CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Political Studies

Department of Security Studies

Master's Thesis

Civil Applications of Autonomous Systems: Implications for the Military Sector

Name: Lorenza Fortunati

Academic advisor: Mg. Anzhelika Solovyeva, PhD.

Study programme: MISS

Year of submission: 2023

Declaration

1. I hereby declare that I have compiled this thesis using the listed literature and resources only.
2. I hereby declare that my thesis has not been used to gain any other academic title.
3. I fully agree to my work being used for study and scientific purposes.

In Rome on July 30th, 2023

Lorenza Fortunati

Reference

FORTUNATI, Lorenza. 2023. *Civil Application of Autonomous systems: Implications for the Military Sector*. Praha. 62 pages. Master's Thesis. Charles University, Faculty of Social Sciences, Institute of Political Science. Department of Security Studies. Supervisor Prof. Mg. Anzhelika Solovyeva, PhD.

Length of the thesis: 62 pages; 21,755 words; 136,077 characters.

Abstract

This thesis aims to examine the growing development of autonomous systems from a civilian and military perspective. The literature analysis has shown that in the military sector there is much access to an interactional-scientific and literary debate, especially on the growing possibility of autonomous weapons development. On the other hand, however, the civilian sector is increasingly advancing the development of autonomous systems. Many of these are already being applied in civilian society: from self-driving cars to medical equipment. Nonetheless, the study showed that although these sectors are going at different speeds, the risks and challenges related to the topic of autonomous machines are very similar, especially with regard to the concept of responsibility and human-machine nexus. For this reason, this thesis aims to analyze the civilian experience in the field of autonomy to determine whether there might be any transferable lessons from the civilian to the military world for the future development of autonomous weapons. After a thorough analysis of the aircraft crashes involving the Boeing 737 MAX and the accidents caused by the self-driving cars, this thesis concludes that the best way to address the ethical and moral challenges of autonomous weapons is to develop and design such systems from a human-centric perspective first.

Key words: Artificial intelligence, Autonomy, Autonomous Systems, Human-Machine Nexus, Responsibility Gap

Title: Civil Applications of Autonomous Systems: Implications for the Military Sector

Abstrakt

Cílem této práce je prozkoumat rostoucí vývoj autonomních systémů z civilního a vojenského hlediska. Analýza literatury ukázala, že ve vojenském sektoru existuje velký přístup k interaktivní. vědecké a literární debatě, zejména o rostoucí možnosti vývoje autonomních zbraní. Na druhé straně však civilní sektor stále více postupuje ve vývoji autonomních systémů. Mnohé z nich se již uplatňují v civilní společnosti: od samořídících automobilů až po lékařské přístroje. Studie nicméně ukázala, že ačkoli tyto sektory postupují různou rychlostí, rizika a výzvy spojené s tématem autonomních strojů jsou velmi podobné, zejména pokud jde o koncept odpovědnosti a propojení člověka a stroje. Z tohoto důvodu si tato práce klade za cíl analyzovat civilní zkušenosti v oblasti autonomie a zjistit, zda mohou existovat nějaké přenositelné poznatky z civilního do vojenského světa pro budoucí vývoj autonomních zbraní. Po důkladné analýze havárií letadel Boeing 737 MAX a nehod způsobených samořízenými automobily dochází tato práce k závěru, že nejlepším způsobem, jak řešit etické a morální problémy autonomních zbraní, je vyvíjet a navrhovat takové systémy nejprve z pohledu člověka.

Klíčová slova: Umělá Inteligence, Autonomie, Autonomní Systémy, Propojení člověka a stroje, Odpovědnost

Title: Civilní Aplikace Autonomních Systémů: Důsledky pro Vojenský Sektor

Table of Contents

List of Abbreviations

ACL	Autonomous Control Level
AI	Artificial Intelligence
AoA	Angle of Attack
API	Additional Protocol 1 to the Geneva Convention
AWS	Autonomous Weapons Systems
CCW	Convention on Certain Weapons
CT	Computer thermography
DARPA	Defense Advance Research Agency
FAA	Federal Aviation Authority
HMT	Human-machine Teaming
ICRC	International Committee of Red Cross
IHL	International Humanitarian Law
MCAS	Maneuvering Characteristics Augmentation System
ML	Machine Learning
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
OODALOOP	Observe, Orient, Decide, Act
TREWS	Targeted Real-time Early Warning System
UAV	Unmanned Aerial Vehicle
VSD	Value Sensitive Design

Introduction

Artificial Intelligence (AI) and robots have the potential to usher in some of the greatest technological revolutions in history, which will have an impact on both the military and the civilian worlds as well as current international security problems. The primary quality that distinguishes contemporary technological systems from older ones is advanced autonomy, or the ability to carry out certain tasks without human assistance and, in some situations, to learn from their own experience or to develop over time in relation to the task at hand.

It is precisely these properties, combined with their increasing complexity, that have raised questions about the effectiveness of existing imputation mechanisms: *who should be responsible for any damages and failure of such machines? To what extent should humans remain in control of such technologies and how this will influence the relationship between human and machines?* These are concerns that interest both civil and military realms.

From one side, the civil domain is making strides in the development and implementation of autonomous systems in civil society. Airplanes have already been equipped for some time with autopilot systems that are becoming increasingly autonomous. Big car and tech companies such as Tesla and Google are continuing to develop increasingly self-driving cars, parts of which are already available on the market (waymo, n.d.; Ajitha & Nagra, 2021). The medical sector is moving toward a 4.0 revolution in medicine through the development of autonomous systems with artificial intelligence that can support the work of doctors, both logistically and therapeutically (Saracco, 2019).

From the other side, when it comes discussing the main challenges and risks of autonomous systems, the debate looks prominent in the military realm, particularly on the question of the development of autonomous weapons systems. Many issues have been brought to public attention: from ethicality to the social impact, to the effect that these systems can have on existing legislation. Even if, civil and military concerns on autonomous systems coincide. Moreover, there is a certain lack of debate when discussing civil-military nexus under a transferable lessons perspective.

Thus, the intent of this thesis is to fill this gap up by bridging existing knowledge on autonomous civilian systems and existing concerns about AWS. Therefore, the scope of this project is to understand what transferable lessons the military sector could learn from the civil autonomous systems applications. During the analysis, this thesis will focus primarily on two prominent issue that are mutual to both civil and military realms: (1) responsibility gap and (2) the human-machine relationship, known as human-machine nexus.

Research Method

The scope of this section is to present the methodology behind this thesis for the purpose of answering the research question of this work: *Are transferable lessons from the civil to the military realm possible in the context of autonomous systems?*

In order to answer the proposed question, the thesis has been divided in different sections. First of all, an extended literature review on the issue of autonomous systems from both civil and military perspectives is presented. Moreover, the focus will shift on how the topic of civil-military nexus has been discussed in literature. Furthermore, the following section aims to conceptualize the most important terms that constitute the pillars of this thesis, from the definition of autonomous systems to the characterization of responsibility gap and human-machine nexus. In order to acquire a comprehensive understanding of the topic, analysis of primary and secondary sources of academic journals from both civil and military perspectives will be utilized.

The third part of the thesis will be dedicated to the empirical analysis of the civilian application of autonomous systems. In order to accomplish this goal, this thesis will follow a multiple case study design. In this sense, the cases of Boeing 737 MAX fatal crashes and the accidents of Tesla S model and Volvo XC 90 self-driving cars will be analyzed in order to demonstrate whether the military realm could grasp any transferable lessons from these events. For this purpose, articles from notable newspapers like New York Times and Washington Post will be examined. Moreover, analysis of secondary documentation and official reports on Boeing 737 Max have been utilized in order to have a general understanding on the airplanes' technical functions. Last but not least, it is important to highlight why multiple cases are taken into consideration. Unlike the single case analysis, the multiple case studies design allows for a more detailed exploration of the complexity of particular phenomena. By analyzing different applications of autonomous systems in the civilian sphere, it is possible to capture the issue of responsibly gap and human-machine nexus from different perspectives and facets. Although the fields of application of autonomous systems in the civil realm are different, for the purpose of this thesis, it was deemed necessary and important to analyze the cases of self-driving cars and Boeing 737 MAX for different reasons. The first motivation lies in the availability of sources. Autonomous driving and autopilot systems are much discussed in the literature, which allows for a more accurate examination of case studies, compared to the analysis of new applications of autonomous systems, for example, in the medical field. Second of all, the accidents of self-driving cars as well as the fatal crashes of the Boeing 737 are real cases that have shaken public opinion the most in recent cases. For these reasons and because of the amount of data available, it is believed that these scenarios may be the most suitable cases for analyzing the impact on responsibility and the human-machine relationship, as they both involve humans in their processes.

1. Literature Review

First, this chapter will scrutinize what are the most common challenges in the application of autonomous systems, both from a civilian and a military point of view. Then, it will analyze how the nexus between the military and civilian sectors has been addressed in the literature and where the contribution of this thesis lies.

In recent years, scientists, researchers and theorists of new technologies have been mainly interested in the development of autonomous systems. Although this topic is approached from both civilian and military perspectives, a careful analysis of the existing literature on the subject shows that recently, the debate has focused primarily on the military sector. The main subject of discussion in this field has as its main focus the development of autonomous weapons (AWS), notoriously known as, killer robots. Theoretically, the debate focuses on what implications these weapons may have on the battlefield, once applied (Daniels, 2022). However, the literature mainly focuses on the ethical, legal, political, and strategic difficulties of such weapons. The fact that, there is a preponderant focus in the military world on autonomous weapons is highlighted by the fact that the international community, both political and civil society, is particularly active on the issue. There have been international campaigns to raise awareness on AWS impact over the past year. For instance, the International Committee for Robots Arms Control (ICRAC, s.d.) has called for banning AWS of all kinds, likewise the famous campaign "Stop Killer Robots" comprised of scientists, NGOs, activists and scholars, has become a global endeavor (Stop Killer Robots, n.d.). Similarly, States are also considered active players in the AWS debate, such that it was necessary to create a group of experts on the topic within the Convention on Certain Weapons in Geneva (CCW, 2019). Hence, a great number of actors are contributing to the discussion on the topic of AWS. The growing and overwhelming debate about AWS has also made it difficult to find common ground in finding a single definition of an autonomous weapon. From a general perspective, according to the International Committee of the Red Cross (ICRC), it is possible to define autonomous weapon as "a weapon that has the ability to detect and attack, thus causing damages and destroy any target without human intervention" (Davison, 2018). However, scholars, scientists and theorists seek to add technical facets to that definition (Scharre & Horowitz, 2015; Scharre, 2018; Congressional Research Service, 2023). Moreover, States also try to give their own definition of the term AWS depending on their political and strategic position in international politics (China, 2018; France, 2016; UK, 2018). As previously announced, several issues are shaping the debate on AWS. From an ethical and legal perspective, scholars believe that it is not possible to delegate the decision of who should live or die to autonomous machines. They point out how the effective lack of human control over these weapons may generate a regulatory vacuum effect regarding the issue of responsibility (Sparrow, 2007; Matthias, 2004; Asaro, 2020; Sharkey N. E., 2012). Whereas, others believe that the lack of human control on AWS could lead to the violation of the International Humanitarian Law (IHL), particularly the principle of distinction[1] and proportionality[2] (Amoroso, Garcia, & Tamburrini, 2022; Human Rights Watch, 2016). The main concern of this legal issue consists of the fact that both principles, in order to be respected, require socio-cognitive abilities that are not typical of autonomous systems, thus there is still a need for human evaluation in the end (Amoroso, Garcia, & Tamburrini, 2022).

Moreover, from a technical and ethical perspective, there is a widespread fear among scholars that inherent software vulnerabilities could lead to unpredictable behavior of weapons that humans cannot take control of (Sharkey N. , 2017; Boulanin, Davison, Goussac, & Carlsson, 2020). When it comes to analyzing the nexus of human and machine, there is a general understanding in literature of the fact that, the more robots are available in battlefield and the more dehumanized warfare becomes (Wagner M. , 2014; Sharkey N. E., 2012). As a result, this would result in emotional detachment on the part

[1] ART. 52 API:" *It is in the power of an adverse Party; he clearly expresses an intention to surrender; or he has been rendered unconscious or is otherwise incapacitated by wounds or sickness, and therefore is incapable of defending himself"*

[2] ART.51 API "*Among others, the following types of attacks are to be considered as indiscriminate: an attack by bombardment by any methods or means which treats as a single military objective a number of clearly separated and distinct military objectives located in a city, town, village or other area containing a similar concentration of civilians or civilian objects; [....] an attack which may be expected to cause incidental loss of civilian life, injury to civilians, damage to civilian objects, or a combination thereof, which would be excessive in relation to the concrete and direct military advantage anticipated"*

of soldiers, making them feel more legitimate in using these weapons. Despites these concerns, there are those who believe that AWS could bring military advantages. There is the possibility of reducing civilian casualties, collateral damages and limiting humans' unpredictable behavior that can occur in high-stress situations like wars (Noone & Noone, 2015).

From a civilian perspective, the existing literature shows how the risks associated with the application of autonomous systems coincide with those demonstrated in the military realm. One of the issues that is mainly debated by scientists and theorists is the issue of liability and responsibility (Yazdanpanah, et al., 2023). Mainly, scholars discuss who should be held legally and morally responsible in case of accidents and malfunctions of autonomous systems underlying how it becomes difficult to attribute legal responsibility to human operators when autonomy increases, especially for the case of self-driving cars (Darling, 2022; Freitas, 2023; Hevelke & Nida-Rümelin, 2015). Second of all, as for the military realm, it is becoming increasingly difficult to apply existing legal framework to new emerging technologies such as autonomous systems. From an overall perspective, scholars and law experts underline how the lack of international regulatory framework could lead to important challenges as these technologies are spreading across the civilian sphere (Fisher, et al., 2021). As it is now, private companies are in charge of setting regulations on autonomous systems. However, these internal policies only apply to the products and services they sell and produce (Vihul, 2020). Thus, this creates an overwhelming and confusing number of regulations. Hence, scholars are urgently asking for countries to harmonize domestic and international politics, in order to have a universal legal framework for the application of autonomous systems (Adler, 2020; Vihul, 2020).

As in the military, ethical risks are also analyzed in the civilian world. The literature mainly focuses on the ability of machines to make morally and ethically right decisions, especially in situations where human life may be put at risk, such as in the medical field or in the case of self-driving cars (Bonnefon, Shariff, & Rahwan, 2015; Farhud & Zokaei , 2021). For example, as in the case of AWS, self-driving cars might find themselves in complex conditions and having to choose between the driver's life or the life of a pedestrian, especially in dangerous situations (Wiseman & Grinberg, 2018). Hence, the need to determine the adequate level of autonomy and ensure a correct balancing with human control in autonomous systems in order to prevent unexpected behavior of autonomous systems (Firlej & Taeihagh, 2020).

Even if ethical and legal concerns surrounding the question of autonomous systems in civil and military spheres coincide, the existing literature shows unbalancing interest on the topic that focuses mainly on the military realm - the overwhelming debate in the international arenas is a valid example. Furthermore, there is consistent lack of communication between military and civil realms when considering the possibilities of transference of ethical and legal lessons from a domain to another. However, this is not always the case. The existing literature shows how the civil-military nexus is strong when it comes to analyzing the possibility of transferable existing technology from a sector to another-as it will be presented below in this section.

First and foremost, it is necessary to outline how technology is conceived in literature for the success of the task at hand. There is not a common definition of what technology is. However, most of the scholars in literature constitute their definition of technology around the word as "knowledge" "know-how" For instance, according to Kumar and his colleagues technology consists of two components: one is physical and the other one is an informational component such as "know-how" that needs to be shared (Kumar, Kumar, & Persaud, 1999). Lan and Young believe that technology is connected to obtain a determined result by employing knowledge (Lan & Young, 1996). Together with the concepts of technology, technology transfer has been discussed a lot in literature. In the past, technology transfer was referred to as the sharing of know-how from local conditions and its diffusion across countries (Chung, 2001). Whereas others refer to it as a process through which knowledge is transmitted to a particular service (Baranson, 1970). Moreover, the concept of technology transfer has been also studied under a more practical perspective. It has been referred to as a process through which concepts and knowledge are moved from labs to commercial sectors (Phillips, 2002), whereas others believe that in order to make technology transfer efficient there should be an interaction between more social entities. Nonetheless, from an overall perspective, it is easy to tell that the concept of "knowledge" persists in technology-oriented literature. Thus, it is extremely important to question whether it is possible to consider knowledge transfer in the same way as technology transfer. Scientists and scholars seem divided by these two terms: from one side, these concepts are conceived under the same meaning, they believe that at the base of both processes there is a continuous exchange of know-how which is embodied in products (Stephenson & Hayden, 1994). From the other side, scientists have tried to analyze the terms separately: Gopalakrishnan and Santoro in their studies pointed out the fact the terms cannot be used interchangeably. According to them, technology is about how things are materially done, whereas knowledge is more about how and why things are projected, thus this process requires a more moral judgment, thinking process that technology does not necessarily need to have (Gopalakrishnan & Santoro, 2004). Nonetheless, most of the scholars agreed that technology transfer and knowledge transfer are connected to each other. The main idea is that without a share of knowledge or know-how, there is no technology transfer and vice versa.

From an overall perspective, it is important to give a brief historical background overview on how the relationship between military and civil technology cooperation – thus technology transfer was conceived and changed throughout the years. David Edgerton, in his project "The Relationship Between Military and Civil Technology: A Historical Perspective" attempts to give an important overview on how technology transfer started to gain more and more importance throughout the decades (Edgerton, 1988). The author stated that in the XIX century it was argued that warfare was aristocratic and feudal and that the gains were minimal. At the time, according to liberals, science and technology could only matter in a liberal society and not one based on war. The progress of science and technology was seen as an integral part of the development of civil society, which goes beyond all types of conflict. David Edgerton still argued that during World War I there was a dominant thought that war had only a destructive value and consequently weapons and related technology also took on the same meaning. However, after the outbreak of World War II, the thinking about technological development seems to have taken on a different connotation. It was precisely in these years that scholars began to look at military technology with a positive outlook: military and civilian technologies could be used in a progressive way through the so-called spill over between the military and civilian realms. In these years, therefore, the idea pervaded that military technological inventions could be used directly in the civilian sphere bringing practical and economic advantages.

Furthermore, the authors believed that at the time, military technology has taken over the civilian technology, not only from a technical point of view but also from an organizational point of view of the society's thinking (Edgerton, 1988).

