

Policy and Ethical Issues related to Artificial Intelligence

Shailendra K. Hajela

Chairman, ITU – APT Foundation of India, India

Policy and Ethical Issues related to Artificial Intelligence

- Technology and Society: Agrarian Age, Industrial Age, Information Age
- Moores law: Rapid advancement in silicon technology and phenomenal rise in computing power;
- Key components of ICTs: Computer chips, lasers, broadband Internet, and software ;
- ICTs: tools for thought that amplify brainpower in the way the technologies of the Industrial age amplified muscle power;
- Evolution of computer technology: Internet and World Wide Web;

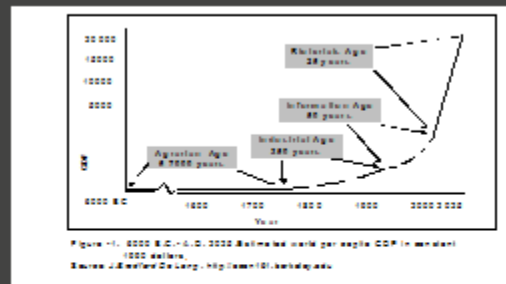
Policy and Ethical Issues related to Artificial Intelligence

- Internet of Things (IoT); Artificial Intelligence: Fuzzy logic, soft computing, cognitive informatics;
- Potential, Possibilities and predictions;
- Ethical Issues of AI;
- Studies by UNESCO (COMEST), ITU, IEEE, Academia, and ICT Industry: Stanford Project
- Public Policy Imperatives;
- Ethical Rules
- Conclusion

Impact of Technology on the Society

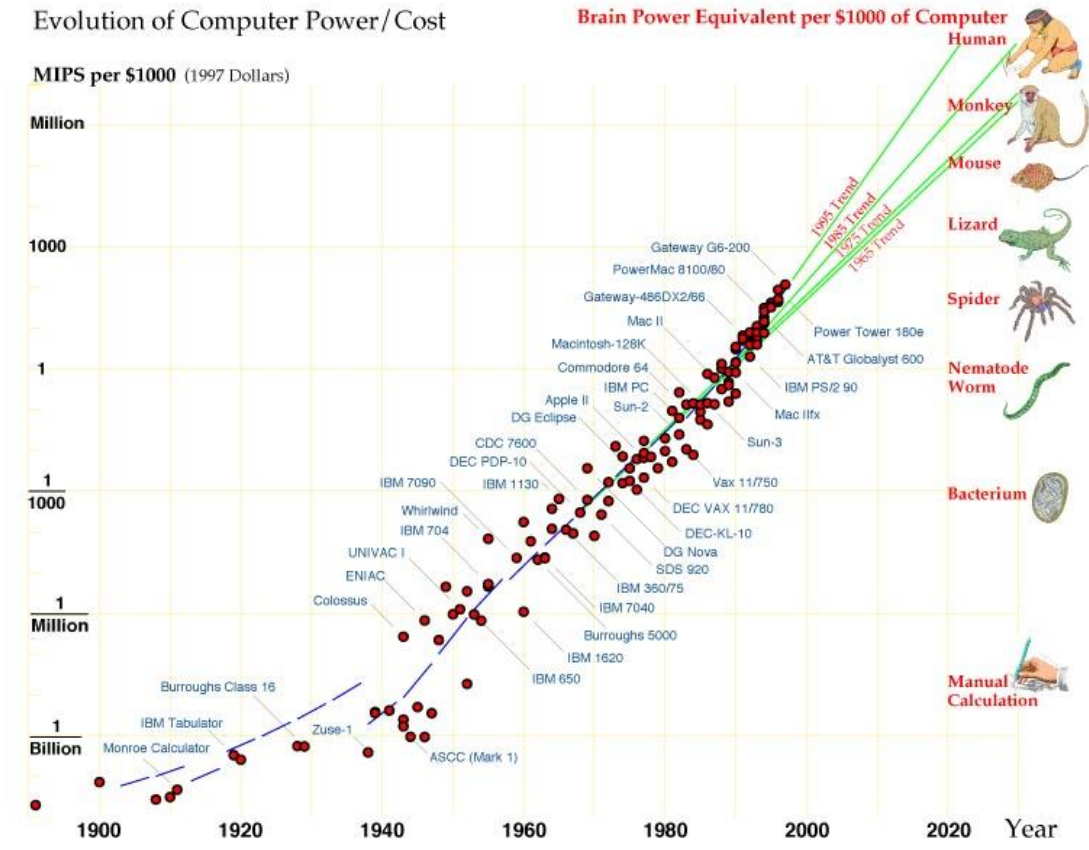
Impact of technology on the society

"Technology creates economic eras"



