CHAPTER

10

Ethics and AI

AI, especially ANI (Artificial Narrow Intelligence), is beginning to occupy a unique place in our lives. There are virtual personal assistants, hosts of intelligent apps on our smart phones, self-driving cars, intelligent home appliances, smart buildings, smart homes, and many more. No other technology has the kind of far-reaching implications that AI has, across so many industries and all spheres of our lives. This is what makes it so fascinating; it is truly incredible and no less than a sci-fi plotline.

For a technology designed to mimic human-like intelligence, understanding the darker side of AI is also equally important and probably requires defining the dos and don'ts at a policy level. There are many questions we need to address, if we want to truly leverage the transformative power of AI.

In the next 20 years, experts predict that AI will continue to make great strides in transforming and in many cases eliminating hosts of human tasks. If this innovation is done in an ethical way and keeping human-machine collaboration as the goal, we can build a future in which humans are not competing with machines. Instead, we will be entering into a new era of jobs that requires no manual effort but higher emotional intelligence.

It is perhaps a way of life now to hail a taxi at the tap of a smart phone screen. Imagine the same thing, extended to getting a self-driven taxi! Imagine the extent of woes that the traffic management officials will have to go through to manage transportation and safety to accommodate the next wave of driverless cars. For sure, the department of transportation has their work cut out for them; they will have to innovate too.

S. Mohanty and S. Vyas, *How to Compete in the Age of Artificial Intelligence*,
https://doi.org/10.1007/978-1-4842-3808-0_10

This is just one of the example that shows how the implications of AI are unprecedented. As we continue to innovate and solve new problems using AI, we must simultaneously pay attention to ethical usage and policy related aspects because there are implications to the society at large, not just a few industries or a few individuals. We must take a democratic approach to the future of technology led disruptions.

Ethical Issues in Artificial Intelligence

AI, in many ways, is pervasive and provocative at the same time. We need guidance and frameworks to ensure ethical usage of AI, and this is the right time in the AI evolution phase to consider ethics in relation to AI seriously.

The ability to be introspective has made us what we are. However, it seems that we are in a hurry to outsource this to algorithms. There are grave consequences for such an approach. We are pleased to see how automation is improving the quality of our lives, but we have not sat down and created a list of activities that we would never delegate. We have not even crafted guiding principles for clear demarcation of responsibilities between humans and machines. These things are important. In the absence of values-based standards and guidelines, the biases of AI manufacturers will take over and dictate the terms and conditions.

Ethics is not about saying, no you can't do this. The point is to ask what are the goals you are trying to achieve and how can you attain these goals without infringing on the cultural values we hold dear. You can do things that are entirely legal yet highly unethical. The Socratic process of always asking for contrarian views is not always fun, but basking in the glory of AI achievements and not paying attention to ethical standards is no longer an acceptable stance.

The following sections contain a list of topics that are subject to raging debates in various forums. The camps are still divided.

Unemployment

What happens when all manual jobs are replaced by AI-assisted machines?

Cultural aspects such as position in society and consequently the division of labor are governed by the magnitude of automation. In the pre-industrial age, one's ability to do physical work determined the paycheck. Gradually we evolved and invented ways to automate jobs, not eliminating the physical work completely, but augmenting human skills with machines so that we could do more with less. The result? Newer jobs like machine operators came into existence and slowly the skills orientation moved from physical labor to the cognitive labor, putting more and more emphasis on judgment-related skills.

For example, trucking as an industry currently employs millions of individuals across the globe. What will happen to trucking jobs if self-driving trucks become a reality in the coming years? On the one hand, self-driving trucks might lower risk of accidents, won't show signs of fatigue, will cover the distance in a predictable way all the time, reduce the cost of operations, and deliver significant efficiency gains. Hence, self-driving trucks seem like an ethical choice. The same scenario could happen to the majority of the workforces across other industries and sectors.

The ethical question before us is, when more and more jobs are automated, what will we do with all the time on our hands? Our current employment contracts are based on one fundamental factor—we sell our time to earn enough to sustain ourselves and our families. So, if the prospect of time-based compensation goes away, we need to find newer ways to earn money.

Inequality

How can we reward the machines?

The majority of our current compensation frameworks are based on hourly wages. There are few notable changes where risk and reward mechanisms are in place, but in general our services always equate to a rate per hour. If AI is going to do most of our work at a fraction of a cost, companies will naturally drift toward a newer workforce mix (an increasing number of AI agents and a smaller human workforce). This means fewer humans will get compensated. However, the company owners and others who invested in the company will take home a major share of the revenues earned.

