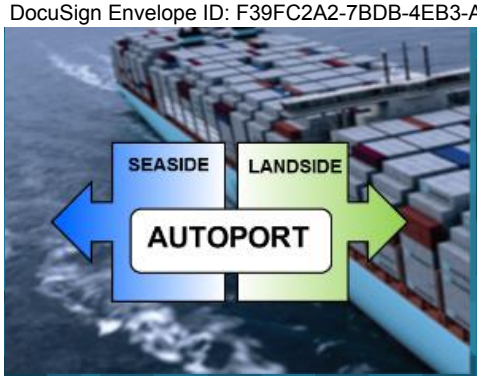
VTT Technical Research Centre of Finland

Safety engineering approaches and system analysis methods for autonomous mobile machinery

Tiusanen, Risto; Malm, Timo; Heikkilä, Eetu; Sarsama, Janne; Alanen, Jarmo; Ahonen, Toni

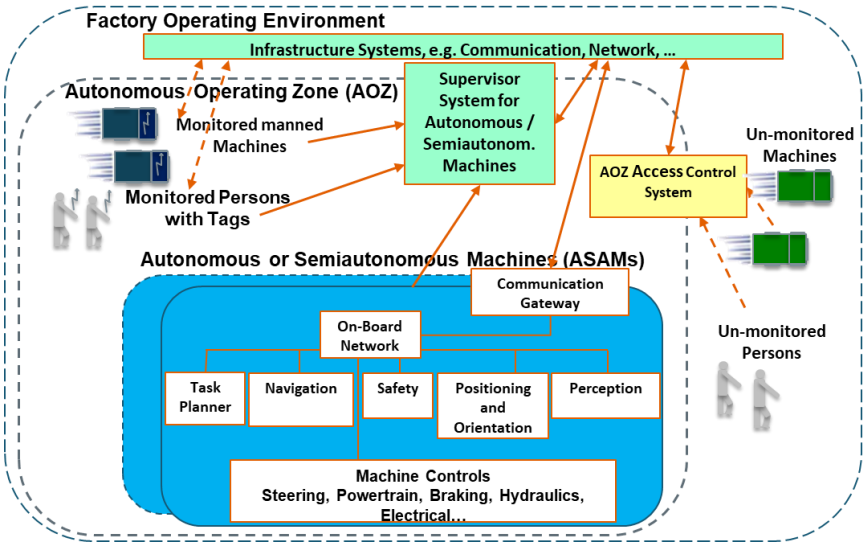
Published: 14/06/2021

Document Version
Publisher's final version



RESEARCH REPORT

VTT-R-00390-21



Safety engineering approaches and system analysis methods for autonomous mobile machinery

Authors: Risto Tiusanen, Timo Malm, Eetu Heikkilä, Janne Sarsama, Jarmo Alanen, Toni Ahonen,

Confidentiality: Public

Report's title Safety engineering approaches and system analysis methods for autonomous mobile machinery	
Customer, contact person, address	**Order reference**
Project name Operational excellence and novel business concepts for autonomous logistic systems in ports (AUTOPORT)	**Project number/Short name** 122480 / AUTOPORT
Author(s) Risto Tiusanen, Timo Malm, Eetu Heikkilä, Janne Sarsama, Jarmo Alanen, Toni Ahonen	**Pages** 31
Keywords Autonomy, mobile machine, safety engineering, methods	**Report identification code** VTT-R-00390-21

Summary

This report is conducted as a part of the Business Finland funded AUTOPORT - co-innovation project 'Operational excellence and novel business concepts for autonomous logistic systems in ports'. The main goal of AUTOPORT project is to pave the way towards business renewal and operational excellence by developing ecosystem level approaches for automated logistic systems. Research activities are focusing on digitalization and novel user-centric ecosystems in ports and terminals and advanced control systems and safety solutions in applications, which have not been automated before.

This report introduces and discusses advanced safety engineering approaches and system theoretic methods that can be used in this context to identify and assess new autonomy related safety risks in mobile machine applications.

The systematic and clearly phased safety engineering approach and risk assessment process supports the development of machine autonomy when selecting operating concepts and technological solutions. New methods and tools are needed when designing interactions between autonomous machines and human operators in safety critical decision-making situations. Advanced safety engineering procedures need to be in place to not only identify and control new safety risks, but also to document and communicate safety-related aspects between all relevant stakeholders.

