

Artin Kiany  
4700 Keele St  
North York, ON M3J 1P3  
(416) 736-2100  
[artin59@my.yorku.ca](mailto:artin59@my.yorku.ca)

Mar.27.2025

Paul Cotton  
Standards Council of Canada  
600-55 Metcalfe Street  
Ottawa ON K1P 6L5

Dear Dr. Cotton,

In accordance with our agreed date, March.27.2025, I am writing to inform you about the technical report you requested, titled *Ensuring AI Safety in Robotic Systems: The Need for Explainable AI and Industry Standards*.

This report provides an analysis of different explainable AI (XAI) techniques that can be used in tandem with potential industry-specific standards, to fill the gap and establish a solution to the lack of safety guidance and best industry practices for AI-embedded robotic systems such as autonomous vehicles, robot-assisted surgery, manufacturing robots and drones. This report focuses more on potential solutions for the safety certifications of autonomous vehicles.

If you have any questions please don't hesitate to reach out and contact me at [artin59@my.yorku.ca](mailto:artin59@my.yorku.ca), or (416) 736-2100. Thank you for the opportunity to contribute to this project, and I look forward to hearing from you.

Sincerely,  
Artin Kiany

# Ensuring AI Safety in Robotic Systems: The Need for Explainable AI and Industry Standards

Artin Kiany #219457969

March 27th, 2025


Eng 2003 Section O

Submitted in fulfilment of the requirements of the PEO Engineering Report

---

**My signature below attests that this submission is my original work**

Following professional engineering practice, I bear the burden of proof for original work. I have read the Policy on Academic Integrity posted on the Lassonde school of Engineering website and confirm that this work is in accordance with the Policy.

**Signature:** 

**Date:** Mar.27.2025

# Executive Summary

The rapid technological advancement in AI-embedded robotic systems has outpaced the development of safety certifications that are able to manage the probabilistic nature of AI algorithms. Missy Cummings, a professor and director of the Humans and Autonomy Lab at Duke University has highlighted her worries about the future of this technology because we currently have no implemented frameworks in place to verify the safety of these systems [1]. This report identifies and analyzes potential solutions to this challenge. It demonstrates how different explainable artificial intelligence (XAI), which are a collection of techniques that help improve transparency, can be implemented into these stochastic systems to ensure reliability. A limitation of these techniques is the tradeoff between performance and transparency. This limitation has been addressed by the use of hybrid architectures that employ higher transparency-based models for safety-critical systems while implementing more simple and efficient models for more complex and computationally intensive systems that require performance.

In addition to XAI, the current gap in industry-specific standards has been highlighted. These certification protocols are essential and beneficial to all stakeholders involved, by ensuring these systems meet safety, ethical and performance standards. These two methods can work in tandem together to create reliable solutions. Regulators and standardization organizations like the International Organization for Standardization (ISO) can leverage the data from the inner thinking/decision-making processes of implemented AI systems to create evidence-based safety and trustworthiness standards.

# Table Of Contents

Executive Summary.....	ii
1.0 Introduction & Background.....	1
2.0 Artificial Intelligence In Robotic Systems.....	2
2.1 Autonomous Vehicles.....	2
3.0 Explainable Artificial Intelligence (XAI).....	4
4.0 Industry Standards.....	7
4.1 ISO.....	7
4.2 ISO/IEC JTC 1/SC 42.....	8
5.0 Discussion.....	9
6.0 Conclusion.....	10
7.0 References.....	11

## **List of Tables**

Table 1. Summary of levels of driving automation with market examples .....	3
Table 2. Comparison of AI verification methods [7, 22, 23, 24] .....	5
Table 3. An overview of different XAI techniques used in robotic systems .....	6

# Introduction & Background

One of the key challenges we are facing in the new age of AI is safety. This issue is especially prevalent in robotics systems that incorporate artificial intelligence such as autonomous cars, drones, surgical and manufacturing robots. Missy Cummings, Professor and Director at Duke University's Humans and Autonomy Lab, has highlighted these growing concerns with AI-driven robotics. While AI algorithms come in many forms, it ultimately boils down to mathematics, statistics, and probability; as uncertainty increases in the world, driven by rapid technological advancements, political turmoil, environmental changes, and economic fluctuations, so does an algorithm's capacity to successfully and safely arrive at a solution.

