

# Artificial Intelligence: Top 9 ethical issues associated with AI

## **Relevancy:**

*GS Mains paper IV, Essay paper, Optional: Sociology*

*Ethics, Issues in Artificial intelligence*

## **Background:**

- *Optimizing logistics, detecting fraud, composing art, conducting research, providing translations: intelligent machine systems are transforming our lives for the better.*
- *As these systems become more capable, our world becomes more efficient and consequently richer.*
- *Tech giants such as Alphabet, Amazon, Facebook, IBM and Microsoft – as well as individuals like Stephen Hawking and Elon Musk – believe that now is the right time to talk about the nearly boundless landscape of artificial intelligence.*
- *In many ways, this is just as much a new frontier for ethics and risk assessment as it is for emerging technology. So which issues and conversations keep AI experts up at night?*
- *The main issues to be discussed are:*
  - 1. Unemployment. What happens after the end of jobs?**
  - 2. Inequality. How do we distribute the wealth created by machines?**
  - 3. Humanity. How do machines affect our behaviour and interaction?**
  - 4. Artificial stupidity. How can we guard against mistakes?**
  - 5. Racist robots. How do we eliminate AI bias?**
  - 6. Security. How do we keep AI safe from adversaries?**
  - 7. Evil genies. How do we protect against unintended consequences?**
  - 8. Singularity. How do we stay in control of a complex intelligent system?**
  - 9. Robot rights. How do we define the humane treatment of AI?**

### **1. Unemployment. What happens after the end of jobs?**

*The hierarchy of labour is concerned primarily with automation. As we've invented ways to automate jobs, we could create room for people to assume more complex roles, moving from the physical work that dominated the pre-industrial globe to the cognitive labour that characterizes strategic and administrative work in our globalized society.*

*Look at trucking: it currently employs millions of individuals in the United States alone. What will happen to them if the self-driving trucks promised by Tesla's Elon Musk become widely available in the next decade? But on the other hand, if we consider the lower risk of accidents, self-driving trucks seem like an ethical choice. The same scenario could happen to office workers, as well as to the majority of the workforce in developed countries.*

*This is where we come to the question of how we are going to spend our time. Most people still rely on selling their time to have enough income to sustain themselves and their families. We can only hope that this opportunity will enable people to find meaning in non-labour activities, such as caring for their families, engaging with their communities and learning new ways to contribute to human society.*

*If we succeed with the transition, one day we might look back and think that it was barbaric that human beings were required to sell the majority of their waking time just to be able to live.*

## **2. Inequality. How do we distribute the wealth created by machines?**

*Our economic system is based on compensation for contribution to the economy, often assessed using an hourly wage. The majority of companies are still dependent on hourly work when it comes to products and services. But by using artificial intelligence, a company can drastically cut down on relying on the human workforce, and this means that revenues will go to fewer people. Consequently, individuals who have ownership in AI-driven companies will make all the money.*

*We are already seeing a widening wealth gap, where start-up founders take home a large portion of the economic surplus they create. In 2014, roughly the same revenues were generated by the three biggest companies in Detroit and the three biggest companies in Silicon Valley ... only in Silicon Valley there were 10 times fewer employees.*

*If we're truly imagining a post-work society, how do we structure a fair post-labour economy?*

## **3. Humanity. How do machines affect our behaviour and interaction?**

*Artificially intelligent bots are becoming better and better at modelling human conversation and relationships. In 2015, a bot named Eugene Goostman won the Turing Challenge for the first time. In this challenge, human raters used text input to chat with an unknown entity, then guessed whether they had been chatting with a human or a machine. Eugene Goostman fooled more than half of the human raters into thinking they had been talking to a human being.*

*This milestone is only the start of an age where we will frequently interact with machines as if they are humans; whether in customer service or sales. While humans are limited in the attention and kindness that they can expend on another person, artificial bots can channel virtually unlimited resources into building relationships.*

