Claim: Software Engineers need to implement software engineering ethics to provide customers with reliable and explainable artificial intelligence-based software products from the perspective of transparency and robustness.

**DOES IMPLEMENTING ETHICS IN ARTIFICIAL INTELLIGENCE-BASED APPLICATIONS PROVIDE TRANSPARENCY TO THE USER SIDE?**

Artificial intelligence (AI) is becoming increasingly widespread in system development endeavors. As AI systems affect various stakeholders due to their unique nature, the growing influence of these systems calls for ethical considerations. Academic discussions and practical examples of autonomous system failures have highlighted the need for implementing ethics in software development. For example, by 2025, all Bosch products are intended to either contain artificial intelligence or to be developed or manufactured with the help of artificial intelligence. The company wants artificial intelligence products to be robust, reliable and explainable. According to the, Michael Bolle of the CDO and CTO Bosch in the 2020, "If artificial intelligence is a black box, people won't trust it, but trust will be essential in a networked world” (as cited in Wild-Raidt, 2020, para.3). And while it is important to go advance in artificial intelligence-based technology it is important to implement software engineering ethics. This essay will advance the claim that software engineers need to implement ethics to provide reliable and explainable artificial intelligence-based applications to their customers with both the ideas that are aligned and not aligned with it.

One of the strongest arguments for the implementing ethics in the artificial intelligence-based applications is the fact that it does not take too long into a conversation about the ethics of a new form of AI technology before concerns about that implementing software-engineering ethics prevents raising of unanticipated consequences. For instance, according to Implementing Ethics in AI: Initial results of an industrial multiple case study paper published in 2019, “The role of ethics in software systems development has dramatically changed following the increasing influence of Autonomous Systems (AS) and Artificial Intelligence (AI) systems. AI/AS systems necessitate ethical consideration due to their unique nature. Following various real-life incidents out on the field, AI ethics has also begun to incite public discussion. This has caused various government, public, and private organizations to react, primarily by producing their own demands and guidelines for involving ethics into AI development. Countries such as Germany [14] have emphasized the role of ethics in AI /AS, and the EU drafted its own AI ethics guidelines [15]. Industry organizations such as Google and IBM1 have also devised their own guidelines.” (Vakkuri, V., & Kmell K., & Abrahamsson P., 2019).

In addition, not implementing software-engineering ethics in autonomous weapon systems could cause several problems which may lead to unanticipated consequences from the perspective of the peoples' safety and software products relationship. According to Jason Borenstein, “The U.S. military is actively seeking to construct and deploy a variety of different types of robotic systems. The different branches of the military certainly see a vision of the future where autonomous robots will better enable them to meet their objectives. Of course, many potential ethical problems are associated with the use of human-controlled robotic systems. For instance, human operators of UAV (Unmanned Aerial Vehicles) s might experience fewer psychological obstacles in their way before human life is taken because a considerable geographical distance is often placed between their decisions and their actions. Yet the monitoring of human operators might prevent some of the abuse.” (Borenstein J., 2008).

Further, from the ethical perspective, the psychological dimensions of war can arguably help to make the case for the use of an AWS (Automated Weapon Systems). According to an Army report that examined the mental health of military personnel serving in Iraq, “Only 47 percent of the soldiers and 38 percent of Marines agreed that non-combatants should be treated with dignity and respect. Unlike their human predecessors, a robotic soldier would not act out of malice or prejudice” (Officers of the Surgeon, 2006, p. 35). Therefore, with the wide-spread use and incrementally production of the artificial intelligence-based autonomous systems, the need for ethics in software engineering raised obviously to prevent unanticipated consequences. Software engineering shall, in their professional role, act only in ways consistent with the public safety dense of their professional judgment and their reputation for such judgment (Gotterban, D. & Miiller, K. & Rogerson, S., 1997).