Despite these challenges, AI-driven robotic systems have already demonstrated remarkable success. In the medical field, AI-assisted robotic surgery has improved precision and reduced complications, enhancing patient outcomes [2]. Sadly, in the U.S., according to the National Highway Traffic Safety Administration (NHTSA), 42,795 people died in vehicle crashes in 2022 [3], where 94% of those crashes are due to human error. AVs have the potential to remove/reduce human error and decrease casualties and accidents. [4, 5].

However, given the inherent uncertainties of the real world, allowing these algorithms to interact with and manipulate physical environments through these robotic systems, makes it exponentially more challenging to guarantee safety. Unfortunately, the absence of universally accepted methodologies or industry standards for testing and validation, further complicates the widespread adoption of this technology, limiting the full benefits AI-driven robotics can offer.

This report will take a theoretical approach and will explore these challenges and examine how explainable AI (XAI) and AI safety standards can enhance reliability, safety and gain public trust. This report will not cover the broader ethical implications of AI beyond safety, nor will it delve into AI applications outside of robotics.

# Artificial Intelligence In Robotic Systems

Across industries, artificial intelligence in robotic systems is revolutionizing efficiency, automation and productivity. Robots have traditionally been used in tasks that are deemed too hard or risky for humans to be able to do such as lifting heavy equipment or working in extreme conditions, or tasks that are too repetitive like assembling vehicles. Currently, most robots are not artificially intelligent and they have very limited functionality, for example: “all industrial robots could only be programmed to carry out a repetitive series of movements” [6]. But AI introduces a whole new world of functionality, with autonomous vehicles and surgical robots being the most notable examples.

If implemented correctly, AI-embedded robots offer numerous benefits, transforming industries by increasing efficiency, precision, and adaptability. These intelligence systems can automate complex tasks, reduce human error, and improve decision making by analyzing vast amounts of data in real-time from a robot's array of sensors. They can also adapt and learn continuously without human intervention. However, successful implementation requires careful consideration of ethical, safety, security, and job displacement concerns.

## 2.1 Autonomous Vehicles

An autonomous vehicle (AV) as defined by the University of Michigan’s Center for Sustainable Systems is one that uses “technology to partially or entirely replace the human driver in navigating a vehicle from an origin to a destination while avoiding road hazards and responding to traffic conditions” [5]. Furthermore, AVs are not only limited to cars, they also include off-road applications such as in agriculture, mining and construction settings, and also in other areas like maritime environments in ports and self-navigating vessels [7]. Artificial Intelligence plays a crucial role in AV development, enhancing perception, and decision-making through complex algorithms, reinforced training and also multi-sensor fusion techniques. However, concerns about the safety of these machines remain. Without a global safety standard, it’s difficult to implement AVs on a large scale.

Anthony Corso, a postdoctoral scholar in aeronautics and astronautics and executive director of the Stanford Center for AI Safety, says that their “main challenge as a field is how do we guarantee that these amazing capabilities of AI systems—driverless cars and pilotless planes—are safe before we deploy them in places where human lives are at stake?” [8]. The first step to guarantee safety and the first step in any process like making laws and stuff are classifications and definitions. That's where J3016\_202104 comes in. In collaboration with the International Organization for Standardization (ISO) whose role in AV safety certification is further discussed in Section 4.1, the Society of Automotive Engineers (SAE) has defined 6 levels of driving automation ranging from level 0 (no driving automation) to level 5 (full driving automation) [9].

AVs have the potential to reduce a lot of traffic incidents; they most noticeably help in reducing accidents caused by human error, fatigue and distraction. Moreover, the safety of AVs depends on their level of automation as shown in Table 1. A study done by the Nature Communications journal analyzed traffic accidents of 2100 AVs and 35,133 human-driven vehicles (HDV). The study found that involving AVs generally occurs less frequently than those involving HDVs in the majority of similar accident scenarios [10]. However, AVs are prone to higher rates of accidents in certain conditions. The study highlights that AVs are 5.25 times more likely to be involved in accidents under dawn/dusk lighting conditions and 1.98 times more likely during turning maneuvers [10]. This suggests that the sensors and algorithms struggle with object detection and identifying potential hazards because of the sun's shadows and reflections.