*Even though not many of us are aware of this, we are already witnesses to how machines can trigger the reward centres in the human brain. Just look at click-bait headlines and video games. These headlines are often optimized with A/B testing, a rudimentary form of algorithmic optimization for content to capture our attention. This and other methods are used to make numerous video and mobile games become addictive. Tech addiction is the new frontier of human dependency.*

*On the other hand, maybe we can think of a different use for software, which has already become effective at directing human attention and triggering certain actions. When used right, this could evolve into an opportunity to nudge society towards more beneficial behavior. However, in the wrong hands it could prove detrimental.*

#### **4. Artificial stupidity. How can we guard against mistakes?**

*Intelligence comes from learning, whether you're human or machine. Systems usually have a training phase in which they "learn" to detect the right patterns and act according to their input. Once a system is fully trained, it can then go into test phase, where it is hit with more examples and we see how it performs.*

*Obviously, the training phase cannot cover all possible examples that a system may deal with in the real world. These systems can be fooled in ways that humans wouldn't be. For example, random dot patterns can lead a machine to "see" things that aren't there. If we rely on AI to bring us into a new world of labour, security and efficiency, we need to ensure that the machine performs as planned, and that people can't overpower it to use it for their own ends.*

#### **5. Racist robots. How do we eliminate AI bias?**

*Though artificial intelligence is capable of a speed and capacity of processing that's far beyond that of humans, it cannot always be trusted to be fair and neutral. Google and its parent company Alphabet are one of the leaders when it comes to artificial intelligence, as seen in Google's Photos service, where AI is used to identify people, objects and scenes. But it can go wrong, such as when a camera missed the mark on racial sensitivity, or when a software used to predict future criminals showed bias against black people.*

*We shouldn't forget that AI systems are created by humans, who can be biased and judgemental. Once again, if used right, or if used by those who strive for social progress, artificial intelligence can become a catalyst for positive change.*

#### **6. Security. How do we keep AI safe from adversaries?**

*The more powerful a technology becomes, the more can it be used for nefarious reasons as well as good. This applies not only to robots produced to replace human soldiers, or autonomous weapons, but to AI systems that can cause damage if used maliciously. Because these fights won't be fought on the battleground only, cybersecurity will become even more important. After all, we're dealing with a system that is faster and more capable than us by orders of magnitude.*

#### **7. Evil genies. How do we protect against unintended consequences?**

*It's not just adversaries we have to worry about. What if artificial intelligence itself turned against us? This doesn't mean by turning "evil" in the way a human might, or the way AI disasters are depicted in Hollywood movies. Rather, we can imagine an advanced AI system as a "genie in a bottle" that can fulfill wishes, but with terrible unforeseen consequences.*

*In the case of a machine, there is unlikely to be malice at play, only a lack of understanding of the full context in which the wish was made. Imagine an AI system that is asked to eradicate cancer in the world. After a lot of computing, it spits out a formula that does, in fact, bring about the end of cancer – by killing everyone on the planet. The computer would have achieved its goal of "no more cancer" very efficiently, but not in the way humans intended it.*

### **8. Singularity. How do we stay in control of a complex intelligent system?**

*The reason humans are on top of the food chain is not down to sharp teeth or strong muscles. Human dominance is almost entirely due to our ingenuity and intelligence. We can get the better of bigger, faster, stronger animals because we can create and use tools to control them: both physical tools such as cages and weapons, and cognitive tools like training and conditioning.*

*This poses a serious question about artificial intelligence: will it, one day, have the same advantage over us? We can't rely on just "pulling the plug" either, because a sufficiently advanced machine may anticipate this move and defend itself. This is what some call the "singularity": the point in time when human beings are no longer the most intelligent beings on earth.*

