Reflective Activity 1: Ethics in Computing

Read Correa et al. (2023) and Deckard (2023). After reviewing the article and reading how different countries across the world deal with the generative AI revolution, discuss your views on the subject and recommend what you think could be a suitable course of action. You should justify your stance by also reviewing any papers included in this study or other relevant literature (additional links to industry have been provided as 'Other Resources' to the module). Your discussion should also highlight the impact your actions would have on applicable legal, social and professional issues. Please note that there is no right or wrong answer here, this exercise is to help you evaluate the legal, social ethical and professional issues that affect computing professionals in industry. The word count is 1,000 for the reflection piece.

Local and international policy is needed for governance of a technology as powerful as Artificial Intelligence (AI), however, this is by no means a simple task (Fjeld et al., 2020). The modern world has become a complex entanglement of global economics, politics, and societies (Malan, 2018). International travel allows people to easily move around the world. Communication systems transfer data and information in almost real-time between distant locations. Multinationals trade effortlessly with customers across the globe. Social platforms bring news pertaining to events in foreign countries into the palm of one's hand. Many international political and technological governance and standards structures exist. Added to these complexities exists the very dynamic and fast-paced technological landscape, which many want to use for varying purposes (Wheeler, 2023). Finally, different cultures require different ethical standards of their people (Corrêa et al., 2023). Considering all these points one can begin to understand the hotbed of ethical challenges to be encountered when attempting to implement international regulations for AI.

So how can sufficient governance be put in place to prevent large-scale problems. No doubt this requires significant intervention from a multitude of parties at various strata in business, government, academia, and other communities such as those pertaining to open source (Roy-Chowdhury, 2023; Fjeld et al., 2020). Furthermore, collaboration

between interested groups will be paramount (Gajjar, 2023; Fjeld et al., 2020). For instance, technology leaders will need to continually educate policy makers as to the reach, uses, risks, and benefits of AI (Fjeld et al., 2020). Additionally, advisory bodies comprised of people who do not have any conflicts of interest will need to be formed with clear channels of communication open with relevant law influencers (Fjeld et al., 2020).

Clarification of a global nominal set of ethics pertaining to AI needs to be established (Corrêa, 2023). Significant effort to this end has already been expended, however, a lot more work still remains (Corrêa, 2023). For example, exposition and research into the nuances of what is deemed ethical within a varied cultural subject group is pertinent (Corrêa, 2023). Consider, the value system differences of those living in the United States of America versus that of individuals in the People's Republic of China.

Overregulation can be a stifling attribute of law systems, and this will need to be avoided (Wheeler, 2023). Therefore, using a risk-based framework which doesn't apply broad rules to all forms of AI will go a long way in assisting (Wheeler, 2023). This will check that AI systems which have a larger impact to society meet stricter requirements (Wheeler, 2023).

Like most advanced systems AI requires a stack of supporting technologies to function, some of which are:

- Internet connectivity

- Compute power – often provided by cloud computing.
- Data sets
- Advanced algorithms

A defence-in-depth approach will consider the components of an AI system (Rae et al., 2024). By considering the various sub-systems necessary to produce the aggregated service, regulators can break the problem down into smaller, more manageable facets (Rae et al., 2024). For instance, consider how an organisation might go about acquiring the large amounts of data required to feed into an AI model. There have been various cases where privacy and intellectual property concerns have been raised around the use of data not attained through sanctioned means (Hao, 2021; Brittain, 2023). Having stricter laws surrounding the aforementioned may prevent the production of unethically created AI products. Another example relates to algorithms which are biased, favouring a particular sex or race (Sutaria, 2022). Although, this is an ongoing area of research, there should still be some level of oversight to prevent any unnatural leaning of a final product when it assesses its inputs (Sutaria, 2022).

Just like in general society, having a higher ethical standard creates more effective governance structures than trying to heavily police every aspect. In the AI community this would need the training and instilling of values in all practitioners who play a role in the sector (Deckard, 2023). Focus areas for such may include:

- A strong technical undergirding (Deckard, 2023).
- An understanding of the impacts of AI (Deckard, 2023).
- Inter-personal communication (Deckard, 2023).