If you have not noticed, this is already happening and it's creating a widening wealth gap. Companies whose business model revolves around this algorithmic economy are employing fewer humans and sharing the wealth among the workforce from the economic surplus they are creating.

The ethical question before us is, if we're truly moving toward a no manual-job society, how do we structure a fair compensation mechanism for the workforce consisting of fewer humans and more machines?

Humanity

How will machines affect human behavior?

On the one hand, we are relishing the fact that we are letting machines think and act on our behalf. On the other hand, we are truly altering human behavior. Here are few examples that illustrate how we are being influenced:

- Websites are designed by taking into the account the minutest levels of detail that would appeal to an individual's liking.
- Recommendation engines push additional products to us by suggesting "people like you have bought these other things".
- While you are driving past a supermall, your phone vibrates with a discount just for you.
- Intelligent apps on your smart phone suggest which route to take to get to your destination faster.

The right information at the right time through the right channel is all good, but the side effect is that we have stopped thinking about anything anymore.

The ethical question before us is this—first, by allowing AI agents to think and act on our behalf, are we progressing toward a world where we are becoming increasingly uneducated about our surroundings? Second, if the AI agents are becoming better and better at modeling human behavior, is there a possibility that the same AI agents will be used to direct human attention and trigger certain actions that are detrimental to the very existence of the human race?

Artificial Stupidity

How do we ensure that machines do not become biased?

Humans develop cognitive capabilities by learning from their environment; machines also go through a similar learning phase to acquire intelligence. Humans learn bad things if they are exposed to bad environments; machine learning goes through a similar risk if the data is incomplete or purposefully distorted.

Machines need to learn continuously and need to be exposed to a wide variety of data sets in order to be prepared to handle real-world occurrences. Just relying on training data sets that are curated based on the data that's available won't suffice if we want machines to take on real-world challenges.

Humans are not always fair and neutral. Machines also can exhibit similar unfair and irrational behavior. AI systems are created by humans, thus there is a high likelihood that humans will introduce judgmental bias into the very machines they build. Bias can also creep into machine behavior in many different ways—data bias and design bias are the most prominent ones.

The ethical challenge before us is, if our future is going to be completely dependent on AI systems, we need to ensure that the machines perform as expected and aren't biased.

Security

How do we keep AI safe from evil intentions?

The more powerful a technology becomes, the more possibilities open up for malicious intent. Autonomous systems require a greater responsibility of making them secure, and it is not just adversaries we need to worry about. What if AI agents become so focused on achieving their goals, that they recommend and implement things that may bring disastrous consequences for us? For example, what if the goal of an AI system is to find solutions for cancer, and after careful considerations of numerous diagnosis results, root causes, treatment plans, and effectiveness of medicines, it realizes that the most effective and best way of solving the cancer problem is to kill everybody on the planet? From a machine's point of view, it has found the solution. From human point of view, it is catastrophic.

The ethical question before us is, how do we ensure that there are enough checks and balances in place before we start using AI systems across all spheres of our life and as ubiquitously as we want to?

Singularity

How do we manage artificial super intelligent systems?

Human evolution is almost entirely due to our intelligence and our ability to adapt to changing conditions. However, in our zest to invent more and more artificially super intelligent systems, we may get into a scenario where the machines are the most intelligent beings on earth, far superior than humans! This state is called "singularity".

The ethical question before us is, even if it is far fetching, someday a sufficiently advanced machine may come to life, and this machine will be able to anticipate what we are anticipating, so how do we stay a step ahead?

Machine Rights

How do we define a legal framework for AI?

We have seen how mechanisms of reward and aversion play a decisive factor in human lives. Reinforcement learning in particular applies a virtual risk and reward mechanism to let AI agents learn and become adaptive to the environment.

The ethical question before us is, once machines as entities attain sufficient maturity levels to see, sense, think, and act autonomously, they will demand a legal framework to protect and manage their share of rights. Should the intelligent machines be treated like humans? Should they have a parallel legal and grievances system? When an AI agent is penalized for wrong doing, how

do we provide a support system to make it understand its mistakes and stay motivated to do well in the future? How do we resolve situations where the machine is doing all the right things but the human-in-the-loop is in the way?

