COMP4920 Group Assessment: AI/ADM Case study

Flora Salim, UNSW Sydney

September 11, 2024

1 Introduction

In this assignment you are to consider, discuss, and present your arguments on the ethical dilemmas around a case study of an AI or Automated Decision-Making (ADM) system that is already deployed or planned to be deployed in the society. You are to work in a group of three. All members of a group must be from the same tutorial. In those cases where the total tutorial numbers do not divide into three perfectly, your tutor will arrange things. The assessment components include a group presentation during your tutorial and a project report.

2 Background

According to a McKinsey's report[1], and a more recent one on the state of AI[2], AI has a huge potential to significantly boost economic productivity. The adoption of AI-based automation, or ADM-S technologies are increasingly prevalent in the private and public sector with a direct impact on our society. With AI and ADM being used in influencing the situations or outcomes where people's lives or livelihoods are at stake, responsible innovation practices and FATE (Fairness, Accountability, Transparency, and Ethics) in AI is becoming more paramount. These ethical frameworks are often not considered during the design of an AI or ADM technology. Although this technology or system might have been designed "for good", it might lead to clashes of values and objectives among the potential stakeholders of the proposed technology.

3 Your Task

Your task is to investigate a particular AI or ADM technology as if you and your team are the ones who are responsible for deciding whether to adopt and deploy this technology, and if so, how to do it ethically and responsibly.

1. Pick a technological case study to be investigated. Describe the technology, the purpose, the scope, the needs, and the benefits. The technology's purpose includes the overarching and the specific goals/objectives of the tech. The scope includes the specific context, time, location, or specific use case. The needs describe the underlying reason/motivation for why it is designed, and explains why the tech is needed. The benefits may encompasses health, societal, economic benefits. Describe them in detail.

[1] Modeling the impact of AI on the world economy, https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy

[2] The State of AI in 2021, https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2021

2. Consider your users and the stakeholders. List them all and describe their attributes, needs and skills, why they are considered a user or a stakeholder. With regard to the user/stakeholder, you can describe a persona. Be as specific as needed. Their attributes can be to do with their private attributes (e.g. age group, gender, etc). Describe their specific needs/goals when interacting with this tech. If none, then say none. Note that they may not be a direct user, but they may be positively/negatively impacted by the deployed tech nonetheless.

3. Using value-sensitive design, conduct a conceptual investigation around the human values in the technology. Pick three main users/ stakeholders (or one per group member in your project group) out of your list of possible users/stakeholders and perform an imaginary role play to discuss the possible impact of the technology to them. You can use the Judgment Call game[3](printable) as a tool to review these technologies from the point of view of the stakeholders. Discuss the impact of the technology along each of the ethical principles (e.g. Microsoft's ethical principles of fairness, privacy and security, reliability and safety, transparency, inclusion, and accountability)[4], as shown in the Judgment Call's cards[5]. Ideally, perform this task before you conduct the research in the next step. Record your discussions (e.g. take photos) and summarize the conceptual investigation[6]. It is highly recommended that you summarise the discussions and findings on the ethical dilemmas in a chart or a table.

4. Conduct research on how a similar technology has been used. Pick two of the issues or concerns in relation to the technology that were discussed by the stakeholders, which might include fairness, transparency, explainability, and/or responsibility issues, ethical and/or discrimination risks. Provide critical analysis on its impact with examples from similar or related real-world technologies.

5. Compare your findings with your initial conceptual investigation. You may want to use a chart or a table to describe, compare, and summarize your findings. Are there any agreements or disagreements? Are the findings complementary? What can you conclude from this exercise?

6. Provide your recommendation. Should you adopt and deploy this technology? If so, do you need to perform some changes or adapt it in a specific way?

7. Prepare a group presentation and project report.

4 Group Presentation

- To be scheduled during your tutorial in Weeks 9-10.
- Each group presentation is 15 minutes, followed by 3 minute Q&A.

[3]Stephanie Ballard, Karen M. Chappell, and Kristen Kennedy. 2019. Judgment Call the Game: Using Value Sensitive Design and Design Fiction to Surface Ethical Concerns Related to Technology. In Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19). Association for Computing Machinery, New York, NY, USA, 421- 433, https://doi.org/10.1145/3322276.3323697

[4]Microsoft's ethical principles, https://www.microsoft.com/cms/api/am/binary/RE4pKH5

[5]Judgment Call, https://learn.microsoft.com/en-us/azure/architecture/guide/responsible-innovation/judgmentcall

[6]In an ideal world, you will follow this up with a user study, but this will require more time and effort to find the users/stakeholders, within a human ethics approved study, which will not be feasible in the time-frame of this course.