Confidentiality	Public

Tampere 14.6.2021

Written by	**Reviewed by**
Risto Tiusanen Senior Scientist	Timo Malm Senior Scientist

VTT's contact address
Visiokatu 4, PL 1300, 33101 TAMPERE

Distribution (customer and VTT)

Project web page, PDF version
VTT Pure, PDF version
Risto Tiusanen, project manager, 1 copy
VTT archive, 1 copy

Approval

Date:	21.6.2021
Signature:	DocuSigned by: Nadezhda Gotcheva E21E683840FD424...
Name:	Nadezhda Gotcheva
Title:	Research Team Leader

Contents

List of abbreviations

ASC	Automated Stacking Cranes
AGV	Automated Guided Vehicle
ALARA	As Low as Reasonably Achievable
ALARP	As low as reasonably practicable (IEC 61508-5:2010)
ANSI	American National Standards Institute
AOZ	Autonomous operating zone
ASAMS	Autonomous or semi-autonomous machine system (ISO 17757:2019)
ASC	Automatic Stacking Crane
CAST	Causal Analysis using System Theory
CBA	Cost Benefit Analysis
CCPS	Center for Chemical Process Safety
CHE	Container Handling Equipment
DF	Disproportion factor
ECS	Equipment Control System
EN	A unique reference code for European Standards which have been adopted by one of the three recognized European Standardization Organizations (ESOs): CEN, CENELEC or ETSI
EMM	Earth-moving machinery
FDIS	Final Draft International Standard (ISO)
FMEA	Failure modes and effects analysis (IEC 60812:2018)
FMECA	Failure mode, effects, and criticality analysis
FRACAS	A failure reporting, analysis, and corrective action system
FRAM	Functional Resonance Analysis Method
GAME	Globalement Au Moins Equivalent (Globally at least as good)
GSN	Goal Structuring Notation
HAAM	Highly Automated Agricultural Machinery
HAZOP	Hazard and operability study (IEC 61882:2016)
HFACS	Human Factors Analysis and Classification System
HSE	Health and Safety Executive
HSWA	Health and Safety at Work Act (New Zealand)
IEC	International electro technical commission
INCOSE	International Council on Systems Engineering
IPL	Independent protection layer
ISO	International organization for standardization
LOPA	Layer of protection analysis
MEM	Minimum Endogenous Mortality

OHA	Operating Hazard Analysis
PFD	Probability of Failure on Demand
PFH	Probability of Failure per Hour
PHA	Preliminary Hazard Analysis
PL	Performance Level
PLC	Programmable Logic Controller
QRA	Quantitative risk analysis
RAMS	Reliability, Availability, Maintainability and Safety
RAMSS	Reliability, Availability, Maintainability, Safety and Security
RMG	Rail Mounted Gantry Crane
SFAIRP	So far as is reasonably practicable
SIL	Safety Integrity Level
STAMP	System-Theoretic Accident Model and Processes
STPA	System-Theoretic Process Analysis
TLS	Terminal Logistic System
TOS	Terminal Operating System
UCA	Unsafe Control Actions
UCSA	Use Case Safety Analysis
UL	Underwriters Laboratories
WHS	Work Health and Safety (WHS Act and regulations, Australia)

1 Introduction

1.1 Background

Machinery suppliers and their ecosystem partners need to improve their capabilities in co-creating novel solutions in order to accelerate the implementation of automated systems and new service solutions for the global customers. The seamless automated operation in the whole logistic chain in ports necessitates a change also in the ship-to-shore operations. In port operations, the new solutions based on open operating environments require system level approaches, information exchange, technological innovations and wide collaboration at an ecosystem level.

This far, higher-level automation has been implemented mostly in the container handling systems of large ports. However, the level of automation could increase in all sort of container handling systems if automated solutions could provide lower lifecycle costs than the manual options. Adoption of automation technologies at smaller ports is still a challenge from the cost-benefit perspective. Re-design is needed for the automation solutions to make them fit for smaller ports. Major share of the new automation installations take place as brownfield investments since investments to greenfield harbours are seldom. When existing installations are automated, special service offering is required to design and plan the replacement in such a way that costs and efficiency are optimized. There is also an increasing demand for thinking the re-use of the previous assets. Processes for installations need to be developed also to minimize interference with the terminal operations. [1]