### **9. Robot rights. How do we define the humane treatment of AI?**

*While neuroscientists are still working on unlocking the secrets of conscious experience, we understand more about the basic mechanisms of reward and aversion. We share these mechanisms with even simple animals. In a way, we are building similar mechanisms of reward and aversion in systems of artificial intelligence. For example, reinforcement learning is similar to training a dog: improved performance is reinforced with a virtual reward.*

*Right now, these systems are fairly superficial, but they are becoming more complex and life-like. Could we consider a system to be suffering when its reward functions give it negative input? What's more, so-called genetic algorithms work by creating many instances of a system at once, of which only the most successful "survive" and combine to form the next generation of instances. This happens over many generations and is a way of improving a system. The unsuccessful instances are deleted. At what point might we consider genetic algorithms a form of mass murder?*

*Once we consider machines as entities that can perceive, feel and act, it's not a huge leap to ponder their legal status. Should they be treated like animals of comparable intelligence? Will we consider the suffering of "feeling" machines?*

*Some ethical questions are about mitigating suffering, some about risking negative outcomes. While we consider these risks, we should also keep in mind that, on the whole, this technological progress means better lives for everyone. Artificial intelligence has vast potential, and its responsible implementation is up to us.*

Source: weforum

## **Topic 2**

### **UP accounted for maximum crimes in India in 2016: NCRB**

#### **Relevancy:**

- GS Prelims, GS Mains paper I, II and IV, Optional: Public Administration, Sociology.
- Crime in India, NCRB (National Crime Record Bureau) report.
- The purpose of this document is not to retain all the facts and figure. The candidate is expected only to have an overview of the data.

#### **Recently:**

- Union Home Minister Rajnath Singh released the annual publication of the NCRB, "Crime in India 2016".
- This year, the NCRB has published data on 19 metropolitan cities with a population of over 2 million and has included a new section on Missing Persons & Children, Fake Indian Currency Notes (FICN), and arms seizure.

#### **Latest Statistics**

- **UP Ttops the list:** Uttar Pradesh accounts for 9.5% of total crimes reported in the country, while Delhi reported the highest crime rate (974.9) for the year 2016, according to the National Crime Records Bureau.
- **Crime against women:** Cases under 'Crime Against Women' have reported an increase of 2.9% in 2016 over 2015.
  - Majority of cases under crimes against women were reported under 'Cruelty by Husband or His Relatives' (32.6%) followed by 'Assault on Women with Intent to Outrage her Modesty' (25.0%), 'Kidnaping & Abduction of Women' (19.0%) and 'Rape' (11.5%).
  - Rape cases have reported an increase of 12.4% from 34,651 cases in 2015 to 38,947 in 2016. Madhya Pradesh and Uttar Pradesh reported the highest incidence of Rape with 4,882 cases (12.5%) and 4,816 (12.4%) followed by Maharashtra 4,189 (10.7%) during 2016.
- **Atrocities against the SC/ST:** They have increased by 5.5% in 2016. Of the 40,801 cases registered, 10,426 were reported from Uttar Pradesh accounting for 25.6% followed by Bihar with 14% (5,701) and Rajasthan with 12.6% (5,134) during 2016.

- **Cyber Crime:** The incidence of 'Cyber Crime' has shown an increasing trend by 6.3%.
- **Human Trafficking:** A total of 8,132 cases of Human Trafficking were reported in the country with West Bengal reporting the highest number of cases (3,579).
- **Fake Currency:** A total of 2,81,839 notes worth Rs. 15,92,50,181 were seized under Fake Indian Currency Notes (FICN) during 2016.
- **Sedition:** There were 35 cases of sedition in 2016, highest from Haryana (12), followed by Uttar Pradesh (6). Sixty cases of custodial deaths were reported in 2016 with 12 from Maharashtra.

