Are Programmers and AI Scientists Ethically Responsible for Their Code?

As the age of technology is in full swing, society has grown reliant on the use of physical and virtual machines. Where a human once operated, a machine is now accomplishing the same work in a more efficient and cost-effective manner. Due to the rise of the Internet and evolution of data storage, the proliferation of data "has now made machine learning a significant component of modern life" (Dickey et al., 2019, p.16), steering the world as we know it towards a modernized high-tech future revolving around computer algorithms and virtual reality. But are we truly moving towards the ideal technologically enhanced utopia that we seek?

Contrarily to many top car companies' beliefs, the implementation of fully autonomous cars into society may be a far more ambitious feat than previously assumed. According to a new study conducted by the Georgia Institute of Technology, the car's machine learning algorithms are more likely to misidentify certain demographic groups than others. The utopian vision of AI such as autonomous cars is that they would be operated subjectively by systems unclouded by the prejudices and emotional factors of their human creators. However, a machine's desired objective outlook and its rational, unemotional decision-making process is precisely what becomes its worst flaw. There is still much ambiguity as to on whom does the responsibility of ensuring the equitability of the machine algorithms lies, exposing a predicament complicating the generation of a sustainable solution. Which organizations should assume the responsibility of ensuring fair and unbiased decisions of machine learning algorithms? What can be done to remedy to the issue at hand to preserve the integrity of AI decision-making processes?

Having joined a Machine Learning engineering design team, we are taught how to clean data and check for inclusivity and correctness as well as biases that may infiltrate the data. Such knowledge can be applied to the issue of AI self-driving cars. From experience dealing with datasets, while a heavier weighing in the dataset could help correct the bias to be more inclusive of minorities and outliers in the sample data, algorithms should constantly be monitored and tested, to ensure the fairness and consistency of decisions. They must maintain transparency, especially in such a situation where public safety may be involved. Doing so could allow observers to share insight and identify possible areas of discrimination that may not be apparent to manufacturers and developers.

As most clients lack basic understanding of the functioning of these algorithms, their inquires rarely touch upon the testing process itself, nor do they demand additional tests, reducing the odds of exposing any issues. As a result, dealers do not bother testing beyond their minimum set standards. The model's users should not be held accountable, as "they typically lack the expertise to evaluate [a machine learning] model" (Sammy, 2019, p.42). Furthermore, internal control principles dictate that "the person who creates a system cannot be impartial evaluators of that same system" (Sammy, 2019, p.42), suggesting that the model developer should not be held responsible for its AI's harmful decisions. After all, simply because a model is trained on biased data does not necessarily indicate that its programmer holds the same views, but rather that the sample data used in training was corrupted or incomplete. According to the Data Science Association's Professional Code of Conduct, developers have the duty to "protect the client from relying and making decisions based on bad or uncertain data quality" and "inform the client of all data science results and material facts known" (Data Science Association, 2019). Thus, once these biases become recognized, the party unwilling to rectify the problem would be held accountable for any further incidents.

Data Science Association. (2019) Data science code of professional conduct. *Data Science Association Code of Conduct*. Retrieved from http://www.datascienceassn.org/code-of-conduct.html

Dickey, G., Blake, S., Seaton, L. (2019) Machine learning in auditing. *CPA Journal (89)*(6), 16-21. Retrieved from http://web.a.ebscohost.com.ezproxy.library.ubc.ca/ehost/pdfviewer/pdfviewer?vid=22&sid=c938c9cd-ea04-42f4-ae7d-ef4d61d23753%40sessionmgr4006

Sammy, Allan. (2019). Bias in the machine. *Internal Auditor, 76*(3), 42-46. Retrieved from http://web.a.ebscohost.com.ezproxy.library.ubc.ca/ehost/pdfviewer/pdfviewer?vid=9&sid=c938c9cd-ea04-42f4-ae7d-ef4d61d23753%40sessionmgr4006