The main container logistic processes are quite the same in port terminals. Container terminals can be described as open systems of material flow with two external interfaces. These interfaces are the quayside with loading and unloading of ships, and the landside where containers are loaded and unloaded on/off trucks and trains. Containers are stored in stacks thus facilitating the decoupling of quayside and landside operation. [2]

After the containers are unloaded from a ship they are transported to yard positions near the place where they will be transhipped next. Containers arriving by road or railway at the terminal are handled within the truck and train operation areas. They are handled by specific container handling equipment and moved to the respective stocks in the yard. Additional moves are performed if sheds and/or empty depots exist within a terminal; these moves encompass the transports between empty stock, packing center, and import and export container stocks (See Figure 1 and Figure 2). [2]

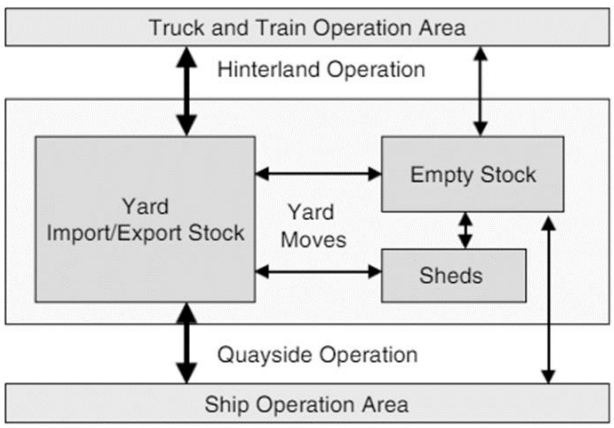


Figure 1 A simplified flowchart of harbour operations [2]

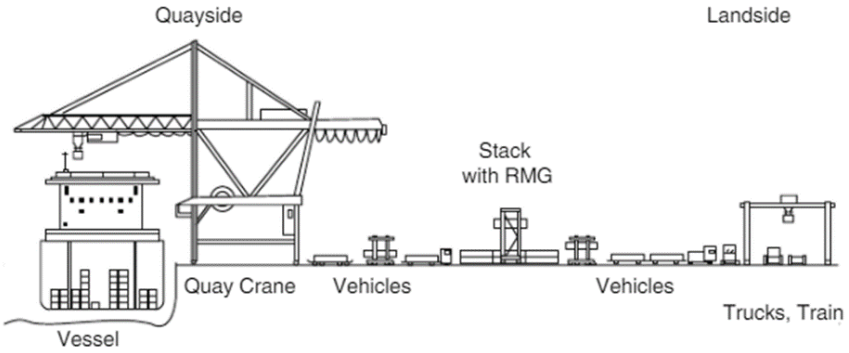


Fig. 6. Container terminal system (schematic side view, not true to size)

Figure 2 Schematic side view of a container terminal system [2]

The unmanned container handling machine or vehicle can be completely controlled by a computer or by using a combination of robotic and remotely operated work phases in sequence. This development follows a similar pattern seen earlier in warehouse automation; the main difference being that the technology required for outdoor conditions has proven to be vastly more demanding. The financial drivers of this development are related to efficiency and economics. A robotized work sequence is more predictable, without human errors. Remotely operated container handling machines and vehicles also make it possible for one operator to control and supervise a large volume of equipment. In extreme cases, 100% of the work cycle has been robotized and the role of the operator is to supervise and handle exceptional situations. [3]

Terminal Operating System (TOS) software controls the logistics of a terminal, including key functions such as vessel planning, container inventory maintenance, job order creation and gate operations. TOS software is provided by several commercial companies and many terminal operators themselves. In a modern container terminal, some Container Handling Equipment (CHE) may be unmanned and operated by a computer and navigation system while most of them are manually operated. There is just a little differences between these modes of operation from TOS point of view. Differences usually occur in situations where drivers improvise container moves as computers don't and this can cause problems with the overall logistic operation. To enable efficient exception handling, the software should be able to handle most common exceptions automatically. [3]

A group of automated vehicles may share a common software control module at equipment level, often referred to as the "Equipment Control System" (ECS). ECS is handling for example safety features and coordination between the vehicle and the terminal infrastructures. Typically, automated vehicles operating on the same tracks or pathways, such as in Automatic Stacking Crane (ASC) applications, where Rail Mounted Gantry Cranes (RMG) are coordinated by such software. ECS is defined here as the software that monitors and controls all events and processes at equipment level, either for a single CHE or group of CHEs. When it comes to coordinating interactions between different types of automated equipment, an ECS is now an essential part of the terminal software landscape. Driverless operation also requires some dedicated software to implement all the actions and decisions previously executed by a driver (such as navigation, traffic rules and deadlock resolution). This is another motivation for a layer of additional functionality (ECS) between the CHE on-board control software system (e.g. PLC) and the TOS software. [3]

Figure 3 shows an example of Kalmar's solution for container terminal automation and its layers which follow the above-mentioned general system hierarchy.