The Board and CEO's Roles in Ethical AI

AI's predictive measures are meant to provide intended outcomes. Out in the real world, this is the only side that matters. However, just as humans are imperfect, subjective, and prone to corruption, machines are too. In other words, despite established laws and protocols, dubious actors in our society find a way to game the system. Similarly, AI's usefulness is open for people to use it for malicious or self-serving purposes. The stakes are getting higher.

Hence, companies, especially those with access to vast pools of data and that develop and deploy AI products and services, should not only demand more responsible work ethic from engineers, entrepreneurs, and executives, but should also establish more assertive boards and committees to make ethical AI their top strategic priority.

The Board's Role

Companies in traditional industries turn to decades of laws, regulations, and litigation for guidance while defining corporate ethics. However, for something as transformative and disruptive as AI, there is no prior knowledge. Norms and standards are still emerging; laws, regulations, and legal precedent are scarce. That is why it is important for boards to become aggressive proponents of corporate ethics for AI, making it a top priority alongside other concerns such as growth, profitability, M&A targets, and succession planning.

Boards should hold CEOs accountable to making AI ethics an enterprise priority. For example, directors should not take it on face value that the company's AI offerings and initiatives do not inadvertently promote biases. Instead, board members should stipulate that AI offerings and initiatives should be continuously be subjected to stress tests to uncover AI biases and take remedial actions.

The CEO's Role: Governance Excellence on AI Ethics

CEOs can build their AI ethics and governance prudence in three ways.

First, CEOs should strengthen and expand their own knowledge, covering issues ranging from AI bias to best practice in transparency and accountability. They should also be deeply aware of the implications of their company's AI products and services, including data sources and nature/type of algorithms.

CEOs should also set up a dedicated think-tank to collect industry wide information about AI ethics-related litigation and concerns raised by customers and views of policymakers. This will help CEOs conduct management briefings across the enterprise and raise awareness levels.

Secondly, CEOs should include senior advisors into their think-tank who can bring in additional perspectives regarding implications of AI to the boardroom. For example, if your company is offering HR-related AI products and services, it is crucial to bring in a senior advisor to critically review whether biases related to diversity, inclusion, labor, and civil rights regulations are adequately marginalized.

Third, CEOs must realize that ethical AI can't be an afterthought. Enough checks and balances should be built right into the core of AI systems development process to ensure that technologists don't get carried away in developing AI systems without paying attention to making AI transparent. We cannot ignore the risks of deploying AI systems that work exceptionally well on the one hand, but are not fully explainable on the other. Hence, CEOs must ensure that the enterprise is equipped with the right tools, methodologies, and review mechanisms.

The Technologist's Role

Technologists and AI experts must pay attention to the purpose and goals of the AI system. It is easier to say that the algorithm is highly sophisticated but the data is biased. Real-world data will be always biased, because it is situational. It is the responsibility of AI experts to fully understand the context behind the data and then design algorithms to support human values.

Conclusion

It is important to decouple the pursuit of fairness in AI from commercial interests and create room for fair and transparent AI solutions. To that end, making AI explainable is of paramount importance.

Companies building their business strategy based on AI's disruptive power are caught midway: What are the costs and benefits of making their AI transparent? How do they protect their IP and competitive advantage?

Admittedly, there are rules of thumb or best practices right now. Perhaps, for every AI system, if we can detail the kind of data used, details of data sources, parameters used in the prediction model development, accuracy levels achieved, and statistics about false-positive and false-negatives, we may achieve a generally agreeable transparency level that may further lead to establishing a commonly acceptable set of standards for ethical AI.

In this chapter, we discussed several ethical related viewpoints of the consequences of AI. In the next chapter, we bring together all the lessons from previous chapters and try to establish certain approaches to building a human-machine collaborative ecosystem.

References

1. https://80000hours.org/articles/ai-policy-guide/
2. https://medium.com/artificial-intelligence-policy-laws-and-ethics/the-ai-landscape-ea8a8b3c3d5d
3. https://www.cs.ox.ac.uk/efai/category/codes-of-ethics/
4. https://www.wired.com/story/ai-research-is-in-desperate-need-of-an-ethical-watchdog/
5. https://mashable.com/2015/10/03/ethics-artificial-intelligence/#kyYeZZHNMsqD
6. https://techcrunch.com/2017/01/22/ethics-the-next-frontier-for-artificial-intelligence/
7. http://www.zdnet.com/article/artificial-intelligence-legal-ethical-and-policy-issues/
8. https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/
9. https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/