## Topic 3

### ***The Transgender Persons (Protection of Rights) Bill, 2016***

#### Relevancy:

- GS Prelims, GS Mains paper I, II, Optional: Sociology, Public Administration
- Polity and governance, legislation, rights of minorities, Transgender Persons (Protection of Rights) Bill, 2016

#### Recently:

- The transgender community and its supporters have agitated against the decision of the Ministry of Social Justice and Empowerment to re-introduce the original Transgender Persons (Protection of Rights) Bill, 2016 in the winter session of Parliament.

#### Why the uproar against the Bill?

- The Supreme Court's landmark decision in NALSA v. Union of India recognised that transgender persons have fundamental rights.
- The judgment was followed by a private member's Bill, the Rights of Transgender Persons Bill, 2014, which was unanimously passed in the Rajya Sabha.
- Instead of introducing it in the Lok Sabha, the Ministry uploaded its own Bill, the Rights of Transgender Persons Bill, 2015, on its website in December for public comments.
- The 2015 Bill, which was largely based on the 2014 Bill, did away with the national and State commissions for transgender persons and transgender rights courts.
- The Bill was fairly progressive since it granted a transgender person the right to be identified as a 'man', 'woman' or 'transgender'. However, the 2016 Bill, that was finally introduced in the Lok Sabha, was a highly diluted version of the previous 2015 Bill.
- It also pathologised transgender persons by defining them as "partly female or male; or a combination of female and male; or neither female nor male".

- Met with backlash, the Ministry set up an expert standing committee on social justice and empowerment to examine the Bill.

## Standing committee report:

The standing committee in its report released on July 22, 2017 made the following recommendations:

1. It criticised the 2016 Bill for its stark deficiencies and recommended re-drafting the definition of a 'transgender person' to make it inclusive and accurate;
2. providing for the definition of discrimination and setting up a grievance redress mechanism to address cases of discrimination; and
3. granting reservations to transgender persons.
4. The law must grant equal civil rights to transgender persons (marriage, divorce and adoption), thus opening the door for the legal system to take steps to undo its oppressive **heteronormative** (the presumption that heterosexuality is the norm) and **cisgendered** (the presumption that people's gender identity is aligned with their anatomical sex) foundation.

The Ministry's decision to re-introduce the 2016 Bill disregards the pre-legislative consultative policy which requires Ministries to grant a minimum of 30 days for public comments and to place a summary of feedback/comments received from the public/other stakeholders on their website.

## Features of the 2016 Bill

- **Definition:** According to the Bill, a transgender person as one who is partly female or male; or a combination of female and male; or neither female nor male.
- In addition, the person's gender must not match the gender assigned at birth, and includes trans-men, trans-women, persons with intersex variations and gender-queers.
- **Identity of a transgender:** A transgender person must obtain a certificate of identity as proof of recognition of identity as a transgender person s to invoke rights under the Bill.
- Such a certificate would be granted by the District Magistrate on the recommendation of a Screening Committee.
- The Committee would comprise a medical officer, a psychologist or psychiatrist, a district welfare officer, a government official, and a transgender person.
- **Welfare of transgender:** The Bill prohibits discrimination against a transgender person in areas such as education, employment, and healthcare.



- It directs the central and state governments to provide welfare schemes in these areas.
- **Punishment for offenders:** Offences like compelling a transgender person to beg, denial of access to a public place, physical and sexual abuse, etc. would attract up to two years' imprisonment and a fine.

### **Detailed analysis of the issues relate to the Bill:**

#### **1. Identifying a transgender:**

- The Supreme Court has held that the right to self-identification of gender is part of the right to dignity and autonomy under Article 21 of the Constitution.
- However, objective criteria may be required to determine one's gender in order to be eligible for entitlements.
- The Bill states that a person recognised as 'transgender' would have the right to 'self-perceived' gender identity.
- However, it does not provide for the enforcement of such a right.
- A District Screening Committee would issue a certificate of identity to recognise transgender persons.

