**Collaborative Learning Discussion 1 - Summary Post**

Several themes emerged during this discussion, including social media content filtering, the impact of filtering on equality, and proactive measures to manage artificial intelligence in an ethical manner.

Paladini (2022) draws parallels between the filtering of content by Blocker Plus and the filtering of posts on social media; both may have good intentions but can result in negative impacts on users. In addition, both utilise opaque algorithms that guide user actions without them understanding the underlying logic (Isic, 2022). In both instances, echo chambers can emerge, resulting in confirmation bias through an enforced partial view of the world (Barrett et al., 2021). In the long term, this can lead to polarization and extreme behaviour.

Manipulation of the Blocker Plus content filter led to exclusion of sensitive information which constitutes indirect discrimination under the Equality Act 2010 (Rashid, 2022; CSIE, 2018). Children, and people in general, should have access to unbiased information in order to form their own views (Chawla, 2022). Further to impacting the development of balanced world views, lack of access to certain sensitive materials, on sexual orientation for example, could impact children’s tolerance and lead to bullying due to lack of understanding.

The integrity threat to the filter should have been proactively identified by a competent IT professional (Küfner, 2022). In some instances, the risks identified in an initial assessment would mean that machine learning is used carefully or not at all (Miller, 2022). If used, there should be human oversight in place to ensure the algorithm is working as expected and to allow for early intervention when issues are detected, thus minimizing potential harm to users.

In conclusion, care should be taken when employing artificial intelligence technologies. Regardless of intention, harm can be caused to end users if algorithms are not risk assessed and competently managed.

References

Barrett, P., Hendrix, J. & Sims, J. (2021) Fueling the Fire: How Social Media Intensifies U.S. Political Polarization - And What Can Be Done About It. Available from: <https://bhr.stern.nyu.edu/polarization-report-page> [Accessed 13 February 2022].

Chawla, V. (2022) Peer Response. Available from: <https://www.my-course.co.uk/mod/hsuforum/discuss.php?d=293538> [Accessed 13 February 2022].

CSIE (2018) Equality Act 2010. Available from: <http://www.csie.org.uk/inclusion/equality-act-2010.shtml> [Accessed 13 February 2022].

Isic, D. (2022) Peer Response. Available from: <https://www.my-course.co.uk/mod/hsuforum/discuss.php?d=293538> [Accessed 13 February 2022].

Küfner, J. (2022) Peer Response. Available from: <https://www.my-course.co.uk/mod/hsuforum/discuss.php?d=293538> [Accessed 13 February 2022].

Miller, S. (2022) Peer Response. Available from: <https://www.my-course.co.uk/mod/hsuforum/discuss.php?d=293538> [Accessed 13 February 2022].

Paladini, S. (2022) Re: Initial Post. Available from: <https://www.my-course.co.uk/mod/hsuforum/discuss.php?d=293538> [Accessed 13 February 2022].

Rashid, A. (2022) Peer Response. Available from: <https://www.my-course.co.uk/mod/hsuforum/discuss.php?d=293538> [Accessed 13 February 2022].