**Collaborative Learning Discussion 1 – Initial Post**

Case Study: Malicious Inputs to Content Filters

Blocker Plus is an internet content filter intended for use in institutions, such as schools and libraries, as well as domestic settings to protect children from harmful online content (ACM, N.D.). The makers employed machine learning to maintain the content blacklist based on user feedback and this resulted in manipulation of the blacklist by activists.

Whilst the premise of the system supports the ethical use of computers and intends to protect children, the maker’s implementation has proved problematic and does not align to the ACM Code of Ethics (ACM, 2018). The product is not sufficiently protected from misuse leading to discrimination and suppression of information. This reinforces the importance of principle 2.5 as powerful computing concepts, such as machine learning, should be used with due care. Whilst there is evidence that testing and risk assessment were used, unethical decisions have still resulted in complaints and the deception of stakeholders.

Another applicable code of conduct is that of The Chartered Institute for IT (BCS, 2021). Blocker Plus is intended to support the public interest, however it has failed in this through reducing access to vital public health information and impacting the digital inclusion of children. In terms of duty to the relevant authority, professional judgment and due diligence has not been properly exercised which has resulted in further poor decisions to take advantage of consumer ignorance and continue to use a vulnerable model based on user feedback.

Further to professional codes of conduct, there are also legal and social issues to be considered. This content filter is intended for institutions to comply with the US Children’s Internet Protection Act (CIPA) (ACM, N.D.). The product achieves this aim but, by restricting legitimate material, it could disadvantage children through enforcing prejudices and reducing important learning opportunities (Thoreson, 2016). Whilst legislation in this area varies by state, restricting access to this legitimate material goes against human rights guaranteed by the International Covenant on Civil and Political Rights (ICCPR).

This case study presents professional, legal and social issues that could impact both the producer of and consumers of this product. Whilst the impacts are not intentional, this could still leave the Blocker Plus organisation and the institutions using the product open to reputational damage and possible litigation on the basis of the harm caused to end consumers.

References

ACM (2018) ACM Code of Ethics and Professional Conduct. Available from: <https://www.acm.org/code-of-ethics#h-2.3-know-and-respect-existing-rules-pertaining-to-professional-work> [Accessed 25 January 2022].

ACM (N.D.) Case: Malicious Inputs to Content Filters. Available from: <https://ethics.acm.org/code-of-ethics/using-the-code/case-malicious-inputs-to-content-filters/> [Accessed 25 January 2022].

BCS (2021) Code of Conduct for BCS Members. Available from: <https://www.bcs.org/media/2211/bcs-code-of-conduct.pdf> [Accessed 25 January 2022].

Thoreson, R. (2016) Discrimination Against LGBT Youth in US Schools. Available from: <https://www.hrw.org/report/2016/12/08/walking-through-hailstorm/discrimination-against-lgbt-youth-us-schools> [Accessed 26 January 2022].