#### **2. Definition of transgender:**

- The definition of 'transgender persons' in the Bill is not as per the definitions recognised by international bodies and experts in India.
- The Bill includes terms like 'trans-men', 'trans-women', persons with 'intersex variations' and 'gender-queers' in its definition of transgender persons. However, these terms have not been defined.

#### **3. Not in-sync with other laws:**

- Certain criminal and personal laws that are currently in force only recognise the genders of 'man' and 'woman'.
- It is unclear how such laws would apply to transgender persons who may not identify with either of the two genders.

## Topic 4

### *Issue of Net Neutrality*

#### Relevancy:

- GS Prelims, GS Mains Paper II and III
- Sci-tech, Internet regulations, issue of net neutrality

#### Recently:

- This week, the Federal Communications Commission- FCC (the US federal communications regulator), announced plans to change the rules governing the Internet that were issued by the Obama administration two years ago.
- The move will confer sweeping powers to Internet Service Providers (ISPs) over the content consumers can access.
- Ajit V Pai, Chairman FCC, appointed by President Donald Trump this January, spearheaded the change in what is broadly referred to as “**Net neutrality**” rules.
- Pai’s draft ruling will be voted on at the FCC’s Open Meeting on December 14 and will almost certainly be cleared.

#### What is Net neutrality?

- It is the concept of content and application providers (ISPs) being treated equally by telecom operators.
- Consumers get access to all websites, nothing is blocked, and speed of access is not differentiated.
- Just as a phone company does not get to decide who a person can call and what she can say on that call, the ISP is not expected to dictate the content that a consumer views or posts online.
- FCC rules issued in 2015 aimed at upholding this broad principle of neutrality, giving consumers equal access to Web content, and barring broadband providers from blocking or slowing access to content, or charging consumers more for certain types of content.

#### What changes does the FCC want?

- Chairman Pai's draft ruling, 'Restoring Internet Freedom Order', seeks to remove what it calls "heavy-handed Internet regulations".
- According to the plan, FCC will no longer regulate ISPs; the job of protecting consumers from detrimental business practices will go to the Federal Trade Commission (FTC), which has neither any telecommunications' expertise nor rule-making regulatory authority in this area.
- The proposal steps away from the responsibility to protect telecom consumers by redefining Internet service as not being a telecommunications service.

### **What do critics say?**

- Net Neutrality will allow ISPs to give preferential treatment to particular sites and apps, and to their own digital content.
- ISPs could block access to sites and change network speeds by segregating the Internet into fast and slow lanes.
- It will slow down competitor's content, block unfavourable political opinions, and charge consumers extra for a faster Net experience.

### **What is Pai's justification?**

- He said that he intends to remove those rules that are holding back investment, innovation, and job creation in telecommunications.

### **Who benefits from the order?**

- Big ISPs such as AT&T Inc, Comcast Corp and Verizon Communications Inc, who want a repeal of the Obama-era rules.

## Topic 5

### *The Delhi air crisis*

#### Relevancy:

- GS Prelims, GS Mains paper I, III, Optional- Geography (Climatology)
- Sci-tech, environment, SAFAR report, dust storms, disaster management.

#### Recently:

- On November 16, the System of Air Quality and Weather Forecasting and Research (SAFAR) under the Ministry of Earth Sciences published a 'Scientific Assessment of Delhi Winter Air Quality Crisis' for the period November 6-16.

#### What's in the report?

- The report listed two "Extreme" events behind the smog that had smothered Delhi and its neighbourhood.
- "Extreme 1", the SAFAR report said, was a "large multi-day dust storm that emerged at Iraq, Kuwait and Saudi Arabia in the last week of October 2017 and continued up to November 3-4".
- "Extreme 2" was the much discussed stubble burning in Punjab, Haryana and Uttar Pradesh.
- The pollution contribution of Gulf dust storm on peak day (November 8) was around 40% and from stubble burning was 25%".

