**Assignment 1: AIDI 2004 Winter 2021 (Group 10)**

**Group members: Gbemisola Banjoko (100766479) and Shrita Gaonkar (100799307)**

**Collaboration plan**

|  |  |  |  |
| --- | --- | --- | --- |
| **Meeting Day** | **Meeting time** | **Agenda** | **Action item** |
| 18/01/2020 | 12-1pm | To understand the assignment and decide a meet time | Workload distribution  Decide meeting timings |
| 19/01/2020 | 12-1pm | Discussion about paper selection (out of 3) | **The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions**  Decided. |
| 21/01/2020 | 12-1pm | Discuss:   1. Summary of the paper 2. Three strong and weak points | Concluded six points each.  Added to the report |
| 23/01/2020 | 12-1pm | Discuss:   1. Main idea of the paper and how it is solving the issue. 2. Research methodology of the paper | Concluded six points each.  Added to the report |
| 24/01/2020 | 12-1pm | Discuss:   1. Three significant questions about the paper 2. Other remarks | Concluded six points each.  Added to the report |
| 25/01/2020 | 12-1pm | Final changes and submission | **Submitted on DC Connect** |

**Workload Distribution**

|  |  |
| --- | --- |
| **Student** | **Work allotted** |
| Gbemisola | Summary of the paper |
| Gbemisola | Three strong and weak points |
| Gbemisola | Main idea of the paper and how it is solving the issue |
| Gbemisola | Research methodology of the paper |
| Shrita | Three significant questions about the paper |
| Shrita | Other remarks |
| Shrita | Report making |
| Shrita | Scheduling meeting |

**Resolving any conflicts or inequitable work distribution.**

* All the team members communicated properly about the work distribution. Hence, there were no conflicts.
* Work was distributed evenly.

**Review paper title: The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions**

**Paper Review**

The main problem addressed by this paper is associated with the fact that even though there has been a huge increase in the principles that guide Artificial Intelligence ethics, however, these principles are too broad and high-level, which makes them impractical to be applicable to Artificial Intelligence ethics as closer look reveals that these principles conflict with one another when applying them to specific situations. The paper therefore focuses on exploring the tensions that always emerge in the process of implementing principles in practice.

To resolve this problem, this paper suggests that Artificial Intelligence ethics should focus on identifying and trying to solve the tensions that emerge from the application of principles to specific situations. The authors of this paper tackle this problem by applying the following:

* A review how principles in Artificial Intelligence ethics have progressed over the last 2 years of the paper.
* Comparing Artificial Intelligence principles with bioethics.
* Consider and review some of the limitations of principles and the conflicts that already exists in relation to the use of Artificial Intelligence in the society, and how to possibly resolve them.
* Debate that all aspects of Artificial Intelligence ethics would gain from a more rigorous exploration of the conflicts that emerge in the application of principles to practical situations.

Some of the strengths of this paper includes:

* A clear explanation of why the authors recommends a focus on understanding the conflicts that are associated with Artificial Intelligence ethics principles and their practical applications.
* Highlighting and discussing the different types of conflicts that exists in the field of Artificial Intelligence ethics including those associated with the collection and use of data.
* Addressing some of core issues like bias and fairness.
* The paper made preposition on an approach to be used in resolving the problem identified by the authors, which is conflicts in the principles that guide Artificial Intelligence ethics.

Some of the weakness of this paper includes:

* This paper could be stronger if its recommendations were applied to a real case study.
* There was no clear definition of the research methodology used in the paper.
* While research papers are written based on the strengths of the research, they need to also identify the limitations of the study and this paper has failed to identify and clearly state its limitations.

Research methodology:

While this paper has failed to specifically identify and define its research methodology, since no primary data was collected, this qualifies it as a secondary research. In addition, the fact that the paper draws on a comparison of Artificial Intelligence principles and Bioethics, makes it is a comparative study. As Bukhari, (2011) defined a comparative study as the analysis and comparison of two or more objects or ideas.

Based on the review on this paper, the following questions have been raised:

* The paper tackled the problems of fairness and bias by prioritisation, costs, and benefits. The older technology can be better than the current technology which can raise some issue. How can we solve these issues?
* Data and Data collection from real-time users are a crucial part of AI. How can both customers and companies be satisfied with the data collection methods?
* Since AI ethics is an emerging field, how can researchers study the more advanced bio-ethics to come to more universal standards? the article mentions many different organizations that created lists of principles in 2017. How can these be clarified in a way that is fair to all?
* How much risk can be tolerated to privacy values when the result would be so beneficial - should there be a general consensus on this or should it be the responsibility of the researchers on a case by case basis?

Other observations of the paper are listed below:

* The paper has been written in clear and easily understandable format using simple vocabulary that can be understood by a broad audience.
* The argument appeared within the first couple of paragraphs of the paper.
* This gave us an idea about the thesis statement of the paper.
* The reasons for performing the study were clearly defined.
* The paper has a higher potential to make a difference.

**REFERENCES**

Bukhari, Syed Aftab Hassan, what is Comparative Study (November 20, 2011). Available at SSRN: <https://ssrn.com/abstract=1962328> or [http://dx.doi.org/10.2139/ssrn.1962328](https://dx.doi.org/10.2139/ssrn.1962328)

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019, January). The role and limits of principles in AI ethics: towards a focus on tensions. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 195-200).