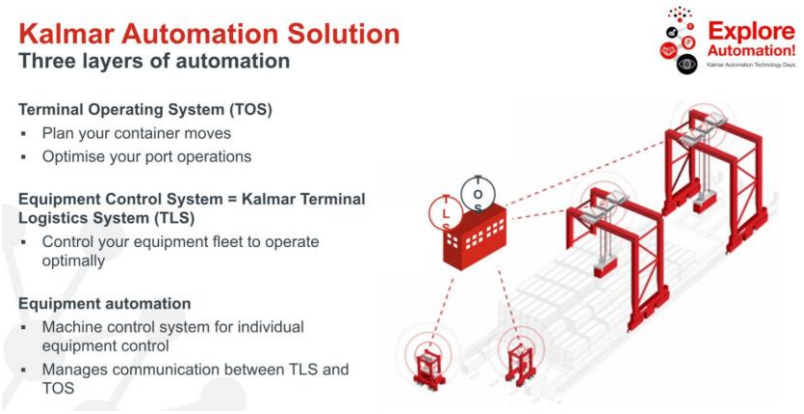


Figure 3 Schematic drawing of the layers of terminal automation systems (Kalmar) [4]

1.2 Needs for new safety solutions

While in large ports and new greenfield ports full automation may be designed right from the beginning, in smaller ports and in existing brownfields the automation investments need to be carefully focused. Small volumes without standard cargo and machine fleet with multi-function work cycles make small and medium sized ports difficult areas for full automation. The most important factors in the development of the solutions for small and medium-sized ports are the scalability and cost-efficiency of the proposed solutions. One interesting and challenging vision of advanced operating concepts in logistic systems in small and medium-sized ports is to enable automated machinery, manual machines and manual workers to operate and collaborate in the same open work area. This is called a ‘Mixed traffic’ or ‘Mixed mode’ operation.

Different operating environments and work processes require different solutions to ensure safe operation of the autonomous system. Safety critical systems are based increasingly on software solutions and safety functions use information from several interacting systems [5]. From safety perspective, full automation is a desired option when compared with mixed traffic where safety regulations, system safety requirements and costly situational awareness systems bring challenges. However, for small and medium-sized ports, mixed traffic solutions are to be emphasized [1].

There are many challenging safety aspects on the way towards more automated and even autonomous mobile machinery. Each step towards full autonomy introduces new safety risks compared to lower levels of autonomy and manual operation. New approaches and methods are needed to identify, assess and manage autonomy related safety risks from system concepts to specific procedures and functions, and to design feasible and acceptable safety solutions. Product orientation should be changed towards system thinking also in safety engineering. According to the general systems engineering approach, safety engineering should be a continuous top-down process in all the system development phases from concept evaluation to detailed design of the safety critical operations and functions [6]. The safety engineering methods should be selected not only to support the objectives of each system development phase but also to support the overall risk assessment process, traceability, and reuse of the analysis and assessment results [7].

Current safety engineering methods developed for automated machinery do not cover or consider real autonomy aspects [8], [9]. Machine’s autonomous behaviour cannot be fully predetermined, because the key element in machine autonomy is adaptability to dynamically changing environment based on the perception of the available information [10]. New methods and tools are needed to identify and assess machine autonomy related safety risks and specific aspects raising from the increased autonomy in safety critical decision-making. New methods are also needed to model and

manage safety requirements from concept level to safety functions and to design feasible and acceptable safety solutions [11], [12].

The first safety standards have recently been published on autonomous mobile machinery. ISO 17757:2019 defines risk assessment process and general safety requirements for autonomous or semi-autonomous earth moving machinery and mining machinery systems [13]. The key point in this standard is that the requirements should be defined from a system-level perspective taking into account site-specific operating conditions and risks. The standard ISO 18497:2018 defines requirements for agricultural machinery and tractors in respect to their automated operation and automatic functions [14]. ISO 3691-4:2020 standard, in turn, defines the requirements for unmanned forklifts, AGVs and associated systems [15].