#### The Storm from 3000 kms west of Delhi:

- The NASA satellite captured images of a massive dust and sand storm over Saudi Arabia and Iraq which on 29<sup>th</sup> october.
- Also on October 31, Moderate Resolution Imaging Spectroradiometers (MODIS) aboard NASA's Terra and Aqua satellites released images of skies over Saudi Arabia and Iraq from two days earlier, red with a thick blanket of dust.
- Other conditions added to the crisis, including stubble burning, which, the SAFAR report said, was "very high on November 6, and upper air winds became north-westerly (towards Delhi) with high speed and started pumping pollution in Delhi".

- These conditions started to change from November 11 onward because “there was no pumping and influence of stubble burning and Gulf storm dust after November 10 night onwards owing to slowing down of upper air winds and change in wind direction”.
- According to the SAFAR assessment, the late October-early November dust storm “was carried (eastward) by relatively cool winds”.

**Reasons behind the storms:**

- While dust storms are normal in Iran, Saudi Arabia, Kuwait and Iraq, there has been an increase in their frequency, especially in Iraq, over the past decade.
- The reasons include:
  - Advancing desertification
  - Climate change— which led to drastic changes in annual rainfall and temperature —
  - Drought
  - Mismanagement of water
- Abandonment of agricultural lands

## Topic 6

### OBC quota

#### Relevancy:

- GS Prelims, GS Mains paper I, II, Optionals- Sociology, Public administration.
- Polity and governance, Social issues, NCBC (National Commission on Backward Classes) Other backward classes, quota.

#### Recently:

- Last week, the government had given indications of its plans to reintroduce the legislation to grant Constitutional status to the **National Commission for Backward Classes (NCBC)** in the forthcoming winter session of Parliament.
- The **Constitution (One Hundred and Twenty-Third Amendment) Bill, 2017** aims to bring the NCBC on par with the National Commission for Scheduled Castes (NCSC) and the National Commission for Scheduled Tribes (NCST).

#### About NCBC:

- National Commission for Backward Classes is a statutory body established under the Ministry of Social Justice and Empowerment (Since 1993).
- It was constituted in compliance to the provisions of the National Commission for Backward Classes Act, 1993.
- In 2017, a bill seeking to grant constitutional status to the National Commission for Backward Classes is scheduled to be passed in Parliament.
- The bill was passed by Lok Sabha in April 2017 and is waiting for the assent of the Rajya Sabha.

#### What will be the effects of this amendment?

- **Strengthening the NCBC:** It will empower the NCBC, which presently has a mere recommendatory role for the inclusion or exclusion of a community in the Central list of OBCs.
- **Powers of civil court to the NCBC:** The Bill gives the NCBC powers on par with the civil court to issue summons, examine all matters regarding the welfare and development of OBCs and investigate complaints pertaining to the non-implementation of reservations in jobs and educational institutions.

- **Powers to the parliament:** Significantly, the Bill also gives Parliament the final authority to make changes to the OBC list. This means that Parliament now gets to decide on the issues pertaining to the Jats, Marathas, Patels, Kapus and the many other communities that have been agitating for inclusion in the OBC list.
- **Increasing income limit of creamy layer:** Alongside, an important definition involved has changed. In August this year, the union cabinet approved a proposal to revise the income definition of the 'creamy layer' within OBCs, increasing this from those earning Rs 6 lakh per annum to Rs 8 lakh per annum.

#### **What is creamy layer?**

- As per the Supreme Court directive in the Indra **Sawhney And Others vs Union of India case** (1993), the creamy layer of OBCs (relatively better-off sections) cannot benefit from reservation facilities.
- Then, the income criteria was Rs 1 lakh per year. The last time the 'creamy layer' ceiling was revised, from Rs 4.5 lakh in 2008, was in 2013, when it was increased to Rs 6 lakh per annum.

#### **Commission to examine the central OBC list:**

- In August 2017, the cabinet approved the setting up of a Commission to examine the sub-categorisation of 5,000-odd castes in the central OBC list.