Strengths and weaknesses in technology transfer have been widely studied in the literature. When analyzing the possibility of civil military collaboration in the technological sphere, scholars tend to divide this cooperation in two categories: spin-ins and spin-offs, mainly. The first term refers to the use of civilian technology innovation such as facilities, tools and know-how to produce arms or/and military products (Cao, Yang, & Zhang, 2020). Whereas the latter refers to the technology transfer of military innovation into the civil realm (Bellais & Guichard, 2006). In recent decades, a substantial divergence has emerged between military and civilian industries. In the years following the world wars, the world witnessed an exploit of military products in the civilian world. There has always been the idea that the United States pioneered the spin-off technology process. There are several military inventions that were first dedicated to the battlefield that eventually were introduced and spread in the civil sphere (Chiang J.-T. , 1991). The most famous inventions are the GPS and the internet which originated from the U.S. department of defense between 1970s and 1990s. These systems were created for the purpose of aiding military logistics in the battlefield and operational settings and are now the basis of every navigation tool implemented in smartphones, as well as the internet (McFadden C. , 2020). According to Chang (1991) many of the U.S. technology programs, like nuclear ones, have been pivotal for leadership in the technology sector, both from a military and civilian perspective. However, even if there is a general understanding of the importance of such spin-offs, many scholars believe that these were not as successful as one might think. In his book, Mario Pianta, argues that even if military R&D represented a starter point for technology innovation, he believes that not all spin-offs have produced satisfactory results in the civilian domain. He states that the nuclear program has been a fiasco when applied in the civil realm. Furthermore, The Boeing SST project proved to be fallacious because the requirements between the two domains - civilian and military ones- were divergent. However, he believes that technology transfer is effective only when there is product compatibility between the civilian and military industries (Pianta, 1988). Furthermore, it is also believed that spin-offs in civil sectors are unable to be successful due to the fact that most of the technological programs tend to exploit existing technology (Chiang J.-T. , 1991).

Most of the scholars and researchers agree on the fact that military investment on technology innovation and related programs are decreasing compared to two decades ago, whereas civilian industries are gaining more and more importance and investments when it comes to technology modernization (Carozza, Marsh, & Reichberg, 2021; Caine, 2019). In fact, more and more military industries are trying to cooperate with civilian ones and the differences between the two realms are getting blurry (Cao, Yang, & Zhang, 2020). The fact that many civilian and private companies often find themselves supplying military companies, causes these companies to become hybrids, which no longer call themselves either civilian or military. However, what matters is the revenue, therefore spin-ins are an increasing phenomenon (Sezal & Giumelli, 2022). Nonetheless, there is still some skepticism on the real efficacy of such spin-ins. Even if the military sector is more interested in improving existing technology and the civilian sector is pushing further for technological innovation, it does not mean that the spin-in will function right away. Scholars believe that there is a need for specific policy to make it happen (Verbruggen, 2019). Verbruggen studies how civil innovation can influence the development of lethal autonomous weapon systems, in his study he finds out how technology transfer will necessarily encounter obstacles to its application. First of all, when it comes to spin-ins from civil to military field, there is a need for a degree of modification for developing technologies. For instance, autonomy is not a defined technology, it is composed of algorithms, software that needs to be trained based on the context where it will be applied, thus there is a need of a strong cooperation between military and civil domain together with bureaucracy and intellectual property issues. Moreover, not all civilian companies want to work for military purposes, mostly because of ideological constraints. Let's take for example the case of project Marven undertaken by Google. In 2018 this company had signed a project with the pentagon called Marven with the aim of developing technologies based on artificial intelligence to interpret images and videos, without saying that facial recognition could then be implemented in autonomous weapons systems. In the face of this project, workers rebelled by writing their own open letter criticizing Google and the project itself, which was considered unethical. From then on, the project was not renewed and Google decided to publish ethical standards on AI (PAX, 2019).

Verbruggen believes that military objectives need to go through a very strict process of testing and training that are not even comparable to the civil processes. For this reason, he suggests that the military research and development will always be in charge of implementing restricted military tools, such as lethal autonomous weapon systems (Verbruggen, 2019). Others believe that spin-in and spin-off processes cannot happen in the short term. The civilian and military markets are profoundly different. The most complicated process would remain spin-ins, where the technology requires commercial adaptation. According to some authors, the best way for civilian-military cooperation to succeed is to increase long-determined collaborations based on consulting and awareness of programs in technology innovation (Bellais & Guichard, 2006). Whereas, others state that innovation technology should mainly depend on the commercial sector, thus national security should not be addressed only through a military perspective. The main reason that supports this thesis is the fact that as the R&D of the civilian realm is speeding up it manages to keep up with the new threats in the military sphere (Stowsky, 2004).

From a geopolitical perspective, it is believed that technology transfer can have repercussions on the international arena. Generally speaking, some authors suggest that the production of military arms, tools and facilities should remain under the control of military industries of major countries such as Russia, China and the USA (Horowitz, 2018). Horowitz studies the issue of technology transfer from the perspective of artificial intelligence, and he finds that the rush, especially the development of these technologies by both the civilian and military worlds may lead to the balance of power crisis in the long run. Moreover, when dealing with AI in technology transfer, the problem lies within the algorithms training required in both fields. It is hard to directly transfer those algorithms, since the civil and military realms treat the subject under different perspectives: The civil sector is more human centered, whereas the military field is focused on AI robustness. Thus, cooperation is difficult to facilitate (Schmid , Riebe, & Reuter, 2022).

After having presented the major contributions regarding the existing literatures on the debate of autonomous systems and the cooperative relationship between the civilian and military domains, it is necessary to establish where the contribution of this thesis lies. First of all, the intent is to give another perspective on the debate on autonomous systems. Hence, there is a need for on going beyond the overwhelming debate about the risks of autonomous weapons in the military and focusing on the legal and moral problems of these technologies from a broader perspective that includes both the civilian and military realms. Second of all, the scope of this work is to fill the gap in the existing literature on the civil-military nexus, examining the possibility of transferable ethical and legal lessons from the civil to the military field. Thus, in order to accomplish this goal, it is important to study the lessons from already existing autonomous civil systems for military systems under development, such as AWS.

2. Concepts and Theoretical Framework

The scope of this first chapter is to define and give a general understanding of the main concepts that characterize the most important elements of this thesis: autonomy, artificial intelligence, responsibility and the human-machine nexus. To this end, the existing literature based on the definitions of the terms, will function as a guide to present how these concepts are conceived for the purpose of this thesis.

2.1 Autonomy and autonomous systems

As the title of this work suggests, autonomous system represent one of the main objectives of this thesis. Thus, it is crucial, for of this research to outline how autonomy is studied in literature and under what perspectives this term is used for the purpose of this thesis.

First of all, the word autonomy comes from the ancient Greek, “autonomia”, “autonomous”. Which is a composed word of “auto”, which means self, and “nomous”, “law”. Thus, autonomous is something that has the capability to govern itself (Swaine, 2016). Even if, there is a common understanding of the etymology of the word, what can be considered autonomous cannot be interpreted in the same way. In literature, there are still many discrepancies on the meaning of the term. First, depending on the science studying the subject, autonomy takes on different meanings. For example, philosophers study the term autonomy from an ethical and moral point of view, of what is right or what can be considered ethically and morally acceptable in the sphere of autonomy (Schermer, 2002). Whereas, from a scientific point of view, the term is viewed by scientists under a more techno-scientific lens of the term, by analyzing, technically to what extent a system can be considered partially or fully autonomous.

Some scientists try to analyze the concept highlighting aspects of human autonomy, like the ability to understand the significance of certain aspects and behaviors (Falcone & Castelfranchi, 2001). Some scholars, refers to it as a system capable of operation in a real-world context without any kind of external control for a long period of time (Bekey, 2005; Stornmont, 2008). Others, tend to refer to autonomous systems as robots that undergo unsupervised activity (Wallach & Allen, 2013). According to Panos Antsaklis, at the base of autonomous system there are two main aspects: first of all, its ability to accomplish certain goals, second of all, an autonomous system is always subject to certain uncertainty and the environment it works in. Thus, the degree of autonomy is determined by the set of uncertainty it can handle without human intervention. Moreover, the author believes that the level of autonomy is determined by the multiplication of "measuring of the set of goals" and "measuring of the set of uncertainties" (Antsaklis, 2020). When it comes to looking at the degree of autonomy, there are different labels that scientists can look at. Some scholars believe that autonomous systems can be considered on a spectrum that goes from "low level attributes" and "higher level attributes". The former grouped together those systems that include attributes such as perception, learning, context awareness and decision making, whereas the latter include also, domain independence, self-motivation and self-identification of goals (Ezenkwu & Starkey, 2019). There are different scales of degrees of autonomy that have been implemented and created by institutions and armies as well. For instance, the Department of Defense of the United States, developed the so-called ACL (Autonomous Control Levels) scale that range from remotely guided systems to fully autonomous one (Chen, Wang, & Li, 2009). However, the Air Force Research Laboratory bases its chart of degree of autonomy on the OODA loop (Observe, Orient, Decide, Act), thus it bases the level of autonomy on the decision-making process (Wu, Zhou, & Shen, 2021).

As one can easily understand, there are still some challenges when it comes to understanding the concept of autonomy itself. Thus, to make the understanding of the topic less convoluted, Paul Scharre's classification and concept of autonomous system will be used as a guide throughout the thesis.

Paul Scharre's idea assumes that autonomy is composed of three main concepts: type of task performed, relationship between human and machine and the process of decision making that the machine undertakes while completing the task (Scharre, 2018). For this reason, based on these principles he characterizes the autonomous system in three categories: semiautonomous operation systems, where the humans are in the loop, supervised autonomous operation (human in the loop) and fully autonomous operation (human are out of the loop). The fact that, in his classification of autonomous systems he considered the interaction of human and machine represents a good starting point for this thesis, as one of the main concerns is founded on the problem of human interaction on autonomous systems (ibidem). Finally, P. Scharre considers autonomy on a spectrum of machine intelligence that goes from automatic, automated up to autonomous. The scholar believes that all three terms cannot be analyzed using the same meaning. For instance, automatic systems are simple and do not have much autonomy in their decision, “they sense the environment, and they act” (ibidem). Whereas automated systems are more complex but at the same time machine are traceable by human users (ibidem). Autonomous systems have another function itself; they understand the task they are meant to do. Humans might specify the goal, but how the system accomplishes the goal is up to the machine (ibidem).

Since it is generally believed that fully autonomous systems are not yet implemented, for the purpose of this thesis supervised autonomous systems will be considered. As well as machines that still need a certain grade of human interaction to operate.

2.1.1 Artificial intelligence and machine learning

Another concept which is strictly related to the concept of autonomous system, from how it is conceived in this thesis, is artificial intelligence. As mentioned before, an autonomous system can self-improve, it can make decisions without having an operator watch it, it can make decisions and move within its environment based on the task it has been given. Thus, it needs to continuously adapt to changing conditions and the intelligence behind it is powered by AI. First, for the purpose of this thesis, it is important to highlight the fact that AI in this case, is understood as an enabler not as a weapon. It represents the power, the electricity through which autonomous systems operate. Thus, AI in this sense, is not designed for a specific purpose, but a general-purpose technology which is a part of something bigger. In this case, it enables autonomous system as a support, by assisting information processing and interpreting it (Horowitz, 2018).

A general overview on how AI has been seen through history is needed. One of the pioneers of artificial intelligence is Allan Turing. In October 1950 Alan Turing began in one of the first scientific contributions on artificial intelligence (Turing, 1950). His main thesis was to discover and evaluate how machines had the ability to think. In the same years in which the first artificial neural network was being built, the distinguished mathematician studied the theoretical possibility of introducing a thinking machine and proposed a test, destined to influence later developments, to verify in operational terms when a machine could be considered to be holding intelligent behavior. The one later known as the Turing Test was based on a famous imitation game: a human and a computer, placed in two separate rooms, were enabled to communicate in writing with a human interviewer who could ask all kinds of questions. At the end of the exchange, the interviewer had to discover the identity of the two conversation partners: if he was unable to correctly discern the results of his interlocutors, the machine would pass the test, proving that it was interchangeable with the human conversant (ibidem). For the official birth of the field of artificial intelligence research, however, it was necessary to wait until the summer of 1956, when the brilliant scholars John McCarhty, Marvin Minsky, Claude Shannon and Nathaniel Rochester invited a number of researchers in automaton theory, neural networks and the study of intelligence to participate in a two-month workshop at the town of Dartmouth (Solomonoff, 2023). The approach by which the first algorithms were designed consisted of the technique of exploring a given search space step by step. Not unlike someone looking for the way out of a maze, the algorithms would try the different possible paths and, in case of error, go back to take new paths until they had reached the solution. A decisive upswing in studies in the field of artificial intelligence came with the emergence of so-called expert. In the early years of the new discipline's development, researchers' attention had focused on creating systems that were capable of reasoning out abstract problems.

The ability to work through problems is also a part of the universe of intelligence. This ability requires the problem solver to possesses some necessary knowledge. Expert systems come to meet this need: they are computer programs that can represent and reason with the knowledge of certain specialized individuals in a given field for the purpose of solving problems or giving advice (Javapoint, s.d.).

Beyond these recent achievements, the most significant fact is that, today, artificial intelligence is widely present in everyday life. It is used to automatically translate from one language to another, to recognize speeches, to perform logistical planning and basic everyday operations that we can conduct with electronic devices such as Google Home and Alexa (Freeze, 2021). Underlying the incremental progress is not so much a change in the techniques used, which are largely similar to those already widespread in the past, but the availability of significantly more computing power and large quantity of data.

Although AI has been widely discussed in literature, here is still not a common definition of what AI is. Some refers to AI as a system that is able to correctly interpretate external data, to learn from them and to use these capabilities to achieve specific goals (Haenlein & Kaplan, 2019). Others refer to Ai "as a simulation of human intelligence process by machine" (Burns, n.d.). Whereas Zuiderveen Borgesius refers to it as a science that make machines smart (Zuiderveen Borgesius, 2019). A study that it is worth mentioning is from Russell and Norvig scholars, identify four schools of thought concerning AI. According to the first school of thought, some scientists believe that the main scope of AI is creating machines that think like humans, thus machines should reproduce human thinking. The second school of thought focuses on the actions of robots that should resemble human actions. The third and the fourth schools, however, are more positive and scholars believe that machines will always act correctly and think rationally (Russell & Norvig, 2010). These. Thus, it is easy to tell that there is not a general understanding of what AI is. The concept itself is very volatile and it changes based on the subject it is applied to. For this reason, it is important to highlight how AI is conceived throughout this research, as it brings together the concept of autonomous machines and knowledge, together. The definition that best suits this scope is the definition given by the European Commission. In the AI Act, the EU Commission states " «*Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the* *information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement earning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber -physical systems)"* (High-Level Expert Group, 2019).

Horowitz distinguishes two main approaches to AI. The first one is symbolic AI, which represents the creation of a set of rules that make the machine able to detect certain kinds of behavior. The second one is referred to computation AI, where machines act through problem recognition and calculation, where autonomous systems stand. At the base of the computational AI, there is machine learning approach, since many computational techniques are at the base to create algorithms for machine learning and decision trees as well (Horowitz, 2018).

Machine learning is at the base of autonomous system processing. ML consists of the study of algorithms that enable an artificial intelligence system to learn (IBM, n.d.). To get into the concept of ML, it is convenient to start with the concept of the algorithm. The term algorithm refers to a procedure used to achieve a specific result. While traditional algorithms are manually programmed to follow explicit steps, the hallmark of ML algorithms is that they learn the program automatically from observing data. Thus, such systems are able to improve their performance in relation to future tasks after making observations about reality (ibidem). In particular, the ability of ML systems is to be able to draw generalizations from the examination of some examples. During the training phase, the programmer subjects the algorithm to several instances, from which the algorithm infers how to behave (Kretz, 2019). For instance, this happens when we receive "spam" emails or phone calls and they get label as "spam". Specifically, the programmer provides the algorithm with some e-mails or phone numbers that have already been recognized as Spam, and from the analysis of these examples the algorithm tries to detect the characteristics that make it plausible to categorize an e-mail or phone calls as junk e-mail (Dada, et al., 2009). Within ML there are several ways through which an algorithm can be trained (Burns, n.d.). First of all, supervised learning which is characterized by the fact that the examples given to the algorithm are composed of a set of input values accompanied by a label, indicating the result or a value judgment. Secondly, unsupervised learning is a process where the programmer provides the algorithm with only a set of input data, devoid of any label. It is therefore the algorithm itself that is in charge of analyzing the information it has, classifying and structuring it, so that it can independently form more general knowledge. Last but not least, reinforcement learning is based on a system of rewards and punishments. In this hypothesis, there is no initial set of data, but they are collected from some simulations. The algorithm does not look for a correlation between the data, but rather for a policy of behaving. The algorithm's course of action is what determines what action to take in light of given circumstances and is learned through the aforementioned system of rewards and punishments. Each of its actions receives a certain number of rewards predetermined by the programmer (ibidem).

Another categorization of AI that is worth mentioning is the difference between strong AI and weak AI. The former, which sometimes is referred to general artificial intelligence, has the scope to develop machines with high intelligence standard like human intelligence (Burns, n.d.). According to Searle, strong AI was intended to create a human-like machine capable of completing multiple tasks at the same time (Searle, 1980). However, it is believed that is quite difficult to create such technology in a short time (Fjelland, 2020).

By contrast, weak AI or narrow AI, is the level of AI implementation where autonomous machines stand. This kind of AI is certainly much less ambitious than strong AI. Weak AI is limited to a single task, thus most of the systems that have this AI implemented, are able to accomplish and handle a single problem the have been programmed for (Burns, n.d.).

It is appropriate to emphasize how this *excur*sus on artificial intelligence is necessary for the purpose of this thesis. Machine learning-based algorithms are also referred to as black-box systems to indicate that the number and complexity of the steps involved in such systems is often such that even the programmers cannot always be sure what led the system to reach certain conclusions (Scharre, 2018) In other words, they are characterized by the fact that it is possible to observe the solution provided by the system, but not how it managed to reach it. This generates several problems: unpredictability of systems, uncertainty of legal responsibility for such technologies.

2.2 Responsibility gap

As presented in the literature review, the concept of the accountability gap is a widely discussed topic from both a civilian and military point of view. However, in order to have a general understanding of the various facets of the concept of legal responsibility, it was decided to analyze the concept from a military point of view.

The perspective of the possibility to incorporate AWS and autonomous systems with high levels of autonomy, and the correspondent decreasing of lack of human control in such systems has shifted the attention of the debate to another aspect. In the current literature there are multiple scientists, theorists and scholars that have been asking the following questions: *Who should be responsible for those AWS when applied in the battlefield? Who should be accountable for killing innocent people also in civil realm? Or who should we blame when AWS start to malfunctioning or behaving in a way that does not correspond to the task the have been programmed for?* Those are only some of the questions that theorists have been trying to answer in the past twenty years. However, it is clear that a common solution on the topic is still hard to find as scholars have still their own interpretation of AWS which is also characterized by their interpretation of responsibility.

As previously mentioned, there are not any fully autonomous weapons systems in today's battlefields, thus, even if non-autonomous systems are questioned for many reasons, when it comes to the problem of responsibility, as these weapons are under human control, it is easy to address accountability of humans operators (Schulzke, 2012). However, when it comes to consider AWS, the solution is not that easy to find.