According to the systems engineering approach, risk management decisions in the system-development phase are made systematically as the system development proceeds. In practice, the decisions to reduce the safety and availability risks are based on comparison of alternative solutions at different layers of protection and prediction. One promising approach for the management and documentation of safety requirements in autonomous mobile machinery is the Goal-based design approach that has successfully been applied and demonstrated in safety qualification e.g. in marine sector [16], [17] and it is also a main element in the new Safety case approach that has been developed for autonomous products and systems [18].

1.3 AUTOPORT project

Research and development work on the system safety-engineering methods and safety requirement management in VTT in Finland has been done among others in a national co-innovation project ‘AUTOPORT’ financed partly by Business Finland, VTT and participating companies (https://autoport.fi/). The AUTOPORT project consortium consists of companies Atostek, Exertus, Intopalo Digital, Kalmar, Solita and Huld and two research organizations Tampere University and VTT. The research work in the AUTOPORT project focuses on autonomous mobile machinery for cargo and container handling in small and medium size terminals.

The main goal of AUTOPORT is to pave the way towards business renewal and operational excellence by developing ecosystem level approaches for logistic robot systems. Research activities are focusing on digitalization, novel user-centric ecosystems and advanced control, and safety solutions in applications, which have not been automated before. The objective is to develop automated operations in terminals and factories by developing cost effective, reliable, safe and secure solutions with novel system engineering methods and digital services for complex systems in emerging open access operational environments. The project scope is limited in logistics chain of internal cargo transport in terminals and factories.

During the preparation of AUTOPORT ecosystem project the following four important enabling factors were identified:

- ✓ Adaptable software platforms and modular control system structures
- ✓ Systematic procedures and tools for design and validation
- ✓ Stepwise automation towards lifetime business
- ✓ Extending the service business to overall machine fleets

The AUTOPORT project addresses the main goal by the following four research themes that reflect the company-specific needs and interests as stated by the participating companies:

- Model-based design flow
- Operational excellence by novel models of information sharing
- Cost-effective logistic robot technologies
- Ecosystem models for shared benefit

The aim of the safety research in AUTOPORT project is to study and implement system-driven safety engineering methods and tools to support requirement specification, design, analysis, and verification and validation activities especially in the early design phases of autonomous mobile machinery.

The target case system in the AUTOPORT project for experimenting new risk analysis and assessment methods is a container terminal application where several (a half dozen) automated machines transport containers at the certain routes and handle them in specific loading and unloading places (). The operation of the case system is based on the "Mixed traffic" operating concept where manual trucks, work machines as well as pedestrians also share the same routes.

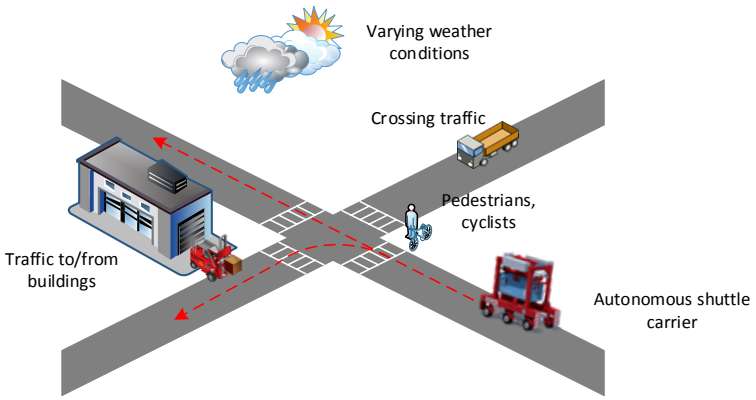


Figure 4 An outline of the AUTOPORT case system

1.4 Objective and content of this report

The objective of this report is to introduce and discuss safety engineering approaches and system-driven methods for the analysis and assessment of logistic systems utilising autonomous mobile machinery. Chapter 1 first describes the background of container handling in ports and needs for autonomous mobile machinery for container handling. Chapter 2 introduces a safety engineering approach for autonomous mobile machinery defined in ISO 17757: 2019 standard and the database centric safety engineering approach for safety critical machine control systems developed in VTT. Chapter 3 introduces system-driven methods like STAMP/STPA, ALARP, and LOPA for hazard identification, analysis and risk evaluation. Chapter 4 introduces goal-based approach for safety requirements management and an ANSI/UL 4600 based safety case approach for autonomous mobile machinery. Chapter 5 summarizes the review results and discusses the safety engineering approach for autonomous mobile machinery based on the experience gained in the AUTOPORT project.

