Implementing AI ethics in a software engineering project-based learning environment - The case of WIMMA lab

Mamia Ori-otse Agbese[1][0000−0002−5479−7153], Marko Rintamaki[2], Rahul Mohanani[1][0000−0001−7018−8836], and pekka.abrahamsson[1][0000−0002−4360−2226]

[1] University of Jyväskylä, Seminaarinkatu 15, 40014, Finland
mamia.o.agbese@jyu.fi,rahul.p.mohanani@jyu.fi,pekka.abrahamsson@jyu.fi
https://www.jyu.fi/en
[2] Jyväskylä University of Applied Sciences, Rajakatu 35, Finland
marko.rintamaki@jamk.fi
https://www.jamk.fi/en

Abstract. Increasing ethical concerns necessitate AI ethics forms part of practical software engineering (SE) foundational educational learning. Using an ethnographic approach and focus group discussions in a SE project-based learning environment, WIMMA lab, we gain insight into how AI ethics can be implemented to enable students to acquire these necessary skills. We propose a framework as an outcome to aid the implementation of AI ethics skills within SE project-based learning environments.

Key words: Artificial Intelligence, Implementing AI ethics, Software engineering, Project-based learning, AI, software business, PBL

1 Introduction

Advances in the engineering of Artificial intelligence (AI) using Big data, the rapid rise of the internet, and machine learning (ML) enables current AI systems to harness vast amounts of data and model solutions used in areas such as retail, health, and transport and warfare [1]. These advances, however, have become associated with ethical concerns such as bias and data privacy from negatively impacting incidents. Incidents such as the failure of the COMPAS AI-based recommender system necessitate that AI systems are engineered with appropriate AI ethical practices to help mitigate these challenges [2]. Jobin and Borenstein [3, 4] explain that implementing AI ethics at the practical educational level can enable it to become foundational and cultural in software engineering (SE). It can also equip SE graduates to gain relevant AI ethics skills needed to adapt to real-life environments requiring AI ethics implementation and aid software businesses to spend less on training recruits and practitioners [5]. However, scarce research exists for implementing AI ethics within practical third-level education, particularly in project-based learning environments, which aids students in competence building and practical experience. Most focus on

kindergarten to middle school level with hardly any representation for third-level education. We conduct anthropological research to gain a holistic perspective in a SE project-based learning environment, WIMMA lab, to help us understand how AI ethics practices can be implemented in a SE Project-based learning environment. Our research question is **RQ: How can AI ethics be implemented in a project-based learning environment?**. We propose a framework to aid the implementation of AI ethics practices in a SE project-based learning environment. This contribution is significant as it can help inform guidance on this topic within SE project-based learning environments.

2 AI ethics in SE project-based learning environment

AI ethics which deals with the moral behavior of humans in the design usage and behavior of machines [6] is mainly targeted at ensuring people, processes, and organizations responsible for AI systems assume responsibility for their engineering [6]. Currently, over 80 AI principles exist from research bodies, governments, and private organizations [7, 3] in response to the growing ethical challenges associated with engineering AI. They highlight the importance of implementing ethics in AI engineering and the negative impact of poorly engineered AI systems [7, 8]. They also help to push the AI ethics agenda to the forefront of SE by guiding the engineering of AI and highlighting the need for effective education of AI ethics for students and practitioners [3]. While they are criticized as insufficient in translating principles to practice [9], they remain the primary measure in implementing AI ethics. There is little research for implementing AI ethics skills in SE engineering project-based learning education. Studies from [10], and [11] create an understanding of technical and ethical AI concepts awareness for kindergarten grade to grade 12 and middle schools. [12] examined the concept of AI stakeholder identification in middle school students. However, no research was identified for SE project-based learning in third-level education.

2.1 WIMMA Lab

WIMMA lab, a SE project-based learning environment at the University of Applied Science Jyvaskyla, Finland, is used for this study. The lab uses project-based learning to enable students to gain DevOps and DevSecOps SE skills using open source materials within four simulated SE virtual companies, IoTitude, Mysticons, Overflow, and Penguin Media. The entire project is managed using an open project framework (OPF) and documented in two guidebooks or playbooks. The Black book records the virtual labs' operation, the types of practices and routines utilized, and the green technical processes. Figure 1 illustrates the current structure at WIMMA lab.

