TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

THOMAS HEATH

AUTONOMOUS INDUSTRIAL MACHINES AND THE EFFECT OF AUTONOMY ON MACHINE SAFETY

Master of Science Thesis

Examiner: Professor Kalevi Huhtala
Examiner and topic approved by the Dean of the Faculty of Engineering Sciences on 28th March 2018

ABSTRACT

THOMAS HEATH: Autonomous Industrial Machines and the Effect of Autonomy on Machine Safety
Tampere University of Technology
Master of Science Thesis, 62 pages
March 2018
Master's Degree Programme in Automation Engineering
Major: Fluid Power
Examiner: Professor Kalevi Huhtala

Keywords: autonomy, industrial machines, road vehicles, machine safety, mining

Autonomous machines and vehicles are an increasing part of everyday life and industrial operations. These machines and vehicles have enjoyed rapid technological advancements in recent years, which has led to increasingly sophisticated functions and functionalities. The advancements in autonomous technologies have, however, given rise to questions and concerns relating to the safety of these machines and vehicles, and on how an adequate level of safety can be ensured when no dedicated operator or driver is present.

This thesis looks at the main areas that affect the overall safety of autonomous industrial machines and civilian road vehicles, and presents the most prominent challenges faced in ensuring the safety of autonomous applications. The goal of the thesis is to give the reader an overview of the safety-related aspects of autonomy and to show what has to be considered when ensuring an adequate level of safety for autonomous machines or vehicles. This is achieved by an extensive literature review on autonomous applications in both industrial and automotive fields, and on the safety-related aspects of autonomy. Additionally, mining is used in the thesis as an example of autonomous machines in practice and on the challenges autonomy can face in industrial operations.

Based on the research carried out, it can be said that the overall safety of machine autonomy is currently hindered by two main aspects: the lack of applicable standards, legislation and guidelines regarding the autonomy of machines and vehicles, and the paradox that arises from balancing the desired level of autonomy with the needed level of safety. This has led to a situation where, in theory, highly complex and sophisticated autonomous machines are possible from a technical standpoint, but they lack a common and thorough method for ensuring an adequate level of safety.

TIIVISTELMÄ

THOMAS HEATH: Autonomiset työkoneet ja autonomian vaikutus koneturvallisuuteen
Tampereen teknillinen yliopisto
Diplomityö, 62 sivua
Maaliskuu 2018
Automaatiotekniikan diplomi-insinöörin tutkinto-ohjelma
Pääaine: Fluid Power
Tarkastaja: Professori Kalevi Huhtala

Avainsanat: autonomia, työkoneet, ajoneuvot, koneturvallisuus, kaivosteollisuus

Autonomiset työkoneet ja ajoneuvot ovat kasvavissa määrin osana arkielämää ja teollisuutta. Lähivuosina nämä laitteet ovat hyötyneet soveltuvien teknologioiden nopeasta kehityksestä, mikä on johtanut hyvinkin kehittyneisiin toimintoihin ja toiminnallisuuksiin. Autonomisten teknologioiden kehitys on kuitenkin nostanut esille kysymyksiä ja huolia näiden laitteiden turvallisuuteen ja sen varmistamiseen liittyen etenkin tilanteissa, joissa laitteella ei ole käytännössä selkeää kuljettajaa.

Tässä diplomityössä tutkitaan tärkeimpiä osa-alueita, jotka vaikuttavat autonomisten työkoneiden ja ajoneuvojen turvallisuuteen, sekä esitellään suurimmat haasteet autonomisten laitteiden turvallisuuden varmistamisessa. Työn päämääränä on tarjota lukijalle kattava katsaus autonomian turvallisuuteen liittyvistä osa-alueista, sekä osoittaa mitä tulee huomioida, jotta voidaan saavuttaa tarvittava turvallisuuden taso autonomiselle laitteelle. Työn päämäärän perustana on kattava kirjallisuustutkimus autonomisiin työkoneisiin ja ajoneuvoihin, sekä näiden turvallisuuteen liittyviin osa-alueisiin liittyen. Lisäksi työssä käytetään kaivosteollisuutta autonomian käytännön esimerkkinä, jonka avulla esitellään suurimpia haasteita, joita autonomia voi kohdata käytännön ympäristöissä.

Tehdyn tutkimuksen perusteella voidaan todeta, että autonomisten laitteiden turvallisuutta jarruttaa tällä hetkellä kaksi pääasiallista seikkaa: sopivien lakien, standardien ja ohjeistuksien puute, sekä ristiriita, joka syntyy tasapainoilusta kehittyneen autonomian ja riittävän turvallisuuden tason välillä. Tämä on johtanut tilanteeseen, jossa teoriassa hyvinkin monimuotoiset ja kehittyneet autonomiset laiteominaisuudet ovat teknologian kannalta mahdollisia, mutta näiden toteuttamista varten ei ole olemassa yhtenäistä ja kattavaa menetelmää, jolla riittävä turvallisuuden taso voidaan varmistaa.

PREFACE

In late 2017, I was offered an interesting new career path at the company I work for. This was a great opportunity, but as I had not started work on my thesis yet, I was reluctant to accept. Luckily, my employer was generous enough to give me a few months off to finish my studies. This thesis is now the result of those dark winter months that were filled with long days and hard work.

I would like to thank Professor Kalevi Huhtala from the Department of Intelligent Hydraulics and Automation for arranging a very interesting subject for my thesis on very short notice. I also give my heartfelt thanks to my family for their support over the years and to my girlfriend Heini. Lastly, I would also like to give separate thanks to my father Peter for proofreading and checking the grammar of this thesis.

Tampere, 30.3.2018

Thomas Heath

CONTENTS

LIST OF TERMS AND ABBREVIATIONS

ASIL	Automotive Safety Integrity Level
AV	Autonomous Vehicle
AVC	Autonomous Vehicle Control
AVO	Autonomous Vehicle Operation
AVP	Autonomous Vehicle Protection
AutoMine	A mine automation system offered by Sandvik Mining & Rock Technology
CPU	Central Processing Unit
DDT	Dynamic Driving Task
ECU	Electronic Control Unit
GPS	Global Positioning System
GPU	Graphics Processing Unit
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
I/O	Input/Output
LHD	Load-Haul-Dump (Machine)
NHTSA	National Highway Traffic Safety Administration
ODD	Operational Design Domain
POSE	Position and Orientation
SAE	Society of Automotive Engineers
SFS	Finnish Standards Association
SIL	Safety Integrity Level

1. INTRODUCTION

The nature of industrial machines and the role of their operators is currently in a state of change. Traditionally, industrial machines have been human-operated machines that perform either manual or automatic functions while requiring almost constant control and monitoring of their actions. Therefore, the machines require the presence of an operator, which at times requires the operator to expose themselves to hazardous environments and other risks. A vision of a machine that can perform these actions autonomously, without the need for an operator, has been in the minds of researchers and manufacturers for the past several decades. Similarly, in automotive fields, the idea of a completely self-driving car has been a vision of the future for a number of years. Due to the advancements in technology in recent years, the idea of self-operating machines and self-driving vehicles is no longer a distant vision, but rather a possibility of the very near future.

The automation of machines and their features is, however, nothing new. Numerous different automatic functions and features have been available for machines and vehicles for years, which have been used to lessen the workload on operators and drivers and in some cases to minimise exposure to hazards they might be faced with. The impact of autonomy on machines, however, is far more complex. Autonomy offers a way for machines or vehicles to gather information on themselves and on their surroundings, and importantly, to use this information to make decisions and actions to fulfil a goal they have been set – without the need for intervention from the operator or driver, and thus eliminating exposure to risks and hazards completely.

Autonomous machines are therefore highly complex machines that are able to perform independent decision-making and to operate without the supervision of an operator. Ensuring the safety and safe operation of such a machine is therefore a challenge that has not been previously faced that requires new methods and new ways of thinking. The safe operation of manned machines has ultimately always been the responsibility of the operators themselves, who have had to control the machine and monitor their environment in a manner that ensures no harm or hazards result from the operation of the machine. In worst-case hazardous situations, the operator could always act as a safety net of sorts if needed, stopping the machine before any harm could occur. However, autonomous machines do not have this advantage, and thus their safety must be ensured by other methods. The importance of these methods cannot be overstated, as in autonomous applications a small error in operation can lead to great consequences, for example, if an autonomous road vehicle encounters a fault in a densely populated area.

Autonomy is a relatively new field of research, which is why most research on machine and vehicle autonomy has centred on proof-of-concepts and on how these systems could be designed and implemented in practice. The safety of such machines has, however, generated far less research, but some previous research is available. Similarly, very few standards and other legislation on autonomous machines or on their safety are available. This has led to a situation where complex autonomous machines and vehicles are possible from a technical standpoint, but manufacturers and developers lack a common, comprehensive and effective way to ensure the safety of the machines. By their very nature, autonomous machines and vehicles operate in varyingly differing areas and around varying types of other machines, vehicles and people, which means they are faced with an essentially infinite number of different operational situations. Without proven methods for ensuring safety, it is a considerable challenge to make sure the autonomous machine or vehicle can operate safely in every operational situation. As the situations are, in theory, infinite in number, Murphy's Law can be used to portray the scope of the problem: any error or fault in operation can lead to a safety incident given enough time, if no precautions are put in place.

As current technology allows for fairly complex and sophisticated autonomous machines, manufacturers are faced with a paradox of sorts. Especially due to common methods not being available, ensuring safety of autonomous machines becomes a balancing act between an adequate level of safety, the level of autonomy and the functionalities the machine can offer. For example, it is relatively effortless to ensure the safety of a fully autonomous machine, if the functionality of the machine is simple and minimalistic. Similarly, it is relatively straightforward to create a fully autonomous machine with complex features, if it does not need to adhere to any safety requirements.

When comparing autonomous civilian road vehicles and autonomous industrial machines, it is clear the former is the more researched and discussed field. This is largely because autonomous road vehicles attract far more interest, as they affect most of the general populous, rather than only a select field. Hence, there is more information available on autonomous road vehicles, such as standards, guidelines and ways of classifying levels of autonomy, than on the equivalent industrial machines. Therefore, many points made in this thesis are originally aimed solely for civilian autonomous vehicles (AV), but the knowledge gained from the research and development in this area will be a benefit for industrial fields, as the challenges and technical hurdles faced by both fields are very similar.

This thesis is based on a thorough literature review on autonomous industrial machines and road vehicles with an emphasis on their safety. The goal of this thesis is to present and discuss the main aspects of autonomous machine and system design that affect overall machine safety. Furthermore, the main challenges of ensuring safety that arise from the increase in autonomy in machines and vehicles will also be discussed. The point of this thesis is not to present a specific practical method of ensuring safety for autonomous machines, but rather to be an overview of autonomous safety and the challenges and hurdles that have to be overcome to create safe autonomous machines and vehicles.

The thesis begins with general information on autonomous machines and vehicles, such as a definition on what constitutes as an autonomous machine. Importantly, the distinction between autonomy and automation is presented because these terms are often used interchangeably even though they imply notably different functionalities in machines and vehicles. Additionally in this chapter, an overview of the current standards and legislation that apply to autonomous industrial machines and civilian road vehicles is presented. This information is then used to present different categorisation methods for machines based on their autonomy. As no official categorisation methods exist that can be directly applied to industrial machines, a closer look is taken at the equivalent categorisations for autonomous road vehicles. Additionally, previous research is used to present alternative methods for categorising the levels of autonomy in industrial machines.

In the next chapter, the main safety challenges that arise from the increase in autonomy are discussed. The first main challenge is to construct system architectures that are suitable for autonomous applications and also ensure effective performance and overall safety. The other main challenges include the position and movement planning characteristics of autonomous machines, which include such topics as localisation, motion planning and situational awareness. Next in the chapter, the risk analysis and verification challenges and methods of autonomous systems, which ensure the safe operation of machines in use, are presented. Lastly, the moral and ethical dilemmas of autonomy, which has been a widely debated topic in recent years as it is possible the actions of an autonomous machine or vehicle results in the death of a person, are discussed. This topic is presented from the viewpoint of road vehicles, as this has not been discussed in industrial applications.

In the final main chapter, the mining industry is used as a practical example of autonomous machines in operation. The chapter begins with an overview of mining and mining operations. It is also discussed how mining can benefit from the increase in autonomy, as mining work tasks are often hazardous and repetitive and are thus well suited for autonomy. Next, the main challenges that are faced in increasing autonomy in mining applications, which stem mainly from the operational environments of mining, are presented. After this, the current developments in autonomous mining are discussed, with an emphasis on load-haul-dump mining machines, the autonomy and automation of which have been researched for over three decades. The last main topic in the chapter is the main safety challenges in autonomous mining, which are mainly the challenges in overcoming the harsh and hazardous operating environments.

2. AUTONOMOUS MACHINES IN GENERAL

Autonomy has been a vision of the future for several decades. For instance, automotive manufacturers, such as General Motors, have shown interest in self-driving cars since the forties and fifties. In the last few decades, however, autonomy has evolved from a vision of the future to actual reality with offerings available in both industrial and civilian fields with varying degrees of autonomy. Completely self-driving cars and self-operated machines are still largely under research and development, but they too are not far in the future.

This chapter begins with the definition of an autonomous machine and what is considered autonomy and what is not is discussed. Next, a brief overview is given of the current state of legislation and standards that apply to autonomous industrial machines and autonomous civilian vehicles. Lastly, as autonomy can be implemented in varying degrees, standards and other sources are used to present different ways to classify the level of autonomy of machines and vehicles from both automotive and industrial viewpoints.

2.1 Definition of an autonomous machine

Autonomy is often defined in a broader sense as meaning: “the ability to self-manage, to act or to govern without being controlled by others” (Baudin et al. 2007, p.5). In a more practical sense, an autonomous machine or system is an entity that is able to gather information on its surroundings and use this information to make decisions and perform actions in order to fulfil an ultimate goal given to it by an outside source. This outside source is usually an operator in industrial applications or a driver in autonomous road vehicles. Such goals given to an autonomous machine can be for example: “travel to this location” or “perform task A when criteria X is met”.

The terms autonomy and automatic are often used interchangeably, as they are both similar in meaning and offer similar functions in machines. There is, however, a clear distinction between the two. An autonomous system has greater complexity and is capable of making decisions based on the information it has gathered, and then acts on those decisions. As the situations where autonomous machines make decisions vary, and no two situations are the same, there is no way to determine accurately how an autonomous system will act in a random and unknown situation in the future. Only broad assumptions can be made. On the contrary to autonomous systems, an automatic system’s behaviour can be determined beforehand, as it is always a predefined function or set of functions in regard to a specific input. (Baudin et al. 2007) For example, a simple cruise control feature could be classified as an automatic function: a set speed is given to the cruise control module by the driver, and the system adjusts the speed of the vehicle to suit this value. Adaptive cruise control, however, is an autonomous feature because the vehicle makes decisions on whether to accelerate or brake in regard to the distance of the vehicle in front. Autonomy is not, however, a binary classification. Machines may have varying degrees of autonomy ranging from full autonomy to mere autonomous features. These levels of autonomy will be discussed later in chapter 2.3.

Autonomous machines rarely operate in complete isolation, but rather operate around other machines and vehicles, both manned and unmanned, people and other dynamic objects. In the literature, these are often called agents. These are entities that act in the same area as the autonomous machine and have their own trajectories, goals and intentions that the autonomous machine must take into account. Another common term found in the literature is the state of an autonomous machine. Put simply, states are the sum of both internal and external variables of the autonomous machine in a specific situation, at a specific point in time. Thus, states range from normal safe operational states to states that can be abnormal and include some form of risk or hazard.

2.2 Standards and legislation

Autonomous machines and vehicles have enjoyed rapid technical advancements in recent years. This has led to numerous plausible applications where autonomy can be utilised. State regulatory establishments and standardising organisations have not, however, been able to keep up with these advancements in technology, which has led to a situation where numerous autonomous functions and features are technically plausible, but they lack a common method for development, verification and for ensuring safety, because of the lack of appropriate standards and legislation.

Some previous standards are available that can be, at least in part, applied to autonomous industrial machines. These include standards relating to the safety integrity of machine control systems, such as IEC 61508: *Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems* and ISO 13849: *Safety of machinery -- Safety-related parts of control systems*. For civilian vehicles, there exists a similar standard - standard ISO 26262: *Road vehicles – Functional safety*. Some more specific and definitive standards for autonomous industrial machines are in the development phase and some, such as ISO 17757: *Earth moving machinery and mining - autonomous and semi-autonomous system safety*, have very recently been released.

In the automotive field, the state of autonomous road vehicle legislation and regulation in general is still a work in progress. Some countries and states are in the stages of preparing and passing legislation on autonomous vehicles, but the work is still very much ongoing.

2.2.1 Standards on safety integrity levels

Several standards are available for ensuring the safety of electrical control systems, and these standards can also be applied to autonomous machines to some degree. The two most prominent standards in this area are IEC 61508: *Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related System* and ISO 13849: *Safety of machinery -- Safety-related parts of control systems*. Furthermore, function-specific standards have been developed based on the aforementioned standards, such as ISO 26262: *Road vehicles – Functional safety*, which is specifically intended for road vehicles.

A system is categorised as safety-related when it performs functions that keep safety-related risks at a tolerable level. Therefore, if these functions do not operate correctly and this corresponds to increased safety-related risks, the system is labelled as safety-related. (SFS IEC/TR 61508-0 2012) As such, autonomous machines can be categorised as safety-related as a whole because they have numerous systems that ensure the safety and correct operation of the machine. If the machine does not operate as intended, a definite safety risk is present. Thus, standards on safety integrity levels (SIL) can be applied to autonomous machines, at least in part.