First of all, it is necessary to define the concepts of responsibility from a general point of view and from a legal perspective thereafter. First, Aristotle defined the agent responsible for its own actions when it is has actual control over its behavior and it is aware of what it is doing. Hence, according to this definition, responsibility requires control and knowledge while in action (Hsieh, 1995). However, if AWS are considered under these two characteristics, it is easy to tell that these requirements cannot be met. Once humans use AWS or any other autonomous systems, they automatically lose control and awareness over them, since autonomous systems are based one AI and machine learning models, their behaviors remain unpredictable. Thus, this is where the responsibility gap comes into play. From a legal perspective, responsibility is the base of any kind of binding laws, *Ius cogens* in jurisprudence.

The concept of responsibility gap is relatively new. The first scholar that began to talk about the implications of AWS in the responsibility sphere is Andreas Matthias, in his research called "*Ascribing responsibility for the actions of learning automotia*" published in 2004 The author was the first one to bring the problem of responsibility in relation to AWS in war into the debate. He argues that traditional laws on accountability are not more applicable, since AWS will reach a certain amount of autonomy from which humans are not in control anymore (Matthias, 2004). Furthermore, he believes that the international community should somehow overcome this gap if the overall intention is to deploy such autonomous systems in war (ibidem).

Wars are settled in a dynamic and chaotic environment where violations of law are inevitable, thus there should be always someone that can bear the legal consequences of transgression. This is why, from Matthias' thesis, scholars, scientists and the international community have been trying to give a concrete solution to this "responsibility gap".

The existing literature on the topic of accountability can be divided into several strands. However, for the purpose of these theses it was deemed necessary to group them into three main groups. There is one line of thought based on the fact that neither AWS nor humans should be held responsible for violations carried out by these machines (Sparrow, 2007; Matthias, 2004). A second strand argues that AWS in the future will gain some degree of awareness and for this reason possess a narrow legal responsibility (Wallach & Allen, 2013). A final strand, on the other hand, is the one that considers AWS to be mere tools in the hands of the operators and for that reason responsibility will always fall on humans, because they will always have a certain meaningful control over those machines, from the programming/assembly phase to the actual use on the battlefield (Arkin R. , 2009). The first strand is characterized by the thinking of Sparrow, who in his research called "Killer Robots" expresses his thinking based on the fact that no human can be held responsible for the behavior of autonomous weapons (Sparrow, 2007). Consequently, even the creation of these systems should be considered unethical and for that reason they should be banned. Underlying his thinking is a rather broad definition and creation of the concept of autonomy. He is convinced that the next generations of autonomous systems equipped with AI will be so intelligent that their actions cannot be predicted.

Moreover, the machines will have their own values, beliefs and ideas. For Sparrow, a weapon is truly autonomous when it is truly autonomous, that is, when it is able to decide for itself and act accordingly. AWS would have, in his view, moral responsibility for what they do. In his view, it will be the humans themselves who, by creating these machines, will keep themselves out of the loop, eliminating any kind of individual responsibility. In this case, applying AWS in battle would undermine the fundamental principle of warfare, that someone will always have to be responsible for the death of the victims. Again, Sparrow makes use of the hypothesis that it will be possible to punish the machines by destroying and reprogramming them, however, they will never have the human capacity to feel guilt for their actions, consequently the punishment would be futile (ibidem). Similarly, it would be inappropriate to attribute blame to the operators and programmers of these machines, because as mentioned earlier, according to Sparrow, AWS are only really responsible if they have true autonomy and therefore no human being possesses control over them, and it would consequently be inappropriate to attribute blame to them (Sparrow, 2007). Similar to Sparrow, Matthias (2004) agreed on the fact that since AWS are based on AI and ML models that are generally unpredictable, humans cannot be considered guilty for their actions, thus either AWS are banned or the international community needs to face these gaps. In contrast to Sparrow and Matthias, Asaro believes that AWS will not be able to think and decide for themselves but will be able to acquire moral capabilities that can replicate human ones (Asaro, 2020). Thus, machines should be responsible in a narrow sense, since there will be always a need of human presence behind such weapons. Wallach and Allen go even further Asaro's idea. They believe that AWS could be considered ethical actors and held responsible for their actions. There will come a point in the future time where these machines will no longer be considered machines but will have some form of personhood (Wallach & Allen, 2013).

Contrary to what has been said so far, some believe that these AWS are like any other weapons in the hands of soldiers and operators who are ready to be used if necessary (Arkin R. , 2009). Created by humans and used at their discretion, although autonomous, they will always be under a meaningful human control. For this reason, responsibility for actions should lie with those who create them and those who use them. Other points out that the responsibility of any actions taken by AWS should be attributed to the operators, designers and developers rather than machines themselves (Arkin R. , 2013). In order to make this happen, he underlines that there should always be human control over these machines at anytime, together with maintaining human judgment and decision making when it is time to deploy them (ibidem). In order to have a better understanding of the discussion of responsibility, it is important to highlight the different aspects of legal responsibility.

In jurisprudence, legal responsibility falls into both civil and criminal law, meaning that "*specific duties imposed upon a person to care or provide for another including liability for personal* *obligations as granted through a Power of Attorney or Court order*"[3]. However, only physical entities possess legal personality, thus can be punished by the law. For instance, in criminal law, in order to apply criminal responsibility, two elements must exist together: an external or factual element, i.e. the criminal conduct (*actus reus*); an internal or mental element, i.e. the knowledge or general intent to determine the conduct element (*mens rea*). When considering AWS under the legal perspective of responsibility, as shown in previous sections, theorists have different interpretation on how one should consider AWS, thus this determines the concepts of those machines under a legal perspective.

However, for the purpose of this research, it is not appropriate to make any speculation about further development in law, as there is not much essential information for undertaking such considerations. Hence, when it comes to analyzing the AWS and any autonomous systems as a legal entities, it is important to take into account the two aforementioned legal aspects: *mens reus and actus rea*. Since they are machines, they are not capable of experiencing emotions such as repentance nor do they have the conscience and control of what they are doing, so the fact of legal responsibility (*mens rea and actus reus*) does not stand. It is also important to mention the fact that in jurisprudence legal responsibility can fall on various actors who will be further analyzed below with regards to the possible application of AWS, but it can be applied to any autonomous system, as the technology is the same.

2.2.1 State Responsibility

The first aspect that needs to be analyzed in this thesis is whether it is still possible to hold the state accountable for humanitarian law torts perpetuated by AWS. First, it is necessary to bring to light how state responsibility is governed by international law. According to art 3 of the IV Hague Convention which states "*A belligerent party which violates the provisions of the said Regulations*

[3] *Legal responsibility definition . Law Insider*. Available at: https://www.lawinsider.com/dictionary/legal-responsibility#:~:text=Legal%20responsibility%20means%20specific%20duties,of%20Attorney%20or%20Court%20order.

shall, if the case demands, be liable to pay compensation. It shall be responsible for all acts committed by persons forming part of its armed forces". and article 91 of the First Addition Protocol of the Geneva Convention (API) states "*A Party to the conflict which violates the provisions of the Conventions or of this Protocol shall, if the case demands, be liable to pay compensation. It shall be responsible for all acts committed by persons forming part of its armed forces*". Both articles recognize the responsibility of the state for all violations committed by state troops and state agents, also they recognize the right of reparation, between interstate relations, in the head of the state itself. However, the real debate falls on who actually commits the wrongdoing; a State does not act by itself, but it employs state agents who act on behalf of their state (UN, 2012). Accordingly, the armed forces of a State enter into war on behalf of their state. Therefore, any violations they produce are the responsibility of the State, even the use of specific weapons. As it is for now, if State agents do not take every precaution and infer, for example, deaths and harm to civilian populations, there is no difference from any other weapon, the state is imputed. Some theorists believe that State responsibility is not a risk, and existing legal frameworks are a good starting point to overcome the responsibility gap under this aspect (Geiss, 2016).

First of all, art. 36 of API guarantees the safeguarding of International Humanitarian Law (IHL) in the face of the creation of new weapons. "*In the study, development, acquisition or adoption of a new weapon, means or method of warfare, a High Contracting Party is under an obligation to determine whether its employment would, in some or all circumstances, be prohibited by this Protocol or by any other rule of international law applicable to the High Contracting Party*" Thus this is also attributed to the case of AWS. He argues that AWS like every other weapon can be unpredictable, like any other tools applied in war, it cannot function and cause unexpected damages that violate IHL or any other international law. As mentioned, AWS are characterized by the fact that, due to the ML systems inside of it they have a higher risk of being unpredictable, moreover theorists, scientists and scholars also cited in this thesis do not know exactly how AWS works and behave because they are still under development. Moreover, Geiss underlines the fact that if a State fielded such weapons and it is aware of such risks, it should be always responsible, like for any other weapons deployed (Geiss, 2016). From another point of view, J. Beard believes that states and individual responsibility are based on human judgement, but since AWS require high degree of autonomy, this requisite is lost and nether states nor individuals can be accountable for any actions committed by it (Beard, 2014). However, it is important to highlight that, in jurisprudence, states do not necessary requires the mental element, thus in the absence of *mens rea*, the state can be held liable (UN, 2001).

2.2.2 Commander responsibility

When it comes to analyzing the concept of responsibility, concerning the application of AWS, indirect responsibility, as known as, command responsibilities, are well discussed in literature. As forementioned, Matthias (2004) and Sparrow (2007) underline how neither operators nor commanders could be held responsible for AWS as they act like individual agent. Nonetheless, it is important to analyze the concept of the responsibility gap and command responsibility under the existing legal framework.

According to IHL, a commander or a supervisor can be held responsible for the actions of their subordinates only if three conditions are met: first, the commander knew or had insights that crimes were about to be committed by subordinates; the commander failed to prevent or stop such crimes from being committed; last but not least, the supervisor did not punished the subordinate for their action (ICRC, 2014). Clearly, these requirements call for a human relationship between commander and subordinate. In addition, the third element is what determines how the punishment is defined to affirm the guilt of the commander. However, as Sparrow (2007) points out, weapons are machines and do not possess moral agents. For such reasons, if we consider the weapon as subordinate to the commander the element of punishment cannot be maintained. Furthermore, since AWS are highly unpredictable, the commander cannot have a precise knowledge on how AWS will behave once deployed. Moreover, the first aforementioned requirement, presented seems at risk. Schulzke underlines the fact that commanders and supervisors do not have direct control over AWS, as they might have with their supervisor. However, he believes, that commanders hold little control over where they sent AWS, regarding the engagements that have been given to the machines, and what kind of enemy they have been instructed to target (Schulzke, 2012). Hence, as it happens between commander and soldier, the blame must be inflicted depending on how the behavior of the commander has influenced the behavior of the autonomous weapons in this case (ibidem).

2.2.3 Individual responsibility

Another very problematic aspect concerning the responsibility gap and perhaps the most complicated to solve is direct responsibility or better known as individual responsibility. First, it is important to remember how the responsibility falls on those who have directly committed wrongdoing and in order to ensure that an individual is considered guilty, one must always subsist the facts of *mens rea* and *actus reus* (Keiler, Panzavolta, & Roef , 2017).

Another aspect that should not be underestimated with regard to individual responsibility is the fact that it acts as a deterrent to individuals. Humans are moral agents who experience feelings, such as pain, anger, repentance. Consequently, being aware that a certain action corresponds to a certain consequence, in this criminal case, the punishment provides a deterrence aspect which, in general, should discourage people of committing future infractions (Shelton, 1999). However, if AWS is assessed as directly responsible for a possible violation of international law, again, under existing law, they cannot be considered criminal. Art. 30 of Rome states argues that "*a person should be criminally responsible and liable for punishment for a crime [...] only if the material elements are committed with intent and knowledge*". In order to be punished by international criminal law, the perpetrator must have directed and attack within the armed conflict realm and should be fully aware of what they were doing[4].

[4] See article 8 of Roma Statue available at: https://www.icc-cpi.int/sites/default/files/RS-Eng.pdf

Again, AWS are lacking any moral agent that could keep the machine accountable for its actions. For these reasons, most of the literature has abandoned the theory that AWS could be directly responsible for unlawful acts. However, the gap still remains.

Because AWS is part of a long building process that brings together manufacturers, developers, and finally commanders, the debate has shifted to the possibility that the responsibility falls on multiple subjects that make up the assembly line of autonomous weapons. Furthermore, it would create an overlap of responsibility between the different actors (Jain, 2016). It would be very difficult to attribute individual responsibility in the case of autonomous weapons given the large number of actors present.

2.2.4 Developers' responsibility and the problem of "Many Hands"

In the previous sections of this thesis, the problem of the responsibility gap has been addressed from an individual, State, and commanders' perspectives closely related to the military and the law in war. However, as pointed out earlier, the fact that AWS have a long assembly line behind them has caused the international debate on accountability to die down beyond those who use the weapons by going on to analyze the key players, as well as the minds behind the autonomous weapons, the developers.

The question then being asked in this section is whether those who create and develop these weapons can be the actual perpetrators of tort committed by AWS in war. The answer to this question is not as simple as it seems; it is not a matter of analyzing the responsibility of one individual, but rather a large number of individuals and organizations involved at all stages of the process from the conception into the deployment into battlefield, that in this thesis are defined as developers in broader sense. Compared to other weapons, here developers have the scope to create AWS according to specific task that will be performed after they are deployed. Once can say that by increasing the level of autonomy of such machines, one decreases the level of control exercised by the operator, whereas the level of control that developers have increases.

However, the issue remains: *to what extent it is possible to blame developers for unlawful acts committed by AWS, if developers have more control over AWS?*

First of all, it would be impossible to attribute individual responsibility to everyone operating behind the creation and deployment of AWS. This issue is often cited in literature as the "Many Hands problem" (Tamburrini & Amoroso, 2016). This issue often comes into play when a group of individuals may be collectively responsible for committing a wrongdoing and when there is a lack of grounds for attributing individual responsibility (ibidem). This term was coined by D. Thompson, who in his paper "moral responsibility of public official: the problem of many hands" published in 1980, pointed out how complicated it could be, legally, to attribute blame for a wrong committed, when there are many hands in a system that come into play (Thompson, 1980). According to Neelke Doorn, responsibility exists when there are four main elements involved altogether: the first is freedom, an identity is culpable when it acts freely and without compulsion. The second aspect is control and awareness over one's actions, the third is that there must be a causal connection between action and negative consequence leading to the fourth presupposition, violation of a norm (Doorn, s.d.).

Having presented these assumptions let us proceed by analyzing what is concretely meant by the problem of many hands. Let us take as an example the case presented by Tamburrini and Amoroso in their paper (Tamburrini & Amoroso, 2016). An autonomous drone was designed to distinguish civilian from military targets on the battlefield. However, in the training phase in the labs it was seen that the failure rate was 5%, yet to incentivize the sale of the product, the company lowered the threshold to 0.5%. At the time the drone was deployed in the war, the commander was aware of a rather low threshold of failure that the weapon can have, so he deployed the drone but made mistake and the drone hits a hospital killing civilians. In this example it is clear that all of the individuals involved had a casual contribution in the realization of the illicit fact, however they cannot all be declared individually responsible for the killing of civilians because none of the four necessary assumptions exist altogether. Not one person, taken individually, can be held guilty for the bombing of the hospital. However, there are those who argue that responsibility should be shared between commander and developer. According to Schulzke responsibility should be attributed based on how developers and commanders set the conditions that can lead AWS to commit unlawful acts (Schulzke, 2012). Hence, the responsibility should be distributed even if the developers are not directly part of the battlefield. They should in any case share some of the responsibility for allowing the machines to violate the rules of just in war and for not taking all the necessary precautions.

2.3 Human-machine nexus: between meaningful human control and human-machine teaming

When it comes to the discussion on autonomous systems, the lack of human control is taken into account, both from civil and military realms. Current debates have shown how the lack of human control in such systems could pose a threat to the existing legal framework, especially in terms of responsibility (Sparrow, 2007).

However, theorists and scholars have tried to come up with a solution that should overcome such issue. During the CWW meeting in 2013, the British NGO, Article 36, proposed there should be a "Meaningful human control" over AWS in order to find a solution to the problem (Article 36, 2016). A theory that has also been applied to any other autonomous systems (Santoni de Sio & Van Den Hoven, 2018). In this sense, for the first time the key to overcoming such issues was assigned to humans. According to Article 36, humans should continue to have sufficient control and a sort of authority in the decision – making process over autonomous systems. They argue that without proper human control these systems could not apply legal rules and might misinterpret legal frameworks and erode civilian protection (ibidem). In order to emphasize the necessity of a meaningful human control, the organization finds different key elements in order to establish such arguments.

First of all, technology should be predictable and transparent so that operators can understand how it works. Second, users should be informed of the precise tasks the machines are able to execute; thus, based on their information, users are able to exercise their judgement and decided whether there are the conditions to activate the AWS. For this reason, humans should be able to interrupt the function of the machine whenever possible. Lastly but not least, since there is a strong need for humans to acquire credible and accurate information on AWS, accountability should also encompass the wider system that develops such machines and those who provide training and information, together with commanders who are responsible for specific attacks (ibidem). This principle has gained a lot of success among scholars in the international arena (Horowitz & Scharre, 2015). States like Colombia, Croatia, Denmark, during CCW meeting underline the facts that there is a need for a meaningful human control that ensure legality in the application of AWS on the battlefield (Human Rights Watch, 2016). Others like, Austria, Chile and Brazil, call for negotiating a legal instrument that is binding aimed at ensuring meaningful human control over critical functions on lethal autonomous weapon systems (Austria, Chile, & Brasile, 2018).

According to Tamburrini and Amoroso, the definition given by Article 36 on significant control is rather ambiguous. It does not give a real explanation of what is meant by truly meaningful control (Tamburrini & Amoroso, 2016). However, it is a constructive ambiguity because it expresses a shared need to find key aspects to overcome the problem of autonomy in law (Crootof, 2016). However, the theorist tries to give an answer to the question. As Tamburrini and Amoroso mentioned, control is significant when human control functions as a catalyst for accountability, ensuring the presence of the conditions that attribute responsibility in case of violation of laws. (Tamburrini & Amoroso, 2016). In addition, human control must intervene as an auxiliary safeguard system (fail-safe) to prevent weapon malfunction or unpredictable behavior from leading to a direct attack (ibidem). In order to do so, human operators should have all of the necessary information about all the actions that the specific AWS is able to undertake, humans must be able to understand the reasons for a decision taken by a weapon (ibidem).

From another perspective, there are theorists and scholars that show some skepticism on the concept of meaningful human control, as there is neither a common definition of what we could consider “meaningful” nor what specific regulation could address such theories (Roff & Moyes, 2016). The authors believes that in the absence of a specific definition of what meaningful human control is it is difficult to achieve such a turning point under a legal and design perspective. According to these authors, the problem also lies in what is meant by control. An operator may be in control of the weapon and be able to act, while he may be in control of the weapon from a theoretical point of view, so they may recognize machines’ capabilities but may not be able to block its operation if necessary. In addition, there may be a problem that the operator who controls an autonomous machine does not have the ability to respond promptly to a hazard (Roff & Moyes, 2016). Elke Schwarz shares a similar position. She believes that it is impossible to acquire meaningful human control as it is generally conceived today for two main reasons: cognitive and speed limitations (Schwarz, 2019). First of all, she considers AWS as a complex system that include sensors, algorithms, software, hardware altogether with human designers and developers. In this sense, a human’s role is not just a controller, but the human is an interactive element within the artificial system (ibidem). In a fast-paced context such as war, soldiers and military carry the psychological burden of using lethal weapons during battle. In addition, the possibility of using new weapons such as autonomous ones do nothing but feed this cognitive stress. Thus, human cognitive abilities represent limits for meaningful control over these weapons.

