| A2  Group 02  Abdalrahman Afifi  Hunter Lavender  Travis Mueller  Zachary Wildasin |
| --- |

# **Part I: Ethics in Traditional Engineering Projects:**

The Therac-25 is a cautionary tale in the history of medical technology, highlighting the critical importance of ethical considerations in the development and deployment of medical devices. This radiation therapy machine, designed for cancer treatment, came to symbolize a series of ethical failures that had catastrophic consequences. Here, we will examine the ethical aspects of the Therac-25's development and the key factors that went wrong. The Therac-25 was run mostly off of software rather than physical mechanisms leaving much more room for edge case errors to mess with its functionality. In 1983 AECL, the company that developed the Therac-25, did a safety analysis on the machine and launched the products to be sold to the public. That safety analysis did not study the software of the machine rather they just looked at the physical mechanisms of it. The software was based off of code written for an older model the Therac-20 that had the physical safety mechanisms that would prevent the operator from doing something to endanger the patients safety, the Therac-25 did not have those physical mechanisms. AECL decided that for the Therac-25s system they would use computer control only. They removed all the manual controls and safety mechanisms that the Therac-20 had. The computer was supposed to handle machine setup and shut down if it was performing dangerously.

After the Therac-25s launch in 1983 the worlds health care professionals started using the device commercially and for two years the machines operated as intended until the first incident in 1985. A woman came in to be treated for breast cancer with an intended dose of 200 Rad. When the machine started up the woman felt like she was being burned. The operator said that that was impossible but little did he know that his patient was just hit with at least 10000 rads, almost 50 times her intended dose. The patient lived; she lost use of her left arm and lost her left breast because of the malfunction.

Later that same year, a 40-year-old patient was undergoing her 24th Therac-25 treatment for cervical cancer. During the treatment, the Therac-25 machine encountered an "H-tilt" error message that alerted the technician to an issue with the treatment process. The operator had encountered this error before and did what the manual said to do, however each time this happened it would throw the same error again. The operator called a technician for assistance, but the technician couldn't find anything wrong with the machine.the patient said she felt an “electric tingling shock” in her hip. She later died of the cervical cancer she was being treated for however had she not died she would have needed a full hip replacement as a result of her overexposure.

Another patient the same year as well fell victim to the Therac-25. After her treatment she developed a “striped burn pattern” on her hip which looked suspiciously close to the beam blocking strips of the Therac-25. The patient survived but needed skin grafts to treat the radiation burns.

In 1986 another patient was scheduled for his 9th treatment with Therac-25. When he laid on the table and the machine started up he felt an intense heat as well as a sharp pain which alerted him that something was wrong because the treatment is supposed to be painless. He got up and yelled for help. He received a massive overdose of radiation that resulted in him dying 5 months later.

Later that same year, Another patient of Therac-25 was receiving treatment for a patch of skin cancer on his ear. When the machine turned on the patient saw a bright flash of light and said he heard something akin to eggs on a frying pan. He said it felt like his face was on fire. He died a few weeks later because of radiation burns on his brain stem as well as his right temporal lobe.

The last incident occurred in 1987. The patient laid on the table and the machine started and stopped with the same “H-tilt” error message. And the operator started it again when it threw the same error. The patient complained about a “burning sensation” after the treatment. The patient died 3 months later because of complications caused by the radiation overdose.

When going over these incidents one main question comes to mind: how were these incidents able to happen over a 19 month time span?

When the first accident happened in 1985, the company that developed the Therac-25, AECL was told about this malfunction and asked whether there was something wrong with the machine. After a few days AECL responded saying that an incident like that of the first was impossible. This confused the hospital because the extent of the burns the patient had showed that she had received somewhere from 15000 to 20000 rads. After these responses the patient sued the hospital and AECL and even after hearing about being sued AECL maintained that such an event was impossible.

When the second incident occurred AECL took it a little bit more seriously. After being informed, AECL sent a service engineer to the hospital where it occurred to figure out what happened. He tried unsuccessfully to reproduce the malfunction but he assumed that the problem was caused by a microswitch that’s function was to control the position of the turntable which controlled the radiation dose. AECL made some hardware changes to fix this possible issue however they did not discern the actual cause.

After the third incident AECL decided to dig their heels in and sent the hospital reporting it a letter that said "after careful consideration, we are of the opinion that this damage could not have been produced by any malfunction of the Therac-25". This trend in AECL’s response was unfortunately continued until the sixth incident only because the FDA got involved.