Table 1. Summary of levels of driving automation with market examples (Adapted from SAE J3016 [9])

Level	Automation Type	Description	Human Role	Market Examples
0	No Driving Automation	The allowed features are only limited to “warnings and momentary assistance” such as blind spot and lane departure warnings [9]	The driver is engaged and in full control all the time.	Traditional Cars (e.g., Toyota Corolla, Ford F-150)
1	Driver Assistance	The vehicle only assists with either steering or acceleration/deceleration, but cannot perform both.	The driver must be engaged at all times and must supervise the automation system and intervene if necessary	Adaptive cruise control in Honda Civic, Toyota safety sense
2	Partial Driving Automation	The vehicle can assist with both steering and acceleration/deceleration.	The driver must monitor the automation system and intervene if necessary.	Tesla Autopilot, GM Super Cruise, Ford BlueCruise [7]
3	Conditional Driving Automation	The vehicle can perform all driving tasks only under specific conditions.	The driver is not driving when the automated system is engaged, however, they must drive if the system requests	Mercedes-Benz Drive Pilot (S-Class, EQS), BMW (7 Series) [7]



Table 1 (continued).

4	High Driving Automation	The vehicle is completely autonomous in limited environments, such as designated highways or urban areas (at low speeds)	The system will not require the driver to drive.	Waymo One, Cruise AV (San Francisco) [7]
5	Full Driving Automation	Same as level 4 however, at this level the vehicle may drive under all conditions	No human driver is needed at any stage.	No mass-market vehicles yet

Despite these challenges, AVs significantly reduce accident severity by helping control the vehicle and alerting drivers to potential dangers. Furthermore, projections suggest that “if AVs were to be introduced with an average safety level ten percent higher than that of the typical human driver, approximately 600,000 fatalities could be averted in the United States over a span of 35 years” [10]. These findings emphasize the value and potential in AVs while also highlighting the need for continuous improvement and testing in sensor technology and decision-making algorithms.

## Explainable Artificial Intelligence (XAI)

AI and machine learning algorithms used in robotics such as deep learning (DL), or convolutional neural networks (CNNs) are extremely computationally intensive and complex models, and they often function as black boxes, which are systems whose internal operations or the way they think, are a mystery to its users. The complexity in these algorithms stems from multi-layered neural networks (working similarly to neurons in our brains) that learn patterns autonomously [7]. Moreover, since robotic systems must handle dynamic environments like roads, or surgeries, they require adaptability and the ability to process vast amounts of data, from a multitude of sensors to make reliable decisions. Given these complexities, robust verification methods are necessary to ensure safety.

AI validation methods fall under 2 categories: Black box and White box testing. Some of the most popular black box verification methods are explained in Table 2. The thing all of them have in common is the fact that they only evaluate inputs and outputs but fail to detect any biases and do not address the reason why a model behaves the way it does. Meanwhile, white box testing seeks to provide insights into the internal framework of the model. These verification methods provide more transparency and higher confidence levels, however, they are more computationally expensive and have scalability issues.

Table 2. Comparison of AI verification methods [7, 22, 23, 24]

Verification method	Description	Strengths	Limitations	White Box	Black Box
Formal Verification	Testing using mathematical proofs	<ul style="list-style-type: none"> <li>- Allows for regulatory compliance</li> <li>- Guarantees safety through mathematical proofs</li> </ul>	<ul style="list-style-type: none"> <li>- Doesn't represent the real world</li> <li>- Cannot be used for complex models</li> </ul>	✓	
Human-in-the-Loop	A human monitors the AI decisions	<ul style="list-style-type: none"> <li>- Can adjust AI behaviours in real-time</li> </ul>	<ul style="list-style-type: none"> <li>- Not scalable</li> <li>- Introduces extra biases</li> </ul>		✓
Benchmarking	Testing models against standardized benchmarks	<ul style="list-style-type: none"> <li>- Scalable</li> <li>- Assures regulatory compliance</li> </ul>	<ul style="list-style-type: none"> <li>- Disregards dynamic environments</li> </ul>		✓
Simulation-based testing	Testing models in a controlled environment	<ul style="list-style-type: none"> <li>- Models can be put through various scenarios</li> </ul>	<ul style="list-style-type: none"> <li>- Not a fully accurate representation of real-life dynamics</li> </ul>		✓
Adversarial Testing	Testing edge cases	<ul style="list-style-type: none"> <li>- Can reveal the vulnerabilities</li> </ul>	<ul style="list-style-type: none"> <li>- Identifies failures but might not explain the root cause</li> </ul>		✓