When it comes to computational system interaction, humans are subject to cognitive limitation, as they rely on what Schwarz defines as deliberative reasoning for making the right decision and automatic reasoning for routine events. However, when humans are interacting with autonomous computer systems, the latter reasoning prevails, as automatic reasoning simplifies complex situations

– as war might look- and ignores contradictory information. Thus, in situations of stress and multiple inputs, humans tend to rely on computer reasoning rather than on a deliberative reasoning.

However, in daily life situation this might be helpful, whereas in situations where people's lives are at risk, this could pose a severe threat. According to Schwarz, the more sophisticated AWS are, the more humans give up on their decision-making authority.

Speed is another limitation closely connected to automatic reasoning. The necessity to act rapidly in war and the stress that falls on the soldier leads the operator to blindly rely on the machine in front of him without having time or the way to evaluate other options. Nonetheless, as decision become more distributed among technology, the capacity of humans to make decisions and maintain control diminishes, thus it is difficult to acquire a meaningful control (Schwarz, 2019).

The fact is that, the more humans rely on digital network and algorithm systems and the more humans abandon moral values in favor of calculative logic. In this sense it is necessary to go beyond the mere debate on human control and analyze human-machine nexus.

First of all, when it comes to dealing with artificial intelligence systems, people tend to overtrust the operations of such systems for different reasons. There is a general understanding that machines can be trusted as they perform their functions effectively and safely (Wagner, Borenstein, & Howard, 2018). Thus, overtrust comes into play when humans, in a specific situation, decide to accept the risk and make robots decide and act for themselves, as they believe that machine could mitigate the risks that humans cannot, or that they feel they are not capable of facing (Borenstein, Wagner, & Howard,, 2018). Placing human trust in machines is a growing phenomenon that does not only interest the military realm but that is already present in everyday life. People tend to rely on AI machines constantly nowadays, from home assistant like Alexa, to search engines like ChatGPT (NYT, 2022). The real problem behind this trend it is the lack of transparency on how these systems work and the complicated designs behind those systems that are other too complicated to be understand by human operators (Daneshjou, Smith, Sum, Rotemberg, & Zou, 2021). Because of this, people tend to take everything that the machines propose to them as the truth. Let's take the case of ChatGPT as an example. The new system created by OpenAi is revolutionizing the way we do searches, now we no longer need to do individual searches to find information we are interested in, instead we ask the machine what we need, and it brings back the answer we want in seconds. However, humans are inclined to believe and take for granted that what the chat reports to us is true. Consequently, humans no longer have the will nor the need to verify the accuracy of the information the chat reports to us.

However, research has shown how OpenAI's tool has often reported untruthful results. The hypothesis was also conferred by the ESMT Berlin research that finds that humans during their decision – making process do not always take into account whether the machine's recommendations or input were correct. Thus, the lack of feedback from human operators creates a bias in machine learning. Moreover, it negatively effects the use of machine itself (ESMT Berlin, 2023). Obviously, this kind of action does not create any physical or moral harm. However, the situation changes when humans apply the same reliability to autonomous systems such as self-driving cars and automobiles in dynamic environments such as warfare and roads, where a wrong action represent important repercussions from a legal and moral point of view.

Another aspect that is it is closely connect to the issue over-trusting is bias within the autonomous systems' software. It is well known that AI systems are characterized by such "prejudices" that are the result of training the algorithm from a large quantity of data that the systems require in order to function (Howard & Borenstein, 2018). Thus, this issue continues to be a large problem as it can affect people lives, and it can even interfere with decision making (ibidem). The problem also lies in the intent and motivation that push software engineers to develop such AI enabled autonomous system and what kind of biased dataset are used for machine learning training (Leslie, 2020). The initial set of algorithms with which learning is initiated is developed through an enormous amount of data collected and subsequently compared by the system in a world that not even the developers can fully understand. Moreover, the types of data that are collected in order to train this system have a role in such issues. There are different reasons that contribute to the development of bias in data, from language barriers to biased database use to train algorithms that are not part of the scope of this project (Kamińska, 2022).

Another aspect that is inherent in autonomous machines is their unpredictability of their behavior. Unpredictability and predictive problems are the basis of machine learning characterizing autonomous systems and difficult to eliminate. First, the behavior of a supervised ML during the operation phase depends on the quality of the data and training procedures. Therefore, the creation of bias in the training phase may depend on whether ML underwent a poorly loaded training phase, based on flawed data and or whether the system was not trained properly. However, the presence of bias in these autonomous systems is very insidious and it is very likely that an autonomous system can hide discrimination without the creator's knowledge. Hence, it is of utmost importance to have a large amount of good quality data especially in a dynamic situation such as warfare (Zuiderveen Borgesius, 2019). According to Buchanan, relying solely on a specific algorithmic model of decision making, especially for autonomous systems turns out to be insufficient for the efficient accomplishment of a given task entrusted to the machine. Every computational model has its weaknesses, of which the operator must always take into consideration (Buchanan B., 2017).

Another problem inherent in AI-equipped systems is what is commonly called the black box. This feature is very important and dangerous due to the fact that ML systems at times and in a less than provable manner behave unexpectedly and produce erroneous results (Scharre, 2018). Several image recognition tests have been carried out in the AI field, and it has often been noted that machines produce error that in a battlefield context would be fatal. For instance, in various tests, it was noted how autonomous recognition machines classify a bus as an image of an ostrich (Amoroso, Garcia, & Tamburrini, 2022).

Furthermore, adversarial attacks could pose a serious threat when AI-enabled autonomous systems are implemented. Adversarial AI happens when someone (such as an enemy in the battlefield) manages to determine a particular behavior of the black box ignored by the developers themselves and exploit it to their advantage (Scharre, 2018). This vulnerability in deep neural networks bring us directly to another phenomenon called poisoning attack. This attack is deployed when someone is able to directly influence the process of self-learning of a ML, thus the training dataset is contaminated by untruthful information and properly constructed in order to induce the system to make mistakes (ibidem). However, unpredictability does not lie on machine behavior. Yet, humans themselves are subjects to high number of psychological factors. Fears, hanger, or strong emotions could push humans to behave in a certain way that it is not coherent for a particular situation (Klincewicz, 2015). This is especially evident in a context such as war and on streets where men are subject to constant impulses and stimulations. Hence, the application of autonomous systems could represent a solution to minimize human unpredictability in certain situations (Noone & Noone, 2015).

For the purpose of this thesis, there is a need to find a balance between human presence and autonomous systems if meaningful human control is necessary to ensure the applicability of autonomous systems. Here, the concept of Human – Machine Teaming (HMT) comes into play.

At the basis of this concept there is the need to ensure collaboration between humans and machines to achieve a common goal reachable by one tool that will make people and machines interact with each other (Stowers, Brady, MacLellan, Wohleber, & Salas, 2021). The scope of this teaming is to render the human presence as useful as possible, in this sense, the main idea is to design machine interfaces that allows operators to interact with machines in the smoothest way possible by making the design easy and understandable to the one that are in charge of controlling the machines (Smith, 2019). The HMT has emerged in different field such as engineering, computer science and psychology as well in order to strike a balance between the importance of human control and judgment and the strong power that machines have to enhance operation capabilities especially on the battlefield (Rickli, 2022). Shared autonomy between humans and machines it is useful in order to achieve a set common goals. In this sense, autonomous systems provide assistance to humans while the latter exercise control over the system (Nikolaidis, Zhu, Hsu, & Srinivasa, 2017). A good combination of human judgement and machine expertise it is required to make AWS work accordingly. As autonomous systems lack the ability to exercise judgement especially under dynamic context such as war, human presence and cognitive capabilities will enhance machines functions overall (Skinner, 2018).

According to some delegation at the CCW meeting on AWS, in order to mitigate the risk of human interaction, they proposed to maintain the user's cognitive involvement in human machine teaming (Puscas, 2022). However, this could be challenging for different reasons. First of all, if humans have supervisory roles, they should be capable of maintaining their attention constantly and be able to intervene when needed. Loss of attention can happen at anytime and if operators are in charge of dealing with emergency situations, this might cause severe damages. Moreover, the more a machine is autonomous, the more humans enhance their tendency to rely on it, thus the more humans are dependent on machines, the likelihood of failures increase. (ibidem).

The first step to acquiring such capabilities is to make the design of such system more transparent and understandable as possible, thus this will make also operator more comfortable to work with AWS (Puscas, 2022). Therefore, a value sensitive design approach is needed to reach the transparency required. The term was first coined by Btya Fridman and his colleagues (Friedman & Hendry, 2019). From a general perspective, Value-sensitive Design (VSD) it is a technology design approach that underlines the fact that, in order to fully operate, technology needs value, thus it involves theories and practices that are gather both human and non- human values (Umbrello, 2020). In this sense, VSD tries t place human values in the engineering process to promote human presence within technological system. In this sense, Fridman proposes a list of thirteen values that should be included in technology system design as, human welfare, privacy, trust and autonomy (ibidem).

However, when it comes to AI systems, ethical values should be added as well. There are three different values conceived in value sensitive design: intended values, realized values and embodied values. The first one refers to the values that guide the AI systems, the second refers to the values expressed by system's operation and the third one refers to the actual values that are embodied in AI systems by designers (Umbrello, 2021). In order to acquire value sensitive design in AI system there should be a continual feedback loop and redesign iteration when assembling AI systems. It should be a loop that involve developers, designers, users that have the scope to bring up any incongruence between their values. Thus, the more feedback there is, the more the machines is trained to gather all the possible values (Umbrello, 2020).

When it comes to the creation of AI system applied on the battlefield, such as AWS, there should be a distinction between civil and military values, thus developers operating in military context should be also considered in the feedback procedure (Wyatt & Galliott, 2021). Moreover, in the context of the military field, AWS should be designed in accordance with value norms, such as IHL and Law of armed conflict (Umbrello, 2021). If developers are able to recognize the importance of adding military values into VSD process, addressing ethical issue, VSD can contribute to maintaining meaningful human control over AWS. According to Santoni de Sio e Van den, values sensitive design represents a good way to maintain a meaningful human control over autonomous systems (Santoni de Sio & Van Den Hoven, 2018). The authors argue that the concept of meaningful human control should not only be considered under the aspect of the operator, rather it should include the design and development process of autonomous system. Thus, in order to achieve meaningful human control, two conditions should be met: tracking and tracing conditions. The first one refers to the fact that decision making process should be aligned with the human operators' intentions, thus it is necessary that the system design follows humans, in order to consider an autonomous system under meaningful control. The tracing aspect represents the condition needed to attribute moral responsibility to a human agent. The fulfillment of such conditions, make the autonomous system able to provide traceability back to responsible human agents (ibidem).

3. Civil Application of Autonomous Systems

This chapter will be focused on autonomous systems applications in the civil sphere. Compared to the military sectors, in the civil realm there already some AI-enabled systems that have been implemented in civil society, from self-driving cars, autopilots and AI systems in aircrafts, and in the medical fields. These systems are capable of operating without a constant human intervention and are mostly able to enhance human performance, efficiency and safety especially in areas like transportation and public services – where we see an increase in autonomous system application. Autonomous systems are already involved in the daily routine of human life. For instance there are many restaurants around the world that have started using robots as waiters (Huang & Liu, 2022) or bars in Asia that are completely run by robots that serve clients and prepares drinks and meals (Avery, 2020). However, even if there is a general understanding that autonomous systems are a great add for civil society, concerns about transparency, responsibility, morality and ethics remains (The Royal Acedemy of Engineering, 2009).

Hence, the scope of this chapter is to analyze the case of self – driving cars, specifically the case of Tesla S model, Uber and the crashes of the Boeing 737 MAX in order to evaluate how the concept of human machine teaming and responsibility gaps are addressed and tackled in today's reality. The main scope is to understand what lessons we could grasp from these common case studies in order to overcome the issues concerning AWS in the military sphere.

3.1 The case of self-driving cars

The idea of developing cars that are able to drive themselves without human intervention is not new. The development of the first self-driving car can be traced back to the 1920s when researchers and engineers began imagining a future where humans would not need to drive their cars thus neither be in control of their motion. In 1925 the American company Houdina Radio Control Developed the first radio-controlled driverless vehicle, the car, which was named Chandler, toured the boroughs of New York without anyone at the wheel (Heinzelman, 2019). In 1956 General Motors designed an automated control system that could control the vehicle's accelerator and brakes. Their car was called the Firebird and was equipped with a cruise control system that could drive along stretches of road without the need for human control (McFarland, 2015). In the early 2000s, the Defense Advance Research Agency (DARPA) played a decisive role in the development of autonomous cars. In fact, the government launched three military competitions for companies with the aim of finding the best autonomous technology for off-road vehicles (Behringer, et al., 2004). However, since the early 2000s, the technology sector has been expanding rapidly, and more and more industries such as Google, Tesla, Uber and Volvo are experimenting with artificial intelligence in the automotive sector (Berk, 2021).

Before moving further, the analysis of the impact of self-driving cars, it is important to underline the differences between self-driving cars and assisted driving cars. The former refers to a vehicle outfitted with cameras, sensors and other AI enabled systems, that allow it to understand the environment. Moreover, the car moves based on what it perceives without the need for human intervention (Thatchan Research, 2018). An assisted driving car has a series of solutions that help the driver while driving, but do not exempt them from fully paying attention. Driver assistance gathers all of those features that help the driver with their maneuvers, for instance, while parking.

Furthermore, it is also useful to highlight the fact that there are different levels of autonomy in cars that have been internationally recognized by the SEA, the Society of Automotive Engineers (SEA, 2021). The Level 0 is no autonomy, which means that the driver has total control over the vehicle; level 1 is where assisted driving can be observed. Level 2 refers to the partial automation of the vehicle, which means that the car is able to accelerate and stop the car when the system understands there is a need. That being said, the driver still maintains control over the environment. Level 3 refers to conditional automation, where all tasks are performed by the vehicle, including the monitoring of the surrounding environment; however, the driver must be able to intervene should the system require it. Automation should be expressly activated by the driver and is deactivated if the driver does not keep his or her hands on the steering wheel or in cases where autonomous driving is not authorized or is unsafe. Level 4 is high automation, according to this level the vehicle can operate all the tasks without human intervention, however only under normal conditions which exclude bad weather. Finally, level 5 refers to the level of fully autonomous car, when the vehicle is capable of operating all assignments in any conditions (SEA, 2021). Moreover, this level requires a real transformation of the vehicle where there are no more pedals and steering wheels. It is easy to tell from the description of the levels, from 0 to 2, the human is still driving, whereas from 3 to 5 the vehicle is in charge of most of the maneuvers. Nowadays, most of the self-driving cars available on the market are on level 2. (Synopsys, 2020). However, there are already level 5 self-driving car circulating in the US, like the taxi developed by Google that does not require human presence in the vehicle (waymo, n.d.). Self-driving cars as they are developed nowadays, require human take-over when there are not optimal conditions for the vehicles to drive itself. However, it is very difficult to make it happen in real scenarios for different reasons. First of all, the tech industries are progressively pushing to build autonomous vehicles designed to keep humans out of the loop, as the Google taxi robot suggests (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015). Nonetheless, the autonomous vehicles available on the market now are mostly designed on level 2 autonomy, thus they still require human take over when needed. However, it is something difficult to acquire when over trust for the machine comes into play. When drivers trust their cars too much, they are immediately biased by their capability of perfectly dealing with any kind of situation and event that might occur in streets (Clancy, 2019).

The more the vehicle is autonomous, the more humans rely on that. According to Skitka and her colleagues, there are two main kinds of errors that are consequences of over-trusting the vehicle and can result in fatal accidents: error of emission and error of commission. As for the first one, humans feel too secure about the safety and machine capabilities and leave the car to do what it thinks is the best for that situation. Whereas, commission error is committed when the drivers do not act according to their judgement but rather, they do whatever the systems might suggest them to do (Skitka, Mosier, & Burdick, 2000).

Dikemen and Burns conducted a survey asking to Tesla drivers how much they trust its autopilot system. They surprisingly found out that most of the people interviewed trusted their AV and that the trust increases when people use these cars frequently, when they are aware of how the system works and when they are easy to master (Dikmen & Burns, 2017). The issue of trust in self-driving cars is just one of many debates on the subject.

The issue of responsibility is another issue that, as in the case of AWS, constitutes the debate and is one of the fundamental doubts as to whether more and more autonomous cars are needed in the automotive market. When it comes to the need to study the possibility of accidents with self-driving cars, scholars have their own opinions on who should be legally responsible. There are those who believe that the car itself is the one to blame (Freitas, 2023); whereas others believe that neither the drivers should be held responsible for the car that they are not actually driving (Darling, 2022), nor manufacturers, because if the blame will always lies on them, this will prevent developer from producing such cars (Hevelke & Nida-Rümelin, 2015). Furthermore, others believe that as long as autonomous vehicles are less safe than traditional cars they should be banned, but when developers get to the point when self-driving cars are safe enough, then all traditional cars should be replaced by them (Sparrow & Howard, 2017).

However, there is one aspect that needs to be pointed out. According to Perves and his colleagues, the issues surrounding AWS are not so far removed from those of self-driving cars (Purves, Jenkins, & Strawse, 2015). When it comes to discussing the reliability or other factors of autonomous weapons, the argument often falls back on the fact that the main purpose of these instruments is to kill people. Consequently, the main concern of the international community is to assess what the consequences of entrusting autonomous machines with this decision might be. What the scientists want to point out is the fact that even self-driving cars would need to be programmed to kill someone under certain circumstances (Purves, Jenkins, & Strawse, 2015). This problem goes beyond any technical solution that the scientists can adopt, but it is a purely moral question of choosing who to let die and who to let live. This dilemma is called "The Trolley Problem". This issue was first represented under a philosophical aspect by the philosopher called Philippa Foot in 1967 (Panahi, 2016). The mental experiment envisages that there are five people on the track being travelled over, tied up or otherwise unable to move, and that the only way to save them is to activate a switch that would divert the vehicle to another track, on which another person is unable to move (ibidem). Therefore, the dilemma concerns a person who is in the vicinity of the switch: *should he activate it and save five lives, at the price of one, or not act and thus condemn the five people to death?* The utilitarian view envisages saving the five people, recognizing a kind of additivity in the value of human lives, and that therefore the most appropriate action is the one that ensures the greatest good for the greatest number of people. With the development of autonomous machines, scientists and scholars have tried to put this problem into the hands of machines as well (Geisslinger, Poszler, Betz, Lütge, & Lienkamp , 2021). The first thing to mention is that, we as humans are conditioned by our beliefs, cultures and priorities that would lead us to assess the case of the trolley problem in a different way than other people who have grown up and lived in a different environment. However, the particularity of autonomous machines is that of having neither personality nor morality. They are given artificial intelligence systems, thus for their decision-making processes, they rely on algorithms within this system. In this sense, data should be trained based on the least harmful situation, however, since there are humans that develop those systems, it is hard to tell and define which values best fit the situation of the trolley problem (ibidem). A study held in 2016 questioned a group of people to rate, on a scale of 0 to 100, how much they agreed that self-driving cars should be programmed to save the most lives, collecting an average score of 70 (Bonnefon, Shariff, & Ra, 2016). On the other hand, when participants were asked to rate, again from 0 to 100, how willing they were to buy cars that minimized the loss of life (thus also subjecting the driver to risk), the average score dropped dramatically to 30 (ibidem). This shows how people believe that cars should be built as safely as possible, but not so much for the safety of others but for the proper safety of the driver (Wired, 2016).