During the Agrarian age the annual GDP per capita worldwide remained around \$ 120, in the Industrial age it increased rapidly in the two centuries up to 1950 from about \$ 180 to \$ 1622. This increase was confined to a relatively small number of geographic regions, Western Europe, North America, and later on Japan. The Industrial age also divided the world into First, Second and Third World. The First or the industrial world was the inventor and the first adopter of industrial technologies. It accounts for 10% of the world's population, but enjoys 70% of its wealth. The Second World's (Eastern Europe, South America) rate of adoption was slower and consequently they fell behind in the standard of living, accounting for 30% of the world's population and 20% of its wealth. The Third World (Rural Asia and Africa) accounts for 60% of the population but enjoys less than 10% of the world's wealth. ICT presents an opportunity to the third world to gain equality in the information Age.

The exponential growth in computing power



(Source: Hans Moravec: "When will computer hardware match the human brain?" *Jour. of Transhumanism*, Vol.1. 1998.)

Information and Communication Technologies

- Key components of ICTs: Computer chips, lasers, Internet, broadband and software;
- Evolution of computer technology: embedded inside – pervasive computing, data analytics, artificial intelligence; and connected outside to form a network of computers, resulting in Internet and World Wide Web; a distributed global database of information accessible 24x7 including through mobile devices
- The extraordinary build-out of the communication networks that link computers together is almost as remarkable as the explosion in computing power.
- Mobile communication, 4G, 5G, coupled with optical fiber communication infrastructure making available huge bandwidths have resulted in faster emergence and rapid diffusion of the Digital Economy, more widely, throughout the world than previous technological revolutions giving rise to a new era in human development.



ICT - Internet of Things and Artificial Intelligence

- ICTs: tools for thought that amplify brainpower in the way the technologies of the Industrial age amplified muscle power;
- Brainpower has a far greater impact on society as it causes and controls all human activity;
- Emergence of smart/ intelligent gadgets/machines/robots
- The Internet of Things (IoT) is built on the seamless interaction between objects.
- As the Internet and smart systems penetrate further each day into our daily lives, concerns are expressed about the benefits to society and potential risks.
- We are on the cusp of totally connected world comprising humans and things, an amalgam of human (natural) and artificial intelligence. The IoT is no longer an emerging trend, it has arrived. But security, safety, and privacy run the risk of becoming an afterthought in the great IoT march.
- Real evidence: IoT platform provider Jasper technologies 1.4 billion dollar acquisition by Cisco in March 2016.

Fuzzy logic - a way for progression toward achieving human level machine intelligence (1)

- Pervasive computing, fuzzy logic, cognitive informatics - an interdisciplinary area, which encompasses informatics, computer science, software engineering, mathematics, knowledge theory, cognition science, neurobiology, psychology, and physiology, tend to make artificial intelligence approach human level natural intelligence.
- Emeritus Prof. Lotfi Zadeh of UC Berkeley in his Keynote at WCECS 2014 had described fuzzy logic as aimed at providing a basis for formalization/ mechanization of two remarkable human capabilities: the capability to communicate, reason and make rational decisions in an environment of imprecision, uncertainty and partiality of truth; and the capability to perform a wide variety of tasks, such as driving a car in heavy city traffic, without any measurements and any computations.
- To this end, fuzzy logic is designed to deal with imperfect information which in one or more respects is imprecise, uncertain, incomplete, unreliable, vague or partially true. In the real world, such information is the norm rather than exception.