While the black box and white box testing methods focus on evaluating the performance and functionality of AI models, they fall short of explaining why a model thinks the way it does. This is where explainable AI (XAI) comes in. XAI is a specialized field within the discipline of AI that “focuses on designing and developing techniques and models that are interpretable and comprehensible to all stakeholders” [11]. It bridges the white box and black box approaches by interpreting the complex structure of a model's thoughts regardless of the underlying model's transparency. XAI also promotes end user trust, and can help to mitigate the deeply rooted pre-existing biases in AI systems. More specifically in AVs XAI can elucidate critical decisions like emergency braking, collision avoidance, and path planning. This not only boosts reliability and safety by helping engineering and developers identify flaws in the system, but also lays the groundwork for industry-wide standardization. In fact, regulatory efforts are already reinforcing this need, as the EU AI Act states that “AV manufacturers and in-vehicle software suppliers are

obligated to fulfill rigorous requirements to ensure compliance with standards related to transparency, traceability, safety, data governance, and accountability” [11].

Table 3 shows some of the most popular XAI techniques that are currently deployed. It's important to select the appropriate or even a combination of them to capitalize on their strengths based on the individual specifications and intended applications. For example: for perception systems “Grad-CAM method can analyze DL detection models by generating heatmaps that visually explain the road semantic segmentation outputs thereby providing a comprehensive understanding of the relevance of their outcomes” [11]. For more complex systems multiple XAI techniques need to be implemented. However, the most challenging issue that needs to be considered is the tradeoff between performance and transparency. Different industries also have different demands, for example in the healthcare industry transparency has a higher priority due to patient privacy, and safety-making techniques like SHAP and LIME as explained in Table 3 are more desirable.

Table 3. An overview of different XAI techniques used in robotic systems (Adapted from [7])

<b>XAI Technique</b>	<b>Type</b>	<b>Implementation Method</b>	<b>Example Uses</b>	<b>Strengths</b>	<b>Limitations</b>
Decision Trees	White-Box	Makes decisions through binary tree splits	Basic path planning decisions	- Easy to understand	- Limited complexity
LIME	Model-Agnostic	Local Interpretable Model-Agnostic Explanations, creates a local approximation of a specific prediction of black box model	Explaining why an AV made a specific lane change	- Efficient - Simple and intuitive	- Sensitive to small disruptions - Limited to precise decisions
SHAP	Model-Agnostic	Shapley Additive Explanations, explains the contribution of specific inputs to the overall output	Comparing current wheel positioning to future vehicle positioning	- Versatile - More accurate than LIME method	- High computational cost

Table 3 (continued).

Grad-CAM	Model-Specific	Gradient-weighted Class Activation Mapping is used for visual explanation of decisions based on image recognition	Creating heatmaps to show road obstacles [19]	- Intuitive and visual	- High computational cost
Saliency Maps	Model-Specific	Highlights which input features have the most significant effect on the output	Understanding the importance of certain features	- Intuitive and visual	- Sensitive to small disruptions

## Industry Standards

As AI-powered robotic systems are increasingly being incorporated into safety-critical industries, the need for international and national, industry-specific standardization and regulatory bodies has never been more urgent. According to the Standards Council of Canada (SCC), “a standard is a document that provides guidelines, characteristics or requirements for products, processes or services. It is developed by a committee or group of stakeholders and approved by a recognized body” [12]. These standards offer significant benefits to all stakeholders, including customers, businesses, and regulators. It gives customers peace of mind knowing the products meet high quality and safety standards. It allows for cross-industry adoption and gives businesses a competitive edge. Finally, it provides regulators with technical references that can be used in policy formulation. Moreover, according to the SCC, standardization has contributed to a significant development of the Canadian economy. Noting that “17% GDP growth is driven by standardization”, and that “68% of companies [they’ve] worked with report increased exports, jobs, or revenue as a direct result of [their] engagement” [17]. This highlights the crucial role that standardization plays in economic growth, and global trade opportunities.