After the FDA deemed the Therac-25 defective stopping all commercial use right after the sixth incident. AECL received a notice from a physicist that worked at the hospital that two of the incidents occurred saying that he figured out how to reproduce the errors that caused the incidents. This allowed AECL to identify the issues and present a plan for fixing all the malfunctions to the FDA. This plan includes changes to over twenty hardware components to add the physical safety mechanisms and to fix all the software bugs for redundancy.

The entity responsible for all the factors that led to the incidents was AECL. They used very bad engineering practices and did not properly test their code. AECL only assigned one programmer to create and test the code. They removed safety mechanisms that previous machines had and they ignored user feedback. AECL violated the IEEE code of ethics. They did not follow rule one “to hold paramount, the safety, health, and welfare of the public, to strive to comply with ethical design and sustainable development practices, to protect the privacy of others, and to disclose promptly factors that might endanger the public or the environment;” by not properly testing their device and only using one programmer. AECL also violated rule number five “to seek, accept, and offer honest criticism of technical work, to acknowledge and correct errors, to be honest, and realistic in stating claims or estimates based on available data, and to credit properly the contributions of others'' by not listening to the feedback on the danger of the device from the hospital. Then the programmer violated IEEE Code of Ethics rule six by accepting a job to work by himself with a machine he didn't have proper knowledge “to maintain and improve our technical competence and to undertake technological tasks for others only if qualified by training or experience, or after full disclosure of pertinent limitations;” however I believe this is the least malicious violation for all we know he needed the job and his bosses ignored his complaints.

**Part II: Ethics Involving Artificial Intelligence Projects:**

As Artificial Intelligence becomes increasingly popular we are beginning to encounter a number of moral and ethical issues. A specific domain where AI is being heavily utilized is autonomous vehicles. Autonomous vehicles are self-driving cars equipped with various sensors and cameras that use AI algorithms which enable them to navigate, make decisions, and interact with their environment without human intervention. AI played a crucial role in enabling autonomous vehicles. AI technologies, such as deep learning for image recognition, sensor fusion, and machine learning for decision-making, have been central to the development of perception, navigation, and control systems. Regulatory and safety challenges have also emerged, prompting discussions on standards and guidelines for autonomous vehicles.

Tesla's autonomous driving system, known as "Autopilot" and "Full Self-Driving" (FSD), relies on a combination of hardware and software to enable advanced driver-assistance features and eventually, full self-driving capabilities. Tesla's AI system relies on neural networks, a type of deep learning algorithm, for image and sensor analysis. These neural networks process the data and use it to understand the environment and identify objects, lane markings, traffic signs, and pedestrians. The AI system uses machine learning to improve over time by continuously updating its algorithms based on real-world driving data collected from all of the Tesla vehicles being driven around the world. The system combines data from all sensors, including cameras, radar, and ultrasonic sensors, to create a comprehensive view of the vehicle's surroundings. This sensor fusion is essential for robust perception and decision-making.

As stated above, numerous ethical and moral issues around autonomous vehicles that have integrated AI systems have been brought to light. While there are several, one of the most significant ethical concerns is the safety of autonomous vehicles. Accidents involving self-driving cars have raised questions about who is responsible when accidents occur, how to prioritize human safety in artificial intelligence algorithms used by these vehicles, and how to ensure the technology is sufficiently reliable.

Another ethical issue surrounding this domain is the privacy of the vehicle owner. Autonomous vehicles collect vast amounts of data, including location information and sensor data. This raises concerns about user privacy, data security, and potential misuse of this data by both manufacturers and third parties.

Diving deeper into the issue of safety, there is the specific instance of decision making in ethical dilemmas. Autonomous vehicles may face ethical dilemmas, such as deciding how to respond in situations where harm to passengers and/or pedestrians is unavoidable. These decisions raise questions about the programming of AI systems and the values they will be programmed to prioritize.

We have brought to light some of the ethical and moral issues surrounding autonomous vehicles integrating AI into their systems, but we want to showcase in what ways autonomous vehicle systems align with the newly established principles of the ‘AI Bill of Rights’. The AI Bill of Rights outlines several principles for AI development and deployment, including transparency, fairness, and privacy.

As far as Tesla’s autonomous vehicles are concerned with transparency, they have done a fairly good job thus far. Tesla has provided some level of transparency regarding its autonomous driving system, particularly through its user manuals and public communications. However, the level of transparency regarding the specifics of the AI algorithms, data collection, and system decision-making is a point of debate. We believe that their transparency could be improved by providing more information about how the AI systems function, how data is used, and how certain ethical dilemmas are handled inside the program. In terms of IEEE code of ethics, this closely relates to rule 5 which is to seek, accept, and offer honest criticism of technical work, while also acknowledging and correcting errors.