Functional safety is described in the standards as the correct operation of the safety-related functions or parts of a system. In other words, if a safety-related control system performs functions that effectively negate the risks posed by the operation of the system, it is called functional safety. An example of this is an electric motor with a temperature sensor that monitors the temperature of the motor. If the sensor senses the motor is about to overheat, it will shut the motor off, thus reducing risk. Here, the system performs actions that correctly minimise safety-related risks, thus performing functional safety. The probability of functional safety, i.e., the probability of safety functions operating as they are intended to operate, is called safety integrity. In standards such as IEC 61508, safety integrities are separated into levels, with each level having its own maximum and minimum limits for the probabilities of failure of the safety-related function. (SFS IEC/TR 61508-0 2012)

Standard IEC 61508 separates safety integrity levels of electrical, electronic, and programmable electronic safety-related systems into four levels ranging from SIL1 to SIL4, with SIL4 offering the highest level of safety integrity (SFS IEC/TR 61508-0 2012). The implementation of the standard has three main goals: to determine the needed safety integrity level of the system, to guide the development process of the system and to verify that an adequate level of safety has been reached. In figure 1, it is demonstrated how this is incorporated into the development phase of a system. (Redmill 2000)

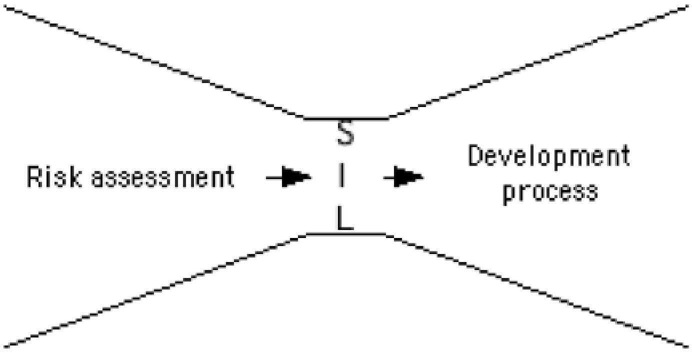

Figure 1. *The role of SIL's in the development process (Redmill 2000)*

The implementation of IEC 61508 and safety integrity levels can be seen as a funnel. The process begins with the risk assessments of the system, with the goal of determining the current risk level posed by the system. If these risks are deemed too great, a suitable risk reduction method is implemented, such as a safety function. The failure probability of the whole system, including the safety function, is then calculated, which corresponds to a specific SIL. If this SIL is too low, the safety function or the rest of the system can be altered to reach the desired SIL. After choosing a SIL, standard IEC 61508 supports and controls the development process of the system by offering guidelines and instructions on how a specific SIL can be achieved. Standard IEC 61508 is based on a life-cycle approach, which ensures the verification of the overall system safety and takes into account the whole life-cycle of the system. (Redmill 1998)

The standard ISO 13849: *Safety of machinery -- Safety-related Parts of Control Systems*, is similar to IEC 61508, but it is a simplified version that is only applicable to machinery control systems. Instead of categorising probabilities into safety integrity levels, the standard uses Performance Levels. Additionally, the equivalent standard for automotive applications is ISO 26262: *Road vehicles – Functional safety*. The standard is a simplified version of IEC 61508 that takes into account aspects important to the automotive field. The standard also uses automotive safety integrity levels (ASIL) instead of traditional safety integrity layers.

The main issue with applying the current standards on safety integrity levels to autonomous machines is that the standards often rely on human intervention in their hazard and risk analyses, which may not be possible in autonomous machines. As such, none of the aforementioned standards can be utilised fully in their current state. Therefore, new standards are needed, or the current standards must be updated for autonomous applications. (Behere et al. 2016, Kaznov et al. 2017)

2.2.2 Other standards for industrial autonomy

To date, only a few standards for autonomous industrial machines have been released by the major standardising organisations. The formation of new committees and work on new standards are, however, currently ongoing.

In mining, standard ISO 17757: *Earth moving machinery and mining - autonomous and semi-autonomous system safety* was released in late 2017 by the technical committee TC127, which is the committee in charge of earth moving machinery. The standard was a joint effort between TC127 and the committee on mining TC82. The standard outlines the safety requirements for autonomous and semi-autonomous machines used for earth moving in mining, such as load-haul-dump machines.

The technical committee for mining, TC82, has not yet itself released any standards regarding machine autonomy. Negotiations are, however, ongoing to form a new subcommittee, SC8, for autonomous mining. This committee will, once formed, prepare new standards for autonomous mining applications. The current problem with forming the subcommittee is the scope and overlap with the previously mentioned standard ISO 17757. (Kempson et al. 2017)

Other developments in autonomous industrial machine standards include ISO 18497: *Agricultural machinery and tractors -- Safety of highly automated agricultural machines,* which is still under development (International Organization for Standardization 2018). No other information is available on this standard as of yet.

2.2.3 The current state of autonomous road vehicle legislation

Currently, the amount of state legislation in effect for autonomous vehicles is minimal both in Europe and in the US. The reason behind this is that legislation has not been able to keep up with the rapid advancements in autonomous technologies. Steps have been recently made, however, to pass legislation and standards for autonomous vehicles in the automotive field.

In regard to autonomous road vehicles, the US has been the forerunner in passing legislation, as several US states have been implementing AV laws since 2011. US states could even be said to be in competition with each other in trying to be the leading state in the implementation of autonomous vehicles and laws, and thus being the forerunner in technological advancement. This is in part due to the push from companies such as Google who will benefit from being able to use autonomous vehicles on the roads as quickly as possible. Most of the legislation passed thus far has allowed the testing of autonomous vehicles on public roads, but few have allowed the actual civilian usage of AV's. (Schreurs & Steuwer 2016) Continuing the trend set by US states, in the latter half of 2017 the US House of Representatives passed a bill entitled the "SELF DRIVE Act"

(2017). The aim of the bill is to create a nationwide framework for the regulation of AV's. The bill, if passed into full legislation, would be the first federal legislation regarding AV's in the US, and thus be a large step for AV legislation in the country.

In Europe, the state of autonomous road vehicle legislation is not as advanced as in the US. On an EU level, autonomous vehicle legislation is almost non-existent, as of 2015. There is also little mention of autonomy in the "EU 2020" strategy – the EU agenda for growth in the coming decade. On a country level, the situation is similar, albeit for a few exceptions. Especially Sweden and Germany have passed legislations for AV's, where, for example, Sweden has allowed the civilian testing of AV's. (Schreurs & Steuwer 2016) The EU has, however, funded a vast number of research programs on autonomous technologies ranging from driver assistance systems to fully autonomous transport systems, thus showing a great interest in autonomy. These projects include the Eureka PROMETHEUS project (Programme for a European Traffic of Highest Efficiency and Unprecedented Safety), which ran from 1987 to 1995, and the ongoing SARTRE project (Safe Road Trains for the Environment), which aims to research vehicle platooning. (European Road Transport Research Advisory Council 2015)

2.3 Classifications for autonomous machines

Due to the varying degrees of autonomous functions and features in autonomous machines and vehicles, different classifications have been conceived to help with, for example, the applicability of standards and other legislation. In this chapter, some of these classifications for both road vehicles and industrial machines are discussed. No common method for classifying autonomous industrial machines, however, currently exists, which is why a close look is taken at the equivalent road vehicle classifications, as these can be used as a guide or starting point for classifications for industrial machines.

2.3.1 Road vehicles

The two most notable classification methods for autonomous road vehicles are the SAE International standard SAE J3016: *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles* (SAE International 2016), originally released in 2014, and the guideline *Preliminary Statement of Policy Concerning Automated Vehicles* issued by the US National Highway Traffic Safety Administration (NHTSA) (2013). The former separates autonomy into six levels and the latter into five.

The SAE J3016 classification is a widely used categorisation method for autonomous road vehicles, and it has been taken advantage of in legislation, for example in the United States (The United States House of Representatives 2017). The classification separates AV's into six different levels ranging from 0 (no autonomy) to 5 (full autonomy). These levels are presented in table 1 with a brief description of each level.

Table 1. *SAE J3016 classifications for AV's (adapted from SAE International 2016)*

SAE Level	Name	Description
0	No Automation	No autonomous features
1	Driver Assistance	Longitudinal or lateral motion autonomy
2	Partial Automation	Longitudinal and lateral motion autonomy
3	Conditional Automation	Full autonomy in certain situations, driver as a fallback
4	High Automation	Full autonomy in certain situations, system as a fallback
5	Full Automation	Full autonomy in all situations, system as a fallback

Currently, vehicles with autonomous functions up to level 2, such as Tesla's Auto Pilot, are commercially available. Level 3 autonomy is predicted to be available in early 2020, while levels 4 and 5 are estimated to be available in late 2020 (European Road Transport Research Advisory Council 2015).

In the standard, a clear distinction between the different levels of autonomy is made. The base level, level 0, is a vehicle without any autonomous features, such as a vehicle manufactured in the previous decade. This level also includes modern vehicles with warning systems, such as lane departure warning systems, that do not affect control of the vehicle. (SAE International 2016)

Next, The Driver Assistance and Partial Automation levels are the first two levels with actual autonomous features. The distinction between the two is that in Driver Assistance the autonomous system controls either the longitudinal or the lateral movement of the vehicle, but not both. In Partial Automation, on the other hand, the autonomous system controls both. In practice, longitudinal autonomy is often adaptive cruise control, where the system maintains a fixed distance to the vehicle in front. Lateral autonomy is lane-keeping assist, where the system keeps the vehicle between lane markers. These autonomous functions are available only in certain situations, generally only when the system or driver deems them fit. The driver on these levels is in charge of monitoring the surroundings of the vehicle and acts as a fallback if needed, i.e., the driver takes back control if the autonomous system encounters an error, fault or a situation where it can no longer operate autonomously. (SAE International 2016)

SAE J3016 makes a clear distinction between the previous levels and levels 3 to 5, which is signified by the thick line in the above table. While on levels 0 to 2, the driver performs most, or all, of the driving functions, described as dynamic driving tasks (DDT) in the standard. However, on levels 3 to 5, the autonomous system performs all of the DDT's and monitors the surroundings of the vehicle, when the system is active. Thus, when the autonomous system is active, the driver releases all control to the autonomous system.

Therefore, the driver can even be removed completely, as on level 5. (SAE International 2016)

On level 3, the vehicle is able to perform fully autonomous behaviour in certain situations. These certain situations are described in the standard as Operational Design Domains (ODD), which are specific situations where the autonomous features are designed to function. Level 3 ODD's and autonomous features could, for example, be self-parking in a parking lot or autopilot on a motorway. When the autonomous system is active, it has complete control of the vehicle, but the driver is still used as a fallback in case of faults or other problems the autonomous system may face, similarly to level 2. (SAE International 2016)

The next level, High Automation, increases the role of the autonomous system. The functionality of the level is the same as level 3, but with the distinction that the driver does not need to be a fallback if the system faces problems. The fallback functionality is performed by the system itself. In such a scenario, the goal of the autonomous system is to achieve a minimal risk condition and keep the system in a safe state. As such, level 4 allows for full autonomy in the scope of an ODD, where the driver can be completely passive and even sleep. (SAE International 2016)

The last level, Full Automation, offers full autonomy of the vehicle in all situations, i.e., the ODD can be said to be infinite. Vehicles of this level perform all DDT's and do not need the input of a driver and, as such, the driver does not need to be in the vehicle. (SAE International 2016)

The other major categorisation method for AV's is the guideline issued by the NHTSA. The categories are similar to the ones in standard SAE J3016, but in the NHTSA classification there are only five levels, from 0 (no autonomy or automation) to 5 (full autonomy), as opposed to six. These levels are presented in table 2 with a brief description of each.

Of note is that the NHTSA guideline does not use the word "autonomous" in its categorisations. The term is only used once in the guideline to describe self-driving cars as autonomous. All other levels of autonomy are described as levels of automation. Thus, the categorisations may be misleading as there is no distinction where the threshold between automation and autonomy lies. While the NHTSA categorisation is discussed in this text, the terms automatic and autonomous will be used according to the definition in chapter 2.1.

Table 2. *NHTSA classifications for AV's*

NHTSA Level	Name	Description
0	No Automation	No autonomous or automatic features
1	Function-specific Automation	One or more autonomous or automatic functions, overall control with driver
2	Combined Function Automation	Autonomy of at least two primary control functions in certain situations, driver to take control on short notice if needed
3	Limited Self-Driving Automation	Full autonomy in certain situations, driver needed to occasionally take control
4	Full Self-Driving Automation	Full autonomy in all situations

The base level, level 0, is similar to the equivalent SAE J3016 level. A vehicle of this level does not have any autonomous or automatic features. Additionally, if the vehicle has warning systems, such as forward collision warning or lane departure warning that do not offer additional control functions, the vehicle is also categorised as level 0. (National Highway Safety Administration 2013)

The next level, Function-Specific Automation, offers one or more autonomous or automatic functions. These functions operate independently from each other and overall control of the vehicle remains with the driver. The driver is thus responsible for the overall operation of the vehicle and must perform all monitoring of the environment. Functions of level 1 are, for example, cruise control and automatic braking. (National Highway Safety Administration 2013) The SAE J3016 counterpart of this level would be level 1, Driver Assistance, but the two have clear differences. The NHTSA classification classifies vehicles with automatic functions, such as cruise control, as level 1, but according to SAE J3016, these would not count as autonomous and the vehicle would thus be level 0. However, if a vehicle has autonomy of one control function, the vehicle would be categorised as level 1 by both SAE J3016 and the NTHSA classification.

Combined Function Automation is the third level in the NTHSA classification. On this level, the vehicle is equipped with autonomy of at least two primary control functions in certain situations. When in such a situation, active control of these functions is given to the autonomous system, but the driver is still tasked with monitoring the environment. The driver must also be available and ready to take control of the vehicle within short notice, if needed. Examples of such autonomous functionalities are adaptive cruise control and lane-keep assist. (National Highway Safety Administration 2013) Level 2 is similar to the SAE J3016 level 2, Partial Autonomy, where instead of two or more autonomous control functions, the vehicle has autonomous control of both longitudinal and lateral movement in certain situations. In both, however, the driver is in charge of monitoring the environment and must be ready to take control if needed.

Limited Self-Driving Automaton is the second to last level of autonomy in the NHTSA classification. Vehicles of this level are able to function autonomously in certain situations. In these situations, the autonomous system takes full control of the vehicle and monitors its surroundings. The driver is not needed for active control but must be able to take control if needed after a transition time. Such a need may arise, for example, if the AV enters a location where autonomous driving is no longer possible. (National Highway Safety Administration 2013) Limited Self-Driving Automation resembles SAE J3016 level 3 Conditional Automation. In both, the vehicle operates autonomously in certain situations, or ODD's. When the ODD is about to end, the driver is prompted to take control. The NHTSA guideline does not, however, state how the system should react if the driver does not act on this prompt. If the system is supposed to reach a safe state in this situation, level 3 of the NHTSA guideline is more in line with SAE J3016 level 4. If not, level 3 is more similar.

The last level is titled Full Self-Driving Automation, which is the highest form of autonomy according to the guideline. In this level, the vehicle is able to operate completely autonomously, with the driver only needed to enter the destination location. (National Highway Safety Administration 2013) This level is thus similar to the SAE J3016 level 5 Full Automation.

2.3.2 Industrial perspective

As discussed in chapter 2.2., only a few standards on autonomous industrial machines have been released thus far. As such, none of the major standardising organisations offer a method to categorise industrial autonomous machines based on their levels of autonomy. This is, however, also likely due to the vast number of different applications for autonomy in industrial fields, whereas in the automotive domain these applications are quite similar. Because of the lack of a standardised way to categorise industrial autonomous machines, more pragmatic approaches are often used to categorise machines, for example, in mining applications.

In mining, a pragmatic approach to categorising autonomous industrial machines is to categorise them by their control method. This categorisation includes both non-autonomous and autonomous machines, as only the most sophisticated level of control is considered true autonomy. Machines are often categorised into six levels: manual operation, remote control, teleportation, blind autonomy, semi-autonomy and full autonomy (Brown 2012, Gustafson 2011).

The base level, manual operation, is a traditional industrial machine that is controlled by an operator from inside or on top of the machine. An example of such a machine is a traditional mining haulage truck, which is controlled by an operator inside the cabin. The first step towards autonomy of such a machine, and thus the second level of categorisation, is remote control of the vehicle. With such a machine, the operator is removed from the machine and the machine is controlled with a remote controller. Importantly, the operator still has a line of sight to the machine at all times and must therefore be situated close by. (Brown 2012, Gustafson 2011)

The next logical step towards autonomy is teleoperation. The clear distinction to the previous level is that the operator no longer has to have a clear line of sight of the machine, but rather operates the vehicle remotely, traditionally via a video feed. (Brown 2012, Gustafson 2011)

The next level, blind autonomy, offers the lowest form of autonomy. Machines of this level can navigate on fixed paths without an operator, but they are “blind”, i.e., they do not have any kind of situational awareness and cannot sense obstacles. (Brown 2012) For example, many mining haul machines used underground are considered blind.

When a machine can operate fully autonomously in only some specific situations, or when it cannot carry out all of the stages of its work cycle independently without an operator, it is considered semi-autonomous. While operating in autonomous mode, these machines gather information on their surroundings and act on this information, similarly to fully autonomous machines. A human operator is, however, needed to ensure safe and correct operation, and to take control when needed. This is traditionally performed via teleoperation. (Gustafson 2011)

Lastly, the final level is full autonomy, where the machine can operate autonomously at all times. The machine has a set goal it has to achieve; it then gathers information on its surroundings and makes decisions using this information to achieve the set goal. An operator is not needed for operation, but traditionally one is required to monitor the machine. (Brown 2012, Gustafson 2011)

This is a rough categorisation, which does not include all aspects of autonomy, such as operator assisting systems, and it can be argued that a fixed path travelling blind machine does not count as autonomy at all. The categorisation is nonetheless a good indication of the steps taken from no autonomy to full autonomy in machines, such as mining haulage trucks or other vehicles, where the main function is not to transport people. More theoretical and general approaches for categorising autonomous machines are also available, as pragmatic approaches are usually specific for only certain applications.

Behere and Liljeqvist argue in the article: *Towards Autonomous Architectures: An Automotive Perspective* (2012) that all autonomous systems can be separated into a 3+1 pattern, which includes all aspects needed for autonomy. They also argue that the pattern can be used to categorise levels of autonomy. The pattern is presented in figure 2 and it includes four portions: User, Environment, Control and Self.