### 4.1 ISO

The International Organization for Standardization (ISO) is an independent organization that is made up of the national standard bodies of 173 countries. Their goal is to research and develop international standards, for anything from product development to process management [13]. Canada is being represented by the Standards Council of Canada (SCC).

The ISO has many standards in regards to AI and robotics. For example, ISO/TS 5083, which is a technical specification (TS) on “Road vehicles — Safety for automated driving systems — Design, verification and validation” [14]. Its goal is to “provide an overview and guidance of the steps for developing and validating an automated vehicle equipped with a safe automated driving system” [14], and more specifically of SAE level 3 and 4 vehicles, described in Table 1. The ISO has many more standards on AI, however, it is lacking in standards on AI-embedded robotic systems. This is a huge gap that needs to be filled. The lack of comprehensive standards on this topic poses significant challenges in ensuring safety, reliability and transparency across industries.

## **4.2 ISO/IEC JTC 1/SC 42**

The International Electrotechnical Commission (IEC) is “the world’s leading organization for the preparation and publication of international standards for all electrical, electronic and related technologies” [15]. In collaboration with the ISO, it established Joint Technical Committee 1, Subcommittee 42, which is responsible for AI standardization. There are many working groups under this subcommittee but the most relevant are JWG 4: Functional Safety and AI Systems, and WG 3: Trustworthiness. There are currently no published documents by these working groups, however, there is currently a standard in the preparatory phase: ISO/IEC TS 22440, which “will be instrumental in ensuring reliability across AI applications in a wide number of areas where safety is paramount, such as for autonomous vehicles, medical devices and industrial control systems” [16]. It is expected to cover risk assessment, hazard analysis, transparency, explainability, validation, verification, and safety analysis [16]. Industry standards like ISO/IEC TS 22440, will be game changers in the near future as they provide a structural framework for implementing new industry practices in safety, and transparency of AI robotic systems for all stakeholders involved.

## Discussion

Due to its inherently probabilistic nature, the safety certification of robotic systems that employ artificial intelligence presents an intricate and complex concern, especially since the technology is progressing day by day. As discussed before, if AI-embedded robotic systems could reach a high level of innovation and development, they have the potential to revolutionize various industries. However, three critical barriers must be addressed: the performance-transparency tradeoff, current industry-specific standardization frameworks, and the complication with the dynamic nature of AI.

The tradeoff between the transparency and performance of AI-driven robotic systems is one of the main challenges engineers face when designing and implementing XAI techniques. In addition, according to [18] “while significant progress has been made in developing techniques that enhance the interpretability of AI models, achieving a balance between high performance and transparency remains a complex task”. The most effective solutions identified thus far are hybrid architectures that allow for a combination of more interpretable systems for safety-critical functions such as SHAP and LIME, and high-performance techniques for systems that call for maximized efficiency and performance.

Current AI standards remain too generic, and since different industries have a wide variety of uses for AI in robotic systems, there needs to be an increased effort in developing specific benchmarks for safety. To help with these efforts, standardization organizations should collaborate with industry innovators and use XAI techniques to comprehend and address the unique challenges of AI safety for each relevant industry. A further complication that comes with AI safety is that, unlike traditional software, AI systems continue to evolve even after deployment. This necessitates a dynamic certification approach that can accredit models without a need to go through a full recertification process. This can be done through continuous monitoring and occasional audits. This approach could leverage embedded explainability logs to track decision reasoning, which enables standardization organizations, regulators and auditors to monitor and verify consistency with the current safety standards. Lastly, future research into the further development of XAI techniques catered to specific industries is recommended.

# Conclusion

Certifying AI-driven robotic systems demands a fundamental shift in the approach from current safety models to a more specialized and dynamic framework that accommodates the stochastic nature of these systems. When successfully implemented, safe AI-enabled robotic systems will transform our industries by paving the way to a new generation of transportation systems, medical procedures, manufacturing, and so much more. This report highlights the use of XAI to foster transparency and end-user trust. While these techniques come with tradeoffs and limitations, they show great promise and can be used in tandem with industry-specific standards to promote safety. The coming decade will reveal and test our ability to define a future for autonomous technologies.