2 Safety engineering approaches for autonomous mobile machinery

Different industrial sectors have different safety strategies and safety engineering approaches and there is also a big difference in safety strategies in industrial environments compared with e.g. passenger cars and public transport. Machine safety engineering approaches have currently rather narrow view to autonomy aspects [9].

Three interesting standardised safety engineering approaches for autonomous mobile machinery already exist and were studied in the project. Firstly, ISO 3691-4:2020 [15] defines requirements for the operation of driverless forklifts in different industrial operating areas and the requirements for on-board safety related functions. The standard requires that any access to the automated area must be controlled. It also defines speed limits for specific operating conditions and functional safety requirements for safety functions. Secondly, ISO 17757:2019 [13] gives guidelines for earth-moving and mining machine sector, how the safety risks should be assessed and how the system safety requirements should be defined in autonomous mobile machine applications. The approach emphasizes the risks related to the actual operating concepts and actual operating environment at the site and the uncertainties related to the safety related functions and technologies. The standard introduces the concept of an autonomous operating zone (AOZ), controlled by the access control system, where monitored manned machines and monitored persons could work at the same time with autonomous machines (See Figure 5). Thirdly, ISO 18497:2018 [14] defines the requirements for the deployment and implementation, monitoring and remote monitoring of highly automated agricultural machinery (HAAM) and their safety systems. The standard specifies requirements for starting, movements and tool movements of a HAAM. It also sets test methods for human detection systems.

In this chapter the safety engineering approach for autonomous mobile machinery described in ISO 17757:2019 [13] and the database centric safety engineering approach for safety critical machine control systems (KOTOTU process and tool) developed in VTT [19] are shortly reviewed and discussed.

2.1 System view in ISO 17757:2019

ISO 17757 standard provides safety requirements for autonomous and semi-autonomous machines (ASAM) used in earth-moving and mining operations, and the related autonomous or semi-autonomous machine systems (ASAMS). It specifies safety criteria both for the machines and their associated systems and infrastructure, including hardware and software, and provides guidance on safe use in their defined functional environments during the machine and system life cycle. It also defines terms and definitions related to ASAMS. [13]

The standard is basically applicable to autonomous and semi-autonomous earth-moving machinery (EMM) defined in ISO 6165 and to mobile mining machines used in either surface or underground applications. However, its principles and many of its provisions can be applied to other types of autonomous or semi-autonomous machines used on worksites in various industrial sectors.

Integration of ASAMS into the site planning process is important. ASAMS are complex systems, because of the complexity of the logistic processes themselves, their relation to people, manned operations and the layers of safety that need to be built into them. Supporting infrastructure and operating area requirements should be identified early in the project, as automation systems can have specific needs (e.g. fuelling facilities, control rooms, communications network).

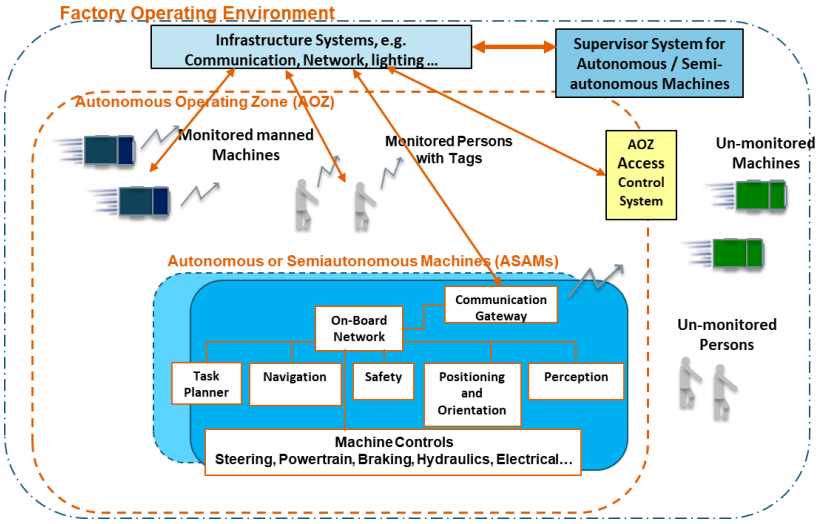


Figure 5 An example of the main system elements in an ASAMS according to [13].