In this part of the thesis, the intent was to demonstrate how the problems surrounding AWS are very similar to the problems in the world of automation. However, what mainly distinguishes the two cases is the fact that self-driving cars have already been tested and have been available on the market for a long time, while the discussion on the implementation of autonomous weapons is still ongoing. This means that the issues that have been raised in the case of autonomous cars are not merely hypothetical scenarios but are based on real facts. In fact, in recent years there have been several accidents involving autonomous cars from different car manufacturers that have also led to tragedies such as the death of a woman in Arizona (Levin, 2020). For the purpose of this thesis, it is useful to report and define how these cases have been dealt with in civil society.

3.1.1 The case of Tesla S Model and Uber

There have been different cars incidents in the past years that involved self-driving cars. However, the scope of this thesis is not to make a list of such cases, rather it is to understand how these cases were in the light of the issues analyzed above.

First of all, the Tesla S model car has often been the focus of attention because it has often been found in accidents of this type. However, in this thesis we will focus on a specific case that occurred in 2016. This accident was much discussed in the news since it was the first fatal accident involving a self-driving car (Yadron & Tynan, 2016). The collision dates back to May 7th, 2016, and occurred between a Tesla Model S and a tractor-trailer truck on a highway in Willinston, Florida. According to initial reconstructions, neither the driver nor the car's computer system would have noticed the presence of the vehicle beside them: the white side of the truck and the bright sky in the background would have misled both of them. Thus, the Tesla would have ended up under the truck's trailer in a situation in which the autopilot would normally have intervened (Clancy, 2019). According to initial investigations by the National Highway Traffic Safety Administration (NHTSA) in charge of cars accident investigations, it was determined that the autopilot was functional at the time of impact. It is important to note in this case, that the Tesla-designed autopilot is designed to assist the driver in the car's speed, steering and braking system in certain situations. However, as Tesla underlined during the investigation process, their cars are not meant to be fully autonomous, thus they believe that the fault should lie on drivers' inability to remain focused on the car behavior and ready to take control when it is needed (NYT, 2016). Tesla's argument was supported by the fact that NHTSA's investigations in 2017 confirmed the fact at the time of the crash, the autopilot was working correctly, however they determined the need for the industry to give clear information to the driver about the risks they are running (NHTSA , 2017). Furthermore, in 2019 the National Transportation Safety Board (NTSB) a U.S. federal agency, has confirmed the fact that the real issue was the fact that the driver was over relying on the autonomous vehicles' skills and this resulted in his lack of ability to take control of the vehicle on time (NTSB, 2019).

Another important case to be mentioned is the case of the VOLVO XC90 equipped with autonomous driving systems. The car, however, was owned by Uber which was carrying out autonomous driving tests in Arizona (Levin, 2020). At the time of the accident the car was being controlled by a woman, who was in charge of analyzing the safety of the car, when she was unable to take control of the car at the moment of impact that led to the death of a 49-year-old woman who was crossing the road. Investigation from the NTSB and police found out that the lady in charge on controlling the car's system within the car was not paying attention of car's system (ibidem). Video shows that at the time of the impact she was focused on watching a video on her phone, which prevented her of stopping the car at the right time. What it is surprising, is the fact the Uber recognition system installed in the car, first showed that it did not recognize the lady crossing the road and did not address operator's automation take over (Hawkins, 2019). Furthermore, the autonomous emergency braking system was not active, which required human operator to function when the car asked them to do so (Allsop & Keeves, 2019). Moreover, investigators found out that there was no evidence connected to Uber and there was no evidence to hold uber criminally responsible (ibidem). On the other hand, there was a common understanding of the fact that the driver was negligent in not maintaining proper control of the vehicle. For this she can be prosecuted for the crime (BBC news, 2020). What it is also important to point out is the fact that the aspects that led to the fatal accident are also the cause of a lack of regulation regarding the circulation of autonomous cars on normal roads in large cities such as those in Arizona (Porter, 2020).

3.2 The case of Boeing 737 MAX

Boeing, together with Airbus, have always been the pioneering companies in the aviation sector (Pilot Institute, 2023). However, from 2018 Boeing was in the spotlight for two aviation disasters involving one of its aircraft, the Boeing 737 Max (Perell, 2019). The flights involved in the tragedies were the JT610 of Indonesian Lion Airlines, in which 189 people lost their lives (Suhartono & Beech, 2019), and the 302 of Ethiopian Airlines, with 157 people on board, both of which crashed within minutes of take-off for no apparent reason (Miriri, Fick, & Maasho, 2019). Evidence has shown that there were many similarities between the two accidents, however, after the second crash that happened five months after the first, Boeing decided to ground its planes as a precautionary measure while investigations developed (Herkert, Borenstein, & Miller, 2020).

Moreover, in order to understand and analyze the case, it is necessary to take a step back and analyze what were the successive events that led to this tragedy. The first 737 was presented in 1967 in the United States with great success among buyers. Several models followed over the years, such as the 737-300, 400, 500 and the best-selling aircraft in history (Olstad, 2022). However, in the face of new challenges and competition brought about by the arrival of the new Airbus A320 neo, which was depopulating at the time, Boeing felt it necessary to have an aircraft that was aligned with the market and competitors (Gelles, Kitroeff, Nicas, & Ruiz, 2019). As a result, the company needed an aircraft that was economical enough for the airlines and efficient enough not to consume too much diesel fuel in a very short time. In order to develop a new generation of Boeing 737s in the shortest possible time, the company opted to redesign the old Boeing to create a new one (ibidem). First, to improve fuel efficiency Boeing increased the size of the engine compared to its predecessor, adding a taller landing gear to provide adequate ground clearance. However, the engine was still too low to the ground, as 737s were generally built too close to the ground compared to other planes (Yglesias, 2019). The center of gravity on previous 737 models was substantially forward of the center of lift. In the case of an in-air stall, due to neutral control inputs, the plane is able to angle downward and recover on its own. In order to counter a natural downward pull, the horizontal stabilizer generates a downward lift. Moreover, this action causes a drag effect and is likely to increase fuel consumption. Boeing appears to have adjusted the relationship between the 737 MAX's center of gravity and center of lift in order to decrease the trim drag effect and maximize efficiency (Holeman, 2019).

During test flight, the 737 MAX models were appeared to have drastically different flying characteristics than prior 737s. In the case of high angles of attack (AoA), when the plane's body lifts from the massive engine nacelles located front of the wings, a significant nose-up pressure is generated and the lift center moves forward. The thrust produced by the low-mounted engines, which act below the center of gravity, also act as a nose-up force which appears specifically at high power levels. Without corrective input, a 737 MAX at a high angle of attack will pitch up even further, resulting in a stall (ibidem)

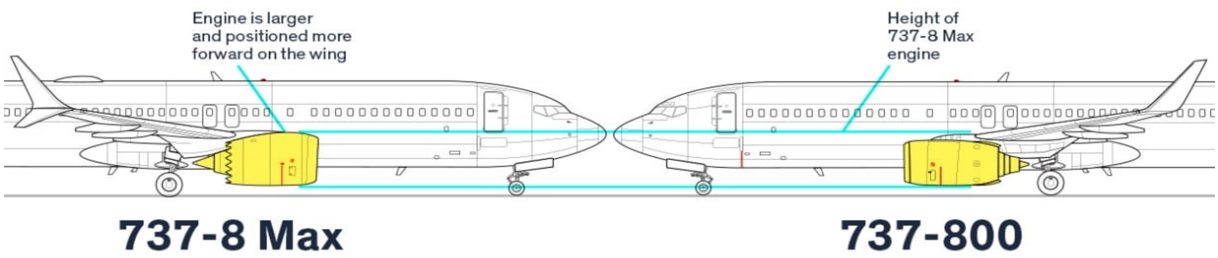

Image of old and new engine in 737 Max (Travis, 2019)

However, this problem would have prevented certification of the aircraft by the Federal Aviation Authortity (FAA) because the stability criteria could not be met (Holeman, 2019) . Hence, in order to overcome the stability problem of the aircraft, the Boeing opted to add an AI autonomous software into the aircraft called Maneuvering Characteristics Augmentation System (referred to as MCAS from now on) that was able to prevent the stall of the airplane (Travis, 2019).

The MCAS receives information from the AoA sensor to determine which direction the aircraft's nose is pointing, also checking airspeed and altitude. When the software determines that the angle of attack is too great, the system comes into play and tries to stabilize the rear of the aircraft by lowering the nose. It then automatically pushes the pilot's yoke down without the pilot being able to take control. According to the MCAS default setting, this is active when the autopilot is off. This effect should be cancelled when the pilot takes control of the stabilizer trim. However, the MCAS used by Boeing is designed so that it does not turn off when a pilot manually pulls the yoke when he sees that the aircraft is going down because it would go against the purpose of MCAS itself, which would be to prevent the pilot from inadvertently entering a stall angle (Johnston & Harris, 2019).

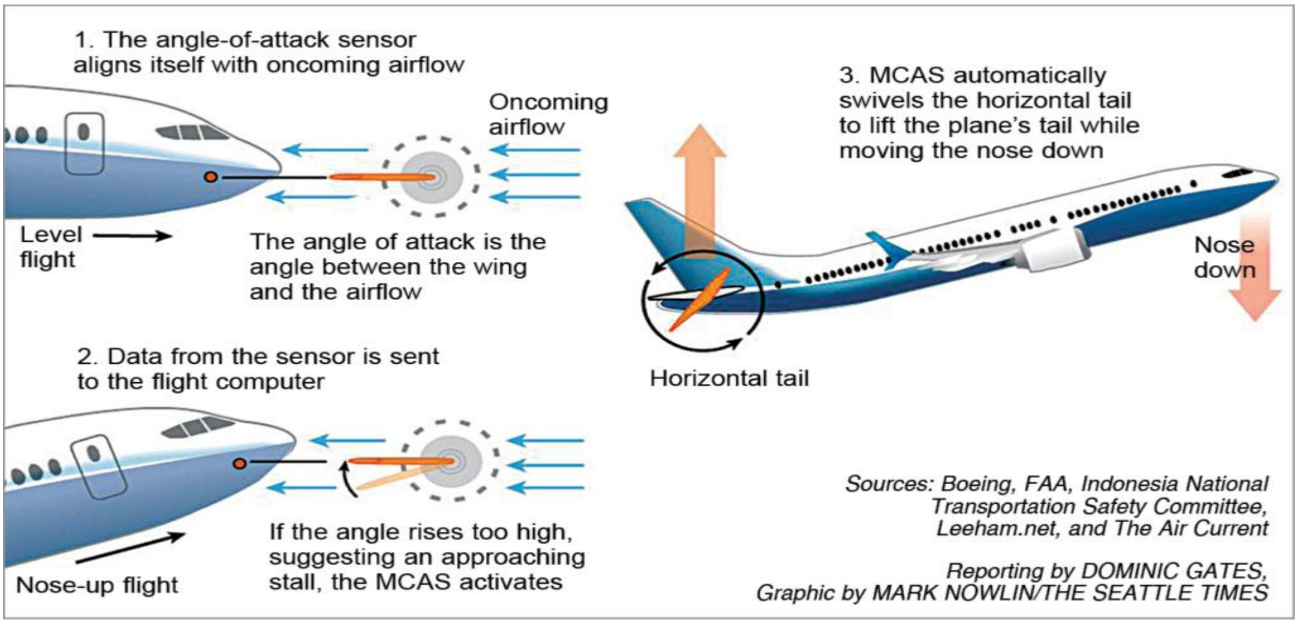

Image of the MCAS function (Theo, 2021)

However, the first problem with this software lies in the fact that Boeing has equipped the 737 with only two angle of attack sensors, whereas normally there are at least three. The MCAS, however, was designed to read data from only one AoA and then acted upon the data from this sensor. Conversely, relying on a single sensor was one of the Boeing's failures, because in the event of a malfunction MCAS would be activated based on faulty data without the pilots being able to deactivate it or take control of the aircraft (Sgobba, 2019).

The fact that the system was meant to work in the background was an indication that Boeing assumed that MCAS could always have data consistent with reality and that in the event of a malfunction they would be able to resolve the situation given that the software was considered non-invasive (NYT, 2019). After a long analysis on plane crashes, it was found that the cause of the accident for both cases was the malfunctioning of the MCAS system, probably due to a flock of birds that hit the attack sensor. The pilots could not resist the will of the software in pushing the plane down, causing the crashes (Burridge, 2020).

In any case, the Boeing 737 MAX had already shown malfunction signals during flight. Although a Lion Air pilot had reported a malfunction in the flight phase of the plane underlining how the plane was behaving in an unexpected way, there were no references of any kind in the Quick Reference Handbook[5] that reported a similar situation. This underlines how Boeing assumed that the MCAS would work without even mentioning its presence in the aircraft (Malunga, Ahmed, Raynard, Stoyanova, & Chisala, 2022).

From the moment of the plane crashes there were many rumors about who should be responsible for the disasters. In the first period of the accidents, there was more focus on the pilots. According to a pilot and journalist William Langewiesche the fault is not primarily on the malfunctioning of MCAS, but he believes that the ultimate blame is on the lack of experience of the pilots involved in the crashes (Langewiesche, 2019). Particularly, he argued that the fact that their airline company had cut the funding on training and due to the inexperience of the crew he suggested that despite the lack of information they had on the presence of MCAS in their aircraft, more experienced pilots would have

[5] The Quick Reference Handbook is a manual which contains simple checklists that are meant to help pilots to manage unexpected and non-normal events.

been able to handle the situation and prevent the accidents from happening. Hence, they did not take all the necessary precautions by flying too fast and at too low an altitude (ibidem). After the first crash in Indonesia, Boeing openly stated that inexperienced pilots should be blamed. The previous chief executive, Dennis Muilenburg, openly said that the pilots in Indonesia and Ethiopia do not have the same kind of training they have in the US, assuming that the problem was bad training or a lack of it, compared to pilots in other parts of the world (Thorn, 2020). However, it is very hard to support these hypotheses given that evidence has shown that pilots were not informed about the presence of the MCAS system within their aircraft, thus appropriate training was unavailable to prepare the pilot for a properly response to this malfunction. Hence, when Boeing sold and presented this product, it did not mention the MCAS in any way (Johnston & Harris, 2019).

Thus, it is very difficult it is difficult to attribute individual responsibility to the pilots given the evident lack of information. As the most famous pilot Chesley Sullenberg states, it is highly impossible to attribute individual moral responsibility to these pilots involved, as they need to have "absolutely mastery" of the airplane at any time. He argues that the lack of experience is also a problem when it comes to fatalities, however, the lack of knowledge cannot be taken as an excuse for Boeing design failure (Schaper, 2019). As Gregory Travis reported, the MCAS system was developed to compensate for the problem of aircraft balancing due to the increase of engine size (Travis, 2019). Hence, MCAS represented a band-aid solution that was needed in order to put the aircraft into the market as soon as possible (ibidem).

However, it is in this situation where multiple agents are involved in fatalities wherein the problem of "Many Hands" comes into play. There was a succession of events leading up to this fateful event: from the decision to replace new engines, mask the balance problem with MCAS system and the decision to rely on two AoA sensors, to not to properly inform the pilot about the software. First of all, evidence was shown during court trials revealing that engineers and pilots showed their concerns about the design and MCAS system (Tara, 2019). A senior supervisor of the Boeing project, Pierson has demonstrated severe problems with the design of the aircraft underling the fact that, the way the aircraft was developed could have posed a severe threat to the security and safety of the flight, asking for a meeting to the general director of 737 project (McFadden, Schecter, Monahan, & Schapiro, 2019). Since Boeing has sold the planes to many of the companies without solving the problem and without listening to their engineers, it has proved its negligence by putting profit first. Furthermore, one of the "Many Hands" involved is the FAA which is in charge of releasing certification of new aircrafts. As the new Boing 737 Max was augmented with new features, it required a new certification, however, again the MCAS was not mentioned, and the aircraft was very similar to the previous ones under FAA's eyes (Reuters, 2019). Moreover, it was able to approve immediately without the need of waiting on long administrative procedures (Pacific ventury, 2019). The FAA produced a 30-page long review about the new aircrafts containing information regarding any particular features, without ever mentioning the presence of MCAS (Campbell, 2019). However, the FAA has been accused for not being severe enough when checking on the needed requirement. (ibidem). It was only after the second crash that the FAA was finally aware of the MCAS software and undertook a risk assessment analysis on the system. However, the company has attributed the "Hazardous Failure" index to the system, which means that in the event of a malfunction, it could have caused only minor damage to some passengers or fatal injury to a small number of people (ibidem). Nonetheless, the second accident demonstrated how the MCAS was not meeting those standards, as the malfunctioning caused the death of almost 300 people.

In the end, The Securities and Exchange has charged the Boeing Company and its former CEO, Dennis A. Muilenburg, with making materially misleading public statements following Boeing aircraft accidents in 2018 and 2019. Thus, they will need to pay $200 million to families' victims. (US Securities and Exchange Commission, 2022). According to SEC orders, after the first accident Boeing and Muilenburg knew that the MCAS was a safety issue with the plane, but they still assured the public that the 737 MAX was as safe as any that has ever flown in the skies (ibidem). After the incidents Boeing upgraded and fixed MCAS and the 737-max returned to operation in 2023. However, it is important to emphasize that in Europe EASA has demanded that there is a continuous update of the MCAS system (Hepher & Vats, 2021).