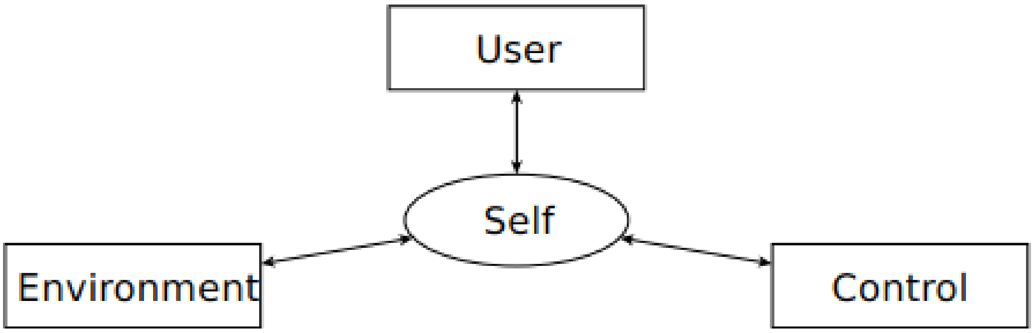

Figure 2. *The 3+1 pattern (Behere & Liljeqvist 2012)*

At the centre of the pattern is the portion Self, which represents the internal decision-making capabilities of the system, that is constantly in interaction with the other portions of the pattern. The Environment portion of the pattern is the situational awareness and world model building functions of the system that build a picture of where the machine is located and what is around it. The User portion contains the interactions with the user of the machine, which can be continuous, or a set of goals given to the machine. Lastly, the Control portion is in charge of controlling the actual machine. (Behere & Liljeqvist 2012)

The 3+1 pattern can be used to categorise the level of autonomy of systems by analysing the complexity of each part of the pattern. In a highly intelligent autonomous machine, all parts of the pattern are present and are highly complex. For example, an autonomous road vehicle utilises all parts of the pattern: Environment is used for localisation, situational awareness and motion planning, while User and Control are used to store the desired destination and to control the vehicle to reach this destination, respectively. Less complex autonomous systems would thus have less complex portions of the pattern. Moreover, if the functionalities that are represented by the portions of the pattern are missing completely, the system is not considered autonomous, but rather automatic. For example, a traditional cruise control system of a road vehicle does not have an Environment portion, as the system does not monitor the operational environment in any way. Therefore, a cruise control system cannot be regarded as autonomy based on the 3+1 pattern, which is also the same conclusion based on the definition in chapter 2.1. (Behere & Liljeqvist 2012)

A standardised method for categorising industrial autonomous machines, similar to the NHTSA guideline or standard SAE J3016, would be greatly beneficial for the development of further autonomous machine standards. Moreover, with a common methodology of categorising autonomous machines, adequate levels of safety would be relatively simple to verify because each level could have specific safety requirements. Lastly, as there is no common way to distinguish between the levels of autonomy in industrial machines, the autonomy and automation of a machine are often used as interchangeable terms, and hence there is a lack of clarity on what the machine is actually capable of.

3. AUTONOMY SAFETY CHALLENGES

Autonomous machines and vehicles are vastly complex and intelligent entities that often operate in highly unstructured environments that include a number of other agents, such as other autonomous machines, manned vehicles and people. This introduces a great number of new safety challenges that have not been an issue in the past, which autonomous machines must overcome. An autonomous machine must operate in these environments effectively and safely, without making errors in operation, that could lead to safety hazards. Errors that an autonomous machine could make include erroneous movement or actions, errors in decision making or systematic errors embedded in the system architecture of the machine itself. The basis of a safe autonomous machine is the definition of safety given in standard IEC 61508, which states safety is: “the freedom from unacceptable risk of physical injury or of damage to the health of people, either directly, or indirectly as a result of damage to property or to the environment." (SFS IEC/TR 61508-0 2012). Autonomous machines include a vast number of safety functions and features, which are tasked with keeping the machine in a safe state. This is an important aspect of safety, but as discussed in chapter 2, an autonomous machine can be labelled as a safety-related system as a whole. because the correct operation of all of the machine’s subsystems is needed to ensure safety and not only the direct safety functions.

In the following chapters, the different aspects of safe operation for industrial autonomous machines are discussed. Topics on the safety of autonomous civilian vehicles are also included, as these issues are more researched, and the challenges faced are often similar to autonomous industrial machines. A study on civilian vehicle autonomy is therefore beneficial because the advancements and findings in autonomous road vehicle technologies can be applied to industrial machines and are indicative of the future developments needed for industrial applications.

The chapter begins with an overview on the nature of the hazards that autonomous machines face in operation. Then, the differences between the challenges faced by industrial and civilian machines and vehicles are discussed. The next part of the chapter deals with system architectures and how they affect overall machine safety, and what challenges are faced in designing architectures for autonomous machines. After this, the main areas that affect the safe operation of an autonomous machine, such as localisation, motion planning, situational awareness and risk analysis, are discussed. Additionally, the moral and ethical dilemmas that arise from autonomy from a road vehicle viewpoint are presented in depth. Lastly, the paradox that arises from ensuring the safety of an autonomous machine while also ensuring effective autonomy is discussed.

3.1 The nature of autonomous safety hazards

Most hazards and risks related to autonomous machines arise from the complex nature of the machines and the varying operational environments where they are used. Most operational state combinations cannot therefore be known beforehand, which may lead to safety issues. The main safety risks posed by autonomous machines are due to both hazardous operation and faults that occurr in the decisional mechanisms of the machine. (Baudin et al. 2007)

Hazards posed by the operation of an autonomous machine can be separated into endogenous and exogenous hazards. Endogenous hazards are caused by faults introduced in the machine itself, such as faults introduced in development or faults due to component failures. (Baudin et al. 2007) These faults may lead to incorrect operation of the machine, and may thus pose a safety risk. In the standard IEC 61508, which is discussed in chapter 2.2.1, these types of faults are labelled as systematic and random faults and the standard outlines how these affect the functional safety of the system. Exogenous hazards, on the other hand, are caused by the operational environment of the machine, rather than by the machine itself. These hazards include faults due to outside interference and unforeseen events due to the environment. Exogenous hazards may also arise from the uncertainty of the environment due to missing environmental information. This may occur, for example, because of unsuitable sensors. (Baudin et al. 2007)

Faults in the decisional mechanisms that autonomous machines may face are separated into internal faults and interface faults, both of which may pose safety risks. Internal faults of the decision making of the machine include situations where the machine makes decisions with incomplete information, resulting in erroneous operation. Internal faults may also arise if the machine is faced with having to make a decision in a situation that was not foreseen by the designer of the machine, and thus the machine cannot act in this situation correctly because it is unsuitable for this situation. Interface faults that decision-making may face are faults due to errors in communication. These include ontological mismatches where one term has different meanings in different parts of the system, leading to errors. Interface faults also occur when human operators interpret information incorrectly, leading to undesired behaviour of the machine. (Baudin et al. 2007)

Additionally, errors faced by an autonomous machine can also be separated into omission errors and commission errors, both of which may result from the faults described previously. Omission errors occur when the autonomous machine does not perform an expected function and the system must then perform a recovery action to keep the machine in a safe state. Commission errors are the opposite and occur when the machine performs an action or chain of actions that were not desired or were otherwise forbidden. Both scenarios may lead to safety hazards. (Baudin et al. 2007)

3.2 Civilian and industrial differences

The safety challenges of autonomous industrial machines are similar to civilian road vehicles. Both types of machine may have to operate in complex environments with several interactions with other vehicles, autonomous and non-autonomous, as well as people. Both types of machine must also do this efficiently and, above all, safely. The challenges machines and vehicles face have, however, some notable differences.

The number of civilian vehicles on the road, the frequency of their usage and the vast distances travelled create a far greater safety challenge than for the equivalent industrial machine which are far fewer in number. As autonomous machines can be considered safety-related systems, as discussed in chapter 2.2.1, the fault tolerance of a civilian AV must be considerably higher because the sum of operational hours is considerable. This leads to a need for a high safety integrity level, which may not be needed for the equivalent industrial autonomous machine, as the number of these machines in use is smaller.

Interactions between other vehicles and people is far more common with civilian AV's than industrial autonomous machines due to the sheer number of vehicles and people in civilian areas. Industrial applications are, on the other hand, far more secluded with less traffic, which lessens the challenge in ensuring safety.

Industrial autonomous machines face their own set of problems that mainly stem from their operational environment. Areas where industrial machines operate are usually harsh with extreme temperatures, large amounts of dust and other disturbances, which affect the reliability of sensors and interfere with the correct operation of the autonomous machine. Areas where industrial machines operate are also often temporary and constantly evolving, which means pre-made maps that could be utilised in navigation, as with autonomous civilian vehicles, are not available. Industrial machines are also much larger than civilian AV's, which increases the risk they pose. (Nebot 2007)

3.3 System architectures

The increase in machine autonomy has brought with it numerous new functionalities to existing machines. This has led to the need to evolve existing system architectures to accommodate these new features, which has, however, introduced numerous challenges, namely in constructing system architectures that are effective and safe. The addition of autonomy to a system architecture cannot be thought of as only a new feature, but rather a from-the-ground-up-approach is needed for safe and effective autonomous system architectures (Kaznov et al. 2017). Architectures that operate correctly are needed for autonomous applications because if an architecture does not allow for the correct operation of an autonomous machine, it may lead to safety hazards due to the nature of autonomous machines and their operational environment. There are, however, no guidelines or standards on designing a system architecture for autonomous machines, so the challenges must be solved by the system designers alone (Kaznov et al. 2017). The architecture challenges are not only technical, but also include the development process and certification phases of system design (Behere et al. 2016).

In the past, the most common type of system architectures used in vehicles and machines were federated architectures, where system parts are separated into self-contained electronic control units (ECU) connected to each other via a communication bus. Each unit has its own function that is controlled by the unit itself. (Behere et al. 2016, Kaznov et al. 2017) An example of a federated system from an aviation application is presented in figure 3 below. Here the system architecture is separated into three parts with their own central processing units (CPU), connected via a communication bus, with one part controlling sensors, the second effectors and the third the interactions with the user.

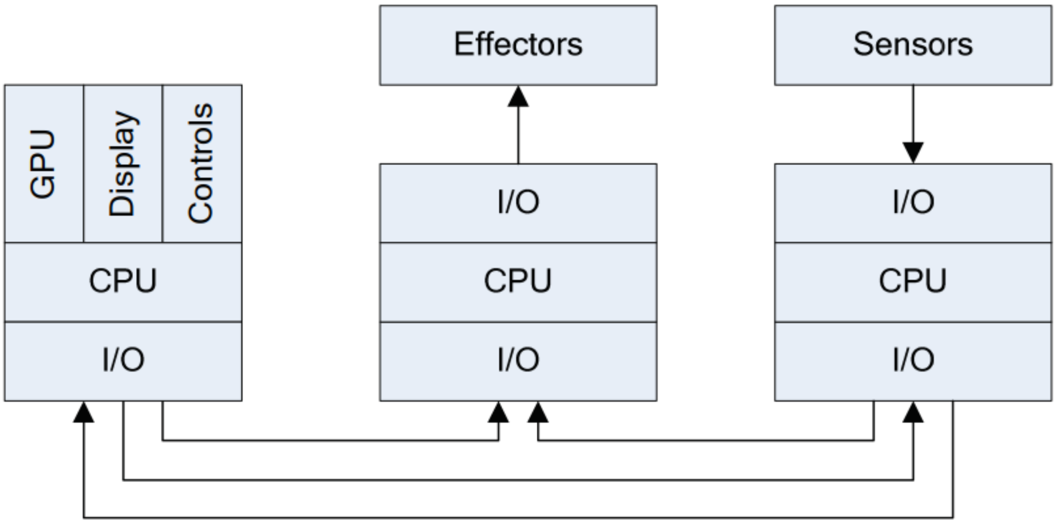

Figure 3. *An example of a federated system architecture (Watkins & Walter 2007)*

Federated architectures are easily expandable and verified due to their modular characteristics. However, as they are expanded, they begin to suffer from high complexity, resource consumption and cost. (Behere et al. 2016, Kaznov et al. 2017)

The limitations of federated architectures has led to the adoption of integrated architectures in both the autonomous industrial and automotive fields. Integrated system architectures differ from federated architectures in that one ECU may control several different functions, or one function may be controlled by several ECU’s. (Behere et al. 2016, Kaznov et al. 2017) An example of an integrated architecture is presented in figure 4, again from an aviation application. In the example architecture, the system is controlled by a single CPU that controls the three functions that were also included in the architecture in figure 3.

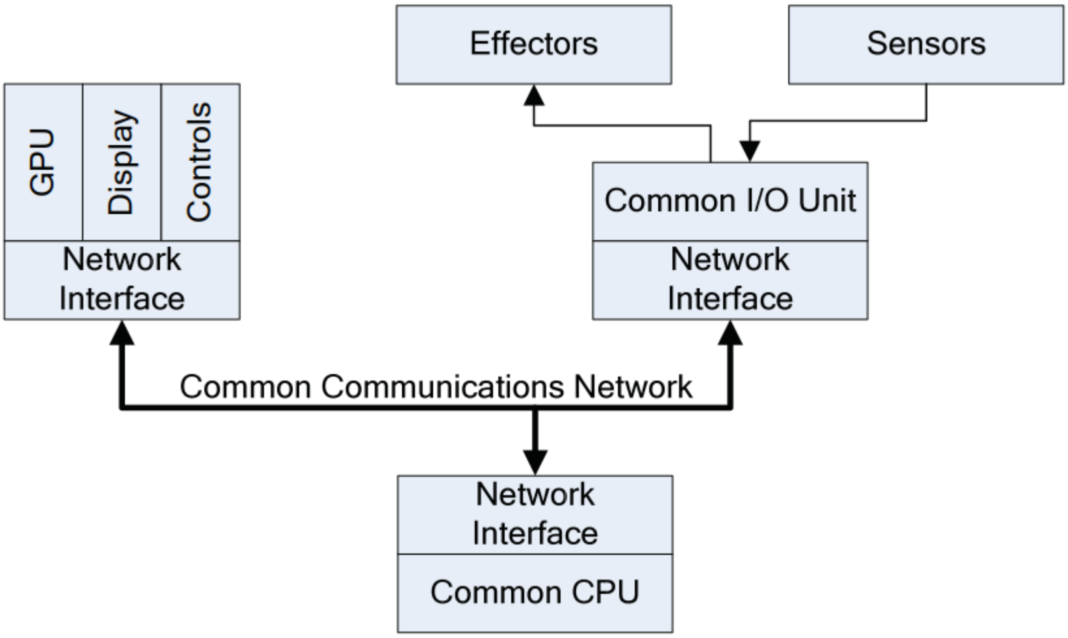

Figure 4. *An example of an integrated system architecture (Watkins & Walter 2007)*

Integrated architectures offer greater functionality, require less space for components and reduce cost. However, because the systems are no longer a group of self-contained functions, integrated architectures are considerably harder to verify and test to assure system behaviour in all scenarios. This leads to the need for new methods for the design and verification of integrated system architectures. (Behere et al. 2016, Kaznov et al. 2017)

Due to the complex nature of autonomous systems, federated architectures are not well suited for autonomous applications. An autonomous system requires considerable communication and functioning between parts of the system, which is why integrated architecture are a better option.

3.3.1 Problem areas

The incorporation of autonomy in integrated system architectures leads to four distinct problem areas in system design. These aspects also have a direct effect on overall machine safety because they affect the operation of the machine. (Behere et al. 2016)

The first major challenge is the implementation and the usage of the world model in the system. The world model is a central part of any autonomous machine because it is in charge of the upkeep and distribution of what the autonomous machine believes is around it and where the machine believes it is located in regard to the world. World model information is needed by several of the autonomous machines subsystems and this leads to the problem of how this information should be gathered, stored and distributed on an architecture level. Traditionally, world data is gathered with sensors, such as radar, laser, machine vision and the global positioning system (GPS), and this information is stored somewhere in the system architecture. The problem is that different subsystems may need this information in varying degrees and formats. Some may need a partial world model at specific moments in time, whereas some may need a complete model at all times. Some may require historical data on location or some may need more accurate data than other subsystems. The question is should all of the varying degrees of information be stored in a central complete world model, or should each subsystem gather and store the more specific data they require and share this data with other subsystems. The former could lead to size issues and questions on which subsystems are allowed to access and write which parts of the world information. The latter, on the other hand, may create needless complexity and have an effect on system efficiency. The challenge is to design the system in such a way that each subsystem receives the information it needs, in the desired format, without affecting the operation of the other subsystems. (Behere et al. 2016)

The second main problem is human interaction. By design, an autonomous system must take some control away from the user as otherwise the purpose of autonomy would be defeated. The autonomous system should operate transparently, relaying all needed information to the user. There is however, no clear distinction on what this transparency should be in practice because no guidelines are available that indicate what information should be given to the user in autonomous operation. Furthermore, it is still a matter of debate what role autonomy should be given in machines in general and what functions should be left to the user. The two main opposite opinions are that autonomy should be left to functions that are not suitable for human operation, and the other, that autonomy should coexist with the user as an equal in control. The differing amounts of information given to the user and the differing degrees in autonomy may lead to situations where similar autonomous systems operate slightly differently to each other. This raises safety concerns when human users are involved. When a user switches from one similar machine to the next, undesired behaviour may occur due to the slight differences in how human interaction is designed in the autonomous system architecture, and in how the machine is intended to be used. (Behere et al. 2016)

Autonomy unavoidably leads to more complex system architectures because it requires considerably more communication between subsystems than in traditional machines. This leads to a situation where the system must simultaneously act as a larger shared system and as isolated subsystems, which all may have different goals. Ultimately, the increase in complexity leads to increased difficulty in the testing, verification and validation of the system in the design phase. This may ultimately also lead to feature interaction, which is a situation where operation of one subsystem affects or counters the operation of another. This can lead to unanticipated behaviour of the system, affecting overall safety. An example of this type of behaviour could be a situation where two self-cancelling operations are performed at the same time, such as acceleration and braking. To eliminate this problem, the possible and probable feature interactions should be eliminated from the system architecture in the design phase and by algorithms while in use. All possible combinations cannot, however, be known beforehand, and thus the autonomous system must have a means to solve these situations independently. (Behere et al. 2016)

The fourth main problem autonomous system architectures face is the effect of autonomy on the systems extra-functional properties, such as redundancy, predictability and above all, safety. For example, most safety critical systems thus far have been designed in such a way that the last-resort failsafe has been for the user to take action by activating an emergency stop. With autonomous machines, this is no longer an option because there may not be a user to take control or the user may face the interaction problems mentioned previously. This means the robustness of the safety-related system must be increased, which is often done by adding redundancy to safety critical sensors, actuators and other components. This, however, leads to an increase in the cost of the system and the need for more space for these components. Therefore, other redundancy methods are needed for autonomous machines. (Behere et al. 2016)

Another area where safety and other extra-functional properties are affected by increased autonomy is the predictability of the system. In general, safety critical systems have to be predictable and deterministic so that the way the system will operate in all situations can be predetermined. With autonomous systems, however, this becomes a problem. Inherently by design, autonomous systems include some degree of intelligence and decision-making capabilities, which leads to operation where only a rough determination can be made on the future actions of an autonomous machine because every scenario the machine may face cannot be known beforehand. This complicates the verification of safety of the system because the machine will have to operate in varying environments and around other heterogeneous machines, where the number of distinct interactions is vast. Some unpredictability is therefore to be allowed for autonomous machines, but the question is how much. (Behere et al. 2016)

3.3.2 Preventing hazards on an architecture level

The main types of hazards that arise from the operation of autonomous machines were presented in chapter 3.1. These hazards stem from internal errors and faults caused by the autonomous system itself and the operational environment of the machine. Autonomous system architectures must have a method to correct these faults and errors to minimise the hazards that arise from operation of the machine.

