involve tens of thousands of students is to hold national contests in each sector. The students will get inspired to work on problems of the country. Out of these contests will emerge technical advances which can then be put to use in service of society through incubated companies or tech transfers to existing companies or open source models. For all this to work there are three important tasks:

1. National missions be selected in each sector. (This report lists several missions in each sectors.) Road map be drawn.

2. For each mission, important technical tasks be identified for contests. (This requires thoughtful work and preparing the training and testing data.) The contests would have to be announced by August 2018 for them to reach fruition by February 2019.

3. Data available with ministries and departments be brought onto common platforms in well defined formats.

Finally, the AI grand challenges be posed as a challenge to catch the imagination of the nation. Road map be prepared for each grand challenge.

8. References

1. Technology Vision 2035, Technology Information, Forecasting and Assessment Council,

2. UN Sustainable Development Programme

3. A National Machine Intelligence Strategy for the United States

Sectors and Line Ministries

A List of Domain Areas vs Line Ministry

Domain Area	Line Ministry
Medical Science & HealthM care	Ministry of Health and Family Welfare
Energy	Ministry of Power • Ministry of Non { ConvenM tional Energy•Ministry of Petroleum and Natural Gas
Water Resources	Ministry of water Resources, River Development
Environment	Ministry of Environment and forests
Food	Ministry of Food Processing Industries
Agriculture	Ministry of Agriculture & Farmer Welfare
Highways & Waterways	Ministry of Road Transport and Highways
Habitat	Ministry of Housing and Urban Affairs
Education	Ministry of HRD
Culture	Ministry of Culture
Transportation (Vehicles etc.)	Ministry of Road Transport and Highways
Information & Communicat tion	Ministry of Electronics Information Technology Technology
Railways	Ministry of Railways
Manufacturing (AutomoM bile, Department of Heavy Industry Machinery etc.)	Ministry of Heavy Industries & Public Enterprises
Metals & Mines	Ministry of Mines
Finance	Ministry of Finance
Legal	Ministry of Law and Justice
Public Safety	Ministry of Home Affairs
Governance	I{
Disaster Management (Preservation	Ministry of Disaster Management Warning, mitiM gation)
Specially Abled	Ministry of Social Justice and Empowerment.

B Mmbers and Meetings Held

B.1 Members

1	Prof. Rajeev Sangal, IIT (BHU)	Chairman
2	Dr. B. K. Murthy, GC R&D in IT , MeitY	Member Convener
3	Prof. Santanu Chaudhury, Dir CEERI Pilani	Member
4	Prof. Hema Murthy, IIT Madras	Member
5	Prof. Anil Gupta, IIMA and Sristy	Member
6	Prof. C.V. Jawahar, IIIT Hyderabad	Member
7	Dr. Narendra Ahuja, DIC	Member
8	Shri Sanjay Sinha, IBM Watson	Member
9	Prof. B. Yegnanarayana, IIIT Hyderabad	Member
10	Shri Sudhanshu Mittal, NASSCOM	Member
11	Rep from NeGD	Member
12	Dr. Hemant Darbari, DG C-DAC	Invitee
13	Shri Ragu Dharmaraju, Wadhwani	Invitee
14	Shri Rahul Panicker, Wadhwani	Invitee
15	Prof. Vadlamani Ravi, IDRBT	Invitee
16	Shri Manoj Jain, Scientist 'F', MeitY	Invitee
17	Shri Anil Kumar Sagar, Scientist 'C', MeitY	Invitee

B.2 Meetings Held

- 1st meeting was held on 28th February 2018
- 2nd meeting was held on 13th March 2018
- 3rd meeting was held on 21st March 2018
- 4th meeting was held on 28th March 2018
- 5th meeting was held on 10th April 2018