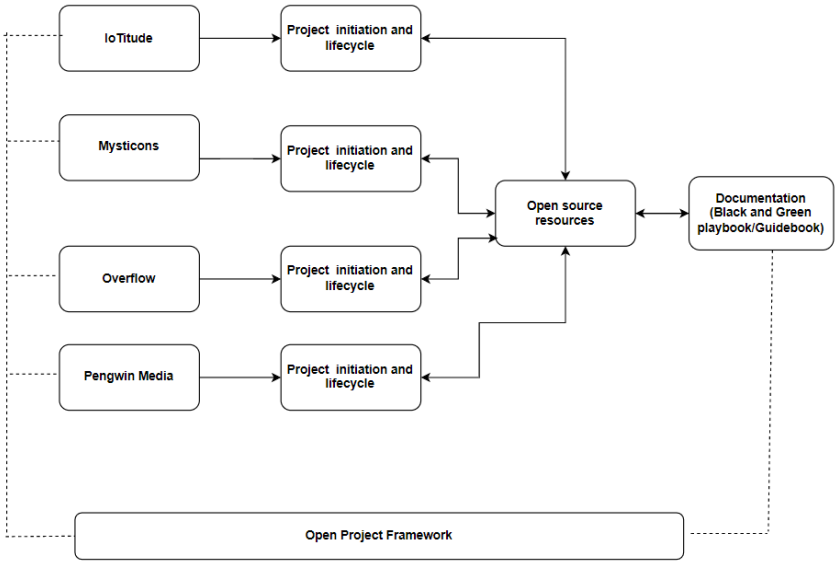

Fig. 1. Current structure at WIMMA lab

3 Methodology

We use an ethnography approach which allows researchers become part of the community and perform the same or relevant tasks to understand better the underlying phenomenon under study.

3.1 Data Collection

The four participants from the information systems field in the university of Jyvaskyla were split into three of teams, Mysitcons, Overflow and IOTitude based on competence in SE practices however, the primary researcher was there as a researcher. The main data collection for the study was in the form of a focus group which is very valuable when in-depth information is needed about how people think about an issue, their reasoning on why things are they way they are and why they hold the views they do [13]. The interview was conducted as a semi-structured interview discussion and recorded over Zoom. The participants from the university of Jyvaskyla took part in the discussions which lasted for 47 minutes. Each participant received adequate information about the study and gave their full consent to participate in the exercise and for the use of the audio recording of the group discussion. For supplementary data, the primary researcher engaged in discussions with other team mates unofficially to understand their stance on the research topic.

3.2 Analysis

The recordings were manually transcribed by replaying the Zoom recording, pausing, and rewinding to ensure the right information was transcribed. The transcribed data was then analysed by coding to identify themes that emerged accordingly. The data was coded under key themes which emerged to help achieve the aim and objectives of the study. The Analysis is categorized into themes derived and suggested practices aimed at achieving the aim and objective of the study. Table 1 illustrates the analysis.

Table 1. Emergent themes and suggested practices

Themes	Practices
Awareness of AI ethics concerns	None, personal ethics, use of security Data Management, Data governance, Security concerns, cybersecurity threats
Current AI ethics practices	None
How AI ethics can be implemented	Simplified framework to enable understanding, At the beginning of the project, Accountability, Stakeholder identification, Security measures to enable efficient implementation in the future like GDPR, Data management, Data governance, documentation/reference (playbooks, guides)

4 Discussion

The discussion is broken down into three main themes that emerged from the data analysis.

4.1 Awareness of AI ethical concerns:

Concerns such as data management, governance, and cybersecurity threats are identified as possible avenues for AI ethical concerns. The students currently employ initiatives such as using normative ethics, creating technical processes to help deploy regulations such as GDPR in the future, and using engineering security practices. Suggesting that current research on SI ethics awareness is impactful and the use of more effective tools and education can be more impactful [3, 9]. It also suggests that targeted effort is needed because engineering AI is different due to data and close interrelation between human and artificial agents, which poses an extra challenge [14].