Conclusion

Throughout this thesis, the issues or responsibility gap and human-machine nexus have been widely discussed both from civil and military perspectives. This thesis was searching for the answer to the question of whether there could be transferable lessons from civil to military realm when considering the challenges of autonomous weapons systems. Thanks to a deep analysis of the civilian cases studies presented, this thesis has found that transferable lessons are possible manly from a legal responsibility perspective. First of all, the cases of Tesla, Uber and Boeing 737 MAX have demonstrated that the autonomous systems themselves cannot be held responsible for any unlawful act committed. Furthermore, these cases shows that humans will always be responsible for the behavior of autonomous systems utilized, either under individual responsibility or collective one. Moreover, there should not be a unilateral application of responsibility, but liability has to be evaluated case by case, as they all have their particularities. For instance, individual responsibility was at the base of Tesla S model and Uber accidents for similar reason. Investigations showed a lack of driver attention while the cars were in motion, as a result they were unable to handle the cars in time at the moment of danger. Hence, causing the death of a pedestrian and the driver itself. Moreover, what distinguishes the Tesla case to the Uber one, is that in the latter, Uber could not be held responsible for the malfunctioning of the recognition system of the car, due to the lack of strong regulation. At the time of the accident, there was a lack of proper regulation on the movement of the self-driving car. Although in 2022, the Vienna Convention on Road Traffic introduced the concept of self-driving cars in article 34-bis (UK Road Traffic Amendment, 2021). However, Article 34 bis merely gives guidelines for the circulation of autonomous cars. It is the Convention itself that states that self-driving cars will only be able to move out of their current experimental or otherwise limited status if they are transposed international laws. Thus, it still remains at the discretion of individual states to adopt the directive and establish legal and insurance responsibilities in the event of an accident (ibidem).

Contrary to the case of self-driving cars, the evidence of the crashes of Boeing 737 MAX demonstrated that it was impossible to hold pilots responsible for the fatal accidents, thus individual responsibility could not be applied in these cases. During courts trials, it was found that Boeing was hiding crucial technical information that led to the deaths of hundreds of people. Thus, this case brings us to the next level of analysis. It underlines how the lack of transparency between developers and human operators can have severe consequences. As previously mentioned, the pilots were not aware of the presence of the MCAS system within their airplane, thus they were not given proper training that would have been necessary to handle the imminent risk.

Looking at the outcome of the analysis of these civilian cases, it does not seem to be some reason to believe that technological advancement in warfare must necessarily endanger the concept of responsibility. The analysis of cases has shown that autonomous systems can be considered neither moral nor legal agents, so even autonomous weapons will not be able to be held responsible for accidents or malfunctions. One of the main lessons that the civilian realm transmits to the military one is to consider the autonomous weapon as any other weapon that has the ultimate goal of annihilating the enemy. However, operators who might use of these weapons must be fully aware of the risks the tools may have once applied. Accordingly, full transparency on the part of developers about the potential risks these weapons may have, is necessary as long as soldiers can make their own calculations in deciding when and why deploying AWS. Consequently, considering autonomous weapons like any other weapon is critical to understanding the fact that the concept of responsibility under international law continues to apply. As in the case of the civilian sphere, events in the military world must be evaluated on a case-by-case basis. But under this view, there are no prerequisites to jeopardizing the possibility of direct, indirect, or State responsibility. Second of all, from a human-machine nexus perspective, the discussion is more delicate and elaborate. From a purely technical point of view, the cases presented were not able to define which specific solution is best suited to address issues on software vulnerabilities. However, the analysis of civil experience has shown how humans still play a critical role for autonomous systems, from the development to the control they exercise while the systems are operating, Thus, the cases analyzed underline the necessity to create the most suitable conditions for ensuring human-machine cooperation. As the civilian cases demonstrate, human presence is still very important in the process of autonomous system application. Thus, there are no evidence, for now, that AWS will replace human presence in the battlefield. What the case of Boeing 737 MAX has mainly demonstrated is the fact that there is an extreme need for putting human safety in the first place. As the value sensitive approach points out, there is a need to ensure human values as part of the design process. In order to make this happen it is important for engineers to prioritize human public safety when design and deploying autonomous systems. In both, cases of self-driving cars accidents and the fatal crashes have shown how machine speed, and human overtrust on autonomous systems make it difficult for humans to act in time when severe events are about to happen. Soldiers can meet similar consequences when are utilizing AWS. Thus, to ensure human judgment as a crucial factor in decision-making process, autonomous systems used in context like warfare, should be designed in a way that could operate at a pace with human capabilities.

As it remains extremely difficult to overcome the issue of black box and unpredictability of autonomous systems, developers and researchers needs to ensure the safest design possible for the very beginning and guarantee constant training to those who are in charge of utilizing autonomous systems. This means that it is also important to test autonomous system in real scenarios.

Future Perspective on the Application of Autonomous Systems

In recent times, discussions surrounding AI and autonomous systems have often focused on their potential negative impacts. While it is crucial to address the risks associated with these technologies, it is equally important to recognize the numerous benefits they bring to human life. Just as the inventions of the Internet and GPS revolutionized society despite inherent risks, AI and autonomous systems have the potential to significantly enhance human capabilities across various spheres. This chapter aims to shed light on the positive contributions of AI and AI-enabled autonomous systems in civil applications, presenting real-life scenarios from fields such as medicine, automation, and waste management. Despite concerns raised, AI has proven to be a valuable ally in supporting human labor. AI systems and autonomous technologies have demonstrated their ability to augment human capabilities, resulting in improved efficiency and outcomes in various sectors, as it will be shown next. By automating repetitive tasks and streamlining complex processes, these technologies allow human workers to focus on tasks that require creativity, problem-solving, and emotional intelligence. This symbiotic relationship between humans and machines empowers individuals to achieve more and drive progress in society.

Medical treatment

The medical field is not absent from the technological changes over the past few years but is becoming one of the most dynamic sectors in the innovation and development of artificial intelligence. According to the research "Artificial Intelligence and life in 2030" the medical sector is considered one of the eight sectors in which in artificial intelligence will have the greatest impact (Stone, et al., 2016). The application of new automated tools for diagnostics as well as robotics applied to medical care, are all products that open the door to medicine 4.0 where the learning and problem-solving capacity of machines and the possibility of real-time information transmission is changing the way medicine is done (Javaid, Haleem, Singh, & Suman, 2021). In recent years, AI has begun to transform numerous medical activities, from automating administrative tasks, developments of apps capable of supporting medical decision-making processes, whether through the implementation of machines capable of aiding medical processes for identifying cancers and diseases (Gregory, 2023).

The role of artificial intelligence and addressing with the pandemic

Pandemics have shaped the course of human history, leaving an indelible mark on societies and reshaping the global landscape. Throughout the ages, infectious diseases have emerged, spreading rapidly and threatening the health, economies, and social fabric of nations. From the devastating Black Death in the 14th century to the Spanish flu in the early 20th century, pandemics have tested the resilience and adaptability of humanity. However, in recent memory, the world has been confronted with an unprecedented challenge—the COVID-19 pandemic drastically and in throughout these years, AI and autonomous systems have helped society to mitigate the virus (European Commission, 2020). As the pandemic required a huge amount of medical expertise, equipment and medicines, AI gained an important role on detecting, monitoring, and predicting covid outbreaks and has contributed to aiding in virus diagnosis (Malik, et al., 2021). Furthermore, AI systems have helped and supported doctors and nurses throughout the peak of the pandemic all around the word. Keeping machines into hospitals and developing new applications systems that could face the amount of help requested from communities has made it possible to decrease the number of medical staff and hospital and clinics (Sarker, Jamal, Ahmed , & Irtisam, 2021).

During the pandemic, hospitals and cure centers represented hot spots for virus transmission, thus having the possibility to implement such AI systems has helped to mitigate the risk of transmission among people. First of all, in order to detect fast covid symptoms, multitude of AI medical equipment was implemented, such as Dr. Spot, a temperature screening system able to capture the temperature of people passing by and capable of carrying out monitoring operation such as keeping track of people's vital symptoms. during a pandemic such as COVID-19, time becomes an important factor, especially in terms of being able to help and save people and figure out what they need in the shortest possible time (ibidem). In this AI along with imaging techniques played a key role. in this case computer thermography (CT) could rapidly show imagines of lungs and matching them with possible COVID conditions (Shamman, Hadi, Ramul, Zahra, & Gheni, 2023), thus it reduced the waiting time of the patients, to be diagnosed for covid, up to 70% (Sarker, Jamal, Ahmed , & Irtisam, 2021). Another machine learning system was implemented to predict the likelihood of symptoms complication on patients, whereas other were developed to support doctors on deciding what medical care was more appropriate (ibidem). Furthermore, AI powered application were used to keep track and identify new hot spot and covid clusters. Apps like *Immuni,* in Italy (Decideurs, 2020) or the French, *TousAntiCovid* (Euroactive, 2020) were developed by the government with the purpose of warning individuals that they have come into contact with an infected person. These kinds of applications have led to streamlining and making the contact racing process much faster, which otherwise would have taken a lot of time and manpower. In order to decrease the presence of staff in highly contagious places such as hospitals and medical centers, the development of robots and autonomous systems has been fundamental especially the delivery and supply chain process. UAVs and robots were used delivery food, medicines and supplies for people in need (Marr, 2020). other robots were used to disinfect places infected by COVID. These presented are just few example that help to understand how, AI could support human labor in critical or emergency were lives can be in dangerous.

During the COVID pandemic we saw an increased use of AI to detect symptoms, as a result of these increments we see a similar trend in that medical professionals are increasingly willing to use such technologies. For instance, the study conducted by Herny and his colleagues has the aim of understanding what the doctors' feelings are towards a new machine based on artificial intelligence, capable of identifying and treat sepsis, called TREWS (Henry, et al., 2022). Even if some of the doctors interviewed showed some doubts about the machinery, surprisingly, most of the clinician’s involved express their willingness to incorporate the feedback provided by the machines into their work. Some urges that even if they do not properly understand how the AI behind TREWS work they feel positive about having another “pair of eyes” to alert them if any change in patients’ symptoms happen. As other believe that any machinery that lightens the emotional and cognitive weight of the doctor is useful in this type of work, regardless of if they learn the mechanics behind the AI system, it would not change their individual decision making.

Waste management

The increase in waste represents a serious threat for human health and natural environment which could get worst when there is an improper handling and disposal of waste. According to the world bank, in 2020, 2.24 billion tons of solid waste was produced, and after COVID-19 the number has increase notably (World Bank, 2022). Waste management is not just putting the litter in the right bin, but it is a complex system were legal, environmental, political, technical and social factors play a role (Sharma & Vaid, 2021). In this sense, AI has emerged as a powerful tool in the process of waste management. AI based technology offers new possibilities for optimizing waste collection, sorting, recycling, and disposal process, by reducing time and human error that characterize manual work.

There are different AI systems that are used to perform waste sorting. For instance, artificial neural networks were used to identify waste fractions, together with deep conventional neural network for classifying and segregate garbage (Hamdy, Darwish, & Hassanien, 2021) . For such purpose, a British start-up has created, and AI system, called Grayparrot, designed able to analyze waste processing in recycling facilities (Grayparrot, n.d.). The systems are able to spot incorrectly classified waste on a conveyer belt faster with greater accuracy than humans (ibidem). It identifies different types of waste by also providing composition information which help facilitate and increase their recycling rates (Hackl, 2020). Furthermore, thanks to AI system enabled by Genetic Algorithm and Geographics information System, waste transportation has been optimized. There are bins that have sensors that are capable of measuring their waste levels and transmitting that information to the main disposal system for processing, thus this will help reducing carbon emissions as well, as garbage truck can be directed to collect waste only from filled bins (Sharma & Vaid, 2021). Finally, AI based system can be also useful to detect illegal dumping and reporting to authorities. The studied proposed by Du (Du, 2020). Highlight how artificial intelligence decision tree process could be helpful to detect illegal trash disposal leading to identify truck that have been involved in unlawful dumping. These examples reported have the intent to show how AI has the potential to revolutionize waste management practices, making them more efficient and accurate, from sorting to disposal process, AI enhance human capability to detect unlawful acts. Thus, AI can contribute to guarantee a more sustainable future.

Autonomous Vehicles

As analyzed in previous chapters, autonomous driving systems are a major innovation especially in the civilian sector. Like any autonomous system, self-driving vehicles represent a double-edged sword with the significant challenges alongside the potential for considerable benefits. As we navigate the future of autonomous transportation, it is essential to give a look also on the advantages that autonomous vehicles could bring in the future. First of all, there is a general understanding that autonomous cars in the future will diminish fatal crashes (Morando, Tian, Truong, & Vu, 2018).

Different studies shows that most of the car accidents are the result of drivers' errors, that can involve alcohol, fatigue and distraction (Fagnant & Kockelman, 2015). Thus, it is thought that replacing human presence within the car, we will be able to contrast human's psychological limits (Milakis, Van Arem, & Van Wee, 2017). Second of all, autonomous car could improve lives of elderly people or people with disabilities that for their condition are not able to drive. Hence, self-driving car can function as a support in these situations, giving these people part of the autonomy, they deserve (Yang & Coughlin, 2014). Furthermore, according to the study held by the University of Cambridge, introducing more and more autonomous cars within the streets will reduce traffic flow and road congestion up to 35%. (University of Cambridge, 2019). However, in order to make this possible, autonomous car should be widespread and able to communicate with each other. This is where the vehicle to vehicle and vehicle to infrastructure technologies come into play (Dey, Rayamajhi, Chowdhury, Bhavsar, & Martin, 2016). The first system allows cars to communicate with each other via WiFi networks, whereas the second one, makes it possible to share information from the vehicle to city's highway system (ibidem). This means that vehicles could gather data from cameras, traffic light and other road components and drive accordingly to this information. Last but not least, it is thought that the possibility of reducing traffic jams by introducing a substantial number of autonomous cars, will reduce fuel emissions (Zewe A. , 2022). According to this research, fuel consumption will be reduced by 18% and carbon dioxide emissions by 25% since autonomous cars will drive according to road limits and from information from road systems.

The scenarios presented in this chapter demonstrate how AI is already contributing positively to society in different fields and towards different directions such technology could go. Even if these are just few of possible examples, they are essential to acknowledge the significant positive impact they can have on human labor, lifestyle and advancement in battlefield. As we continue to harness the potential of AI responsibly and address concerns, we can unlock further benefits for human labor and create a future of collaboration and progress, while always being aware of their risks.

Bibliography

Adler, R. (2020). *Which laws, standards, and research initiatives exist to make Artificial Intelligence and Autonomous Systems safe?* Retrieved from Fraunhofer IESE: https://www.iese.fraunhofer.de/blog/which-laws-standards-and-research-initiatives-exist-to-make-artificial-intelligence-and-autonomous-systems-safe/

Ajitha, P. V., & Nagra, A. (2021). An Overview of Artificial Intelligence in Automobile Industry – A Case Study on Tesla Cars. *Solid State Technology, 64*(2), pp. 503-512.

Allsop, J., & Keeves, P. (2019). *UBER NOT CRIMINALLY LIABLE FOR ARIZONA CRASH*. Retrieved from HSF: https://hsfnotes.com/cav/2019/03/06/uber-not-criminally-liable-for-arizona-crash/#:~:text=While%20the%20vehicle%20was%20operating%20in%20autonomous%20mode,as%20such%20formal%20charges%20will%20not%20be%20brought.

Amoroso, D., Garcia, D., & Tamburrini, G. (2022). *The Weapon that Mistook a School Bus for an Ostrich.* doi:https://doi.org/10.1126/scidip.ade6750

Antsaklis, P. (2020). Autonomy and metrics of autonomy. *Annual Reviews in Control, 49*, 15-26.

Arkin, R. (2009). *Governing Lethal Behavior in Autonomous Robots.*

Arkin, R. (2013). The Robot didn't do it: A position paper for the Workshop on Anticipatory Ethics, Responsibility and Artificial Agents. *Workshop on Anticipatory Ethics, Responsibility and Artificial Agents* .

Article 36. (2016). Key elements of meaningful human control. *BACKGROUND PAPER*, 1-5.

Asaro, P. (2020). Autonomous Weapons and the Ethics of Artificial Intelligence. *Ethics of Artificial Intelligence*, 1-20. doi:https://doi.org/10.1093/oso/9780190905033.003.0008

Austria, Chile, & Brasile. (2018). *submitted to the Group of Governmental Experts on lethal autonomous weapons of the CCW*.

Avery, D. (2020). *AI-powered robots take over South Korea's No Brand Burger restaurant by collecting orders, cooking and serving customers to limit human contact amid the coronavirus pandemic'.* Retrieved from Dailymail: https://www.dailymail.co.uk/sciencetech/article-8734717/AI-Robots-serve-restaurant-customers-South-Korea.html

Baranson, J. (1970). Technology Transfer Through the International Firm. *The American Economic Review, 60*(2), 435–440. Retrieved from http://www.jstor.org/stable/1815842

BBC news. (2020). *Uber's self-driving operator charged over fatal crash*. Retrieved from BBC News: https://www.bbc.com/news/technology-54175359

Beard, J. M. (2014). Autonomous Weapons and Human Responsibilities. *Georgetown Journal of International Law*(45), 618-681. doi:https://core.ac.uk/download/pdf/77934296.pdf

Behringer, R., Sundareswaran, S., Gregory, B., Elsley, R., Addison, B., Guthmiller, W., . . . Bevly, D. (2004). The DARPA grand challenge - development of an autonomous vehicle. *EEE Intelligent Vehicles Symposium*, pp. 226-231. doi:10.1109/IVS.2004.1336386.

Bekey, G. (2005). *AUTONOMOUS ROBOTS, From Biological Inspiration to Implementation and Control.* MIT press.

Bellais, R., & Guichard, R. (2006). DEFENSE INNOVATION, TECHNOLOGY TRANSFERS AND PUBLIC POLICY. *Defence and Peace Economics, 17*(3), 273-286. doi: https://doi.org/10.1080/10242690600645274

Berk, B. (2021). *Driven: Google, Uber, and the Battle to Build an Autonomous Car*. Retrieved from Car and Driven : Driven: Google, Uber, and the Battle to Build an Autonomous Car

Bonnefon, J. F., Shariff, A., & Rahwan, I. (2015). Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars? *ArXiv*, pp. 1573-1576.

Bonnefon, Shariff, & Rahwan. (2016). The social dilemma of autonomous vehicles. *Science Magazine*. Retrieved from Science: https://www.science.org/doi/10.1126/science.aaf2654

Borenstein, J., Wagner, A. R., & Howard,, A. (2018). Overtrust of Pediatric Health-Care Robots: A Preliminary Survey of Parent Perspectives. *IEEE Robotics & Automation Magazine, 25*(1), 6-54.

Boulanin, V., Davison, N., Goussac, N., & Carlsson, M. P. (2020). *Limits on autonomy in weapon systems Identifying Practical Elements of Human Control.* SIPRI. ICRC.

Buchanan B., M. T. (2017). Machine Learning for Policy Makers. *Cyber Security Project, Belfer Center*.