Exogenous hazards can be minimised by adding robustness to the autonomous system architecture. This can be facilitated by increasing the monitoring of the system and of the operational environment. Increased monitoring allows for greater knowledge of the state of the autonomous system, which alleviates the effect of outside interference. Robust monitoring also allows for greater sensing of the outside environment. This increases the probabilistic evaluation of the environment, lessening uncertainties, which minimises unforeseen situations the autonomous machine may face. (Baudin et al. 2007)

Endogenous hazards are also minimised by increased robustness, which is the main method of fault-tolerance of the system. The autonomous system architecture must be able to prevent the hazardous operation of the machine due to errors or faults and keep the system in a safe state at all times. This is carried out by both avoiding unsafe states and by bringing the system back to a safe state if needed. Increased monitoring is a benefit as it allows for the sensing of faults before they become an issue. (Baudin et al. 2007)

3.3.3 Separate safety layers

Another approach to construct system architectures, which ensure safety, is to implement a separate safety layer into the autonomous system. Several different approaches have been proposed on how a safety layer can be integrated, ranging from simple layers to full autonomous control units.

Simple safety layers can be added to autonomous system architectures to ensure safety. These layers can monitor the machine and its surroundings and control safety functions when necessary. They often also include decision-making properties, which, for example, are used to allow or cancel certain functions. (Toben et al. 2012) Complete independent safety systems have also been proposed that have increased control of the overall system. Independent safety systems can monitor and observe the overall system and check each hazardous planned function and stop them if necessary, and thus keep the system from entering an unsafe state. The safety systems also monitor internal data and try to detect faults. The safety systems are independent from the rest of the control system, which leads to simpler verification, and thus a greater level of safety. (Baudin et al. 2007)

Separate complete autonomous control units have also been proposed. These control units perform all autonomous operations and functions, as well as safety monitoring. In essence, an autonomous control unit would take the place of a human operator and would thus perform all control and monitoring functions without the need for the human operator to take control in any situation. A clear benefit of such a control unit is that many of the architecture problems discussed in the previous chapters could be avoided because the autonomous control unit would be separate from the rest of the system, and thus would only require certain inputs and outputs to operate. (Molina et al. 2017)

In the paper by Molina et al. entitled: *Assuring Fully Autonomous Vehicles Safety by Design: The Autonomous Vehicle Control (AVC) Module Strategy* (2017), the proposed autonomous control unit is an autonomous vehicle control (AVC) module. The AVC module is separated into two parts: the autonomous vehicle operation (AVO) and the autonomous vehicle protection (AVP) submodules. The AVO submodule performs all functions needed for the operation of the machine, such as navigation and motion planning. The AVP submodule in turn acts similarly to a safety layer; it monitors the state of the overall system and the environment and keeps the system in a safe state by, for example, deploying safety functions or cancelling hazardous actions. The AVP submodule also carries out internal fault and error detection. The AVC module thus carries out all functions that are needed for safe autonomous operation. The complete system diagram of an autonomous machine with an AVC module is presented in figure 5. (Molina et al. 2017)

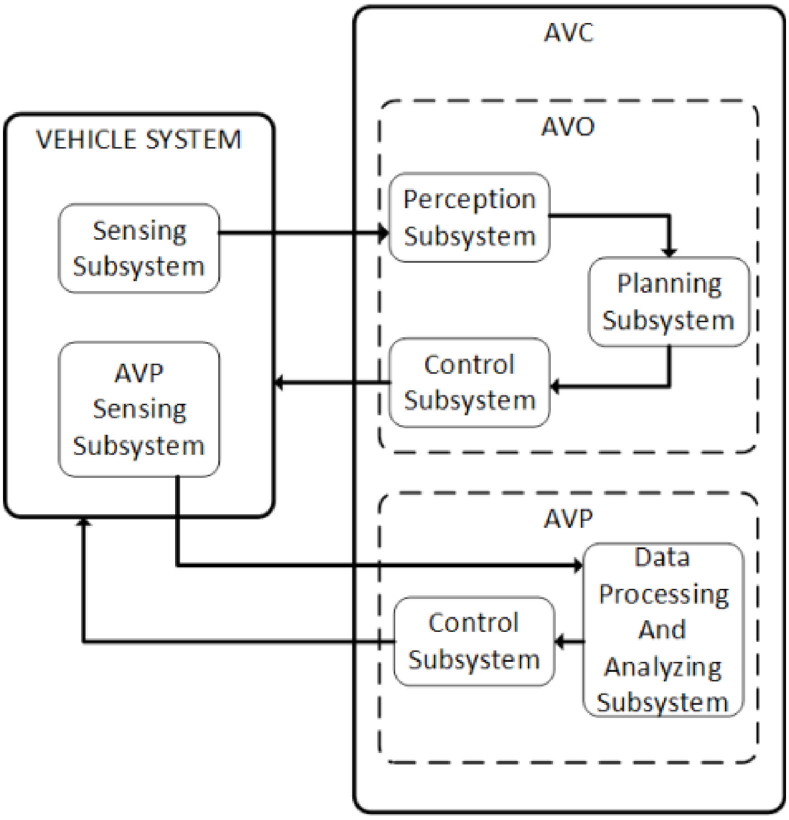

Figure 5. *Diagram of an autonomous system with an AVC module (Molina et al. 2017)*

In practice, the AVO submodule uses its own sensing subsystem, and it uses this information to determine its current location and the current state of the environment. This information is then used to plan an adequate and safe trajectory for the machine, which is put into action by the control subsystem that directly controls the machine. The AVP submodule has a sensing subsystem that is separate to the AVO subsystem. This is used to monitor the machine and its environment and to keep the system in a safe state. As there are two sensing subsystems, a level of robustness is added to the system in case of faults. Both submodules can send orders to the machine’s main systems when needed to ensure safe and effective operation. (Molina et al. 2017) Molina et al. do not, however, mention whether the two submodules can control or communicate with each other directly and not only through the main machine system. This may lead to problems in highly intelligent autonomous machines, as it may be needlessly complex for the two subsystems to only communicate through the machine's own system. Unforeseen, unsafe operation or feature interactions may also result if both submodules control the machine system simultaneously.

As the AVC module is separate from the rest of the vehicle's systems, the module and its submodules can be tested and verified independently, which leads to ensured safety of the system. Moreover, because the modules are separate, in theory, they can be implemented into any existing autonomous machine system. (Molina et al. 2017) This offers greater flexibility in the design of autonomous systems and their architectures, as any means to reduce complexity and to minimise the challenges of verification are clear benefits.

3.4 Localisation and motion planning

Localisation and motion planning of an autonomous machine comprises determining the location of the autonomous machine in regard to the world and the planning of a suitable set of actions to perform the tasks given to the machine.

Localisation of an autonomous machine is the act of determining the longitudinal and lateral position of the machine in regard to the world, and the direction it is facing. Several different methods have been used that include GPS navigation and vision and map-based methods.

Motion planning of an autonomous machine can be separated into two distinct parts: route planning and trajectory planning. Both of these must be computed by the autonomous system when movement is desired. The aim of route planning is to create a plausible route from A to B for the autonomous machine to traverse. Trajectory planning, on the other hand, calculates the exact motions the machine must take to achieve the desired route calculated by the route planner or the desired action it must take. (Benenson et al. 2008) Route planning is a greater challenge in automotive autonomy, where distances and different route options are greater. However, both aspects of motion planning must be solved in both industrial and automotive applications.

3.4.1 Localisation

The correct localisation of an autonomous machine is an important aspect of safety. If the localisation of the machine is incorrect, all future actions and motions of the machine may be incorrect, which may lead to clear safety hazards.

A wide variety of methods exist to determine the location of an autonomous machine in regard to the world. These include the usage of satellite positioning systems, which are generally used whenever a GPS signal is available, odometry-based methods, where position is calculated from the movement of the machine, and vision odometry methods, where position is determined with visual references.

Traditionally in outside applications, for example in autonomous road vehicles, GPS based localisation systems are the most common. In these methods, GPS signals are used to determine the current location and orientation of the machine. The reliability and accuracy of GPS is, however, not always adequate, which is why situational awareness and indoor positioning techniques are often used to enhance the accuracy of positioning. (Han 2008)

Indoor localisation methods depend largely on the nature of the operating environments. In fixed areas that do not change over time, such as factories and warehouses, separate infrastructure can be used for the localisation and navigation of autonomous machines. These are traditionally beacons and other signals that the machine can follow to determine its location and to stay on route (Mäkelä & von Numers 2001). The problem with these systems is the cost and difficulty of constructing the needed infrastructure and the considerable effort needed to make alterations in later use.

In evolving indoor environments, such as in underground mines, GPS signals are unavailable, and the usage of separate beacons or other infrastructure is not economically viable. In these environments, other methods of positioning are needed. These methods are most commonly dead reckoning or vision-based odometry methods, or a combination of the two. (Mäkelä 2001; Faralli et al. 2016; Aldibaja et al. 2017).

Dead reckoning is the practice of calculating a relative position of the machine in relation to a determined starting point via calculating movement. Wheel revolutions during movement of the machine are calculated, which is then used to determine the distance the machine has travelled from the starting point. A gyroscope, or other similar sensor, is used to determine the direction of travel and the sum of these two measurements is used to determine the location of the machine. The drawback of dead reckoning is that measurement error accumulates during movement, which may lead to a considerable position error if a long distance is travelled. Additionally, dead reckoning has to account for wheel slippage during movement, which can also affect positioning accuracy. (Gustafson 2011)

Vision-based odometry methods utilise visual landmarks that the autonomous machine uses for navigation and localisation. These visual landmarks can be, for example, a topological map of the area or a scan of the wall profile in a tunnel. The autonomous machine is fitted with a camera or sensor that is able to detect these visual landmarks. While in motion, the machine scans its surroundings and determines its location in relation to the visual landmarks it has been given in advance. (Aldibaja et al. 2017, Gustafson 2011) The drawback with visual odometry methods is that the visual landmarks must be determined in advance, and without them the machine cannot locate itself or navigate.

3.4.2 Motion planning

The safe motion of an autonomous machine comprises three aspects: the perception of the machine's surroundings, trajectory planning, and the correct control of the machine. Perception of the surroundings of the machine is a combination of effective situational awareness and an adequate world model. When a suitable trajectory has been chosen, it must be put into action by controlling the machine accurately. If errors are made in actuation, it may lead to erroneous movement and hazards. (Benenson et al. 2008)

To ensure a safe trajectory, three areas must be taken into account: the motion of the machine itself, the surrounding environment, and the infinite number of possible states or, in other words, the infinite nature of the time horizon. The first area is self-explanatory; the autonomous system must choose a trajectory that does not directly lead to a collision. The second point acknowledges that a collision can also result from the actions of other agents, not only the machine itself. Lastly, it is important to consider that the time horizon of an autonomous machine and other agents is infinite because, given enough time, it is certain that a collision can happen. Therefore, inaction of the machine itself does not ensure safety because in an infinite time horizon a sequence of trajectories made by another agent will inevitably result in a collision. (Benenson et al. 2008) In other words, a similar way of thinking is to apply Murphy's law to the state space of autonomous machines: any possible collision will happen, no matter how improbable, if enough time is given.

Traditionally in robotics, the safety of planned trajectories of a robot's movement has been ensured by the real time analysis of unavoidable collision states. An unavoidable collision state is a state of the robot where a collision is completely certain, irrespective of what actions the robot tries to make to remedy the situation. Thus, if a robot at all times ensures that it is not in an unavoidable collision state, no collisions will ever happen due the robot's own actions. In practice, this means that safe trajectory planning is a chain of states where none is an unavoidable collision state. (Fraichard & Asama 2003, Benenson et al. 2008) This methodology has also been applied to the trajectory planning of autonomous machines, but it is not enough to ensure safety in autonomous applications because this approach takes only the machine itself into account and not the actions of outside agents, ultimately ignoring the infinite time horizon and the trajectories of other agents. (Benenson et al. 2008)

An autonomous machine has only a limited comprehension of its surroundings, as there is a limit to what the on-board sensors can observe. Thus, some areas around the machine are not visible to the machine, as demonstrated in figure 6. The machine does not have any information on what is outside of the observed area: the unobserved area may include other agents with their own trajectories, static hazards or nothing at all. The information the machine has on the observed area may also include uncertainties, as it is possible the on-board sensors of the machine have made errors or have not gathered correct information due to faults regarding interference. (Benenson et al. 2008)

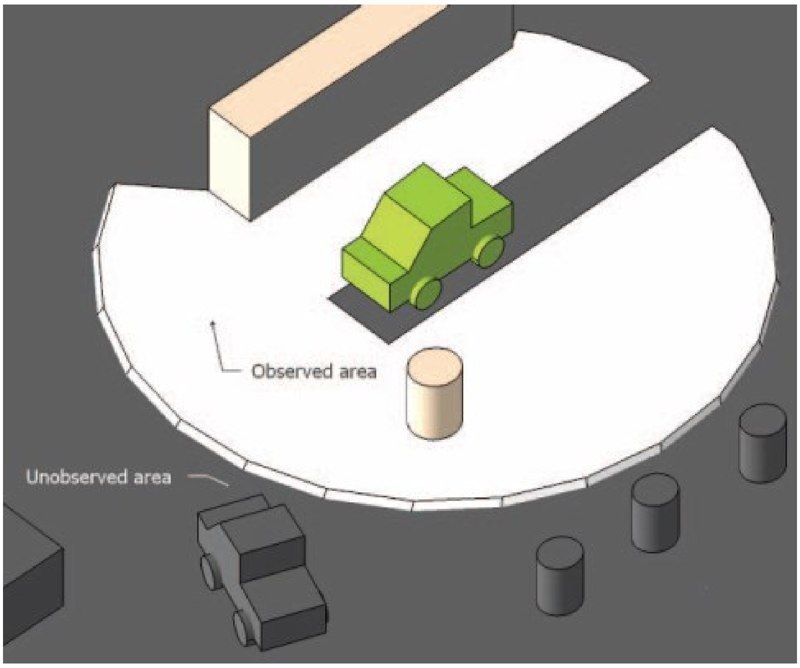

Figure 6. *Observed and unobserved areas around a machine (Benenson et al. 2008)*

Due to the missing or uncertain information, the world model of the autonomous machine is always incomplete, which poses a challenge for trajectory planning. The world model must either be completed, or a trajectory must be planned with an incomplete world model. Building a complete world model that includes all agents and their trajectories is not practical. Therefore, a method to plan trajectories with an incomplete world model is needed. Moreover, an autonomous machine in a dynamic environment must make decisions quickly, as the environment is constantly evolving and inactivity can lead to safety incidents (Laugier et al. 2007). Several different methods on how an autonomous machine should navigate and make decisions in an incomplete world have been studied. These include the use of occupancy grids (Laugier et al. 2007), Markov models (Seward et al. 2007), maximum velocity profiles (Alami et al. 2007), and Temporal logic methods (Jha & Raman 2016), among others. One effective and often used method of navigation with an incomplete world model is the usage of partial motion planning (Benenson et al. 2008, Laugier et al. 2007). First, a conservative estimation is made of the incomplete world model that can then be used in partial motion planning. (Benenson et al. 2008)

The aim of partial motion planning is to create a safe trajectory in the observed area around the machine that takes the machine roughly towards its end goal, without necessarily reaching it. As the machine moves, it gathers new information on its surroundings and is able to plan a more accurate route towards its destination. (Benenson et al. 2008, Laugier et al. 2007) To ensure these trajectories are safe, each step of the trajectory must be a collision-free state, that is not an inevitable collision state, and additionally, the final step must have a speed of zero. This does not mean the machine must continually stop after each trajectory, but rather the machine is able to continuously create new partial trajectories during movement, and if no suitable new partial trajectory is available, then the machine must be able to stop. This implies the machines entire trajectory is safe, and if the machine senses a hazard nearby, it will alter its speed so that it is able to stop at the end of its partial trajectory without a collision. (Benenson et al. 2008)

3.5 Situational awareness

For modern autonomous machines, situational awareness is one of the key areas in ensuring safety. In the past, this was not such an issue, as most autonomous vehicles were blind and thus operated in separate areas where other vehicles and people could not enter. This is not the case for modern autonomous machines that have to operate around other machines and people, often in unstructured environments. Thus, the role of situational awareness in ensuring safety cannot be overstated.

In practice, situational awareness of an autonomous machine is the knowledge of what is located and what is happening around the machine during operation at all times. This includes three aspects: observation of what is happening at the moment, assessment of how this affects operation and lastly, prediction on how the observed may change in the near future. This information is used to determine the level of safety of the machines current and future states. Thus, situational awareness is a method of risk assessment of the current and future states of the machine in regard to its surroundings. To determine the risks involved in a particular state, the autonomous system must analyse the information it has at its disposal. The two main areas to assess in the situational risk assessments of the machine are trust and completeness of information. (Wardziński 2006)

Trust is an attribute given to outside agents that the autonomous system has perceived to be operating around it. The attribute indicates the amount of confidence, or trust, the autonomous system places on the agent that it will operate as expected and in accordance to set rules. For example, set rules govern the operation of road vehicles and they must be followed. Therefore, an autonomous road vehicle can place a fair amount of trust on normal road vehicles that are on an adjacent lane to the AV: the AV can be relatively confident the other vehicles will stay in their own lanes, and therefore the AV can travel without slowing down and without the risk of an accident. However, if a learner driver is observed to be in an adjacent lane, less trust will be placed on it because it is not as clear if the learner will follow all traffic rules. For example, it is possible the learner will veer into the AV’s lane without indicating, the probability of which the AV must account for. This leads to reduced speed and larger safety margins. Therefore, a low level of trust leads to a high assessment of risk, which in turn necessitates the need for risk reduction methods in the current or future state of operation of the machine. Thus, if the environment is given a high amount of trust, the machine may perform to its full potential and with confidence that its actions will not create a safety hazard. But if trust is minimal, the autonomous machine may be unable to operate to its full potential or may be unable to operate at all. Therefore, trust has a direct effect on motion planning and overall safety of the machine. (Wardziński 2006)

For effective situational awareness, the evaluation of the completeness of the information the machine has gathered on its surroundings is equally important as the evaluation of trust because all situational risk assessments are based on this information. The autonomous system must assess the completeness and validity of the gathered information to ensure the risk assessments made are correct because the information may be incomplete or non-valid for a number of reasons. Usually, these are due to technical limitations or faults that can cause missing or erroneous data on the machine's surroundings. For example, weather conditions may have a great impact on visibility and thus on the ability of the machine to sense its surroundings. Missing information leads to an incomplete picture of the machine's surroundings and operational state that will require safety precautions to prevent hazards similarly to situations of low trust. The completeness of information must be ensured, but of importance is also that the autonomous system must have a means to determine when information is missing or if the information is uncertain. If this is not the case, the autonomous system may act hazardously if it makes decisions based on incomplete knowledge. Alternatively, if the autonomous system knows the information on its surroundings is incomplete, it can make assumptions what this information could be and continue to operate safely. (Wardziński 2006)

