Responsible AI

Responsible AI is an emerging area of AI governance and use of the word "responsible" is an umbrella term that covers both ethics and democratization. Now that software programs with artificial intelligence (AI) features are becoming more common, it is increasingly apparent that there is a need for standards in AI beyond those established. The technology can be misused accidentally (or on purpose) for a number of reasons.

AI failures

They are instances where AI has failed :

-Amazon Axes their AI for Recruitment Because Their Engineers Trained It to be Misogynistic

-Amazon's Facial Recognition Software Matches 28 U.S. Congresspeople with Criminal Mugshots

-Apple's Face ID Defeated by a 3D Mask

- Microsoft's AI Chatbot Corrupted by Twitter Trolls

-Uber Self Driving Car Fatality

Implications of AI failures

Increasingly, AI is used to help support major decisions, such as who receives a loan, the length of a jail sentence, and who gets health care first. The hope is that AIs can make decisions more impartially than people often have, but much research has found that biases embedded in the data on which these AIs are trained can result in automated discrimination en masse, posing immense risks to society.

In 2016, a Tesla Model S car on autopilot collided with a truck that was turning left in front of it in northern Florida, killing its driver— the automated driving system's first reported fatality. According to Tesla's official blog, neither the autopilot system nor the driver "noticed the white side of the tractor trailer against a brightly lit sky, so the brake was not applied."One potential way Tesla, Uber, and other companies may avoid such disasters is for their cars to do a better job at calculating and dealing with uncertainty.

Responsibility of organizations

1. Developing AI is a multidisciplinary problem. AI challenges and products can be technical or based on human behavior, and often are a blend of the two.
2. An AI application affects more than just end-users. Input from stakeholders is essential to helping us structure the AI's objectives to increase adoption and reduce potential undesired consequences.
3. Our assumptions shape AI. There is no such thing as a neutral, impartial, or unbiased AI. Our underlying assumptions about the data, model, user behaviors,

and environment affect the AI's objectives and outcomes. We should remember that those assumptions stem from our own, often subconscious, social values, and that an AI system can unintentionally replicate and encode those values into practice when the AI is deployed.

4. Documentation can be a key tool in reducing future failures. When we make a good product, end-users and consumers will want to use it, and other AI developers may want to repurpose it for their own domains.

5. Accountability must be tied to an AI's impact. When using the data or AI could cause financial, psychological, physical, or other harm, we must consider if AI offers the best solution to a given problem.