4.2 Existing AI ethics practices in WIMMA lab:

The findings reveal virtually no implementation of AI ethics practices. AI ethics framework or tools were not identified in the open source resources used by the students. This agrees with the review of literature where AI ethics practices could not be identified in SE third level education. Implying more research needs to be carried out to help inform these practices.

4.3 How AI ethics can be implemented:

The findings reveal that simplified frameworks or tools can improve understanding of AI ethics and aid implementation. Such tools or frameworks can help guide how AI practices are fused with the SE processes. Blackman [15] explains that simple and flexible frameworks and tools that suit the context of its application can improve awareness of AI ethics and aid implementation. Progress has begun in AI ethical tools and methods [16], but a gap still exists in transitioning them to practice [3]. The findings recommend tools and frameworks to be used at the beginning of the project to enable students to have a better understanding of AI ethics and aid their development of the necessary skills associated with it. A study by [17] revealed the use of an ethically aligned design tool within a classroom learning environment enabled students to elicit ethical non-functional requirements. [15] also explains that operationalizing AI ethics using existing infrastructure and an AI ethics framework at the onset can aid implementation and help mitigate AI risks. Documentation of practices as playbooks for guidance and reference also formed part of the findings. [18] explains that documentation improves capabilities by making them more accessible, usable, and available for the learning community. Documentation can also serve as a point of reference to enhance clarity in areas of confusion.

4.4 Framework:

We present our proposed for implementing AI ethics in SE project-based learning environment in figure 2. The proposed framework will provide students with the use of an agreed-upon AI ethics tool or framework to serve as a reference or guide to consult on how AI ethical practices can be implemented at the beginning of the project. The practices within the framework can also be infused with open source resources and serve as a reference throughout the project lifecycle to enable the students to access it for clarity and guidance on the best practices to imbibe. As explained earlier, documentation of best practices helps improve implementation and inform knowledge [15]. Therefore, current practices as they unfold can be documented or infused with the project management system, with the best practices of AI ethics recorded in playbooks to be used by students that come on board in subsequent years for them to leverage for their projects.

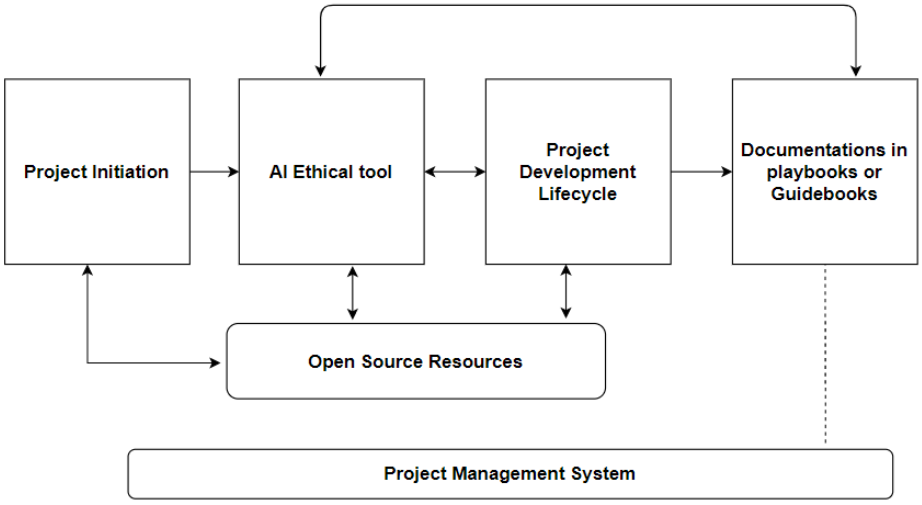

Fig. 2. Framework for implementing AI ethics in SE project-based learning environment

5 Limitation

Focus group investigations typically involve six to ten homogeneous strangers in a formal setting [13], however, we follow recommendation by [19] for less structured discussions using smaller groups to better suit the purpose of the study. Another limitation is the discussion of findings relevant to the narrow setting of the WIMMA lab however, it enabled for the generation of descriptive results for the study. The use of only one project-based learning environment may also pose a validity threat to the study, [20] explains that that for novel research areas such as AI ethics with the scarce literature in its implementation, a low number of case studies can be acceptable.

