Mariam Adegbindin
Buff ID: 1133249
Assignment: Ethics for Data science

There exist three cases with which the technology industry should acquaint itself. The initial scenario pertains to the phenomenon of feedback loops, specifically focusing on the detrimental effects observed in the context of YouTube, as discussed in the article titled "The Toxic Potential of YouTube Feedback Loops." Feedback loops might manifest when the model assumes control over the subsequent iteration of acquired data. The expeditiously retrieved data is prone to being compromised by inherent flaws within the software. YouTube has faced criticism for its alleged promotion of numerous conspiracy theories, resulting in significant backlash. The individuals in question have faced allegations of endorsing paedophilia by selecting inappropriate content from innocent home videos. The intended meaning of their message was not aligned with the interpretation it received.

The second case pertains to a software application utilised for the assessment of healthcare benefits for individuals of low socio-economic status throughout more than fifty states. A software malfunction occurred in the state of Arkansas, resulting in erroneous termination of coverage for those affected by cerebral palsy and diabetes. A prevalent concern observed was the absence of mechanisms to identify and rectify errors in the installed systems.

The final case study pertains to Latanya Sweeney, who currently holds the position of Dean at the Privacy Lab at Harvard University. The individual discovered that upon conducting a Google search of her name, one of the initial search results prominently displays the phrase "Latanya Sweeney Arrested." This particular search result suggests the existence of a criminal record, which is factually inaccurate. Conversely, when conducting a similar search for other individuals, such as Kirsten Lindquist, the search results differ significantly, presenting less incriminating information and omitting any indication of her having been arrested on three separate occasions. Latanya Sweeney, a researcher with a PhD in Computer Science, conducted an extensive analysis by examining a sample of over 2000 names. Her findings revealed that individuals with African American names are more likely to encounter search results that imply a criminal record when their names are searched on Google, in contrast to white individuals who tend to receive neutral search results. This phenomenon gives rise to the notion that advertising systems exhibit discriminatory tendencies based on race and gender, which is a subject of significant concern.

What is the significance of this matter? The utilisation of technology can have detrimental consequences, as exemplified by the historical association between IBM and Adolf Hitler during the 1930s. IBM supplied computers to the Nazi regime for utilisation in coding operations at concentration camps. Technology has the potential to be utilised in manners that deviate from the original intentions of its creators, as seen by the application of facial recognition technology. Numerous instances of technological misuse might be observed. Unintended consequences may arise from the utilisation or misuse of technology by individuals engaging in trolling and harassment, authoritarian regimes, or for the dissemination of propaganda and disinformation. This highlights the absence of ethical considerations in the aforementioned situations.

Ethics is a scholarly field concerned with discerning and evaluating concepts of goodness and badness, encompassing a framework of moral principles. Ethics can be distinguished from religion, law, social conventions, and emotions, as it does not adhere to a rigid framework of regulations. Ethics refers to a set of established principles and values that serve as a solid foundation for determining the distinction between right and wrong, and afterwards provide guidance on the appropriate actions that individuals should undertake. The examination and advancement of an individual's moral principles. Recourse and accountability can be considered as components within the broader domain of ethics, as exemplified by the instance of a system flaw. No individual, including the algorithm's author, assumed responsibility upon the discovery of the defect. Instead, the blame was shifted to the policy makers, who in turn attributed fault to the software engineers. Consequently, there was a collective effort to evade accountability for the incident.

The utilisation of feedback loops and metrics might have certain drawbacks. Placing excessive emphasis on metrics can result in many issues, such as manipulation, gaming, a narrow concentration on immediate objectives, and unanticipated adverse outcomes, among other potential effects. A significant focus within the field of Artificial Intelligence and Machine Learning revolves around the optimisation of metrics. In the context of the United Kingdom's healthcare system, an attempt was made to prioritise the assessment of pertinent factors by implementing targets pertaining to emergency room (ER) wait times. However, this approach resulted in the cancellation of pre-scheduled operations in order to allocate additional personnel to the ER. Consequently, patients were compelled to endure prolonged waiting periods within ambulance queues, while the utilisation of hallways as makeshift accommodations necessitated the conversion of stretchers into beds. The concept of "Blitzscaling" posits that once a company achieves significant growth, it will thereafter generate profits. The primary emphasis is placed on velocity rather than effectiveness, which poses the possibility of encountering a highly unfavourable outcome. There are several drawbacks associated with this phenomenon. Firstly, the attainment of rapid and exponential development necessitates the implementation of automation and a heavy reliance on measurements. Secondly, the prioritisation of speed above all other considerations leaves little room for ethical reflection. Lastly, it is important to acknowledge that challenges and issues tend to arise on a significant scale while pursuing such accelerated growth.

Bias can arise from various sources. Representation bias is a form of bias that pertains to the portrayal of black women with regards to their skin colour in various products, such as IBM and Microsoft. Historical bias is a significant concern in the first stage of the data collection process and can persist despite the implementation of flawless sampling techniques and feature selection methods. Numerous studies have extensively examined racial bias, consistently highlighting the systemic disadvantages faced by black individuals across various contexts. For instance, research findings have demonstrated that when presented with identical patient files, clinicians exhibited a reduced likelihood of recommending cardiac catheterization for black patients. Another instance pertains to the negotiation process for purchasing a pre-owned vehicle, wherein individuals of African descent were presented with initial price offers that were $700 more compared to their counterparts, and were granted significantly lesser discounts.