An adequate level of trust in other agents and an adequate level of information on the surrounding environment are enough to ensure the safety of an operational state in the normal operation of an autonomous machine. Problems arise, however, when irregularities arise, such as sudden hazards. For example, an autonomous system may attribute a high level of trust on another agent with which it is in close operation, i.e., the autonomous system has assessed that the probability of this other agent continuing on the course it is currently on as high. But if this other agent notices a hazard on the outside of the autonomous system's perception, the agent may have to alter its actions considerably, and the autonomous system has no way of knowing this. This may lead to a collision between the two or other similar incidents. If the area of perception of the autonomous system is too small, these situations are far more common. (Wardziński 2006) To circumvent the limitations of situational awareness and perception of a single autonomous machine, communication between agents could be increased. This would allow vehicles and machines to communicate to each other their current perception of the surroundings, dramatically increasing the range of perception. Sharing information would also increase trust between agents, as an agent could notify other agents on the actions it is going to take. Different viewpoints on the same situation also reduce erroneous sensing and missing information, reducing hazardous actions made based on incomplete information. (Wardziński 2006, Benenson et al. 2008)

3.6 Risk assessments

There are two approaches on how risks that arise from the operation of autonomous machines can be assessed and mitigated: the predetermined risk assessment approach and the dynamic risk assessment approach. These methods are used both to minimise the risks posed by the machine itself and, to some extent, minimise the risks the machine may face in operation due to the environment. (Wardziński 2008)

For simpler systems, a predetermined risk assessment approach can be used. In this static approach, the designers of the autonomous system conduct hazard analyses in the design phase of the system, where all possible hazardous states and sequences leading to accidents are determined and analysed. When the sequences of events that can lead to accidents are determined, barriers are designed to stop the autonomous system from entering these hazardous sequences. Barriers can be traditional physical barriers, or they can be software constraints based on sensors or location, or a constraint based on a need for a specific function before continuing, which all stop the machine from operating hazardously. The predetermined risk assessment approach is a linear method that can be visualised and analysed by an event tree analysis, as shown in figure 7. In the event tree, the autonomous system is faced with a potentially hazardous situation, where a hazardous event occurs. To minimise this hazard, the system deploys a barrier that may succeed or fail in mitigating the hazard of the event. If it succeeds, the system enters a safe state. If not, the hazard may increase, or an accident can occur unless another barrier is used. (Wardziński 2008)

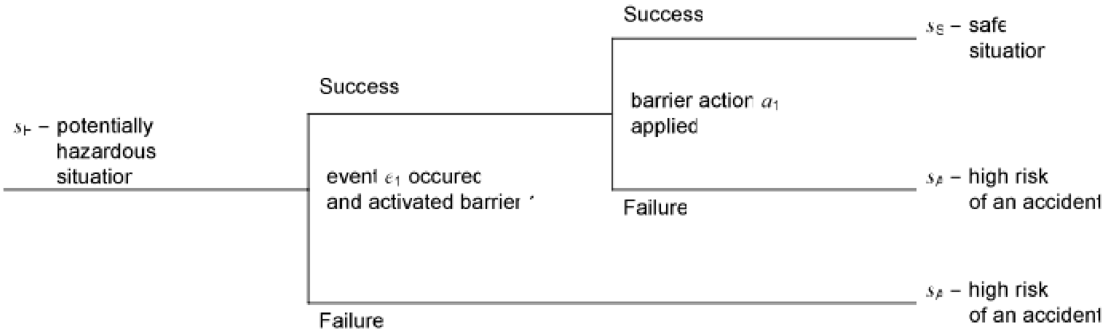
Success
s_S – safe situatior
Success
barrier actior a_1 applied
s_H – potentially hazardous situatior
event e_1 occurec and activated barrier
Failure
s_A – high risk of an accident
s_A – high risk of an accident
Failure

Figure 7. *An example of an event tree analysis as a part of a predetermined risk assessment (Wardziński 2008)*

The predetermined risk assessment approach is a straightforward method for simple systems, such as blind autonomous machines, as it only recognises two states: a safe state and an unsafe state. If a machine is in a safe state, it is allowed to operate, but if it is faced with an unsafe state, a barrier is applied. An adequate level of safety is simple to verify with this method because safety assurances are based on numerous separate cause-and-effect measures that can be analysed and verified separately. For more complex autonomous machines that can make independent decisions and act on them, however, this method would not work because the number of potentially hazardous situations and events would be vast and to analyse all of them would require considerable effort. Moreover, ensuring safety by applying barriers that essentially limit the operation of the machine would hamper the autonomy and intelligence of the autonomous machine considerably. Hence, the dynamic risk assessment approach is needed. (Wardziński 2008)

The dynamic risk assessment approach is suitable for more intelligent autonomous machines. Unlike the previous method, dynamic risk assessments are not carried out by the designers in the development phase, but rather continuously by the autonomous system itself in usage. The method is based on the notion that risks are not binary, as states can be safe, unsafe or anything in between. The autonomous machine may judge each situation independently and choose a suitable action based on internal dynamic risk assessments. This places a great emphasis on the situational awareness abilities of the system because they are needed to sense and determine the safety of the current and future states of the machine. (Wardziński 2008)

Simple event tree diagrams cannot be used to visualise how risks are mitigated in dynamic risk assessment methods because there are no clear cause-and-effect relationships. Rather, the autonomous system can decide on which actions to take which can lead to a varying degree of either safer or more hazardous states. This is presented in figure 8, where the autonomous system is faced with two hazardous situations (S_{H1} and S_{H2}). The autonomous system has several different possible actions it can take, which may lead to safe states (S_{S1} and S_{S2}), or accidents (SA), or anything in between (S_{B1} and S_{B2}). (Wardziński 2008)

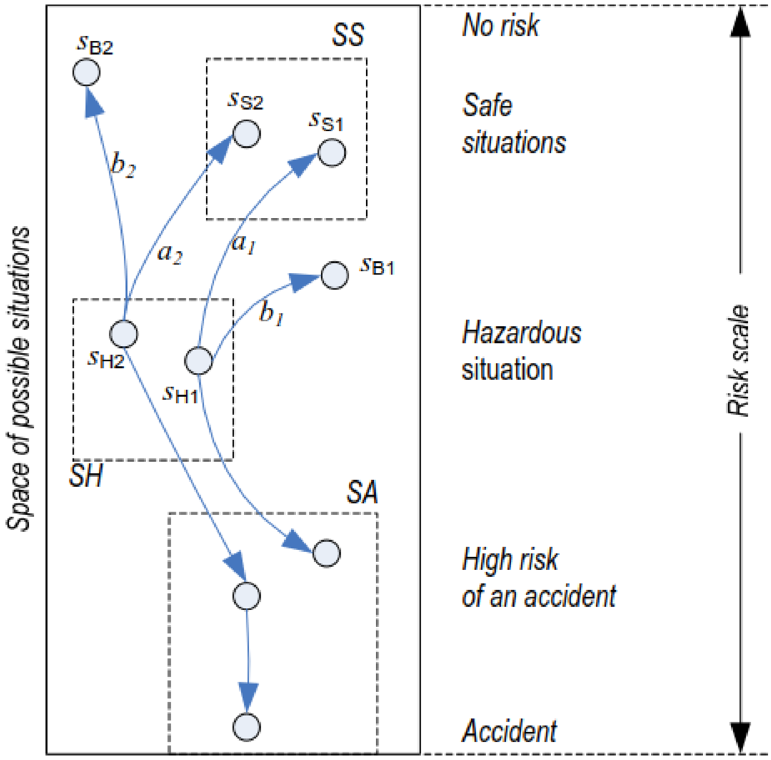

Figure 8. *Possible actions for a dynamic autonomous system in regard to risk (Wardziński 2008)*

The difficulty of the dynamic risk assessment approach is to design and verify such a system that can perform adequate risk assessments continuously in usage. This requires highly intelligent system architectures and situational awareness which, in turn, require precise sensing capabilities. To verify such a system is also challenging, as there is no common methodology or tools available. (Wardziński 2008)

A combination of the two previous approaches can also be used for the risk assessment of an autonomous machine. For example, it is possible to identify the main hazards the system can face as in the predetermined risk assessment approach. This information can then be used to compose specific safety rules for the autonomous system. These rules are a set of guidelines that can be applied to the internal decision making of the autonomous machine. These then ensure the autonomous system remains in a safe state, without breaking the safety rules and performing hazardous actions. (Baudin et al. 2007)

3.7 System verification challenges

Autonomous machines include various safety-related systems and functions, the correct error-free operation of which must be ensured for the safe operation of the machine. Many of the possible errors originate from the design of the system and the implementation of its functions, as discussed in chapters 3.1 and 3.3.2. Methods must be put in place to ensure an adequate level of safety for the machine, both in the design phase and when the machine is in use. These methods can be separated into offline and online techniques.

Offline techniques are used for the elimination of hazards that originate from the design phase, while online techniques minimise hazards that arise in operation. The challenge in verifying autonomous systems is that the state space of an autonomous machine is in essence limitless. In practice, this means not all state combinations can necessarily be verified, which creates a larger emphasis on online verification techniques. (Baudin et al. 2007)

During the development of the autonomous system, a model of the system can be created, which can then be checked and tested offline to ensure the system works correctly and dependably. Offline model checking is an automated method, where the model is given a set of behavioural properties and the full scope of different states are gone through. The model checker then goes through the different states in the state space, searching for contradictions in regard to the given properties, which include safety and liveness properties. The drawback of model checking is that it only offers an estimation of the final system in practice, as checking is conducted on a model of the system. Therefore, the method is not a comprehensive technique and other techniques must also be used. (Baudin et al. 2007)

Testing is another method for the offline verification of an autonomous system. Unlike model checking, testing can be carried out on the system itself, or parts of it, rather than a model. The limitation of system testing is that because the state space of an autonomous system is in practice limitless, not all situations can be covered or covering all situations may require a considerable amount of time and effort. (Baudin et al. 2007)

Offline verification techniques do not offer complete verification of an autonomous system, which is why online verification methods must also be utilised. Online techniques are tasked with eliminating hazards that occur in operation, including exogenous and endogenous hazards, and residual hazards that the offline techniques did not solve. (Baudin et al. 2007)

Online verification techniques can be separated into fault-tolerance and robustness methods. Fault-tolerance methods are traditionally used to ensure the system remains operational even if it is faced by faults, which are usually endogenous hazards, while robustness methods ensure the system avoids faults due to exogenous hazards. Fault-tolerance methods are usually based on adding redundancy to the autonomous system. These methods are used for error detection and for recovery of the system, which include error and fault handling. Robustness methods are separated on how they handle the erroneous states of the system and environment. These can be either implicit or explicit, where implicit handling applies the same methods to all states, and explicit handling only applies methods to specific sensed hazards. (Baudin et al. 2007)

3.8 Moral and ethical challenges

The moral and ethical dilemmas presented by autonomous vehicles have been widely discussed in recent years, although mainly regarding road vehicles. The discussion has mainly centred around how the autonomous vehicle should act in accidents and other emergency situations, where a collision is unavoidable. The main question being should the AV faced with an unavoidable accident be programmed to choose a trajectory based on some predetermined criteria and if so, what should this criterion be. Even though the discussion has revolved around road vehicles, this dilemma also applies to industrial autonomous machines and is, as such, something to be considered by manufacturers. Although the moral and ethical dilemmas have been discussed widely, a uniformly accepted approach to programming some sort of moral code in AV's has not yet been agreed upon.

The moral and ethical questions arise when the AV faces unavoidable accidents, and other hazardous emergency situations, that are not a part of its normal operation, and when deciding how it should react when faced with such situations. The classic example is an AV carrying a passenger that is about to be in an unavoidable fatal accident that includes other road users. This could be due to, for example, an unavoidable object in the way of the AV. In this example, another vehicle has blocked the road in front of the AV and the AV cannot stop in time to avoid a collision. The programming of the AV now has three choices: either manoeuvre to the left and hit person A, manoeuvre to the right and hit group B or finally, do nothing and hit the other vehicle, saving both person A and group B, but killing the passenger of the AV. This resembles the classic Trolley Problem thought experiment where a number of people are tied in front of a speeding train with one person controlling a lever that controls the train tracks. The person controlling the lever can either do nothing and have the train hit group A, or they can pull the lever and have the train alter its course and hit person B, thus saving more lives but ultimately directly causing the death of person B.

Similarly to the trolley problem, the inherent problem in designing an AV is that the AV must be programmed to choose one of these options, i.e., someone has to program this behaviour of the AV beforehand. This burden falls on the manufacturer of the vehicle and the software designers working on the vehicle who must somehow decide which is the correct action for the AV to take in situations such as the previous example. This is no easy task as there are no obvious right answers.

The root of the moral and ethical dilemma is that killing another person is almost uniformly illegal in all parts of the world. This is, however, exactly what has to be programmed in some fashion in the AV's code: in certain extreme situations killing a human being. As such, it is proposed that the answer to the moral dilemma should be based on the Doctrine of Necessity, which is a term recognized by the Anglo-American judicial system. According to the doctrine, in an emergency, extreme situation or extreme conditions, if there is no other option, something illegal can be carried out, and it can be regarded as legal in this specific situation. This translates to AV's in situations similar to the Trolley Problem mentioned above, where the only option is to cause a person's death. Therefore, this could be, from a legal standpoint, regarded as a non-illegal action. This does not, however, solve the original problem of choosing the right option in situations similar to the example given previously. (Santoni De Sio 2017)

The first ethical problem is the question of blame and consequence. In law, intentionally killing an innocent is in almost all circumstances illegal, and the person responsible is prosecuted for the crime. However, the relationship of responsibility and prosecution is not as clear in situations where the AV has taken an action that has resulted in a person's death. In essence, the AV has been programmed by the programmers to make decisions in some way in emergency situations, and to choose who or what to hit in a collision. It could thus be said that if an innocent life is lost due to the AV, this was ultimately due to the actions of the programmers of the AV. It can be argued, however, that the programmer is not to be held accountable because the programmer did not program the AV to kill a specific person, but rather programmed a wider range of guidelines for the AV for a wide range of different scenarios. Therefore, the manufacturer cannot be held accountable in most situations. (Santoni De Sio 2017)

According to studies, most people would choose a utilitarian approach to the AV Trolley Problem: they would simply have the AV in all situations choose the option that results in the fewest number of casualties. This approach, however, leads to several ethical problems, one of which is the problem of incommensurability, i.e., the value of different people is impossible to determine by comparing them to each other, as the value of a person is completely subjective. This is the most significant problem with the utilitarian approach to the Doctrine of Necessity: there is no objective way to compare the value or worth of a person or persons, and thus it cannot be said that choosing the option with the fewest fatalities is somehow objectively the right decision. Moreover, material damage is excluded from this because it is not comparable to the loss of life, and an AV should always choose material damage rather than fatalities. (Santoni De Sio 2017)

Further problems arise from the contractual obligations of the manufacturers of AV's. In law, it is stressed that manufacturers and service providers have a contractual obligation to keep their customers safe. Santoni De Sio uses a court case as an example of this in the article: *Killing by Autonomous Vehicles and the Legal Doctrine of Necessity* (2017), where sailors threw travelling customers off a ship to save the ship from sinking. The sailors where held accountable and prosecuted for this act because, according to the court, they should have sacrificed themselves because they had a contractual obligation to keep their customers safe. This is even though the utilitarian approach here would have been to sacrifice a few customers to save everyone else. This dilemma is also present in AV's, but it is also more complex. The manufacturers of AV's have a contractual obligation to keep to their customers' passengers safe. However, unlike the sailors, AV manufacturers cannot sacrifice themselves to save their passengers, but rather might have to sacrifice a third party in an accident, such as other road users, to uphold their contractual obligations if there is no other option. These parties are, however, entirely innocent in this situation and it would be morally questionable to have the AV choose to hit them. Thus, stating that choosing the AV to hit a non-customer rather than killing the passenger, due to a contractual obligation, is false. Therefore, it could be said that manufacturers also have an extra-contractual obligation to the third parties. This leads to the conclusion that contractual obligations are not enough to choose the appropriate behaviour of an AV in a fatal accident. To circumvent this, manufacturers could, in theory, sign a contract with the customers stating that in an extreme situation the AV might cause the death of the passenger. This is, however, something few people would willingly sign. (Santoni De Sio 2017)

Another aspect to consider in the programming of the AV is the responsibility held by road users. In many court cases throughout the years, great emphasis has been put on the responsibility of drivers of road vehicles, as it is seen they operate the means to harm another. This leads to the fact that even if a pedestrian or cyclist were in a fatal accident with a vehicle due solely to their own negligence, the driver of the vehicle would still be most likely prosecuted. A similar, or even greater, burden would fall on AV's and AV manufacturers as well. Because of this, AV's should always avoid hitting third parties, such as pedestrians and cyclists. However, in situations where the only options are to injure the passenger of the AV or to injure a third party, a clear contradiction can be seen with the earlier point, which states manufacturers have a contractual obligation to their customers. The responsibilities of road users are therefore not a suitable basis for the decision-making of AV’s either. (Santoni De Sio 2017)

Lastly, it is a matter of debate whether decisions of this calibre, i.e., of life and death, are even suitable for the manufacturers of vehicles and the AV’s themselves. As such, a higher authority in the decision-making would be beneficial. Vehicle manufacturers could, for example, be either given a set of binding legal guidelines that the AV's must follow in the case of an accident, or in the future the decision-making could be centralised into a separate automated system that chooses the right outcome in each situation. (Santoni De Sio 2017)

In summary, the moral and ethical dilemmas of AV decision-making are complex, but some guidelines can be drawn from the points mentioned above. Firstly, the AV should never choose to hit third parties, which are not part of the accident otherwise, and the AV should always choose material damage before human fatalities. Secondly, manufacturers have a contractual obligation to keep their customers safe, but this should not come at the expense of other road users. Lastly, the AV's should not target pedestrians or cyclists if there is an option to hit another vehicle, regardless of who is at fault. (Santoni De Sio 2017)

The previous examples are mainly for road vehicles, but the same problems and challenges exist in industrial fields as well. Industrial autonomous machines must also operate around people and other manned vehicles, and thus may cause harm to these other agents with their own actions. Ultimately, industrial machines may also face their own Trolley Problems and, as such, the moral and ethical considerations of decision-making apply.

The operational environments of industrial AV's are, however, not as complex as with autonomous road vehicles. Interactions with humans and other non-autonomous vehicles are not as frequent as with road vehicles, and thus situations where trolley type decisions must be made are rarer. The speeds of industrial machines are also generally slower, which shortens stopping distances, leading to fewer unavoidable collisions. Lastly, a major benefit of industrial autonomous machines is that the goal of industrial autonomy is often to situate the operator into a control room or to eliminate their presence completely, thus reducing the risk of human fatalities and eliminating the contractual obligations of autonomous machine manufacturers and the moral dilemmas they bring. Nonetheless, the moral and ethical implications of the decision-making of autonomous industrial machines is something to consider and something that must be accounted for in the design of such machines, even though situations were these problems arise may be rarer than in equivalent road vehicles.