The overall fairness of Tesla’s systems up to this point seem as good as we would all expect them to be, and as they should be. Fairness is crucial in autonomous vehicle decision-making. Tesla should continue to implement rigorous testing and validation processes to detect and rectify biases and ensure that its AI system treats all equally, whether it be the vehicle owner or pedestrians. The IEEE code of ethics rule 7 also states that all persons need to be treated fairly and with respect, which means not engaging in any form of discrimination.

Perhaps one of the most important areas of AI development, privacy of user data is crucial and Tesla’s AI driven vehicles are no exception. These systems collect a significant amount of data, including location data and sensor information. Protecting user privacy and ensuring data security are of the utmost importance. Above all else, Tesla must have robust data security measures in place to safeguard user data, obtain consent for data usage, and be transparent about data collection and retention policies. Seeing that this is one of the most important areas of AI development, it comes as no surprise that in the IEEE code of ethics the very first rule is “...protect the privacy of others…” and to uphold high standards of safety, health, and general welfare of the public.

## **Part III: Ethical Issues that are related to your own project**

Algorithm research is a dynamic field poised for significant technological advancements. This article delves into ethical considerations concerning data security, intellectual property rights, and the integration of free open-source software (FOSS) for educational purposes in algorithm research. It explores how these ethical concerns impact algorithm research and provides guidance on upholding responsible and ethical practices.

Data security is paramount in algorithm research. Protecting data from collection to storage and analysis is essential to prevent unauthorized access or breaches. Neglecting data security can result in severe consequences, such as privacy breaches and data leaks.

While pursuing algorithmic innovation, researchers must not compromise data security for cost-saving measures. Ethical considerations should always outweigh financial concerns. A thorough cost-benefit analysis is crucial to strike the right balance.

Establishing and adhering to baseline security standards for research networks and systems is vital. This ensures the protection of sensitive data. Robust security measures, including firewalls, access controls, and regular security audits, are necessary to safeguard research data.

Regular security audits and assessments are essential to identify vulnerabilities and maintain data integrity and confidentiality. Ethical researchers understand that data security is an ongoing process requiring continuous improvement.

Data encryption is fundamental to enhancing data security. Implementing encryption protocols for data storage and transmission makes it challenging for unauthorized parties to access sensitive information, ensuring privacy and confidentiality.

Ethical concerns extend to intellectual property rights. Researchers should clearly define ownership and rights for novel algorithms and methodologies developed during projects, preventing disputes and protecting creators.

Recognizing the contributions of individuals or teams to intellectual property is imperative. Proper attribution upholds ethical standards and fosters a culture of trust and collaboration within the research community.

Comprehensive documentation of intellectual property arrangements prevents disputes and ensures the ethical handling of ownership rights. Maintaining records of agreements and contributions is essential.

Stressing compliance with intellectual property laws is crucial for the protection of creators and collaborators. Ethical researchers adhere to these laws and respect others' rights.

Promoting transparency in intellectual property discussions and agreements maintains trust and fairness in research collaborations. Open communication about intentions, expectations, and ownership arrangements fosters integrity and openness.

FOSS plays a vital role in ethical algorithm research. It provides cost-effective access to advanced algorithmic tools, fosters collaboration, and aligns with ethical principles of transparency and openness. FOSS projects often enforce ethical use and redistribution through licenses, contributing to responsible sharing within the community.

In algorithm research, ethics must be at the forefront of every researcher's mind. Prioritizing data security, conducting ethical cost-benefit analyses, adhering to security standards, safeguarding intellectual property rights, and leveraging FOSS for ethical research are pivotal steps in ensuring responsible and ethical contributions to the field. These principles uphold the trust and integrity of algorithm development while driving technological advancements.

The IEEE code of ethics rules we have to watch out for during development are one and seven. When it comes to the cybersecurity part of our project we have to follow rule one: “to hold paramount, the safety, health, and welfare of the public, to strive to comply with ethical design and sustainable development practices, to protect the privacy of others, and to disclose promptly factors that might endanger the public or the environment;” and make sure we have no vulnerabilities that can be exploited to protect our users private information. When it comes to regulating the use of our platform we have to follow rule number seven: “to treat all persons fairly and with respect, and to not engage in discrimination based on characteristics such as race, religion, gender, disability, age, national origin, sexual orientation, gender identity, or gender expression;” by making sure our users understand that their data sets might be biased and help them to realize that and making them aware of this as well as giving them the tools to correct such an error.zs