Of course, the possibility of creating thinking machines raises a host of ethical issues. Also, these questions relate both to ensuring that such machines do not harm humans and other morally relevant beings, and to the moral status of the machines themselves as it was touched above. Another strong argument for implementing ethics in artificial intelligence-based software products is that they provide transparency and robustness from the software products and customers relationship perspective. According to The Ethics of Artificial Intelligence paper published in 2011, “It will become increasingly important to develop algorithms that are not just powerful and scalable, but also transparent to inspection—to name one of many socially important properties. It involves new programming challenges, but no new ethical challenges. But when AI algorithms take on cognitive work with social dimensions-cognitive tasks previously performed by humans—the AI algorithm inherits the social requirements. It would surely be frustrating to find that no bank in the world will approve your seemingly excellent loan application, and nobody knows why, and nobody can find out even in principle.” (Bostrom N. & Yudkowsky E., 2011).

Regarding the fact that artificial intelligence is derived from software engineering, the software transparency should be taken into consideration to create reliable and explainable applications. The software transparency is stated as flows in the Software Transparency paper published in 2010, “Software transparency is a new and important concern that software developers must deal with. As society moves towards increased automation, if citizens wish to exercise their right to know, the transparency of public services and processes acquires fundamental importance” (Sampaio C. S., Capelli C., 2010). Thus, software transparency in implementing ethics will play an important tole from the perspective of creating explainable machines for customers.

From the robustness perspective of ethics, it will also become increasingly important that AI algorithms be robust against manipulation. The following sample case explains well the relationship between the artificial intelligence-based application and manipulation of them, “A machine vision system to scan airline luggage for bombs must be robust against human adversaries deliberately searching for exploitable flaws in the algorithm-for example, a shape that, placed next to a pistol in one’s luggage, would neutralize recognition of it. Robustness against manipulation is an ordinary criterion in information security; nearly the criterion” (Bostrom, N. & Yudkowsky, E., 2011).Thus, software engineers shall, insofar as possible, assure that the software on which they work is useful and of acceptable quality to the public, the employer, the client, and the user, completed on time and at reasonable cost, and free of error.

Even though in the above paragraphs the ethics lead software engineers to produce reliable and explainable machines have been discussed, still one of the strong opposing ideas related to implementing ethics is not the only solution to provide reliable and explainable AI-based products if we will not be able to eliminate bias in AI. According to the Whittaker et al AI now report published in 2018, “Though artificial intelligence is capable of speed and capacity of processing that’s far beyond that of humans, it cannot always be trusted to be fair and neutral. It should be kept in mind that the AI systems are created by humans, who can be biased and in a judgemental manner. Several tech companies already offer tools for bias mitigation and fairness in machine learning. In this context, Google, Microsoft and Facebook have issued the “AI Fairness 360” tool kit, the “What-If Tool”, “Facets”, “fairlern.py” and “Fairness Flow”, respectively” (Whittaker et al. 2018). Also according to The Cathedral of Computation paper published in 2015, “As our reliance on artificial agents continues to grow, so does the risk. A better understanding of our attitudes and interactions with algorithms is essential precisely because of the aura of objectivity and infallibility our culture ascribes to algorithms” (Bogost, 2015). Therefore, eliminating bias in artificial intelligence-based applications is an essential property to implementing ethics in a full meaning.

In Conclusion, as it has been stated generally all above that the role of ethics in software systems development is important to provide reliable and explainable artificial intelligence-based applications in the purpose of covering transparency and robustness. The harm potential of these systems, as well as actual real-life incidents of AI system failures and misuse, have resulted in a growing demand for AI ethics as a part of software engineering (SE) endeavors (Vakkuri V., Kemell K., Abrahamson P., 2019).Overall software engineering ethics has several perspectives like product, public, judgment, client and employer, management, profession, colleagues, self which prevents raising unanticipated consequences and provides transparency & robustness to the user. But on the other hand, there are still some doubts about that how much of the biasing in the artificial intelligence-software products can be eliminated? We can spend much more time to continue make researches about software engineering ethics provide more reliable, explainable, and robust artificial intelligent-based software products, and eliminate biases.