Fuzzy logic - a way for progression toward achieving human level machine intelligence (2)

- To deal with imperfect information what is needed is the fuzzy-logic-based methodology of Computing with Words (CWW). In CWW, the objects of computation are words and propositions are drawn from natural language. There are two stages of computation and deduction.
- In the first stage, the meaning of words and propositions is made more precise through translation into the so-called Generalized Constraint Language (GCL) and in the second stage, computation/deduction is carried out through propagation and counter-propagation of generalized constraints.
- By providing a system for dealing with imperfect information expressed in a natural language, fuzzy logic provides a basis for achievement of two basic objectives. First, it opens the door to a wide ranging enlargement of the role of natural languages in scientific theories; and second, it points to a way for progression toward the achievement of long-sought human level machine intelligence.

IoT and AI – Dawn of New Technological Era

- Siri (Apple's digital assistant), Alexa (Echo, Amazon), Google translate, smart houses and smart cities; Google, Tesla and other manufacturers' chauffeur less vehicles, autonomous Uber and Singapore Taxis, and Airplane taxis as reported recently, to name a few, manifest that the new age of IoT and AI is already upon us.
- The positive impact of intelligent vehicles would be enormous. Consider the potential ecological savings of using highways so much more efficiently and from the safety aspect reduce the injuries and deaths in road accidents. Besides the reduction in accidents, the traffic throughput is estimated to increase up to three times. Elderly and disabled people would be able to travel around in such cars on their own. People could dispatch their cars to the automated parking warehouse autonomously instead of using surface land for parking and then recall them later. Truly, gains in this area are enormous and most encouraging.
- But this presupposes robust, reliable and secure networks, devices, and systems, in which individuals and society can repose complete trust.

IoT and AI – Dawn of New Technological Era

- By most indications from practical product developments and integrated systems, AI is now mature, both as a science and as an engineering discipline. Many opportunities exist for AI to have a positive impact on our planet's environment. AI researchers and development engineers have a unique perspective, tools and the skills required to contribute practically to address concerns of global warming, poverty, food production, arms control, health, care of the elderly, education, and demographic issues.
- The embedded, ubiquitous, distributed intelligence in the World Wide Web and other global computational networks as an amalgam of human and artificial intelligence can be seen to be evolving to become a World Wide Mind, according to UC Berkeley Prof. Bradford Delong.
- The kinds of agents we build, and the kinds of agents we decide to build, will change us as much as they will change our society; we should make sure it is for the better. Many of our old social and ethical codes may be broken if the interplay of NI and AI does not work in harmony in this upcoming new heterogeneous world where humans with natural (biological) intelligence and objects with artificial intelligence would coexist.

Ethical Issues of AI

- The digital economy now enters into too many domains of our daily life, affects too many economic interests;
- In planning for the future, we need to consider the potential benefits of the dramatic progress in technological capabilities.
- It would be tragic if the potential benefits failed to materialize because of technophobia and unnecessary restrictions. On the other hand, it would also be tragic if intelligent life were to become extinct because of some disaster or war involving advanced technologies.
- As the developments of AI tend to move more and more from the realm of science fiction to reality, its role and importance in shaping the future of mankind has sensitized governments, UN/ international organizations (e.g. UNESCO, ITU), academia, R&D Labs, professional bodies, e.g. IEEE, public and private sector corporations, civil society, and others to study and debate the ethical issues.

Asimov's Laws

- Alan Turing in 1950 asked whether machines could think and that same year writer and academician Isaac Asimov contemplated what might happen if they could, and he propounded what we recognize as the first clear laws of Robotics to govern the design of intelligent machines:
 - A robot may not harm humanity, or, by inaction, allow humanity to come to harm.
 - A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
 - A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

Later on Asimov added another law that concerned the humanity as such:

- A robot may not harm humanity, or, by inaction, allow humanity to come to harm.
- Science fiction became fact in 1997 when IBM's chess-playing Deep Blue computer beat world champion Garry Kasparov. But then it was domain specific AI pertaining to the play of game of chess. Deep Blue could have failed in other simpler games.