# References

- [1] B. Lufkin, “50 grand challenges for the 21st Century,” BBC News, <https://www.bbc.com/future/article/20170331-50-grand-challenges-for-the-21st-century> (accessed Mar. 10, 2025).
- [2] J. E. Knudsen, U. Ghaffar, R. Ma, and A. J. Hung, “Clinical applications of Artificial Intelligence in robotic surgery,” *Journal of robotic surgery*, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10907451/> (accessed Mar. 4, 2025).
- [3] NHTSA, “Early Estimates of Motor Vehicle Traffic Fatalities And Fatality Rate by Sub-Categories in 2022,” NHTSA’s National Center for Statistics and Analysis , <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813660> (accessed Mar. 11, 2025).
- [4] NHTSA, “Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey,” NHTSA’s National Center for Statistics and Analysis , <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812506> (accessed Mar. 11, 2025).
- [5] University of Michigan, “Autonomous vehicles factsheet,” Center for Sustainable Systems, <https://css.umich.edu/publications/factsheets/mobility/autonomous-vehicles-factsheet> (accessed Mar. 10, 2025).
- [6] A. Owen-Hill, “What’s the difference between Robotics and Artificial Intelligence?,” What’s the Difference Between Robotics and Artificial Intelligence?, <https://blog.robotiq.com/whats-the-difference-between-robotics-and-artificial-intelligence> (accessed Mar. 27, 2025).
- [7] D. J. Yeong, K. Panduru, and J. Walsh, “Exploring the unseen: A survey of multi-sensor fusion and the role of Explainable AI (XAI) in Autonomous Vehicles,” *Sensors*, vol. 25, no. 3, p. 856, Jan. 2025. doi:10.3390/s25030856
- [8] A. Myers, “How AI is making Autonomous Vehicles Safer,” Stanford HAI, <https://hai.stanford.edu/news/how-ai-making-autonomous-vehicles-safer> (accessed Mar. 27, 2025).
- [9] SAE, “Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016\_202104,” SAE International , [https://www.sae.org/standards/content/j3016\\_201806/](https://www.sae.org/standards/content/j3016_201806/) (accessed Mar. 28, 2025).
- [10] M. Abdel-Aty and S. Ding, “A matched case-control analysis of autonomous vs human-driven vehicle accidents,” *Nature Communications*, vol. 15, no. 1, Jun. 2024. doi:10.1038/s41467-024-48526-4
- [11] D. J. Yeong, K. Panduru, and J. Walsh, “Exploring the unseen: A survey of multi-sensor fusion and the role of Explainable AI (XAI) in Autonomous Vehicles,” *Sensors*, vol. 25, no. 3, p. 856, Jan. 2025. doi:10.3390/s25030856