Based on the risk assessment, the ASAM shall be capable of adapting to environmental conditions as long as any changes in the environmental conditions are within identified constraints. These may include the use of human operators or automated systems to make speed adjustments, disable operations, close off areas or other adjustments as needed to maintain safe operation.

It is emphasized that an ASAM can introduce hazardous situations not normally encountered on a conventional manned worksite. The standard points out that the effective management of the risks associated with operating an ASAMS requires input from the system integrator, system operator and site manager, and potentially from diverse operational groups, ranging from system integrator, researchers, design engineers, project managers, team leaders and control room operators to safety and health representatives and other workers involved in the tasks, as well as emergency response personnel [13].

The risk assessment process for an ASAMS shall be completed according to the principles of ISO 12100 [20]. Risk control options and safety measures should be evaluated according to the risk assessment results by identifying uncertainties, failure modes and risks and assessing what level of measures should be taken to be prepared for risks from the perspective of the whole system. The functional characteristics of the safety related systems and the required safety integrity levels should be based on the actual risks in the specific operating environment. Safety-related parts of control systems shall comply with the appropriate functional safety performance level (ISO 13849-1:2015 [21], ISO 19014-1:2018 [22], IEC 62061:2021 [23] or IEC 61508-1:2010 [24].

Recommended methods for hazard identification are among others HAZOP, LOPA and FMEA and workplace inspections. At the risk analysis stage, the following three viewpoints should be considered:

- the operational environment (scale, complexity and physical environment of operations and activities),
- the operational processes (work processes, maintenance procedures, internal and external interactions)
- the autonomous machine systems (functionality and performance, safety and resilience features)

Risk evaluation and risk management is advised to be accomplished by applying a hierarchy of risk controls and safety measures:

- Primary controls aiming to avoid, remove or change the risk type
- Contingency controls aiming to minimize the effects in case of an incident following the LOPA principle
- Prevention and management controls aiming e.g. to minimize interactions with ASAM fleet and to develop safe work procedures

The standard also emphasizes the importance of human-technology interaction and consideration of human factors to ensure overall safety, which fits well with the general system safety principles that have been applied for years in complex safety critical applications in aviation, transportation and energy sectors. It would be useful if the standard could give more guidance how the intelligence required by autonomous operation should be divided between the "On-board" system and the "Supervisor" system. In addition, more guidance for the design of interactions between the "Access control system" - "Mission planner" - "Supervisor system" could be helpful.

ISO 17757:2019 [13] does not give any fixed SIL or PL requirements for any subsystem or safety related function. It guides system designers to define safety requirements according to the application-specific risk assessment results. This systems engineering approach developed for earth-moving and mining applications differs from the approach that has been developed for driverless trucks and AGV type of machinery in ISO 3691-4:2020 [15].

2.2 Database centric safety engineering approach

A safety engineering process model for the design of machine control systems has been developed in VTT since 2012 to support the design of automated machinery applications [19]. The so called KOTOTU reference model follows the appropriate safety standards as ISO 12100:2010 [20] for risk assessment and ISO 13849:2015 [14] standard family for the design of safety related control systems. The safety engineering process model is closely connected to the development process of a machine control system so that safety related tasks will be carried out systematically and in the right phase according to the systems engineering process.

Parallel to the safety engineering process model development a safety demonstration data model has been developed in VTT [25]. The systematic process and data model aims to help to identify gaps in fulfilling the process and product safety requirement [26].

2.2.1 KOTOTU reference model

The KOTOTU reference model has been demonstrated using web browser techniques (Figure 6). It includes process stages from Preliminary Hazard Analysis (PHA) (Figure 7) to the system safety validation process. The name KOTOTU comes from the Finnish name '*Koneiden ohjausjärjestelmien Toiminnallinen Turvallisuus*' (Functional Safety of Control Systems of Machinery). The other process stages included in the tool are Use Case Safety Analysis (UCSA) (Figure 8) (which is an application of Operating Hazard Analysis, OHA), Function and Communications Analysis and Performance Level (PL) evaluation according to ISO 13849-1 [27].