3.9 Autonomy-safety-paradox

The level of autonomy of a machine is a double-edged sword, as the increase of autonomy may affect the safety of the machine. This is called the autonomy-safety-paradox (Matsuzaki & Lindemann 2016), where the increase of autonomy may come at the expense of safety, and similarly the increase in safety may come at the expense of autonomy.

In the past, autonomous machines were blind, as per the categorisation in chapter 2.3.2, and therefore operated in cordoned off areas where they followed predetermined routes with minimal interactions with other machines or people. This ensured an adequate level of safety for these machines. As technology has progressed, modern autonomous machines can sense their surroundings and do not need to operate in cordoned off areas or along predetermined paths. Therefore, the safety precautions set in place for blind autonomous machines will not suffice for modern autonomous machines, as they would interfere with the autonomous capabilities of the machine. A similar problem occurs if the risks of operation of a highly intelligent autonomous machine are mitigated based on predetermined risk assessments and barriers, as discussed in chapter 3.6. This is the essence of the autonomy-safety-paradox: an adequate and necessary level of safety must be achieved by the precautions put in place and by the design of the system, but they should not interfere with the autonomous operation of the machine considerably, as this would negate the purpose of the machine and its usage. (Matsuzaki & Lindemann 2016)

The autonomy-safety-paradox can also be thought of as a triple constraint for the design and implementation of the autonomous machine, as illustrated in figure 9. An advanced and safe autonomous machine can be seen as a sum of three parts: the level of its autonomy, the complexity of the machine's features and functions, and lastly, the machine's overall safety. To create such an advanced machine requires considerable effort and high sophistication of all three parts of the constraint. For example, a highly advanced autonomous machine would require a high level of autonomy, highly advanced features and a high level of safety.

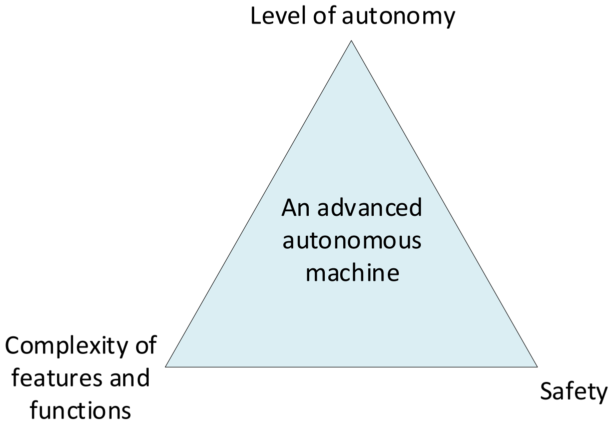
Level of autonomy
An advanced autonomous machine
Complexity of features and functions
Safety

Figure 9. *The triple constraint affecting autonomous machine design*

If any of the three constraints need to be changed in the design phase of the machine, it also necessitates changes in the two other constraints. Therefore, if, for example, the level of autonomy is increased in a machine, safety must be ensured for this new level of autonomy, and similarly, the features of the machine must be updated to take advantage of the higher level of autonomy.

Considerably less effort is needed in the design phase of the machine if only two of the three constraints need to be considered. For example, it is relatively simple to create a machine that offers highly complex and advanced functions with a high level of machine safety, but with no included autonomy. Similarly, a machine with a high level of safety and high level of autonomy is simple to design if the actual features of the machine are minimal and simple.

To create a framework to ensure the safety of an autonomous machine, binding legislation and standards are needed, as discussed in other chapters. This would allow manufacturers to create machines that have highly autonomous functionalities but still offer an adequate level of safety, as the machine would conform with the given standards and legislation, thus eliminating most problems brought on by the autonomy-safety-paradox (Matsuzaki & Lindemann 2016).

4. AUTONOMOUS INDUSTRIAL MACHINES IN PRACTICE: MINING

Mining is a field of industry that can benefit greatly from the implementation and increase in autonomy. This is due to the hazardous operating environments, monotonous tasks and the scale of operations affiliated with mining. This is why in this thesis, mining is chosen as a suitable example for industrial autonomy in practice.

In this chapter, the nature of mining is presented and the main benefits of increasing autonomy are discussed. Next, the main challenges autonomy can face in mining operations are presented, which often also affect operational safety. The current developments in mining autonomy are also presented, with an emphasis on Load-Haul-Dump (LHD) machines, the autonomy of which has been researched extensively in the last few decades. After this, the main safety issues autonomy brings to mining are discussed. Lastly, brief examples on autonomous machines from other fields of industry are given.

4.1 Mining and the benefits of higher autonomy

Mining is generally separated into two different categories: surface mining and underground mining. As the names imply, surface mining is mining above ground where ore deposits are accessed by removing the top layers of soil and rock. In underground mining, on the other hand, the ore deposits are accessed by digging underground tunnels. Both are quite similar in operation, but differ in some key areas, and these differences also effect autonomous operations to some degree.

Regardless of which type of mining is in question, the lifecycle of a mine is generally the same for both. A traditional life cycle of a mine is separated into five distinct phases: prospecting, exploration, development, exploitation and reclamation. In the first two phases, the location for the mine is determined by searching for and verifying ore bodies. In the development phase, the needed infrastructure for the mining operations which includes roads, access tunnels and so forth, is built. Next, the actual mining is conducted in the exploitation phase, where the desired ore bodies are extracted from the earth. Lastly, when the ore has been fully exploited, the reclamation phase begins where the mine is closed, and the environmental impact is minimised by restoring vegetation and water supplies.

In underground mining, several different machines are needed, each of which is used to achieve the common goal of extracting ore deposits from the ground. The machines range from tools for drilling, to tools for shaping tunnels to tools for transporting rock matter.

The most common underground mining method is drilling and blasting, where drill rigs are used to drill deep holes in the rock face. Explosives are then placed in to these holes and detonated. The created rock matter is then hauled from the tunnel with dedicated LHD machines and dump trucks. Other methods and tools used for underground mining are, for example, raise borers, cutting machines, among others. (Heiniö 1999)

In surface mining, drilling and blasting is also one of the most common mining practices. Here, similarly to underground mining, drill rigs are used to drill holes for explosives, which are detonated. Then, the rock matter is loaded into dedicated haulers that transport the material, for example, to be crushed by dedicated rock crushers. (Heiniö 1999)

Mining has been in the past a highly hazardous form of industry, with numerous lost time injuries and fatalities happening each year. This is due in part to the hazardous areas where mining takes place, but also due to the high-risk work methods used in mining. Risks in mining include falling rock, other vehicles and machines, poor visibility and the environment itself. In recent years, great effort has been put on mine safety, which in turn has decreased fatalities and injuries greatly. However, mining is still regarded as a high-risk industry in regard to safety. For example, there were 72 fatalities in US mines in 2004 (Dhillon 2010), and in the same year 6300 people were killed in mines in China alone (Kumar 2010).

Previously, the philosophy in mining was that to increase the amount of ore mined, mining companies would merely deploy more and/or larger machinery to achieve the demand for ore. This expansion unfortunately often came at the expense of safety. More recently, productivity and effectiveness with an emphasis on safety has become the driving force in mining, with mining companies monitoring these areas closely. This has created a need for smarter, more effective mining methods, and thus autonomy. (Marshall et al. 2016)

The safety and productivity of mining can benefit greatly from the increase in autonomy. This is mainly due to the hazards involved and the somewhat repetitive work tasks associated with mining, which can be carried out without an operator with autonomous machines. Mines are also generally located in isolated places with mining companies having simultaneous operations in different countries and continents. Thus, the relocation of personnel is a significant expenditure for mining companies, and therefore the increase in autonomy can reduce cost considerably (Nebot 2007). On a closer level, haulage is an area where both surface and underground mining can gain benefits from the increase in autonomy. In surface mining, for example, haulage accounts for 40% to 50% of operational costs, and haulage vehicles are in many instances a part of mine accidents (Nebot 2007). In addition to the safety and productivity gains, haulage tasks are often repetitive, and can therefore receive great benefits from the increase of autonomy (Marshall et al. 2016). Lastly, mining companies have suffered labour shortages in recent years, which is due to the shifting attitudes of the current generation of workers in regard to physical work. Modern workers are more accustomed to technology than hard labour, and thus both parties would benefit from the increase in autonomy. (Marshall et al. 2016)

4.2 Autonomy challenges in mining

The challenges of autonomy in mining are very similar to the challenges faced by other fields of industry, but the hazardous environments of mining add an additional challenge. For example, as stated previously, haulage accounts for a great portion of the operational expenses in mining, hence the interest in autonomous haulage vehicles. These vehicles suffer from the same situational awareness and localisation problems as other autonomous vehicles, but the operational environment adds to these problems considerably.

Mining operations and environments differ greatly between surface and underground applications, which is why different types of autonomy are called for and different technologies are needed. For example, in underground applications machines navigate in underground tunnels where GPS signals are not available, and other methods of localisation must be used, whereas in surface mining GPS signals are available, but the operational environment is far less structured than the equivalent underground tunnels.

In the article: *Surface Mining: Main Research Issues for Autonomous Operations*, Eduardo M. Nebot (2007) outlines the main issues surrounding the development of autonomous surface mining machines, some of which also apply to underground mining. Some of the issues brought up in the article have already been overcome, or the proposed solutions have already been implemented due to advancements in present-day technology. Nonetheless, the issues brought up in the article still have a great effect on autonomous mining, even if the issues have been solved.

The main issue with the operational environment in mining is the unstructured and unpredictable nature of mines. The layouts of mines are constantly evolving, and many structures and roads may be temporary. As such, predetermined maps and layouts are not readily available, so they cannot be utilised in the localisation and route planning of autonomous machines as effectively as with autonomous road vehicles. Thus, autonomous machines in mining rely heavily on sensory data and other means of positioning. (Nebot 2007)

The constantly evolving and changing nature of mines also means setting up separate fixed infrastructure for autonomous vehicles, such as beacons for guidance, is often not economically viable (Marshall et al. 2016). On-board sensors for localisation and situational awareness are thus a better option but have their own issues. The rugged operational environment of mines can have a negative effect on the performance of sensors due to extreme heat, vast amounts of dust and other factors This degrades the quality of the data acquired by sensors and can thus have a great impact on machine safety, if the autonomous machine performs tasks with missing or incomplete data. The rugged environment also has an effect on the overall health of the vehicle, which may deteriorate at a greater pace than in other industries. This is something the autonomous system must be able to monitor, as maintaining an operational state is important for the overall effectiveness of the machine, as well as for safety. This may, however, be challenging to implement in practice if there is no human operator present because this would require system-wide integration. A human operator, for example, is able notice, without issue, a strange sound coming from the vehicle, indicating a fault. Implementing such a feature as part of the autonomous machine is, however, an entirely different matter. Similarly, the harshness of the environment affects the state of the roads in addition to the state of the machine. For example, a haul route deemed safe and traversable the previous day may have degraded to such a degree that it cannot be used any longer. Therefore, an operator of a haulage machine must continuously monitor and asses the condition of haul roads and determine whether a particular route can be taken. Again, effectively implementing such a feature in a fully autonomous machine may be challenging to achieve in practice because this would require highly accurate sensing and decision making. (Nebot 2007)

Mining applications also require a vast number of interactions with manned machines and other vehicles. These situations are, for example, the loading and dumping phases of a haul vehicle’s work cycle. Interactions between autonomous machines and non-autonomous machines are difficult and potentially hazardous to perform because they require precise situational awareness and control of the machine, and any errors in either may lead to safety incidents. This is why all of the machines in a mine must be monitored and controlled effectively to ensure an adequate level of safety and efficiency. Such systems are already in place in numerous mines, which will be discussed later in chapter 4.3.4. (Nebot 2007)

4.3 Current developments in autonomous mining

Mines of the past have evolved from places with inferior occupational hygiene, and numerous safety hazards and high risks to highly monitored and efficient production systems that utilise a number of state-of-the-art technologies. A key area of interest in mining is autonomy, which is being implemented in all aspects of mining, with often the end vision of a completely autonomous mine. Mining automation and autonomy has been studied comprehensively in the last few decades, with most of the effort having gone into the automation of mining haulage vehicles, especially underground LHD machines. To facilitate autonomous mining operations, a mine must be closely monitored to ensure different parts of the operation work together effectively and safely. This has led to the usage of mine-wide monitoring and control systems that have turned modern mines into complex systems-of-systems.

4.3.1 Autonomous underground haulage

Autonomous mining haulage machines are the most researched aspect of mining autonomy, with several being available commercially for both surface and underground applications. The most studied application has been the automation of underground LHD machines, the first simulations of which were conducted in the 1980's and the first prototypes were in operation in 1999 (Gustafson 2011). Currently, semi-autonomous LHD machines are commercially widely available.

A mining LHD machine is a machine used in the loading and transportation of rock matter in underground mining. A typical LHD machine is presented in figure 10. Traditionally, the machines are used after the blasting of the tunnel face to load the created rock matter and to transport it to a location where it can be loaded onto a haulage truck.

Autonomous underground haulage trucks have also been researched extensively in the last few decades (Gustafson 2011), which has led to manufacturers recently offering these machines commercially. The trucks are used to transport rock matter from the LHD machines dump location to the outside of the mine. Therefore, their operating environment and the challenges faced are similar to underground LHD machines, which is why these machines will not be discussed in more detail in this thesis.

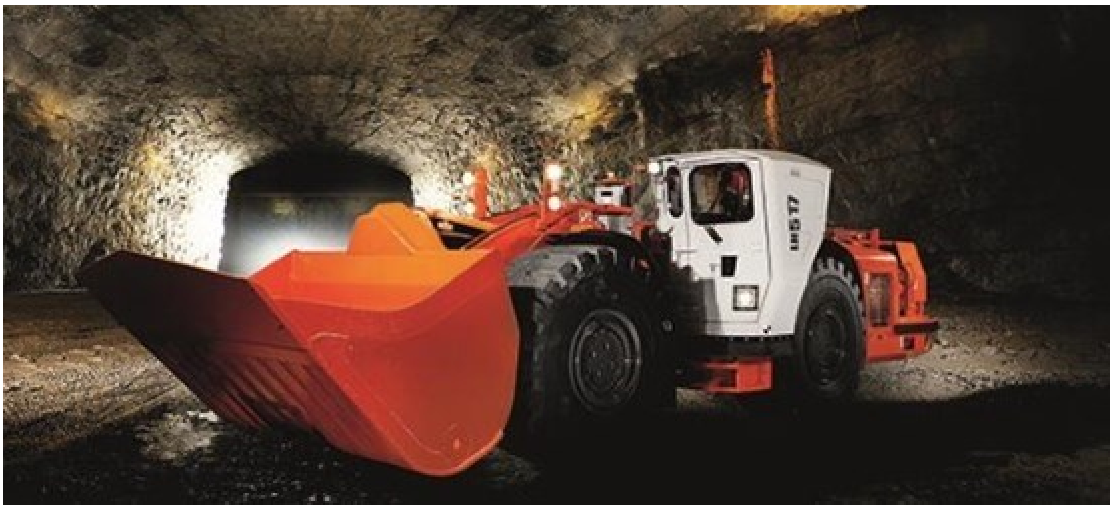

Figure 10. *A typical LHD machine (Sandvik 2018a)*

Traditionally, LHD machines are centre-articulated vehicles that utilise either diesel or electric power. They weigh between 20 and 75 tonnes and are 8 to 15 meters long. Normal operational speeds for LHD machines are roughly 20 km/h to 30 km/h. (Gustafson 2011) A normal work-cycle of a LHD machine consists of first loading the rock matter formed from blasting with the bucket on the front of the machine. After loading, the machine transports the rock matter through mining tunnels to a specified location, where the matter is loaded onto a haulage truck, which is also called the dump phase. This interaction between machines is seen in figure 11.

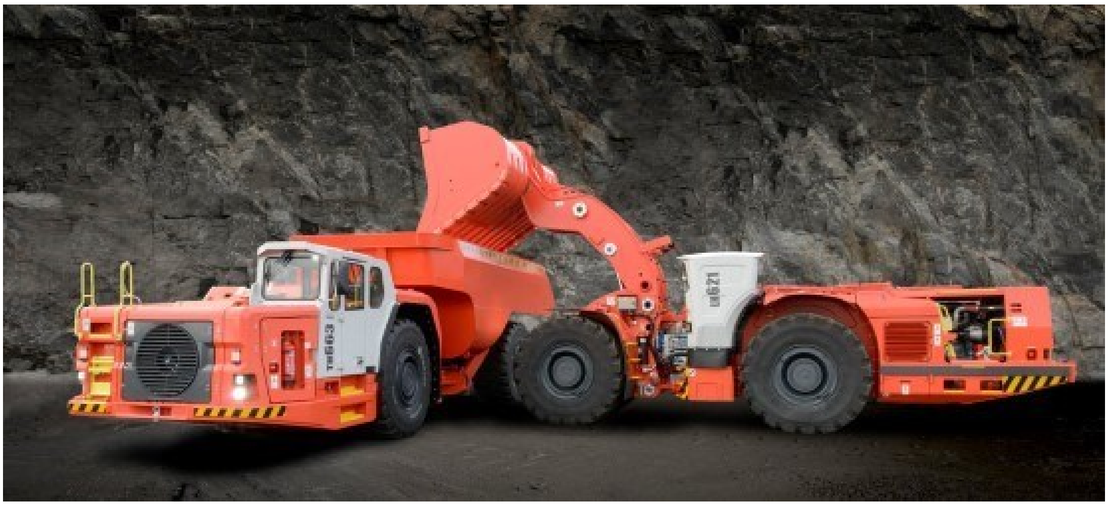

Figure 11. *Machine interaction between a LHD machine and an underground haul truck (Sandvik 2018b)*

Current LHD machines can be operated with a traditional human operator, via teleoperation, or semi-autonomously, but not fully autonomously. The haul and dump phases of the LHD machine's work cycle are fully autonomous, but the loading phase must be carried out by an operator manually or via teleoperation. This is due to the difficulty of automating the loading of rock matter, as the intricate differences in the densities and forms of rock necessitates the precise and skilled control of the bucket, which has not been able to be performed sufficiently using automation. Research and development has been carried out to automate this phase of the work cycle, but this functionality has not yet been able to be incorporated in commercial machines (Gustafson 2011).