- [12] Admin, “Standards,” Standards Council of Canada, <https://scc-ccn.ca/standards> (accessed Mar. 27, 2025).
- [13] “About ISO,” ISO, <https://www.iso.org/about> (accessed Mar. 27, 2025).
- [14] “ISO/TS 5083,” ISO, <https://www.iso.org/standard/81920.html> (accessed Mar. 27, 2025).
- [15] Who we are, <https://iec.ch/who-we-are> (accessed Mar. 28, 2025).
- [16] “IEC and Iso Launch Working Group to advance functional safety of AI Systems,” IEC, <https://www.iec.ch/blog/iec-and-iso-launch-working-group-advance-functional-safety-ai-systems> (accessed Mar. 28, 2025).
- [17] Bianca.fernandes, “Canada’s standardization system,” Standards Council of Canada, <https://scc-ccn.ca/about-us/canadas-standardization-system> (accessed Mar. 27, 2025).
- [18] A. S. Adebayo, O. O. Ajayi, and N. Chukwurah, “Explainable AI in robotics: A critical review and implementation strategies for transparent decision-making,” *Journal of Frontiers in Multidisciplinary Research*, vol. 5, no. 1, Jan. 2024. doi:10.54660/ijfmr.2024.5.1.26-32
- [19] V. Buhrmester, D. Münch, and M. Arens, “Analysis of explainers of Black Box Deep Neural Networks for Computer Vision: A survey,” *Machine Learning and Knowledge Extraction*, vol. 3, no. 4, pp. 966–989, Dec. 2021. doi:10.3390/make3040048
- [20] “Engineering report guide,” Professional Engineers Ontario, <https://www.peo.on.ca/sites/default/files/2019-11/Engineering-Report-Guide.pdf> (accessed Mar. 28, 2025).
- [21] “Applications of AI in robots: An introduction,” University of San Diego Online Degrees, <https://onlinedegrees.sandiego.edu/application-of-ai-in-robotics/> (accessed Mar. 27, 2025).
- [22] S. Anjomshoe, A. Najjar, D. Calvaresi, and K. Främling, “Explainable agents and robots: Results from a systematic literature review,” in \*Proc. 18th Int. Conf. Auton. Agents MultiAgent Syst. (AAMAS’19)\*, Montreal, QC, Canada, 2019
- [23] T. Sakai and T. Nagai, “Explainable autonomous robots: A survey and perspective,” *Advanced Robotics*, vol. 36, no. 5–6, Feb. 2022. doi:10.1080/01691864.2022.2029720
- [24] D. Das, S. Banerjee, and S. Chernova, “Explainable AI for Robot Failures: Generating Explanations that Improve User Assistance in Fault Recovery,” *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, Mar. 2021. doi:10.1145/3434073.3444657
- [26] Bianca.fernandes, “Benefits of applying standards,” Standards Council of Canada, <https://scc-ccn.ca/standards/benefits-applying-standards> (accessed Mar. 27, 2025).
- [27] “ISO/IEC AWI TS 22440-1,” ISO, <https://www.iso.org/standard/89535.html> (accessed Mar. 28, 2025).

[28] “ISO/IEC JTC 1/SC 42 - artificial intelligence,” ISO, <https://www.iso.org/committee/6794475.html> (accessed Mar. 27, 2025).

# Revision summary

## **Peer 1 feedback:**

“The full report is not completely done. The only portion completed is the introduction. Formatting of section is good and concise in illustrating the problem the report is trying to address with references. The introduction explicitly states what is and is not going to be covered in the report. Overall the introduction is good, but paragraphs should have indentations at the beginning of all of them. Reference structure is good as well.”

## **Peer 1 response and revision summary:**

I have decided to not indent my paragraphs. There is no mention of a need to indent paragraphs in the Professional Engineers Ontario (PEO) engineering report guide [20]. Furthermore, there is no mention of indentation in any course material related to technical writing. Additionally, I got a very low score on the logic of my proposal because I hadn't developed my ideas further than the introduction and background section. So I took some time to research and brainstorm to further develop my logic and finish the main topics of my technical report which allowed me to express my proposal and expand it (especially in the discussion section).

## **Peer 3 feedback:**

“First of all, the report's outline is looking well organized. The breakdown of the letter of transmittal meets the outline's standards, as well as the title page, which includes all the requirements, and I appreciate the addition of academic integrity. The title page also meets the requirements, but I did have to count the pages to make sure it is accurate; therefore, I suggest numbering your pages for the report. As it appears you haven't done much of the writing regarding the other sections of the report except the introduction, due to this inconvenience, I won't be able to give you a full feedback regarding your report. Regarding the introduction, it was well written; I was able to understand the topic of discussion and its relevance to our day-to-day lives. I also appreciate the brief explanation of the robotics systems that incorporate artificial intelligence, as well as the examples you provided for it. Credibility is established from the beginning of the report, starting from the introduction, because of the good/relevant information like the mentions of effects of human error within the topic as well as the positive impact of AI within our world with numeric values that were researched and mentioned along with their citations.”

## **Peer 3 response and revision summary:**

I added page numbers to my technical report in accordance with [20] “Starting with the introductory chapter/section, the remainder of the report is numbered with Arabic numerals beginning with 1 and running consecutively to the end of the report”. Also based on the feedback I received on the rubric, I got a very low score in visual elements. So I added 3 tables to my report to be able to reference. These elements also simplify my explanations.

Overall I got very positive feedback which gave me confidence in my report, and allowed me to further develop my ideas, by knowing its strengths.