Concerns about AI

- Alongside the growing interest in exploiting the potential of AI for human good, there are also concerns expressed about its unethical use. Entrepreneur/inventor Elon Musk of Tesla and SpaceX, who spends considerable time on the cutting edge of technology, is reported (Feb. 2015) to have warned that AI could be an existential threat to humanity.
- Bostrom and Yudowsky have studied and analyzed at length the ethics of AI, and emphasized that thinking machines are not as versatile as humans and can have only domain specific intelligence as per its design, suitable for the assigned task, but unsuitable for others. However, the possibility of developing super intelligent machines cannot be ruled out. They go further on to discuss the moral status of such machines themselves.
- AIs with sufficiently advanced mental states may count as persons—though maybe persons very much unlike us and perhaps to be governed by different rules?

Concerns about AI

- Satya Nadella, Microsoft CEO, is reported to have expressed the view that AI research may follow to keep society safe. The AI must be designed to assist humanity; maximize efficiencies without destroying the dignity of people; be transparent so that humans understand how they work; be designed for trust in privacy and guarding their information; have algorithmic accountability” so that humans can undo unintended harm and guard against bias, so that they do not discriminate among people.
- To create a standard of ethics around the creation of artificial intelligence, five major companies: Alphabet, Amazon, Facebook, IBM and Microsoft are backing the Stanford project, named the “One Hundred Year Study on Artificial Intelligence”, which would produce a detailed report on the impact of A.I. on society every five years for the next century.

Concerns about AI

- The recently released (September 2016) first Stanford report titled “Artificial Intelligence and Life in 2030,” states that the basic objective is to ensure that A.I. research is focused on benefitting people. It may be extremely difficult to regulate A.I., since there is no clear definition of A.I. It isn’t universal but generally, domain or application-specific, and the risks and issues are different in different domains. However, it does not rule out regulation totally, but underlines when it would be good to do so.
- The Stanford report attempts to define the issues in the context of a North American city in intelligent systems that mimic human capabilities and explores eight aspects of modern life, including health care, education, entertainment and employment, but specifically does not discuss the issue of warfare and military AI, as being outside their current scope and expertise, but they did not rule out focusing on weapons in the future.
- The report also does not consider the belief of some computer specialists about the possibility of a “singularity” that might lead to machines that are more intelligent and possibly become a threat to humans.

UNESCO's World Commission of the Ethics of Scientific Knowledge and Technology (COMEST)

- The WG on Robot Ethics of COMEST, at its meeting in May 2016 was to examine the major ethical issues surrounding the development and application of machines that encompass both physical robots and software agents, which are designed to function independently from direct human oversight and can learn by themselves new process or behaviours.
- The rapidity in the growth of the autonomous robots, both for civil and military purposes, leaves a gap between the effective use of the technology and its ethical application such as human well-being, safety or social benefits.
- Although the outcome of the meeting is awaited, its agenda envisaged to discuss the emerging ethical issues from two perspectives. First, the ways in which engineers and researchers design, build and use machines/robots in accordance with human morality and ethics, and second to find answers to such questions as:
 - Do we want machines/robots to make morally important decisions? Are we then abdicating our responsibility to machines?
 - What is the ethical understanding given to machines/robots and whether they should be considered as moral agents with artificial intelligence or even living creatures?

Furthermore, considering that in the future, intelligent machines/ robots will be sharing the world with humans, to explore the extent to which developments in robotics, as an aspect of a broader paradigm of technological convergence, point towards potential new understandings of “human” in respect of neurological implants and enhancing technologies.



Public Policy Imperatives

- AI is already changing our daily lives, in ways that improve human health, safety, and productivity.
- Internet of Things (IoT) research is devoted to the idea that a wide array of devices, including appliances, vehicles, buildings, and cameras, can be interconnected to collect and share their abundant sensory information to use for intelligent purposes.
- The policy should be broad enough to cover AI systems that can work collaboratively with other AI systems and with humans in applications that can utilize the complementary strengths of humans and machines—for humans to help AI systems to overcome their limitations, and for agents to augment human abilities and activities.
- As with all technological innovations, we need to ensure that robots are introduced from the beginning in a way that is likely to gain public trust and confidence; maximize the gains for the public and commerce; and proactively take care of any potential unintended consequences.
- Public policies should facilitate society's adaptation to AI applications, extend their benefits, and mitigate their inevitable errors and failures.