The autonomy of LHD machines is mostly based on the safe and effective navigation of the machine. Traditionally, two types of navigation methods have been implemented in autonomous LHD machines: absolute navigation and reactive navigation. The former is a method where machines are blind to their surroundings and navigate along fixed paths to the desired destination that has been determined beforehand. Absolute navigation is thus a method for automatic navigation, rather than autonomous navigation, as the machine does not gather any information on its surroundings. The more modern reactive navigation methods, on the other hand, are based on the machine gathering information on its location and making decisions on navigation based on this information. As such, a predefined route is not needed. For example, the machine can sense its surroundings by analysing the tunnel face around the machine and then use this for positioning and navigation. (Gustafson 2011)

As GPS signals are not traditionally available underground, other methods for positioning and localisation must be used. Manufacturers have taken different approaches to solve this problem. The applied methods, however, are all based on scanning the machines surroundings for information that can be used to determine the position of the machine.

In addition, movement-tracking methods, such as dead reckoning, are also used. (Gustafson 2011) These visual and movement-based methods were discussed in more detail in chapter 3.4.1.

For example, the Sandvik AutoMine autonomous navigation system is based on absolute navigation, and it utilises both visual odometry and dead reckoning. In practice, the system works by first manually teaching the machine a suitable reference trajectory by traversing the tunnels of the mine and simultaneously calculating the machines position by dead reckoning and by scanning the distance from the machine to the tunnel walls. After this, the machine can traverse independently by determining its location by dead reckoning and by comparing its location to the information gathered in the teaching phase. The benefit of this method is that no separate infrastructure or premade maps of the mine are needed for the navigation of the machine because the teaching-phase gathers all needed information for navigation. (Gustafson 2011, Mäkelä 2001)

Other similar methods of navigation have also been researched. For example, a similar method to the previous AutoMine system has been proposed by other parties. In this method, the machine is similarly taught by traversing the mining tunnels manually, while data are gathered with lasers and by articulation and speed sensors. The data are then used to create a metric map of the mine and a suitable route profile. In autonomous mode, the machine uses this route profile and metric map to navigate by ensuring with the on-board sensors that the machine stays on the desired route. (Marshall et al. 2008) Other methods for navigation include the usage of premade maps together with sensor information (Larsson et al. 2006) and vision-based methods that recognise mine intersections and other visual clues (Gustafson 2011).

4.3.2 Safety of autonomous haulage machines and standard ISO 17757

To ensure the safety of haulage machines, especially LHD machines, the standard ISO 17757: "*Earth-moving Machinery and Mining – Autonomous and Semi-Autonomous Machine System Safety*" has recently been released. The standard outlines the general requirements and main risks for all aspects of an autonomous haulage machine that manufacturers of the machines must adhere to in the future. (ISO 17757:2017)

The general safety requirements for an autonomous underground haulage machine are for the machine itself to comply with ISO 12100 (*Safety of machinery -- General principles for design -- Risk assessment and risk reduction*), while the control system must comply with IEC 61508 or a similar other functional safety standard, which were discussed in chapter 2.2.1. More specific requirements are given in the standard for the main aspects of the machine system that include positioning and orientation, digital terrain maps, perception and task planning. (ISO 17757:2017)

In chapter 3, some of the main safety challenges for autonomous machines and vehicles were presented. These included localisation, motion planning and situational awareness. These aspects are also included in standard ISO 17757 that outlines the main requirements and risks associated with each aspect. Even though it was stated in previous chapters that system architectures are an important aspect of autonomous machine safety, they are not directly dealt with in the standard.

In the standard, localisation is classified as positioning and orientation (POSE). The POSE system is in charge of the calculation and monitoring of the machines position and orientation in regard to the world. The standard outlines the main possible failures associated with the POSE system that include inaccurate determinations of position or orientation, or the complete lack thereof, which can lead to collisions with the environment or other machines. The standard requires that the autonomous system must be able to sense the aforementioned faults and that the system must remain in a safe state even if faced with such faults. The standard also requires a certain amount of robustness in the POSE system because the machine must be able to determine its position and orientation even if one part of the POSE system encounters a fault. (ISO 17757:2017)

The aspects of motion planning are discussed in the standard under chapters on navigation systems and task planning. The navigation system of the autonomous haulage machine ensures the machine navigates effectively and safely to the desired location, while the task planner plans the actions that are needed to reach this location, based on internal risk assessments, and then puts them into action. The main risks associated with the navigation system are possible collision with the environment or other machines, which can result from erroneous POSE information or insufficient control of navigation. To minimise these risks, the navigation system must be able to notice if it no longer has a safe heading or velocity and remain in a safe state in these situations. Risks associated with the task planner, in turn, are hazardous tasks that may lead to damage or injury if they are put into action. For example, an erroneous task may lead the machine directly onto a hazardous route, or the route taken may lead to hazards for others. To minimise these risks, the task planner must be able to detect and avoid hazardous actions before they are put into action. (ISO 17757:2017)

Lastly, situational awareness is discussed in the standard under digital terrain maps and perception. Some machines may utilise a digital terrain map which is used for both situational awareness, task planning and navigation. Operational risks arise if the map is inaccurate or otherwise erroneous. Therefore, the POSE system must be monitored closely when the map is created, or the area surveyed, to minimise errors. Perception, as described in the standard, is similar to situational awareness as discussed in chapter 3.5. A perception system is used to detect what is around the machine at all times and to gather relevant information for navigating without an operator. Possible risks and failures perception systems can face include the failure to detect an object completely or the failure to detect the object in time. Other errors in detection include erroneous locations, misclassifications and false-alarms. Ultimately, the main goal of the perception system is to keep the machine in a safe state at all times. Therefore, if the system is not operating correctly, the operator must be notified, and the machine to be kept in a safe state even if detection errors are encountered. (ISO 17757:2017)

4.3.3 Other autonomous mining machines

Other areas of mining have also made advancements in the adoption of autonomy. These vary from autonomous haulage vehicles to drills with autonomous functions. Many of which have come to fruition due to the advancements made in underground haulage autonomy.

Another highly researched and commercially available area of mining autonomy is surface mining haul trucks that have followed in the footsteps of underground haulage machines. An example of such a machine is presented in figure 12. In general, surface haulage machines are similar to their underground counterparts, but differ mainly in size and operational environment. Surface trucks are also tasked with moving rock matter around the mine site, but usually the distances are considerably longer than in underground applications. Surface haul trucks are also generally far larger, as their carrying capacity is usually several hundred tonnes.

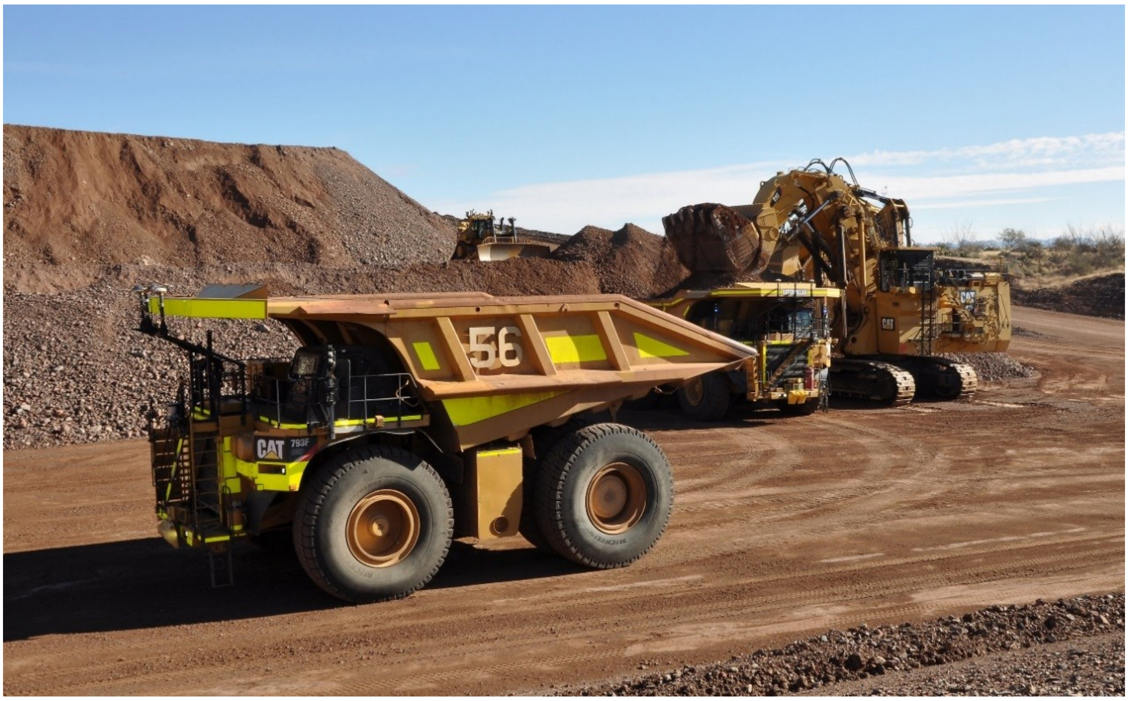

Figure 12. An autonomous surface mining haul truck (Caterpillar 2018)

Surface haul trucks offer varying degrees of autonomy, from mere driver assist systems, such as collision-alert and auto-spot systems to fully autonomous operation. Semi-autonomy, including teleoperation, is also possible. (Brown 2012) The machines generally use a combination of GPS and radar for localisation and obstacle detection, and they also include a communication system that connects it to the mines central command system. (Marshall et al. 2016) Due to the benefits autonomous haul trucks introduce, they have increased in adoption in recent years. For example, there are currently one hundred autonomous trucks similar to the one in figure 12 in active mining operations. (Watkins 2017)

Other areas of mining autonomy include autonomous functions for drill rigs and semi-autonomous bulldozers. Blasthole surface drill rigs are machines that are used for drilling holes for explosive charges in surface mines. The autonomy of these machines is under active development and autonomous models will be available commercially in the near future. Functions that these machines will offer vary from teleoperation to partial autonomy, where the machine can drill a row of holes autonomously after an operator has drilled the starting hole. (Watkins 2017, Leach 2015) Autonomous surface mine bulldozers are also commercially available. At the moment, these machines offer semi-autonomous operation, where an operator is needed to set up the task for the machine. After this, the machine is capable of carrying out the task independently. Such a task is, for example, push-to-edge functionality, where the machine pushes matter over the edge of the mine pit. (Watkins 2017, Jensen 2016)

Other areas of autonomous mining under research include diggers, rock breakers and draglines. Most of these tasks are highly repetitive and suffer from the same hazards as other forms of mining machines.Therefore, these machines are well suited for the application of autonomy. However, some aspects are still under research, which is why these machines are not available commercially at the moment. For example, autonomous diggers are still under research as they suffer from the same bucket control problems as underground LHD machines: the intricate control of the bucket is challenging to perform autonomously due to the heterogeneous consistency and size of the rock matter. Autonomous diggers also require precise situational awareness to determine the terrain around the machine and the location and orientation of the bucket. Similar problems have been faced with the autonomy of rock breakers that are used to shatter large fragments of rock with a hydraulic hammer on the end of a boom. Precise situational awareness is also needed in the automation and autonomy of draglines, which are large machines with a bucket and mast, that are used to drag rock matter in open coal mines. Research and development on these machines are still ongoing. (Marshall et al. 2016)

4.3.4 Mining systems and the mine of the future

A mine consists of a vast number of different machines and vehicles, which all have varying degrees of autonomy and operate in the varying stages of the mining processes. Simultaneously, mines are also under pressure to work as efficiently and effectively as possible. These requirements have led to the widespread adoption of mine-wide control and monitoring systems, which cover all aspects of the mines operation to ensure productivity and efficiency, that have effectively turned modern mines into complex systems-of-systems. A need for such a system was already recognised in the early stages of the modernisation of mining, as especially the introduction of autonomy in machines necessitates a central control and monitoring platform (Nebot 2007). The challenge with these systems is, however, that for them to work effectively they have to be implemented in all aspects of mine operations. This leads to an all-or-nothing approach that requires considerable financial investments.

Mine systems are offered by several manufacturers and offer similar functionalities ranging from central communication systems to fleet management. The systems also integrate teleoperation functionalities, mine mapping functionalities and other tools. Traditionally, the systems are monitored and controlled from a separate control room, which may be, for example, a van on the mine site or an office completely separate from the mine premises.

Fleet management is the most central part of the mine system in regard to the operation of machinery. It is traditionally tasked with three aspects: position monitoring, production monitoring and equipment task management. Position monitoring allows for the real time monitoring of each machine and vehicle in the mine, including the weights and types of material they are transporting, or the current drilling depth, for example. This allows for the effective performance tracking of the mine operations and precise task planning. Above all, this increases safety by minimising collision risks because the positions and routes of machines and vehicles can be actively monitored. The internal state of each machine can also be monitored by production monitoring, which gathers data such as machine cycle times, failures, payloads and so forth. The data can then be used to plan production accordingly. Lastly, assigning tasks for the equipment in use is a central part of fleet management. Production and position data are used by control room personnel to determine what tasks need to be carried out by the machines in operation. These tasks are then sent to each machine, which can be both manned or unmanned, through the fleet management system. The task is then put into action and it can be monitored in real time by the control room personnel. Effective fleet management has a direct effect on mine safety because it allows for real time monitoring of all operations, and thus allows for the elimination of risks related to collisions, as these are responsible for most safety incidents in modern mining. (Marshall et al. 2016)

Standard ISO 17757 also sets requirements for the mining system, or as titled in the standard, the autonomous machine supervisor system. The standard recognises that risks may arise if the supervisor system sends out erroneous tasks to the machine, either due to operator error or a fault. To minimise risks, the connection between the supervisor system and the machine itself must be periodically verified. If there is found to be a problem in the communication system, the machine must be able to keep itself in a safe state without input from the control room. (ISO 17757:2017)

The ultimate goal for mining companies in implementing mine systems and mine autonomy is a mine that does not need personnel at the actual mine site at all. For example, in 2008, the mining company Rio Tinto initiated the Mine of the Future programme in its iron ore mines in Western Australia. The aim of the programme was to find new ways of mining that increase efficiency, safety and sustainability, while minimising environmental impacts. Autonomy has been a central part of the programme, with a vast number of autonomous machines operating in the Rio Tinto mines, which are all controlled by a central operations centre that is 1500 kilometres away. (Jensen 2016)

4.4 Safety challenges in autonomous mining

The increase in machine autonomy has numerous advantages in mining applications, ranging from increased productivity and efficiency to improved safety. For example, the safety benefits of autonomy are clear: with autonomy, the need to situate personnel in hazardous environments is minimised, or in some cases completely eliminated. Nonetheless, the increase in autonomy in mining applications brings with it safety challenges that must be overcome to ensure operational safety. Most of these challenges are similar to the challenges autonomy itself faces in mining applications, which were discussed in chapter 4.2. The safety challenges stem mostly from the operational environment of mining, which includes hazardous weather conditions and frequent interactions between machines and vehicles.

The harsh environmental conditions of most mines has an adverse effect on sensors and their performance. The conditions range from extremely high or low temperatures to high humidity and to considerable amounts of dust. Especially underground mines are harsh environments, as they are dark, damp and humid. All of these aspects can possibly degrade sensory data or render it unavailable. As most autonomous machine functions are based on sensing and sensory data that are acquired by sensors mounted on the machine, the data are critical for the safe and effective performance of the machine (Nebot 2007). Methods are needed to ensure the machine remains in a safe state even if there are lapses in the gathered sensory data. These methods include added robustness to the sensory systems and the ability for the autonomous system to sense missing or erroneous data.

Another major challenge for safe autonomous mining machines stems from the interactions with other machines, vehicles and personnel. As most safety incidents in modern mines are the result of different forms of collisions, it is also a problem autonomous machines must face. Most of the incidents occur when machines collide while traversing the mine or in situations where two or more machines must interact together, for example, during dumping or loading rock matter. The root cause in these incidents is often poor visibility or problems in communication. (Marshall et al. 2016) Central mine command systems and the situational awareness of machines can minimise these hazards, but a great level of system integrity and fault tolerance is needed, as a great weight is put on these systems to ensure safety in interactions with autonomous machines. For example, errors in task assignment or position monitoring can lead to high-speed collisions of machines if they are ordered to traverse the wrong route. Moreover, precise control and monitoring is also needed for situations where autonomous machines must interact with other machines or vehicles to ensure no collisions occur.

A highly challenging form of interaction occurs when autonomous and non-autonomous machines must work together to perform a task. For example, such interaction can occur when an autonomous LHD machine dumps rock matter onto the bed of a manned mining truck, or when a manned and unmanned machine share a road together. These interactions require highly precise situational awareness and control with narrow margins for error to minimise the safety risks for the operator of the manned vehicles. Traditionally, these interactions have been eliminated by having the autonomous mining machines operate in cordoned off areas, where other machines and personnel cannot enter. For example, autonomous LHD machines often operate in separate tunnels that are closed off with gates, such as in the example in figure 13. If a person or machine enters through such a gate while autonomous operation is active in the tunnel, every machine is stopped automatically. In surface mining applications, the same functionality can be achieved by creating GPS perimeters that other machines cannot enter (Gustafson 2011).

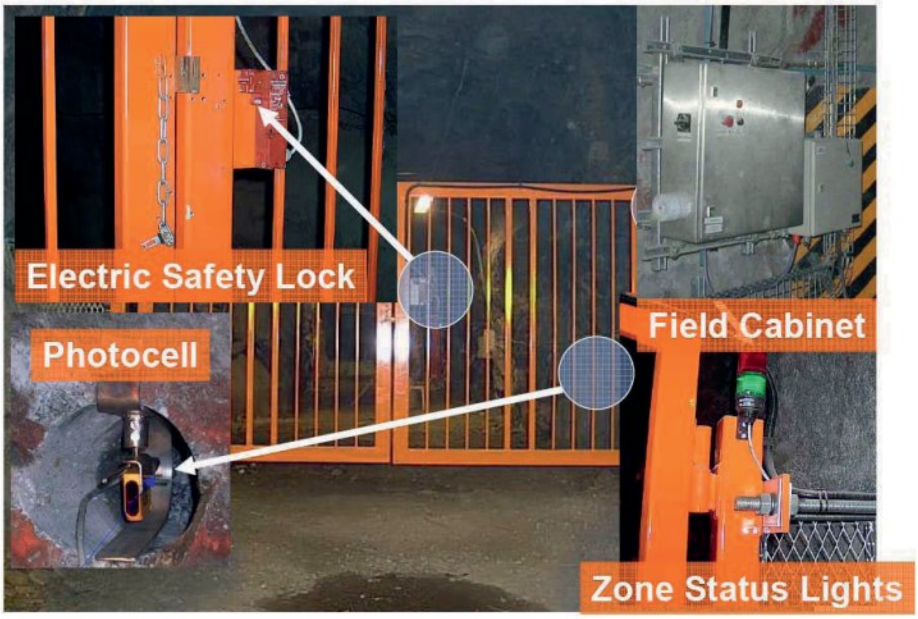

Figure 13. *An example of a gate cordoning off an area of autonomous operation (Gustafson 2011)*

Separating autonomous and non-autonomous mining machines leads to the autonomy-safety-paradox, which was discussed in chapter 3.9. By eliminating the interactions between different machines, an adequate level of safety is achieved. However, this limits the autonomous capabilities and effectiveness of the machines because they are only allowed to operate in certain areas and to carry out certain tasks. For example, in surface mining some routes may only be traversed by autonomous machines, which means manned machines must wait for their turn or find another route. High situational awareness and intelligence of the machines would eliminate the need for such arrangements and would allow for machines with heterogeneous levels of autonomy to operate in the same area. Another simple solution would be an all-or-nothing approach to autonomy: if a mine is completely autonomous, all interactions with manned machines are eliminated and no risk is present. Such an approach, however, is not possible with current technology and, moreover would lead to considerable financial investments which mining companies may not be willing to partake in (Brown 2012).

