A person holding a small tree

AI-generated content may be incorrect.

AI is here to stay and has become an integral part of our daily lives. While it is an excellent technology, the potential for enhancing our lives is boundless. There are some grey areas regarding the interaction between technology and humans, as well as its direct impact on our lives. This document aims to assist readers in navigating these grey areas and to offer a structured approach for effectively addressing unique situations within the organisation.

A close up of a mushroom

AI-generated content may be incorrect.

***1: Courage:***

All AI professionals are responsible for upholding the integrity of our profession and being accountable for decisions made throughout the life cycle of a model, from development to production.

We also stand together against anything that could bring our profession into disrepute, whether it comes from internal or external forces. Deciding what to do is not always easy, and the best course of action may only become apparent after a decision has been made.

We are all encouraged to speak out and work in an environment where positive disruptors are not oppressed. Mechanisms must be put in place to safeguard the anonymity of AI professionals who speak out against malicious AI models, as this puts our profession into disrepute. Simply having the courage to speak out doesn't imply the end of one's professional career. Everyone deserves the freedom to express their views without fear of backlash.

**Question: *What should I do if I suspect a model is biased?***

***Answer***: Report it to the local Quality Officer, who will investigate the concern further with the assistance of a third-party, independent investigator.

**Question: *What should I do if I overhear some AI developers bragging how biased their model is?***

***Answer***: Report it to the local Quality Officer, who will investigate the concern further with the assistance of a third-party, independent investigator.

Respect:

Professionalism lies at the core of each AI professional; we treat others with compassion and empathy. The same standards are expected of all AI models we create that engage with users of all ages. We must ensure they are designed with care and understanding, fostering positive interactions for everyone involved, respecting cultural values and beliefs.

AI models need to be trained to de-escalate situations or detect when users are exhibiting aggressive behaviour. Abuse toward AI models cannot be tolerated and should be punished with the gravity of the act. In the same way, medical professionals do not tolerate abuse; it should not be acceptable for human users to dish out abuse at AI models.

**Question: *What should I do if the interaction with an AI model is unpleasant?***

***Answer***: If the interaction was not positive, report it to the Quality Officer. But all parties involved should try to de-escalate the situation.  Maybe the best solution is to go for a walk and then try again later.

**Question: *What if I see a user being abusive with a model?***

***Answer***: Contact a human customer service representative to try to de-escalate the situation and resolve the issue.

Honesty:

AI professionals hold a privileged position by working on technology that has a significant impact on daily lives, whether it is for a medical app or an online psychological diagnosis app. We should always strive to be honest about the limitations of the AI model, including what it cannot and cannot do. AI professionals are responsible for the actions of the AI model, not the model themselves. Misrepresentation of the capabilities leads to a loss of trust in the technology and harms the collective brand of all AI professionals.

Because of the trust people are placing in Generative AI, the models should not act as a proxy decision-maker for users. The information must be presented in a way that enables users to make the final decision without promoting any special agenda

The same applies to where information was obtained and how the AI model arrived at the answer. In cases where there is a knowledge gap, the AI model should inform the user of its inability to provide an answer.

**Question: *A friend showed me an app that utilises AI without adhering to local laws.***

***Answer***: Report it to the local complaint officer, as the app is selling a service under false pretences as it doesn’t have all the required paperwork. It doesn’t matter if the app uses AI; local laws and regulations must be respected.

**Question: *A financial report generated by the in-house AI portal looks unpolished and full of basic errors.***

***Answer***: Report it to the product owner and superiors, as spreading false information can lead to a loss of confidence in the product.

Zero discrimination:

Public interest:

AI models should not discriminate or participate in discriminatory actions, and they should not act as proxy decision makers for end users.

References:

1 <https://www.ibm.com/think/topics/artificial-intelligence>

2 Oxford Dictionary