5.1 Further work

Future research can employ a quantitative approach to provide breadth to the research. This may provide a holistic view with divergent findings that can lead to a reexamination of the assumptions made in the study. The feasibility of the framework will also be tested to examine its efficacy.

6 Conclusion

We studies how AI ethics can be implemented in SE project-based learning environments using an ethnographic approach in WIMMA lab using focus group discussion in the investigations. The outcome reveals that AI ethics practices

are not implemented in the SE practices. Based on the findings, a framework is proposed to help implement AI ethics.

References

1. Coeckelbergh, M.: AI ethics. Mit Press (2020)
2. Gupta, M., Parra, C.M., Dennehy, D.: Questioning racial and gender bias in ai-based recommendations: Do espoused national cultural values matter? Information Systems Frontiers (2021) 1–17
3. Jobin, A., Ienca, M., Vayena, E.: The global landscape of ai ethics guidelines. Nature Machine Intelligence **1**(9) (2019) 389–399
4. Borenstein, J., Howard, A.: Emerging challenges in ai and the need for ai ethics education. AI and Ethics **1**(1) (2021) 61–65
5. Ali, M.R.: Imparting effective software engineering education. ACM SIGSOFT Software Engineering Notes **31**(4) (2006) 1–3
6. Müller, V.C.: Ethics of artificial intelligence and robotics. (2020)
7. Morley, J., Floridi, L., Kinsey, L., Elhalal, A.: From what to how: an initial review of publicly available ai ethics tools, methods and research to translate principles into practices. Science and engineering ethics **26**(4) (2020) 2141–2168
8. Hleg, A.: A definition of ai: main capabilities and disciplines. Brussels. https://ec. europa. eu/digital-single (2019)
9. Mittelstadt, B.: Ai ethics–too principled to fail. arXiv preprint arXiv:1906.06668 (2019)
10. Williams, R.: How to train your robot: Project-based ai and ethics education for middle school classrooms. In: Proceedings of the 52nd ACM Technical Symposium on Computer Science Education. (2021) 1382–1382
11. Lee, I., Ali, S., Zhang, H., DiPaola, D., Breazeal, C.: Developing middle school students' ai literacy. In: Proceedings of the 52nd ACM technical symposium on computer science education. (2021) 191–197
12. Jordan, B., Devasia, N., Hong, J., Williams, R., Breazeal, C.: Poseblocks: A toolkit for creating (and dancing) with ai. In: Proceedings of the AAAI Conference on Artificial Intelligence. Volume 35, number:17. (2021) 15551–15559
13. Bell, J., Waters, S.: Ebook: doing your research project: a guide for first-time researchers. McGraw-hill education (UK) (2018)
14. Ozkaya, I.: What is really different in engineering ai-enabled systems? IEEE Software **37**(4) (2020) 3–6
15. Blackman, R.: A practical guide to building ethical ai. Harvard Bus. Rev **15** (2020)
16. Vakkuri, V., Jantunen, M., Halme, E., Kemell, K.K., Nguyen-Duc, A., Mikkonen, T., Abrahamsson, P.: Time for ai (ethics) maturity model is now. arXiv preprint arXiv:2101.12701 (2021)
17. Halme, E., Vakkuri, V., Kultanen, J., Jantunen, M., Kemell, K.K., Rousi, R., Abrahamsson, P.: How to write ethical user stories? impacts of the eccola method. In: International Conference on Agile Software Development, Springer, Cham (2021) 36–52
18. Marshall, R., Pardo, A., Smith, D., Watson, T.: Implementing next generation privacy and ethics research in education technology. British Journal of Educational Technology (2022)
19. Morgan, D.L.: Focus groups. Annual review of sociology **22**(1) (1996) 129–152
20. Eisenhardt, K.M.: Building theories from case study research. Academy of management review **14**(4) (1989) 532–550