Burns, C.& Dikmen, M., (2017). Trust in autonomous vehicles: The case of Tesla Autopilot and Summon. *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1093-1098. doi:10.1109/SMC.2017.812275

Burns, E. (n.d.). *Artificial Intelligence*. Retrieved from TechTarget: https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence

Burridge, T. (2020). *Boeing's 'culture of concealment' to blame for 737 crashes.* Retrieved from BBC: https://www.bbc.com/news/business-54174223

Caine, P. (2019). *The defence sector needs to leverage civil technologies to win the innovation battle*. Retrieved from Linkedin: https://www.linkedin.com/pulse/defence-sector-needs-leverage-civil-technologies-win-innovation

Campbell, D. (2019). *The many human error that brought down the Boeing 737 Max*. Retrieved from The verge: https://www.theverge.com/2019/5/2/18518176/boeing-737-max-crash-problems-human-error-mcas-faa

Cao, X., Yang, X., & Zhang, L. (2020). Conversion of Dual-Use Technology: A Differential Game Analysis under the Civil-Military Integration. *Symmetry, 12*(11), 1-18. doi:https://doi.org/10.3390/sym12111861

Carozza, I., Marsh, N., & Reichberg, G. (2021). Dual-Use AI Technology in China, the US and the EU. *PRIO*, 1-50. Retrieved from https://www.prio.org/publications/13150

CCW. (2019). Report of the 2019 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects*, 1-14. Retrieved from https://documents.unoda.org/wp-content/uploads/2020/09/CCW_GGE.1_2019_3_E.pdf

Chen, H., Wang, X., & Li, Y. (2009). A Survey of Autonomous Control for UAV. *International Conference on Artificial Intelligence and Computational Intelligence*, 267-271. doi:10.1109/AICI.2009.147

Chiang, J.-T. (1991). Technological "Spin-Off": Its Mechanisms and National Contexts. *Technological Forecasting and Social Change, 41*(4), 365-390. doi:https://doi.org/10.1016/0040-1625(92)90044-T

Chiang, J.-T. (1991). The Potential of "Spin-off' from a system perspective. *Technology in Society, 13*(3), 279-287. doi:https://doi.org/10.1016/0160-791X(91)90004-G

China. (2018). Convention on certain conventional weapons: Positiion Paper. doi:https://unog.ch/80256EDD006B8954/(httpAssets)/E42AE83BDB3525D0C125826C0040B262/$file/CCW_GGE.1_2018_WP.7.pdf

Chung, W. (2001). Identifying Technology Transfer in Foreign Direct Investment: Influence of Industry Conditions and Investing Firm Motives. *International Business Studies*(32), 211–229. doi:https://doi.org/10.1057/palgrave.jibs.8490949

Clancy, J. P. (2019). Breakdowns in Human-AI Partnership: Revelatory Cases of Automation Bias in Autonomous Vehicle Accidents. Chapel Hill: University of North Carolina.

Commission, E. (2020). *Horizon 2020 – Work Programme 2018-2020.* Retrieved from https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/2018-2020/annexes/h2020-wp1820-annex-ga_en.pdf

Congressional Research Service. (2023). Defense Primer: U.S. Policy on Lethal Autonomous Weapon System. *Congressional Research Service*, (pp. 1-3). Retrieved from https://crsreports.congress.gov/product/pdf/IF/IF11150

Crootof, R. (2016). A Meaningful Floor for 'Meaningful Human Control. *Temple Journal of International & Comparative Law, 30*, 53-62.

Dada, E., Bassi, J., Chiroma, H., Abdulhamid, S., Adetunmbi, A., & Ajibuwa, O. (2009). Machine learning for email spam filtering: review, approaches and open research problems. *Helyon, 5*(6). doi:https://doi.org/10.1016/j.heliyon.2019.e01802

Daniels, O. (2022). *The AI "Revolution in Military Affairs": What Would it Really Look Like?* Retrieved from CSET: https://cset.georgetown.edu/article/the-ai-revolution-in-military-affairs-what-would-it-really-look-like/

Darling, K. (2022). Human Drivers Should not be Responsible for Accidents Caused by Autonomous Vehicles. Retrieved from Science focus: https://www.sciencefocus.com/news/human-drivers-should-not-be-responsible-accidents-caused-autonomous-vehicles/

Davison, N. (2018). A legal perspective: Autonomous weapon systems under International Humanitaria Law. *UNODA Occasional Papers*(30), 1-18.

Decideurs. (2020). Immuni: Italy’s Coronavirus-tracking App. Retrieved from Decideurs: https://www.leadersleague.com/fr/news/immuni-italy-s-coronavirus-tracking-app

Dey, K. C., Rayamajhi, A., Chowdhury, M., Bhavsar, P., & Martin, J. (2016). Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) Communication in a Heterogeneous Wireless Network - Performance Evaluation. *Transportation Research Part C: Emerging Technologies, 68*, 168-184.

Doorn, N. (n.d.). *The Problem of Many Hands: Who is responsible*. Retrieved from Tudelft: https://ocw.tudelft.nl/course-lectures/1-4-problem-many-hands-responsible/

Du, X. (2020). Application of traceability system for dumping garbage of north canal embankment based on ai artificial intelligence recognition. *Mater Sci Eng*. doi:10.1088/1757-899X/740/1/012121

Edgerton, D. (1988). The Relationship between Military and Civil Technology: A Historical Perspective. In P. R. Gummett, & N. A. series (Ed.), *The Relations between Defence and Civil Technologies* (Vol. 46, pp. 106-113). doi:https://doi.org/10.1007/978-94-015-7803-5_7

ESMT Berlin. (2023). *New research from ESMT Berlin: Workplace machine learning improves accuracy, but increases human’s workload, too*. Retrieved from ESMT Berlin: https://esmt.berlin/press/new-research-esmt-berlin-workplace-machine-learning

Euroactive. (2020). tousanticovid. Retrieved from Euroactive: https://www.euractiv.com/topics/tousanticovid/

European Commission. (2020). *Danish disinfection robots save lives in the fight against the Corona virus*. Retrieved from Shaping Europe’s digital future: https://digital-strategy.ec.europa.eu/en/news/danish-disinfection-robots-save-lives-fight-against-corona-virus

Ezenkwu, C., & Starkey, A. (2019). Machine Autonomy: Definition, Approaches, Challenges and Research Gaps. *Intelligent Computing - Proceedings of the Computing Conference, 997*, 335–358. doi:https://link.springer.com/chapter/10.1007/978-3-030-22871-2_24

Fagnant , D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice, 77*, 167-181. doi:https://doi.org/10.1016/j.tra.2015.04.003

Falcone, R., & Castelfranchi, C. (2001). The human in the loop of a delegated agent: the theory of adjustable social autonomy. *transactions on systems, man and cybernetics., 31*(5), pp. 406-418. doi:doi:10.1109/3468.952715

Farhud, D. D., & Zokaei , S. (2021). Ethical Issues of Artificial Intelligence in Medicine and Healthcare. *Iran J Public Health, 50*(11). doi:10.18502/ijph.v50i11.7600

Firlej, M., & Taeihagh, A. (2020). Regulating human control over autonomous systems. *Regulation and Governance, 15*(4), pp. 1071-1091. doi:https://doi.org/10.1111/rego.12344

Fisher, M., Mascardi, V., Rozier, K. Y., Schlingloff, B. H., Winikoff, M., & Yorke-Smith, N. (2021). Towards a framework for certification of reliable autonomous systems. *Autonomous Agents and Multi-Agent Systems, 35*(8). doi:https://doi.org/10.1007/s10458-020-09487-2

Fjelland, R. (2020). Why general artificial intelligence will not be realized. *Humanities and Social Sciences Communications, 7*(10). doi:https://doi.org/10.1057/s41599-020-0494-4

France. (2016). "Characterization of A Laws". In Meeting of experts on lethal autonomous weapons systems (LAWS). *Working Paper*. Retrieved from https://unog.ch/80256EDD006B8954/(httpAssets)/5FD844883B46FEACC1257F8F00401FF6/$file/2016_LAWSMX_CountryPaper_France+CharacterizationofaLAWS.pdf

Freeze, J. (2021). Alexa, What Does The Smart Home Teach Us About AI? Retrieved from Forbes: https://www.forbes.com/sites/forbescommunicationscouncil/2021/02/19/alexa-what-does-the-smart-home-teach-us-about-ai/?sh=367ea74d1e53z

Freitas, J. d. (2023). Will We Blame Self-Driving Cars? *The Wall Street Journal.* Retrieved from gahttps://www.wsj.com/articles/will-we-blame-self-driving-cars-11674745636

Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design Shaping Technology with Moral Imagination.* The MIT Press.

Geiss, R. (2016). Autonomous Weapons Systems: Risk Management and State Responsibility. *Third CCW meeting of experts on lethal autonomous weapons systems*, (pp. 1-3). Geneva.

Geisslinger, M., Poszler, F., Betz, J., Lütge, C., & Lienkamp , M. (2021). Autonomous Driving Ethics: from Trolley Problem to Ethics of Risk. *Philosophy & Technology, 34*, 1033–1055. doi:https://doi.org/10.1007/s13347-021-00449-4

Gelles, D., Kitroeff, N., Nicas, J., & Ruiz, R. (2019). *Boeing Was 'Go, Go, Go' to Beat Airbus With the 737 Max*. Retrieved from The New York Times: https://www.nytimes.com/2019/03/23/business/boeing-737-max-crash.html

Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in Automation – Before and After the Experience of Take-over Scenarios in a Highly Automated Vehicle. *Procedia Manufacturing, 3*, 3025-3032. doi:https://doi.org/10.1016/j.promfg.2015.07.847

Gopalakrishnan, S., & Santoro, M. (2004). Distinguishing Between Knowledge Transfer and Technology Transfer Activities: The Role of Key Organizational Factors. *IEEE TRANSACTIONS ON ENGINEERING MANAGEMENT, 51*(1), 57-69. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1266854

Grayparrot. (n.d.). *You can't optimise if you don't measure*. Retrieved from Grayparrot: https://www.greyparrot.ai/en-gb/applications

Gregory, A. (2023). *New artificial intelligence tool can accurately identify cancer.* Retrieved from The Guardian: https://www.theguardian.com/society/2023/apr/30/artificial-intelligence-tool-identify-cancer-ai

Hackl, C. (2020). *How 4 Companies Are Using AI To Solve Waste Issues On Earth And In Space*. Retrieved from Forbes: https://www.forbes.com/sites/cathyhackl/2020/07/18/how-4-companies-are-using-ai-to-solve-waste-issues-on-earth--in-space/?sh=4982fa5b35fa

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review, 61*(4), 5-14. doi:https://doi.org/10.1177/00081256198649

Hamdy, W., Darwish, A., & Hassanien, A. (2021). Artificial Intelligence for Sustainable Waste Management and Control During and Post COVID-19 Crisis: Critical Challenges. In A. Hassanien, A. Darwish, & B. Gyampoh , *The Global Environmental Effects During and Beyond COVID-19* (pp. 81–91). doi:https://doi.org/10.1007/978-3-030-72933-2_5

Hawkins, A. J. (2019). *Serious safety lapses led to Uber's fatal self-driving crash, new documents suggest*. Retrieved from The Verge: https://www.theverge.com/2019/11/6/20951385/uber-self-driving-crash-death-reason-ntsb-dcouments

Heinzelman, G. (2019). Autonomous Vehicles, Ethics of Progress. pp. 1-24.

Henry, K., Kornfield, R., Sridharan, A., Linton, R., Groh, C., Wang, T., . . . Saria, s. (2022). Human–machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. *digital medicine*(97). doi:https://doi.org/10.1038/s41746-022-00597-7

Hepher, T., & Vats, R. (2021). *Europe lifts safety ban on Boeing 737 MAX jet.* Retrieved from Reuters: https://www.reuters.com/article/us-boeing-737max-idUSKBN29W17U

Herkert, J., Borenstein, J., & Miller, K. (2020). The Boeing 737 MAX: Lessons for Engineering Ethics. *Science and Engineering Ethics, 26*, pp. s2957–2974. doi:https://doi.org/10.1007/s11948-020-00252-y

Hevelke, A., & Nida-Rümelin, J. (2015). Responsibility for Crashes of Autonomous Vehicles: An Ethical Analysis. *Science Engineer Ethics 21,*(21), pp. 619–630. doi:https://doi.org/10.1007/s11948-014-9565-5

High-Level Expert Group. (2019). A definition of AI: Main Capabilities and Disciplines. European Parliament. Retrieved from https://www.europarl.europa.eu/italy/resource/static/files/import/intelligenza_artificiale_30_aprile/ai-hleg_policy-and-investment-recommendations.pdf

Holeman, D. L. (2019). *The Boeing 737 MAX: A Case Study of Systems Decisions and Their Consequences*. Retrieved from The Systems Perspective: https://dennisholeman.com/the-boeing-737-max-a-case-study-of-systems-decisions-and-their-consequences/

Horowitz, M. (2018). Artificial Intelligence, International Competition, and the Balance of Power. *Texas National Security Review, 1*(3), 37-57. Retrieved from https://tnsr.org/2018/05/artificial-intelligence-international-competition-and-the-balance-of-power/

Horowitz, M., & Scharre, P. (2015). Meaningful Human Control in Weapon Systems: A Primer. *Project On Ethical Autonomy Working Paper*, 1-16. doi:https://www.files.ethz.ch/isn/189786/Ethical_Autonomy_Working_Paper_031315.pdf

Howard, A., & Borenstein, J. (2018). The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and SOcial Inequity. *Science Engineering Ethics*(24), 1521-1536.

Hsieh, D. M. (1995). *Aristotle on Moral Responsibility.* t-Louis: Washington University.

Huang, H., & Liu, S. (2022). Are consumers more attracted to restaurants featuring humanoid or non-humanoid service robots? *International Journal of Hospitality Management, 107*. doi:https://doi.org/10.1016/j.ijhm.2022.103310

Human Rights Watch. (2016). *Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban, report.*

Human Rights Watch. (2016). Killer Robots and the Concept of Meaningful Human Control. *Memorandum to Convention on Conventional Weapons (CCW) Delegates*. Retrieved from https://www.hrw.org/news/2016/04/11/killer-robots-and-concept-meaningful-human-control

IBM. (n.d.). *What is machine learning?* Retrieved from IBM: https://www.ibm.com/topics/machine-learning

ICRAC. (n.d.). *About ICRAC*. Retrieved from ICRAC: https://www.icrac.net/about-icrac/

ICRC. (2014). Command responsibility and failure to act. *DVISORY SERVICE on Intenrational Humanitarian Law*. Retrieved from file:///C:/Users/loryc/Downloads/command-responsibility-icrc-eng.pdf

Jain, N. (2016). Autonomous weapons systems: New frameworks for individual responsibility. In *Autonomous Weapons Systems.*

Javaid, M., Haleem, A., Singh, R., & Suman, R. (2021). Substantial capabilities of robotics in enhancing industry 4.0 implementation. *Cognitive Robotics*(1), 58-75. doi:https://doi.org/10.1016/j.cogr.2021.06.001

Javapoint. (n.d.). *What is an Expert System?* Retrieved from Javapoint: https://www.javatpoint.com/expert-systems-in-artificial-intelligence

Johnston, P., & Harris, R. (2019). The Boeing 737 MAX Saga: Lessons for Software Organizations. *Software Quality Professional, 21*(3), pp. 1-12. Retrieved from https://embeddedartistry.com/wp-content/uploads/2019/09/the-boeing-737-max-saga-lessons-for-software-organizations.pdf

Kamińska, J. (2022). *8 types of data bias that can wreck your machine learning models*. Retrieved from Statice: https://www.statice.ai/post/data-bias-types

Keiler, J., Panzavolta, M., & Roef , D. (2017). *Introduction to Law.* Springer, Cham.

Klincewicz, M. (2015). (2015), 'Autonomous Weapons Systems, the Frame Problem and Computer Security,' Journal of Military Ethics), pp.162–. *Journal of Military Ethics* , 162-176.

Kretz, A. (2019). *A Simple Explanation of the Machine Learning Workflow*. Retrieved from Medium: https://medium.com/plumbersofdatascience/a-simple-explanation-of-the-machine-learning-workflow-c5d43d9f5b1c#:~:text=THE%20TRAINING%20PHASE&text=A%20phase%20where%20you%20are,you%20are%20training%20the%20algorithm.

Kumar, V., Kumar, U., & Persaud. (1999). Building Technological Capability through Importing Technology: The Case of Indonesian Manufacturing Industry. *journal of Technology Transfer*(24), 81-96. doi:http://dx.doi.org/10.1023/A:1007728921126

Lagerwall, A. (2015). *Jus Cogens*. Retrieved from Oxford BIbliographies: https://www.oxfordbibliographies.com/display/document/obo-9780199796953/obo-9780199796953-0124.xml

Lan, P., & Young, S. (1996). International Technology Transfer Examined at Technology Component Level: A Case Study in China. *Technovation, 16*(6), 277-286. doi:https://doi.org/10.1016/0166-4972(96)00005-3

Langewiesche, W. (2019). *What Really Brought Down the Boeing 737 Max? Malfunctions caused two deadly crashes. But an industry that puts unprepared pilots in the cockpit is just as guilty.* Retrieved from The New York Times Megazine: https://www.nytimes.com/2019/09/18/magazine/boeing-737-max-crashes.html

Leslie, D. (2020). Understanding bias in facial recognition technologies. 1-49. Retrieved from https://arxiv.org/abs/2010.07023

Levin, S. (2020). *Safety driver charged in 2018 incident where self-driving Uber car killed a woman*. Retrieved from The Guardian: https://www.theguardian.com/us-news/2020/sep/16/uber-self-driving-car-death-safety-driver-charged

Malik, Y., Sircar , S., Bhat, S., Asari, M., Pande, T., Kum, P., . . . Dhama, K. (2021). How artificial intelligence may help the Covid-19 pandemic: Pitfalls and lessons for the future. *Rev Med Virol, 31*(5). doi:10.1002/rmv.2205

Malunga, M., Ahmed, R., Raynard, A., Stoyanova, M., & Chisala, K. (2022). Key lessons from the Boeing 737 MAX 8 accidents. *Institution of Chemical Engineers*, pp. 24-26. Retrieved from https://www.icheme.org/media/19013/lpb287_pg24.pdf

Marr, B. (2020). *Robots And Drones Are Now Used To Fight COVID-19*. Retrieved from Forbes: https://www.forbes.com/sites/bernardmarr/2020/03/18/how-robots-and-drones-are-helping-to-fight-coronavirus/

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and information technology*, 175-183.

McFadden, C. (2020). *9 Military Spin-Off Technologies We Use Almost Everyday*. Retrieved from Interesting Engeneering: https://interestingengineering.com/innovation/9-military-spin-off-technologies-we-use-almost-everyday

McFadden, C., Schecter, A., Monahan, K., & Schapiro, R. (2019). *Former Boeing manager says he warned company of problems prior to 737 crashes*. Retrieved from NBC News: https://www.nbcnews.com/news/us-news/former-boeing-manager-says-he-warned-company-problems-prior-737-n1098536

McFarland, M. (2015). *In 1956, here's how GM imagined the self-driving car would work*. Retrieved from The Washington Post: https://www.washingtonpost.com/news/innovations/wp/2015/06/24/in-1956-heres-how-gm-imagined-the-self-driving-car-would-work/

Milakis, D., Van Arem, B., & Van Wee, B. (2017). Policy and society related implications of automated driving: A review of literature and directions for future research. *Intelligence Trasportation system, 21*, 324–348. doi:10.1080/15472450.2017.1291351

Miriri, D., Fick, M., & Maasho, A. (2019). *Ethiopian Airlines flight crashes, killing 157.* Retrieved from Reuters: https://www.reuters.com/article/us-ethiopia-airplane-idUSKBN1QR07V

Morando, M. M., Tian, Q., Truong, L. T., & Vu, H. L. (2018). Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures. *Journal of Advanced Transportation*. Retrieved from https://doi.org/10.1155/2018/6135183

NHTSA . (2017). *ODI resume*. Retrieved from NHTSA: https://static.nhtsa.gov/odi/inv/2016/INCLA-PE16007-7876.PDF

Nikolaidis, S., Zhu, Hsu, D., & Srinivasa, S. (2017). Human-Robot Mutual Adaptation in Shared Autonomy. *HRI '17: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 294–302.