5 Individual Project Report

- Word length: 4000-4500 words (about 1500 words per team member) excluding references and appendix (which can include photos). If you use a table or a chart, this corresponds to 300 words per table/chart. If this is done collectively, please annotate it accordingly.
- There is no prescribed template. You are suggested to follow the structure of the tasks.
- Ideally, prepare an interim report prior to your presentation. Then incorporate feedback from your class and tutor to improve your report.
- Given that the group report is **marked individually**, you are required to clearly state your contributions and annotate your individual parts in the report, and the parts done together as a group (e.g. a table, figure, etc).
- It is essential that you work collaboratively as a group, as although you have individual responsibilities on your contributions to the group work, it is also your responsibility to ensure that the report is a well integrated deliverable from your group. If this doesn't appear so, it may have an impact on the marking of the overall structure of the report (which will have an impact on individual scores).
- The report is due at the end of Week 11: **Friday November 22nd, 18:00** via Moodle/turnitin.
- Bonus Mark are available for individuals. Check the guidelines below.

6 Course Learning Outcomes

This assignment will contribute to the following Course Learning outcomes:

- Students will discover and learn to define ethical values, principles, and practices towards responsible research and innovation of technological and computing advances.
- Students will have an increased awareness of the ethical issues and pitfalls in their own professional practice of developing novel technologies, including AI (e.g. fairness, transparency, accountability), and learn about existing efforts to mitigate these issues.
- Students will learn to discuss ethical dilemmas around specific technological case studies openly and robustly.

7 Assessment Weight

- Presentation - 20% (marked collectively)
- Report - 30% (marked individually)

8 Marking Guidelines

8.1 Presentation: 20 Marks

- Conceptual investigations: Provides a succinct and clear context for the investigation, clear introduction of the technology and the stakeholders

- Analysis/findings: Provides a sophisticated and critical analysis, clear and succinct assessment of the chosen stakeholder and the value, backed by research. Appropriate technological comparison, analysis.
- Creative and critical thinking: Thinks outside a box. Draws a deep and thorough investigation. Provides a well justified recommendation.
- Delivery/Engagement: clarity of delivery, engaging, appropriate expression for oral presentation; confidence; good timing; suitable pauses to enhance understanding; engages the audience well
- Slides / props: Excellent slides or props that contribute effectively to the oral presentation
- Teamwork: a clear contributor to the team, supportive, the presentation is well integrated as whole with a very smooth flow.

8.2 Individual Project Report: 30 marks

- Introduction and background of tech: A clear focus, and a comprehensive explanation of the AI/ADM technology. States the case study, and provides a clear context for the investigation and the broad stakeholders and users.
- Conceptual investigation of the chosen stakeholders: the stakeholders are well-described, with the values and the frictions across the chosen principles to be well discussed and elaborated across each stakeholder. A clear summary is provided at the end of the detailed descriptions.
- Research, analysis, and comparison: a sophisticated critical analysis of the possible positive and negative impact of the studied AI/ADM technology, based on similar technology in the literature or commercial systems that are already deployed. Discuss the possible mitigation techniques or approaches to deal with the potential risks and ethical concerns of the studied AI system. Well-justified usage and selection of the used data sources, figures, and/or references.
- Overall creative and critical thinking: Thinks out of the box, creates or extends to a novel or unique investigation, incorporating feedback from the presentation session, a well-rounded discussion and recommendation.
- A well-structured report, contains all the required tasks, impressive flow, appropriate use of headings and relevant sections.
- Compliance: adhering to specifications (word limit), correct use of references, appropriate use of grammar.
- **Bonus marks**: If you wish to score up to 5 extra mark in the individual portion of the group project, you can choose to give a short 5-min talk recapping one of the guest lectures or Flora's lectures during the tutes. Please contact your tutor to get this organised.

9 Referencing Guidelines

Your essay must follow the Harvard style in-text referencing system as described here:
https://www.student.unsw.edu.au/harvard-referencing

10 Plagiarism Rules

It is very, very important that you do not fall afoul of UNSW's plagiarism rules. You need to read through and understand all of the following:
https://www.student.unsw.edu.au/plagiarism

11 ChatGPT

I have read through the recent document from UNSW's Pro-Vice Chancellor on the issue of ChatGPT. **I find that this assessment in SENG/COMP4920 falls under the following rules for simple editing assistance**:

For this assessment task, you may use standard editing and referencing software, but not Generative AI. You are permitted to use the full capabilities of the standard software to answer the question (e.g. you may wish to specify particular software such as Microsoft Office suite, Grammarly, etc.). If the use of generative AI such as ChatGPT is detected, it will be regarded as serious academic misconduct and subject to the standard penalties, which may include 00FL, suspension and exclusion.