**Codes of ethics and professional codes-Case study: Malicious Inputs to Content Filters.**

**Introduction:**

The rapid advancement in computer science has made ethical considerations relevant in deployment of technology for daily use. Computing professionals have ethical obligations to clients, employers, other professionals, and public, in fulfilling their professional responsibilities and these obligations are expressed in codes of ethics, which can be used to make decisions about ethical problems (Loui.M. and Miller.k,2007).

**Definition of AI ethics in computing environments**:

According to Richard Deckard (2023), AI ethics are those principles and values that guide the development and deployment of AI technologies in a system. This means the design should be guided by principles of: Fairness, Transparency, Accountability, Privacy, Explainability and Respect for human values etc and an AI system should be built on ethical frameworks.

The understanding of the impact of AI on the society is important in developing and application of the ethical guidelines in the system bearing in mind that AI system works with information being fed into an algorithms whether the outcome results in fair judgment or portrays discrimination(Nguyen.L,2022).These technology companies ensure that AI ethics are prioritized in the decision-making process and that they are being used properly without bias consideration of cultural differences (Yilmaz.H,ND).

**Summary of the “Malicious Inputs to Content Filters” case study**.

This is a case where U.S. Children’s Internet Protection Act (CIPA) mandated public schools and libraries to deploy a content filter in their network system which will block inappropriate contents that deemed harmful to minors. An automated Internet content filter tool called Blocker Plus was deployed by the schools and libraries to meet the CIPA requirements. This tool was designed with a centrally controlled blacklist maintained by the software maker, and it provided a user-friendly interface for home use by parents.

In order for the software developer to meet up with continual updating of the blacklist, they opted for a machine learning technology to automatically identify the inappropriate contents and blacklist them. To achieve the AI algorithm goal, the developer obtained input from the home and library users to aid in the classification of contents and this was a continuous practice to collect input from users to refine their learned models. However, a user complaints review was done and there were complains on inappropriate blocking of topics such as content regarding gay and lesbian marriage, vaccination, climate change, and other topics not covered by CIPA .

A further investigation into complain list showed that there were number of activist groups that had exploited Blocker Plus’s feedback mechanism to provide input that corrupted the classification model and the Blocker Plus’s leadership easy way out was to disable accounts linked to the activist groups to avoid inappropriate input to the AI algorithm.

**Comparison of the case study professional conduct to the British Computer Society (BCS) Code of conduct:**

1. Public Interest: The developer had interest of the minor’s wellbeing by blacklisting internet contents that are inappropriate to the minors and this is inline with the CIPA requirement, professional and BCS code of conduct. Though the developer violated the BSC code of no discrimination of sexual orientation, but the professional conduct to protect the right of children suffices. Also, the blocker plus violated both professional and BCS code by not disclosing the limitation of software in terms of automated user’s input.
2. Duty to relevant authority: The Blocker plus developer accepted the professional responsibility for their work by reviewing the complains, investigating all the facts of the complains and acting by blocking the activist group. This showed good stewardship to the relevant authorities and stakeholders.
3. Duty to the profession: Blocker Plus failed to build a secured system or provide security tools to protect the system from inappropriate user’s feedback that will corrupt the machine algorithms which will end up impacting negatively to the minors. This is not in line with both professional and BCS code of conduct.

**Conclusion**

Artificial Intelligence plays a vital role in our daily lives now, and it is the developer’s responsibility to ensure that software is built under the guidance of ethics code of conduct in a computing environment.

**Reference:**

* ACM Committee on Profession Ethics*. 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 10 August 2024].
* BSC. The Chartered Institute for IT *BCS Code of Conduct*. Available from:<https://www.bcs.org/membership-and-registrations/become-a-member/bcs-code-of-conduct>[Accessed 16 August 2024].
* Deckards.R. MBCS (2023). BSC. The Chartered Institute for IT*.* What are ethics in AI? Available from: <https://www.bcs.org/articles-opinion-and-research/what-are-ethics-in-ai/> [Accessed 8th August 2024].
* Loui.M, Miller.K.(2007). Online Ethics centre For ENGINEERING AND SCIENCE *Ethics and Professional Responsibility in Computing* .Available from: <https://onlineethics.org/cases/ethics-and-professional-responsibility-computing> [Accessed 19th August 2024].
* Nguyen.L(2022).Orient. What *Is Ethical AI & Why Is It Essential for Business?* Available from: <https://www.orientsoftware.com/blog/ethics-in-ai/>[Accessed 14th August 2024]
* Yilmaz.H(N.D). Plugger.ai. *Is AI Ethic?: Ethical Concerns About AI Development* .Available from: <https://www.plugger.ai/blog/is-ai-ethic-ethical-concerns-about-ai-development> [Accessed 17th August 2024].