4.5 Other industrial autonomous machines

In addition to mining, autonomy has been widely researched in other industrial fields including such areas as agriculture, transport, maritime applications and ship port automation, for example. Traditionally, fields that have either repetitive or hazardous work tasks have had the most to gain in adopting autonomy. For example, the autonomous possibilities in agriculture have been studied for a number of years (Torii 2000), which has led to commercial offerings being currently available. Agriculture is an area that can greatly benefit from autonomy because the work tasks are often repetitive and relatively simple in nature. Agriculture is relatively hazard-free, so no great benefits are gained from autonomy in this regard. However, autonomy does have the benefit of minimising operator contact with the poisonous insecticides that are used in agriculture (Pushpavalli et al. 2015).

Work tasks in agriculture usually consist of traversing a field in a straight line, while spreading seeds or pesticide, ploughing, harvesting and so forth. As the fields are large, a single machine would have to traverse end-to-end several times to cover one field, which is why in agriculture fleets of machines often operate at the same time. This has led to the adoption of follow-me based autonomous and automatic systems, where one master vehicle, operated manually, is followed by one or more slave vehicles that are autonomous or automatic. (Zhang et al. 2010, Bedord 2017) Agriculture machines that operate independently of a master vehicle are also under research and some are available commercially (Torii 2000, Agriculture News 2008). These machines navigate with the use of GPS and on-board sensors. A remote operator is, however, still needed to monitor the machine.

Other industrial fields that have utilised autonomy have used similar approaches. Most autonomous machines are tasked with performing straightforward functions independently, while an operator is used for monitoring and the execution of more complex tasks. Technology is, however, advancing rapidly, which can lead to more complex work tasks for autonomous machines with decreasing amounts of input and monitoring needed by machine operators.

5. CONCLUSIONS

Recent advancements in technology has allowed for the development of increasingly complex and intelligent autonomous industrial machines and civilian road vehicles. These machines and vehicles are able to act and make decisions independently and to perform functions that were a mere vision not too long ago. The increase in autonomy has, however, led to increased safety concerns. As autonomous machines and vehicles are able to operate without the supervision or control of a separate operator or driver, concerns arise on how to ensure an adequate level of safety in all operational situations.

In this thesis, the main aspects that affect overall safety of autonomous machines, vehicles and the design of systems were presented. Additionally, the main safety challenges of increasing autonomy in machines and vehicles were also discussed. The ultimate goal of the thesis was to give an overview on what safety-related aspects have to be considered to achieve an adequate level of safety for autonomous machines and vehicles.

The safety of an autonomous machine or vehicle is the combination of both safety-specific functionalities and overall correct operation of the machine or vehicle in all situations. To ensure an adequate level of safety, these aspects have to be considered in the design phase of the machine, as autonomy cannot be regarded as a mere feature, but rather as an all-encompassing aspect of the machine.

The main safety challenges that arise from the increase of autonomy in machines and vehicles include areas such as building suitable system architectures, creating effective situational awareness capabilities, as well as ensuring the correct localisation and motion planning of the autonomous machine. Moreover, the increase in autonomy introduces questions and concerns relating to the internal risk assessment and decision-making capabilities of the machines and vehicles, which ultimately can lead to moral and ethical dilemmas. All of these aspects are equally important for the safe operation of an autonomous machine, as any errors or faults in these areas can lead to undesired and erroneous behaviour, possibly leading to safety hazards and incidents, or even fatalities.

Ultimately, the difficulty in creating safe and effective autonomous machines and vehicles is due to two main correlating aspects. First, increasing autonomy in machines and vehicles can be seen as a paradox of sorts. The intent of the manufacturers of autonomous machines and vehicles is often to create an autonomous system that is as advanced as possible, so that they can compete with other manufacturers effectively. Simultaneously, the autonomous system must also be as safe as possible, which can ultimately limit the autonomous capabilities of the system. In other words, the increase in safety can often come at the expense of autonomy and vice-a-versa, the increase in autonomy can often come at the expense of safety. Designing a safe, effective and sophisticated autonomous machine or vehicle can also be said to be controlled by a triple constraint that comprises safety, the level of autonomy and the complexity of the machine’s functions. Therefore, changing any of these three aspects necessitates alterations in the other two, as for example, the increase in the level of autonomy necessitates an increase in safety and changes in the machine’s corresponding functions, if a safe and sophisticated machine is desired.

The second reason for difficulties in designing safe autonomous machines and vehicles is the lack of appropriate legislation, standards and other guidelines on how autonomy should be implemented in practice. Both autonomous civilian road vehicles and industrial autonomous machines are hampered by bottlenecks in this regard. In the field of civilian road vehicles, some legislation and standards have been passed that offer guidelines on how autonomous vehicles should be designed and what steps must be taken before they can be used on public roads. Guidelines have also been put in place on ways to categorise the levels of autonomy of road vehicles, such as by the NHTSA and SAE International. These categorisations can then be used to create more specific legislation and guidelines on ensuring safety for specific levels of autonomy. Currently, however, only a select few of such legislation and guidelines are available. In industrial fields, on the other hand, the situation is difficult. At the time of writing, only two standards by the major standardising organisations are known to exist regarding the autonomy of industrial machines: the very recently released standard ISO 17757 on the autonomy of mining haulage machines, and the standard ISO 18497 on autonomous agriculture, which is still under development. Therefore, very little information and few guidelines are available for industrial manufacturers on how autonomy should be implemented in practice and how safety of such machines should be ensured. Moreover, as no guidelines are available, there is currently no common method for classifying the levels of autonomy in industrial machines. Therefore, future standards and guidelines can be challenging to apply in practice to machines, as there is no commonly accepted way to separate machines based on their autonomy. Ultimately, the resonsibility for ensuring the safety of autonomous industrial machines in practice is on the shoulders of manufacturers alone – at least at the moment. Therefore, the advances made in the automotive field in this regard could be used as a guide of sorts for industrial applications. For example, a classification method for autonomous industrial machines could be produced based on the guidelines by the NHTSA and SAE International.

As this thesis was conducted as a literature review, no real-world tests or practical examinations were carried out on specific autonomous machines. Rather, mining autonomy was used as an example of industrial autonomy in practice, which was also used to demonstrate the types of problems and challenges safe autonomy can face in practice. Continuing this research, real-world tests could be carried out on differing autonomous machines in usage and data could thus be gathered on the different safety-related aspects of autonomous machines presented in this thesis. This data could then be used to compare how different safety methods affect overall machine safety in autonomous applications.

In the future, it is certain that autonomous machines and vehicles will enjoy continued rapid advancements in technology. As the field of autonomy matures, new standards and other legislation will become available for manufacturers and developers to ensure autonomous machines and vehicles operate safely and correctly, without introducing hazards. It is to be hoped, however, that this happens sooner rather than later, so that the full benefit of autonomous technologies can be taken advantage of as soon as possible. Ultimately, autonomy will be an increasing part of everyday life and operations both in civilian and industrial applications because the vision of self-driving cars and self-operating machines is very near.

REFERENCES

Alami, R., Krishna, K.M. & Siméon, T. (2007). Provably Safe Motions Strategies for Mobile Robots in Dynamic Domains, in: Laugier, C. & Chatila, R. (ed.), Autonomous Navigation in Dynamic Environments, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 85-106.

Aldibaja M., Suganuma N., Yoneda K., Yanase R. & Kuramoto A. (2017). On autonomous driving: Why holistic and feature matching must be used in localization? 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), pp. 133-134.

Baudin, É, Blanquart, J., Guiochet, J. & Powell, D. (2007). Independent Safety Systems for Autonomy: State of the Art and Future Directions, LAAS-CNRS.

Bedord, Laurie (2017). How Automation Will Transform Farming, Agriculture.com, website. Available (accessed 23.2.2018): https://www.agriculture.com/technology/robotics/how-automation-will-transform-farming

Behere, S., Asplund, F., Söderberg, A. & Törngren, M. (2016). Architecture challenges for intelligent autonomous machines: An industrial perspective, Advances in Intelligent Systems and Computing, pp. 1669-1681.

Behere, S. & Liljeqvist, B. (2012). Towards Autonomous Architectures: An Automotive Perspective, KTH Royal Institute of Technology, Stockholm, 16 p. Available: http://kth.diva-portal.org/smash/get/diva2:567440/FULLTEXT01.pdf
http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-104803.

Benenson, R., Fraichard, T. & Parent, M. (2008). Achievable safety of driverless ground vehicles, 2008 10th International Conference on Control, Automation, Robotics and Vision, IEEE, pp. 515-521.

Brown, C. (2012). Autonomous vehicle technology in mining: how it works and how it's applied from user assist to full autonomy, E&MJ - Engineering & Mining Journal, Vol. 213(1), pp. 30.

Caterpillar (2018). Caterpillar to Develop Autonomous Mining Truck Technology for Additional Models and Brands. Available (accessed 30.1.2018): https://www.cat.com/en_US/news/machine-press-releases/caterpillar-to-develop-autonomous-mining-truck-technology-for-additional-models-and-brands.html

Design News (2008). Deere Takes Next Step Toward Driverless Tractor, Design News, website. Available (accessed 23.2.2018): https://www.designnews.com/automotive-0/deere-takes-next-step-toward-driverless-tractor/101595126750426?dfpLayout=article&dfpPParams=ind_184%2Cindustry_auto%2Caid_219094&doc_id=219094

Dhillon, B.S. (2010). Mine Safety: A Modern Approach (Springer Series in Reliability Engineering), 1. Aufl.; 1 ed. Springer Verlag London Limited, London.

European Road Transport Research Advisory Council (2015), Automated Driving Roadmap, European Road Transport Research Advisory Council ERTRAC.

Faralli, A., Giovannini, N., Nardi, S. & Pallottino, L. (2016). Indoor real-time localisation for multiple autonomous vehicles fusing vision, odometry and IMU data, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 288-297.

Fraichard, T. & Asama, H. (2003). Inevitable collision states. A step towards safer robots? Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453), IEEE, pp. 393 vol.1.

Han, M. (2008). DGPS for the Localisation of the Autonomous Mobile Robots, EKC2008 Proceedings of the EU-Korea Conference on Science and Technology (Springer Proceedings in Physics Series Volume 124), pp. 163-170.

Heiniö, M. (1999). Rock Excavation Handbook, Sandvik Tamrock Oy, Tampere.

ISO 17757 (2017). Earth-moving Machinery and Mining – Autonomous and Semi-autonomous Machine System Safety, International Organization for Standardization, Switzerland.

ISO/FDIS 18497: Agricultural machinery and tractors -- Safety of highly automated agricultural machines, International Organization for Standardization, website. Available (accessed 30.1.2018): https://www.iso.org/standard/62659.html

Jensen, S. (2015). The Growing Potential for Fully Autonomous Mines, OEM Off-Highway, website. Available (accessed 21.2.2018): https://www.oemoffhighway.com/electronics/smart-systems/automated-systems/article/12243110/autonomous-mining-equipment

Jha, S. & Raman, V. (2016). Automated synthesis of safe autonomous vehicle control under perception uncertainty, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 117-132.

Kaznov, V., Svahn, J., Roos, P., Asplund, F., Behere, S. & Törngren, M. (2017). Architecture and Safety for Autonomous Heavy Vehicles: ARCHER, in: Watzenig, D. & Horn, M. (ed.), Automated Driving: Safer and More Efficient Future Driving, Springer International Publishing, Cham, pp. 571-581.

Kempson W., Skinner T., Steenhof P. (2017). Autonomous Mining and International Standards, SMART Meeting, April 30th 2017, Smart Mining. Available (accessed 30.1.2018): www.smartmines.com/minutes/april2017/05-ISOAutonomy.pdf

Kumar, D. (2010). Emerging Tools and Techniques for Mine Safety and Disaster Management, in: Jha, M.K. (ed.), Natural and Anthropogenic Disasters: Vulnerability, Preparedness and Mitigation, Springer Netherlands, Dordrecht, pp. 332-365.

Larsson, J., Broxvall, M. & Saffiotti, A. (2006). A navigation system for automated loaders in underground mines, Springer Tracts in Advanced Robotics, pp. 129-140.

Laugier, C., Petti, S., Vasquez, D., Yguel, M., Fraichard, T. & Aycard, O. (2007). Steps Towards Safe Navigation in Open and Dynamic Environments, in: Laugier, C. & Chatila, R. (ed.), Autonomous Navigation in Dynamic Environments, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 45-82.

Leach, A. (2015). Drilling down on data: Sandvik AutoMine drilling system, Mining technology, website. Available (accessed 21.2.2018): https://www.mining-technology.com/features/featuredrilling-down-on-data-sandvik-automine-drilling-system-4517095/

Marshall, J., Barfoot, T. & Larsson, J. (2008). Autonomous underground tramming for center-articulated vehicles, Journal of Field Robotics, Vol. 25(6-7), pp. 400-421.

Marshall, J.A., Bonchis, A., Nebot, E. & Scheding, S. (2016). Robotics in Mining, in: Siciliano, B. & Khatib, O. (ed.), Springer Handbook of Robotics, Springer International Publishing, Cham, pp. 1549-1576.

Matsuzaki, H. & Lindemann, G. (2016). The autonomy-safety-paradox of service robotics in Europe and Japan: a comparative analysis, AI & SOCIETY, Vol. 31(4), pp. 501-517. Available (accessed ID: Matsuzaki2016): https://doi.org/10.1007/s00146-015-0630-7.

Molina, C.B.S.T., Almeida, J.R.Jr., Vismari, L.F., González, R.I.R., Naufal, J.K.Jr. & Camargo, J.B.Jr. (2017). Assuring Fully Autonomous Vehicles Safety by Design: The Autonomous Vehicle Control (AVC) Module Strategy, 2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W), IEEE, pp. 16-21.

Mäkelä, H. (2001). Overview of LHD navigation without artificial beacons, Robotics and Autonomous Systems, Vol. 36(1), pp. 21-35.

Mäkelä, H. & von Numers, T. (2001). Development of a navigation and control system for an autonomous outdoor vehicle in a steel plant, Control Engineering Practice, Vol. 9(5), pp. 573-583.

National Highway Safety Administration (2013). Preliminary Statement of Policy Concerning Automated Vehicles, National Highway Safety Administration.

Nebot, E.M. (2007). Surface Mining: Main Research Issues for Autonomous Operations, Robotics Research: Results of the 12th International Symposium ISRR (STAR: Springer Tracts in Advanced Robotics Series Volume 28), pp. 268-280.

Pushpavalli, M., Ramya, K. & Chandraleka, R. (2015). Autonomous Robots for Agriculture Field, i-Manager's Journal on Embedded Systems, Vol. 4(2), pp. 1-4.

Redmill, F. (1998). IEC 61508 Principles and use in the management of safety, Computing and Control Engineering Journal, Vol. 9(5), pp. 205-213.

Redmill, F. (2000). Understanding the use, misuse and abuse of safety integrity levels. Proceedings of the Eighth Safety-critical Systems Symposium, Redmill Consultancy.

SAE J3016 (2016). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, SAE International.

Sandvik Mining and Rock Technology (2018a). Sandvik LH517. Available (accessed 22.1.2018): https://www.rocktechnology.sandvik/en/products/underground-loaders-and-trucks/advanced-underground-lhds/lh517-underground-lhd/

Sandvik Mining and Rock Technology (2018b). Underground loaders and trucks. Available (accessed 30.1.2018): https://www.rocktechnology.sandvik/en/products/underground-loaders-and-trucks/

Santoni De Sio, F. (2017). Killing by Autonomous Vehicles and the Legal Doctrine of Necessity, Ethical Theory and Moral Practice: an international forum, Vol. 20(2), pp. 411-429.

Schreurs, M.A. & Steuwer, S.D. (2016). Autonomous Driving—Political, Legal, Social, and Sustainability Dimensions, in: Maurer, M., Gerdes, J.C., Lenz, B. & Winner, H. (ed.), Autonomous Driving: Technical, Legal and Social Aspects, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 149-171.

Seward, D., Pace, C. & Agate, R. (2007). Safe and effective navigation of autonomous robots in hazardous environments, Autonomous Robots, Vol. 22(3), pp. 223-242. Available (accessed ID: Seward2007): https://doi.org/10.1007/s10514-006-9721-0.

SFS-EN ISO 13849-1 (2008). Safety of machinery. Safety-related parts of control systems. Part 1: General principles for design. Finnish Standards Organisation SFS, Helsinki.

SFS IEC/TR 61508-0 (2012). Functional safety of electrical/electronic/programmable electronic safety-related systems. Part 0: Functional safety and IEC 61508, Finnish Standards Organisation SFS, Helsinki.

Toben, T., Eilers, S., Kuka, C., Schweigert, S., Winkelmann, H. & Ruehrup, S. (2012). Safe Autonomous Transport Vehicles in Heterogeneous Outdoor Environments, Verification, and Validation, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 61-75.

Torii, T. (2000). Research in autonomous agriculture vehicles in Japan, Computers and Electronics in Agriculture, Vol. 25(1), pp. 133-153.

The United States House of Representatives, (2017). SELF DRIVE Act, H.R.3388, The United States Congress. Available (accessed: 25.1.2018): https://www.congress.gov/bill/115th-congress/house-bill/3388

Wardziński, A. (2008). Safety Assurance Strategies for Autonomous Vehicles, Computer Safety, Reliability, and Security, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 277-290.

Wardziński, A. (2006). The Role of Situation Awareness in Assuring Safety of Autonomous Vehicles, Computer Safety, Reliability, and Security, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 205-218.

Watkins, C. (2017). On the Path to Automation, Canadian Mining Journal, Vol. 138(7), pp. 30.

Watkins, C.B. & Walter, R. (2007). Transitioning from federated avionics architectures to Integrated Modular Avionics, 2007 IEEE/AIAA 26th Digital Avionics Systems Conference, IEEE, pp. 10.

Zhang, X., Geimer, M., Noack, P.O. & Grandl, L. (2010). A semi-autonomous tractor in an intelligent master-slave vehicle system, Intelligent Service Robotics, Vol. 3(4), pp. 263-269.