Noone, G. P., & Noone, D. C. (2015). 'Debate Over Autonomous Weapons Systems. *Case Western Reserve Journal of International Law, 47*(1), 25-35.

NTSB. (2019). Highway Accident Report Collision Between a Car Operating With Automated Vehicle Control Systems and a Tractor-Semitrailer Truck., (pp. 1-53). Retrieved from https://www.ntsb.gov/investigations/accidentreports/reports/har1702.pdf

NYT. (2016). *As U.S. Investigates Fatal Tesla Crash, Company Defends Autopilot System*. Retrieved from The New York Times: https://www.nytimes.com/2016/07/13/business/tesla-autopilot-fatal-crash-investigation.html

NYT. (2019). *Boeing Built Deadly Assumptions Into 737 Max, Blind to a Late Design Change*. Retrieved from New York Times: https://www.nytimes.com/2019/06/01/business/boeing-737-max-crash.html

NYT. (2022). *The New Chatbots Could Change the World. Can You Trust them?* Retrieved from NYT: https://www.nytimes.com/2022/12/10/technology/ai-chat-bot-chatgpt.html

Olstad, A. (2022). *A Historical Overview of the Boeing 737 Family*. Retrieved from Aviation Source: https://aviationsourcenews.com/history/a-historical-overview-of-the-boeing-737-family/

Pacific ventury. (2019). *The 737 MAX 8 Case: What happened to the Leadership compass of Boeing?* Retrieved from Pacific Ventury: https://www.pacificventury.com/post/the-737-max-8-case-what-happened-to-the-leadership-compass-of-boeing

Panahi, O. (2016). *Could There Be A Solution To The Trolley Problem?* Retrieved from Philosophy now: https://philosophynow.org/issues/116/Could_There_Be_A_Solution_To_The_Trolley_Problem

PAX. (2019). Don't be Evil? A survey of the tech sector's stance on lethal autonomous weapons. *PAX*, pp. 1-59. Retrieved from https://paxforpeace.nl/wp-content/uploads/sites/2/import/import/pax-report-killer-robots-dont-be-evil.pdf

Perell, D. (2019). *Why Did The Boeing 737 Max Crash?* Retrieved from David Perell: https://perell.com/essay/boeing-737-max/

Phillips, R. G. (2002). Technology business incubators: how effective as technology transfer mechanisms? *Technology in Society, 24*(3), 299-316. doi:https://doi.org/10.1016/S0160-791X(02)00010-6

Pianta, M. (1988). Technological Strategies. In M. Pianta, *New technologies across the Atlantic: US Leadership or European Autonomy?* (p. chapter 4.3). Wheatsheaf: Harvester. Retrieved from https://archive.unu.edu/unupress/unupbooks/uu38ne/uu38ne00.htm#Contents

Pilot Institute. (2023). *The Difference between Airbus and Boeing.* Retrieved from Pilot Institute: https://pilotinstitute.com/airbus-vs-boeing/

Porter, J. (2020). *Uber backup driver charged in fatal 2018 self-driving car crash*. Retrieved from The Verge: https://www.theverge.com/2020/9/16/21439354/uber-backup-driver-charged-autonomous-self-driving-car-crash-negligent-homicide

Purves, D., Jenkins, R., & Strawse, B. J. (2015). Autonomous Machines, Moral Judgment, and Acting for the Right Reasons. *Ethical Theory and Moral Practice*, pp. 851-872. doi:https://www.jstor.org/stable/24478765

Puscas, I. (2022). Human-Machine Interface in Autonomous Weapons Systems Consideration for Human Control. 9-15. Retrieved from https://www.unidir.org/sites/default/files/2022-07/UNIDIR_Human-Machine%20Interfaces.pdf

Reuters. (2019). *FAA failed to properly review 737 Max jet anti-stall system: JATR report*. Retrieved from CNBC: https://www.cnbc.com/2019/10/11/faa-failed-proper-review-of-boeings-737-max-mcas-system-jatr-report.html

Rickli, D. J.-M. (2022). Human-Machine Teaming in Artificial Intelligence-Driven Air Power Future Challenges and Opportunities for the Air Force. *The Air Power Journal*, pp. 1-12.

Roff, R., & Moyes. (2016). Meaningful Human Control, Artificial Intelligence and Autonomous Weapons. *Technical Report Briefing Paper for the Delegates at the Convention on Certain Conventional Weapons Informal Meeting of Experts on Lethal Autonomous Weapons Systems*, 1-6.

Russell, S., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach. Upper Saddle River.: Prentice-Hall.

Santoni de Sio, F., & Van Den Hoven. (2018). , Meaningful Human Control over Autonomous Systems ems: A Philosophical Account. *Frontiers in Robotics and AI, 5*, 1-14. doi:https://doi.org/10.3389/frobt.2018.00015

Saracco, R. (2019). *Autonomous systems in healthcare.* Retrieved from IEEE: https://cmte.ieee.org/futuredirections/2019/07/21/autonomous-systems-in-healthcare/

Sarker, S., Jamal, L., Ahmed , S., & Irtisam, N. (2021). Robotics and artificial intelligence in healthcare during COVID-19 pandemic: A systematic review. *Robotics and Autonomous System*, 1-14. doi:10.1016/j.robot.2021.103902

Shamman, A., Hadi, A., Ramul, A., Zahra, M., & Gheni, H. (2023). The artificial intelligence (AI) role for tackling against COVID-19 pandemic. *Material Today: procideedings, 80*(3), 3663-3667. doi:https://doi.org/10.1016/j.matpr.2021.07.357

Schaper, D. (2019). *Pilots Criticize Boeing, Saying 737 Max 'Should Never Have Been Approved'.* Retrieved from NPR: https://www.npr.org/2019/06/19/734248714/pilots-criticize-boeing-saying-737-max-should-never-have-been-approved

Scharre, P. (2018). *Autonomous Weapons System and the future of War: Army of None.* New York: Norton & Company.

Scharre, P., & Horowitz, M. C. (2015). an introduction to Autonomy in Weapon System. *PROJECT ON ETHICAL AUTONOMY WORKING PAPER*, 1-23. Retrieved from https://www.files.ethz.ch/isn/188865/Ethical%20Autonomy%20Working%20Paper_021015_v02.pdf

Schermer, M. (2002). Philosophical and Ethical Perspectives on Autonomy. In: The Different Faces of Autonomy. *Library of Ethics and Applied Philosophy, 13*, pp. 1-22. doi:https://doi.org/10.1007/978-94-015-9972-6_1

Schmid , S., Riebe, T., & Reuter, C. (2022). Dual-Use and Trustworthy? A Mixed Methods Analysis of AI Diffusion Between Civilian and Defense R&D. *Science and Engineering Ethics, 28*(12), 1-23. doi:https://doi.org/10.1007/s11948-022-00364-7

Schulzke, M. (2012). Autonomous Weapons and Distributed Responsibility. *Philosophy & Technology, 26*(2). doi:10.1007/s13347-012-0089-0

Schwarz, E. (2019). *The (im)possibility of meaningful human control for lethal autonomous weapon systems*. doi:https://blogs.icrc.org/law-and-policy/2018/08/29/im-possibility-meaningful-human-control-lethal-autonomous-weapon-systems/

SEA. (2021). *SAE J3016TM LEVELS OF DRIVING AUTOMATION*. Retrieved from SEA: https://www.sae.org/binaries/content/assets/cm/content/blog/sae-j3016-visual-chart_5.3.21.pdf

Searle, J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences, 3*(3), pp. 417–457. doi:https://doi.org/10.1017/S0140525X00005756.

Sezal, M., & Giumelli, F. (2022). Technology transfer and defence sector dynamics:the case of the Netherlands. *European Security*, 1-18. doi: https://doi.org/10.1080/09662839.2022.2028277

Sgobba, T. (2019). *B-737 Max And The Crash Of The Regulatory System.* Retrieved from Space Safety Magazine: https://www.spacesafetymagazine.com/aerospace-engineering/lessons-from-earth/b-737-max-and-the-crash-of-the-regulatory-system/

Sharkey, N. (2017). Why robots should not be delegated with the decision to kill. *Connection Science, 29*(2), pp. 177-186. doi:10.1080/09540091.2017.1310183

Sharkey, N. E. (2012). The evitability of autonomous robot warfare. *International Review of the Red Cross, 94*(886), 787-799. doi:10.1017/S1816383112000732

Sharma, P., & Vaid, P. (2021). Emerging role of artificial intelligence in waste management practices. *IOP Conf. Ser.: Earth Environ. Sci*, 1-10. doi:10.1088/1755-1315/889/1/012047

Shelton, D. (1999). *Remedies in International Human Rights Law.* Third Edition. Oxford

Skinner, B. (2018). Chapter 4 – Human-machine teaming. In U. M. Defence, *Human-Machine Teaming (JCN 1/18)* (pp. 39-55). Retrieved from https://www.gov.uk/government/publications/human-machine-teaming-jcn-118

Skitka, L., Mosier, K., & Burdick, M. (2000). Accountability and automation bias. *International Journal of Human-Computer Studies, 52*(4), 701-717. doi:https://doi.org/10.1006/ijhc.1999.0349

Smith, C. J. (2019). Designing Trustworthy AI: A Human-Machine Teaming Framework to Guide Development. Retrieved from https://arxiv.org/ftp/arxiv/papers/1910/1910.03515.pdf

Solomonoff, G. (2023). *The Meeting of the Minds That Launched AI*. Retrieved from IEEE Spectrum: https://spectrum.ieee.org/dartmouth-ai-workshop

Sparrow, R. (2007). Killer Robots. *Journal of Applied Philosophy, 24*(1), 62-77.

Sparrow, R., & Howard, M. (2017). When human beings are like drunk robots: Driverless vehicles, ethics, and the future of transport. *Transportation Research*, pp. 206-215.

Stephenson, K., & Hayden, F. (1994). Corporate Networks: A US Case Study. In J. Groenewegen, *Dynamics of the Firm Strategies of Pricing and Organisation* (pp. 53-92). Retrieved from https://www.researchgate.net/publication/329573216_Corporate_Networks_A_US_Case_Study

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., . . . Teller, A. (2016). Artifical Intelligence and Life in 2030 The One Hundred Year Study on Artificial Intelligence. *Computers and Society*, 1-52. Retrieved from https://ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/ai100report10032016fnl_singles.pdf

Stop Killer Robots. (n.d.). *Take Action*. Retrieved from Stop Killer Robots: https://www.stopkillerrobots.org/take-action/

Stornmont, D. (2008). Analyzing Human Trust of Autonomous Systems in Hazardous Environments. pp. 27-32. Retrieved from https://www.researchgate.net/publication/228457545_Analyzing_Human_Trust_of_Autonomous_Systems_in_Hazardous_Environments

Stowers, K., Brady, L., MacLellan, C., Wohleber, R., & Salas, E. (2021). Improving Teamwork Competencies in Human-Machine Teams: Perspectives From Team Science. *Frontiers in Psychology, 12*, 1-6. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2021.590290/full

Stowsky, J. (2004). Secrets to shield or share? new dilemmas for military R&D policy in the digital age. *Research Policy, 33*(2), 257-269. doi:https://ideas.repec.org/a/eee/respol/v33y2004i2p257-269.html

Suhartono, M., & Beech, A. (2019). *In Boeing Lion Air Crash, Indonesians Learn What Took Their Loved Ones.* Retrieved from The New York Times: https://www.nytimes.com/2019/10/23/world/asia/boeing-737-max-lion-air-crash.html

Swaine, L. (2016). The Origins of Autonomy. *History of Political Thought, 37*(2), 216-237. Retrieved from https://www.jstor.org/stable/26228692

Synopsys. (2020). *Where is my autonomous car?* Retrieved from Synopsys: https://www.synopsys.com/automotive/autonomous-driving-levels.html

Tamburrini, G., & Amoroso. (2016). *The convergence of deontological and consequentialist reasons for banning Autonomous Weapons System.* Napoli. Retrieved from http://www.fedoa.unina.it/11141/1/AWS_Preprint.pdf

Tara, R. (2019). *Boeing 737 MAX Pilots Had No Idea What They Were Up Against.* Retrieved from www.engineering.com: https://www.engineering.com/story/boeing-737-max-pilots-had-no-idea-what-they-were-up-against

Thatchan Research. (2018). *Assisted and Automated Driving*. Retrieved from Thatcham Research: https://www.abi.org.uk/globalassets/files/publications/public/motor/2018/06/thatcham-research-assisted-and-automated-driving-definitions-summary-june-2018.pdf

The Royal Acedemy of Engineering. (2009). Autonomous Systems: Social, Legal and Ethical Issues . *The Royal Academy of Engineering*, pp. 1-19. Retrieved from https://raeng.org.uk/media/tommwqvv/autonomous_systems_report.pdf

Theo, S. (2021). *AI In Aviation: Are You Ready To Fly Without A Human Pilot?* Retrieved from Electronicsforu.com: https://www.electronicsforu.com/technology-trends/tech-focus/ai-aviation-fly-without-human-pilot

Thompson, D. (1980). Moral Responsibility of Public Officials: The Problem of Many Hands. *American Political Science Review*, 905-916. doi: doi:10.2307/1954312

Thorn, A. (2020). *Boeing ceo blames pilots for 737 Max Crashes.* Retrieved from World of Aviation: https://worldofaviation.com/2020/03/boeing-ceo-blames-pilots-for-737-max-crashes/

Travis, G. (2019). *How the Boeing 737 Max Disaster Looks to a Software Developers.* Retrieved from IEEE Spectrum: https://spectrum.ieee.org/how-the-boeing-737-max-disaster-looks-to-a-software-developer

Turing, A. (1950). computing Machinary and Intelligence. *Mind*(49), 433-460. Retrieved from https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf

UK. (2018). Unmanned aircraft systems. *JDP*. Retrieved from https://www.gov.uk/government/publications/unmanned-aircraft-systems-jdp-0-302

UK Road Traffic Amendment. (2021). *Proposal of Amendment to Article 1 and new Article 34 BIS of the 1968 Convention on Road Traffic.* UK Government. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/10

Umbrello, S. (2020). Meaningful Human Control over Smart Home System: A Value Sensitive Design Approach. *Humana.mente Journal of Philosophical Studies,, 13*(37), 40–65. Retrieved from https://www.humanamente.eu/index.php/HM/article/view/315/270

Umbrello, S. (2021). Towards a Value Sensitive Design Framework for Attaining Meaningful Human Control . *Phd Dissertations Of The Northwestern Italian* , 129-143.

UN. (2001). Draft articles on Responsibility of States for Internationally Wrongful Acts. 31-143. Retrieved from https://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf

UN. (2012). Materials on the Responsibility of States for Internationally Wrongful Acts. *United Nations Legislative* .

University of Cambridge. (2019). *Driverless cars working together can speed up traffic by 35 percent*. Retrieved from University of Cambridge: https://www.cam.ac.uk/research/news/driverless-cars-working-together-can-speed-up-traffic-by-35-percent

US Securities and Exchange Commission. (2022). *Boeing to Pay $200 Million to Settle SEC Charges that it Misled Investors about the 737 MAX.* Retrieved from https://www.sec.gov/news/press-release/2022-170

Verbruggen, M. (2019). The Role of Civilian Innovation in the Development of Lethal Autonomous Weapon Systems. *Global Policy, 10*(3), 338-342. doi: https://doi.org/10.1111/1758-5899.12663

Vihul, L. (2020). *International Legal Regulation of Autonomous Technologies*. Retrieved from Centre for International Governance Innovation: https://www.cigionline.org/articles/international-legal-regulation-autonomous-technologies/

Wagner, A. R., Borenstein, J., & Howard, A. (2018). Overtrust in the Robotic Age. *communication of the ACM, 61*(9), 22-24. doi:10.1145/3241365

Wagner, M. (2014). The Dehumanization of International Humanitarian Law: Legal, Ethical, and Political Implications of Autonomous Weapon Systems. *Vanderbilt Journal of Transnational Law, 47*(5), 1372-1424.

Wallach, W., & Allen, C. (2013). Framing robot arms control. *Ethics Inf Technol, 15*, 125–135. doi:https://doi.org/10.1007/s10676-012-9303-0

waymo. (n.d.). *The World most experienced Driver*. Retrieved from waymo: https://waymo.com/

Wired. (2016). *People Want Self-Driving Cars That Save Lives. Especially Theirs*. Retrieved from Wired: https://www.wired.com/2016/06/people-want-self-driving-cars-save-lives-especially/

Wiseman, Y., & Grinberg, I. (2018). The Trolley Problem Version of Autonomous Vehicles. *The Open Transportation Journal*(12), pp. 105-113. doi:10.2174/18744478018120100105

World Bank. (2022). *Solid Waste Management*. Retrieved from World Bank: https://www.worldbank.org/en/topic/urbandevelopment/brief/solid-waste-management

Wu, W., Zhou, X., & Shen, B. (2021). Comprehensive evaluation of the intelligence levels for unmanned swarms based on the collective OODA loop and group extension cloud model. *Connection Science, 34*(1), 630-651. doi:https://doi.org/10.1080/09540091.2022.2026293

Wyatt, A., & Galliott, J. (2021). An Empirical Examination of the Impact of Cross-Cultural Perspectives on Value Sensitive Design for Autonomous Systems. *Information, 12*, 1-21. doi:10.3390/info12120527

Yadron, D., & Tynan, D. (2016). *Tesla driver dies in first fatal crash while using autopilot mode*. Retrieved from The Guardian: https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk

Yang, J., & Coughlin, J. F. (2014). In-vehicle technology for self-driving cars: Advantages and challenges for aging drivers. *International Journal Automotive Technology*, 333–340.

Yazdanpanah, V., Gerding, E. H., Stein, S., Dastani, M., Jonker, C. M., Norman, T. J., & Ramchurn , S. D. (2023). Reasoning about responsibility in autonomous systems: challenges and opportunities. *AI and Society*(38), 1453–1464. doi:https://doi.org/10.1007/s00146-022-01607-8

Yglesias, M. (2019). *The emerging 737 Max scandal, explained*. Retrieved from Vox: https://www.vox.com/business-and-finance/2019/3/29/18281270/737-max-faa-scandal-explained

Zewe, A. (2022). *On the road to cleaner, greener, and faster driving*. Retrieved from MIT News: https://news.mit.edu/2022/ai-autonomous-driving-idle-0517

Zuiderveen Borgesius, F. (2019). Discrimination, artificial intelligence, and algorithmic decision-making. 1-94. Retrieved from https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73