Public Policy Imperatives

- While encouraging innovation, policies and processes should address ethical, privacy, and security implications, and ensure that the benefits of AI technologies will be spread broadly and fairly. Further, research on the fairness, security, privacy, and societal implications of AI systems should be encouraged by removing impediments and increasing private and public spending to support it.
- Policies should be evaluated as to whether they democratically foster the development and equitable sharing of AI's benefits, or concentrate power and benefits in the hands of a fortunate few.
- As future AI technologies and their effects cannot be clearly foreseen, the policies will need to be continually re-evaluated in the context of observed societal challenges and evidence from fielded systems. The issues have moved from the narrowly technical through the narrowly legal into fundamental questions of how AI would reorganize our markets and our society.

Ethical Rules

- Policymakers should recognize that to varying degrees and over time, various industries will need distinct, appropriate, regulations that touch on software built using AI or incorporating AI in some way. The government will need the expertise to scrutinize standards and technology developed by the private and public sector, and to draw up regulations where necessary.
- Robots themselves are not where responsibility lies. They are simply tools of various kinds, and the responsibility for their behaviour must always lie with human beings. Accordingly, rules for real robots, in real life, must be transformed into rules advising those who design, manufacture, sell and use robots about how they should act.
- The five simple ethical rules for robotics as proposed by a group of experts (Ref. Johanna J. Bryson <http://www.cs.bath.ac.uk>) from various disciplines including social sciences, law, philosophy and the arts, intended as a starting point for further debate are as follows:
 - Robots are multi-use tools. Robots should not be designed solely or primarily to kill or harm humans, except - in the interests of national security.
 - Humans, not robots, are responsible agents. Robots should be designed; operated as far as is practicable to comply with existing laws & fundamental rights & freedoms, including privacy.
 - Robots are products: as with other products, they should be designed to be safe and secure.
 - Robots are manufactured artefacts. They should not be designed in a deceptive way to exploit vulnerable users; instead their machine nature should be transparent.
 - The person with legal responsibility for a robot should be attributed. It should be possible to find out who is responsible for any robot.

Ethical Rules

- Legal responsibility: Logically, a robot is never legally responsible for its action or result of its malfunction. It is a tool not a person. If it malfunctions and causes damage, a human will be to blame. Responsibility might be practically addressed in a number of ways, for example, by a license and register (just like for cars) that records who is responsible for any robot. This might apply to all or only operate where that ownership is not obvious (e.g. for a robot that might roam outside a house or operate in a public institution such as a school or hospital). Alternately, every robot could be released with a searchable online license which records the name of the designer /manufacturer and the responsible human who acquired it. More debate and consultation would be required on legal liability issue.
- Benchmarking AI: To date human intelligence has no match in all biological species, and artificial worlds for its sheer versatility in diverse fields, with the abilities “to reason, achieve goals, understand and generate language, perceive and respond to sensory inputs, prove mathematical theorems, play challenging games, synthesize and summarize information, create art and music, and even create histories.” This makes human intelligence a natural choice for benchmarking the progress of AI. Notwithstanding this, the human mind suffers from moral dilemma about the choices for action in complex situations. Errors of judgement do occur in humans.
- Certification: The very notion of “artificial intelligence” suggests a substitution for human skill and ingenuity. And in many contexts, ranging from driving to performing surgery or practicing law, a human must attain some certification or license before performing a given task. Accordingly, law and policy will have to develop yardsticks to determine level of competency in an AI system.

Conclusion

- Recognizing that Policy and Regulation that stifles innovation, or relocates it to other jurisdictions, would be counterproductive, we need to create forums where people can rationally debate what needs to be done and a social order where responsible decisions can be implemented. As the subject matter pertains to ICT, ITU is the natural and unambiguous choice to lead and provide such a platform to all stakeholders for study, information sharing, and coordinated effort for arriving at plausible solutions. Like other technologies, AI has the potential to be used for good or evil purposes.
- A vigorous and informed debate about how to best steer AI in ways that enrich our lives and our society, while encouraging creativity in the field, is an urgent and vital need.
- ITU may bring together all stakeholders for study and exchange of ideas and information with a view to finding plausible solutions and direction for future work.

Thank you for your attention







CCITT / ITU-T