**References**

1. Borenstein, J. (2008), The Ethics of Autonomous Military Robots. Studies in Ethics, Law, and Technology. *Georgia Institute of Technology-Main Campus.* (Volume 2, Issue 1).Retrieved from <https://www.degruyter.com/view/j/selt.2008.2.1/selt.2008.2.1.1036/selt.2008.2.1.1036.xml>.
2. Wild-Raidit, C. (2020), AI code of ethics: Bosch sets company guidelines for the use of artificial intelligence*. Bosch Media Service*. Retrieved from <https://www.bosch-presse.de/pressportal/de/en/ai-code-of-practice-bosch-sets-company-guidelines-for-the-use-of-artificial-intelligence-208384.html>.
3. Office of the Surgeon Multinational Force-Iraq and Office of the Surgeon General United States Army Medical Command. (2006). The Mental Health Assessment Team (MHAT) IV Operation Iraqi Freedom 05-07: Final Report. *Office of the Surgeon General (Army), Washington, DC*. Available at http://www.armymedicine.army.mil/news/mhat/mhat\_iv/MHAT\_IV\_Report\_17N
4. Bostrom N. & Yudkowsky E., (2011), The Ethics of Artificial Intelligence. *NewYork: Cambridge University Press*. Available at

<https://nickbostrom.com/ethics/artificial-intelligence.pdf>

1. Sampaio C. S., Capelli C., (2010), Software Transparency. *Softwaretransparenz.WIRTSCHAFTSINFORMATIK.* Available at <https://link.springer.com/article/10.1007%2Fs12599-010-0102-z>
2. Whittaker, J., Meredith, R., Crawford, K., Dobbe, R., Genevieve, F., Elizabeth, K., Varoon M., West S. M., Richardson, R., Schultz, J., and Schwartz, O., (2018), AI now report. *Biodiversity Research Group, School of Geography and the Environment, University of Oxford.*
3. Bogost, I., (2015), The Cathedral of Computation. *The Atlantic.* Available at http://www.theatlantic.com/technology/archive/2015/01/ the-cathedral-of-computation/384300/
4. Vakkuri V., Kemell K., Abrahamson P., (2019), Implementing Ethics in AI: Initial results of an industrial multiple case study. *The University of Jyväskylä.* Available at <https://arxiv.org/abs/1906.12307>
5. Osoba O., Welser W.,( 2017), An Intelligence in Our Image, The risks of bias and errors in artificial intelligence*. RAND Corporation, Santa Monica, Calif.* Available at

<https://www.rand.org/content/dam/rand/pubs/research_reports/RR1700/RR1744/RAND_RR1744.pdf>

1. Gotterban D. & Miiller K. & Rogerson S., (1997, Vol. 40, No. 11), Software Engineering code of ethics. Available at <https://dl.acm.org/doi/10.1145/265684.265699>
2. Anderson M.&Anderson L. S., 2007, Machine Ethics: Creating an Ethical Intelligent Agen ( Vol. 28, No. 4 ). *Communications of the ACM.* <https://www.aaai.org/ojs/index.php/aimagazine/article/view/2065>
3. Etzioni A., Etzioni O., (2017), Incorporating Ethics into Artiﬁcial Intelligence. *The Journal of Ethics.* <https://link.springer.com/article/10.1007/s10892-017-9252-2>
4. Ehrhart, C., 2020, In brief: Bosch code of ethics for AI. *Corporate Department Communications & Governmental*

*Affairs.* <https://assets.bosch.com/media/en/global/stories/ai_codex/bosch-code-of-ethics-for-ai.pdf>

1. Ethics Commission's complete report on automated and connected driving (2017). *The Federal Ministry of Transport and Digital Infrastructure, Germany.*

https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.html

1. Ethics Guidelines for Trustworthy Ai (2019). *European AI Alliance.* https://ec.europa.eu/digital-singlemaket/en/news/ethics-guidelines-trustworthy-ai