- Ability to work with other professionals (Deckard, 2023).

It would also be important for companies and government to expect registration with ethical bodies such as the British Computer Society (BCS) for those professionals situated in the United Kingdom (UK) (BCS, N.D.a). This would aid in ensuring the continuous development of professionals as well as their awareness of various ethical standards (BCS, N.D.b).

Reiterating, there are many areas that will need to be considered to furnish a wholistic approach (Marr, 2023). Other points of significance include:

- Preventing largescale surveillance operations targeting civilians, foreign or domestic.
- Hindering the spread of misinformation and disinformation, whether deliberate or not (Marr, 2023).
- Stopping the gap between the rich and poor from growing even further apart (Marr, 2023).
- Maintaining jobs within industries where AI will be able to replace the need for a human in certain roles (Marr, 2023).

There are some industry leaders who are even calling for a pause in the development of sophisticated AI systems (Marr, 2023). This may be exactly what is required to allow all stakeholders to recuperate and focus on ethical concerns, without having to madly race the competition.

References:

BCS. (N.D.a) Policy and Influence. Available from: https://www.bcs.org/policy-and-influence/ [Accessed 04 May 2024].

BCS. (N.D.b) About us. Available from: https://www.bcs.org/about-us/ [Accessed 04 May 2024].

Brittain, B. (2023) OpenAI, Microsoft hit with new author copyright lawsuit over AI training. Available from: https://www.reuters.com/legal/openai-microsoft-hit-with-new-author-copyright-lawsuit-over-ai-training-2023-11-21/ [Accessed 15 May 2024].

Corrêa, N, et al. (2023) *Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance*. Patterns 4(10). DOI: https://doi.org/10.1016/j.patter.2023.100857

Deckard, R. (2023) What are ethics in AI. Available from: https://www.bcs.org/articles-opinion-and-research/what-are-ethics-in-ai/ [Accessed 04 May 2024].

Fjeld, J. et al. (2020) *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI*. Berkman Klein Center Research Publication 2020(1). Available from: https://papers-ssrn-com.uniessexlib.idm.oclc.org/sol3/papers.cfm?abstract_id=3518482

Gajjar, D. (2023) Artificial Intelligence: An explainer. Available from: https://researchbriefings.files.parliament.uk/documents/POST-PB-0057/POST-PB-0057.pdf [Accessed 05 May 2024].

Hao, K. (2021) Deleting unethical data sets isn't good enough. Available from: https://www.technologyreview.com/2021/08/13/1031836/ai-ethics-responsible-data-stewardship/ [Accessed 13 May 2024].

Malan, D. (2018) The law can't keep up with new tech. Here's how to close the gap. Available from: https://www.weforum.org/agenda/2018/06/law-too-slow-for-new-tech-how-keep-up/ [Accessed 05 May 2024].

Marr, B. (2023) The 15 Biggest Risks of Artificial Intelligence. Available from: https://www.forbes.com/sites/bernardmarr/2023/06/02/the-15-biggest-risks-of-artificial-intelligence/?sh=20f70a292706 [Accessed 14 May 2024].

Rae, C. et al. (2024) Architect defense-in-depth security for generative AI applications using the OWASP Top 10 for LLMs. Available from: https://aws.amazon.com/blogs/machine-learning/architect-defense-in-depth-security-for-generative-ai-applications-using-the-owasp-top-10-for-llms/#:~:text=Defense%2Din%2Ddepth%20security%20best,workloads%2C%20data%2C%20and%20assets [Accessed 05 May 2024].

Roy-Chowdhury, R. (2023) Why open-source is crucial for responsible AI development. Available from: https://www.weforum.org/agenda/2023/12/ai-regulation-open-source/ [Accessed 05 May 2024].

Sutaria, N. (2022) Bias and Ethical Concerns in Machine Learning. Available form: https://www.isaca.org/resources/isaca-journal/issues/2022/volume-4/bias-and-ethical-concerns-in-machine-learning [Accessed 13 May 2024].

Wheeler, T. (2023) The three challenges of AI regulation. Available from: https://www.brookings.edu/articles/the-three-challenges-of-ai-regulation/ [Accessed 05 May 2024].