From the web-based user interface it is possible to see the safety engineering tasks related to different process stages. All source documents that are necessary for carrying out the safety engineering tasks can be found under each task. Result document templates and references to relevant safety standards are also available from the user interface. One special feature in the tool is that the reference safety standards can be opened from the particular page where the safety requirements relating to the corresponding tasks or design phase are given. [27]

From the KOTOTU process diagram, which works as a graphical user interface, the tools and instructions relating to the safety engineering tasks can be opened. In addition, all the documents relating to the safety process can be opened from the process diagram of the interface for viewing

and editing. For example, Preliminary Hazard Analysis and other worksheets can be opened from the process diagram of the interface. Through the web-based user interface it is easy and illustrative to show how the requirements accordant with the relevant safety standards have been followed within the safety process.

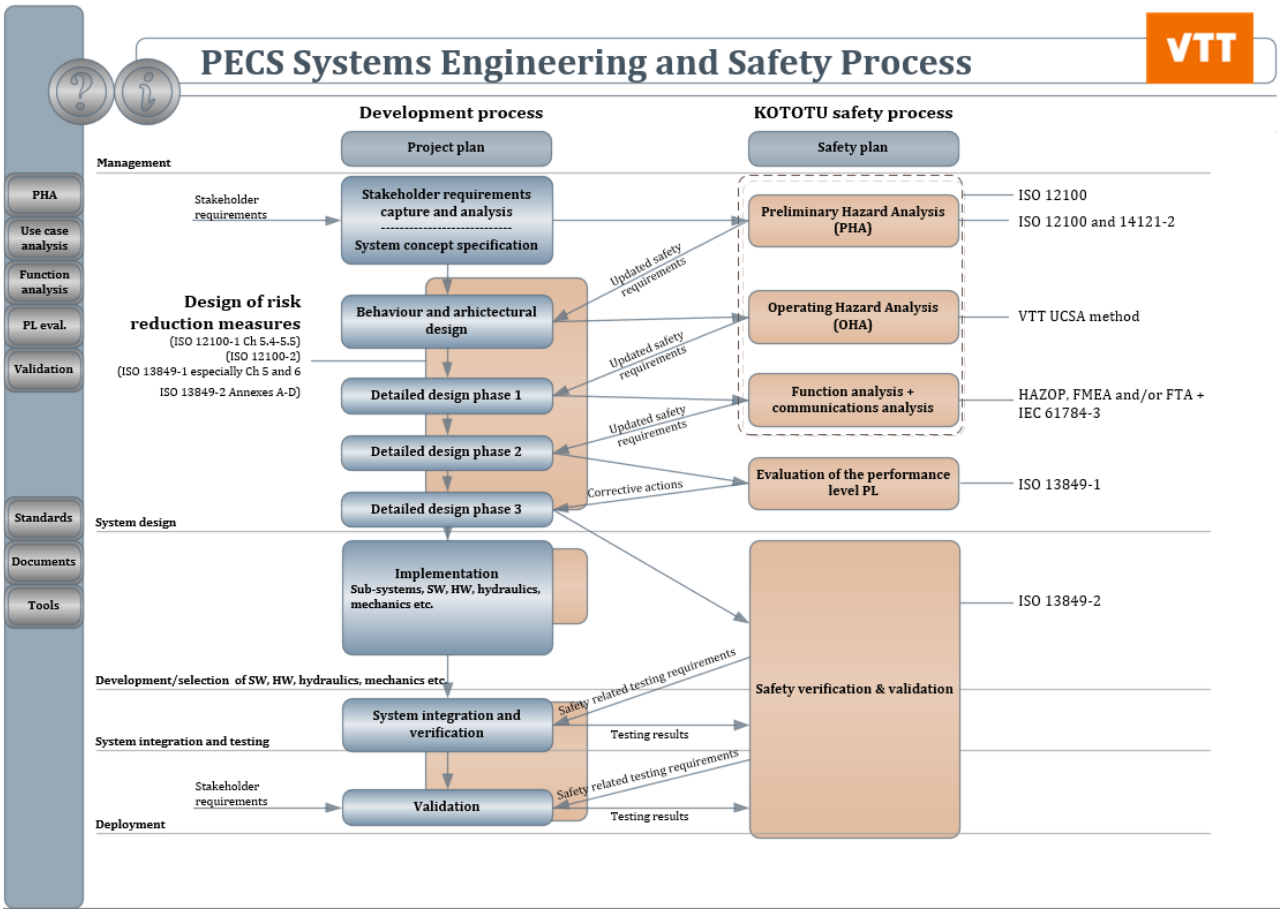


Figure 6 The main page of the KOTOTU tool interface.

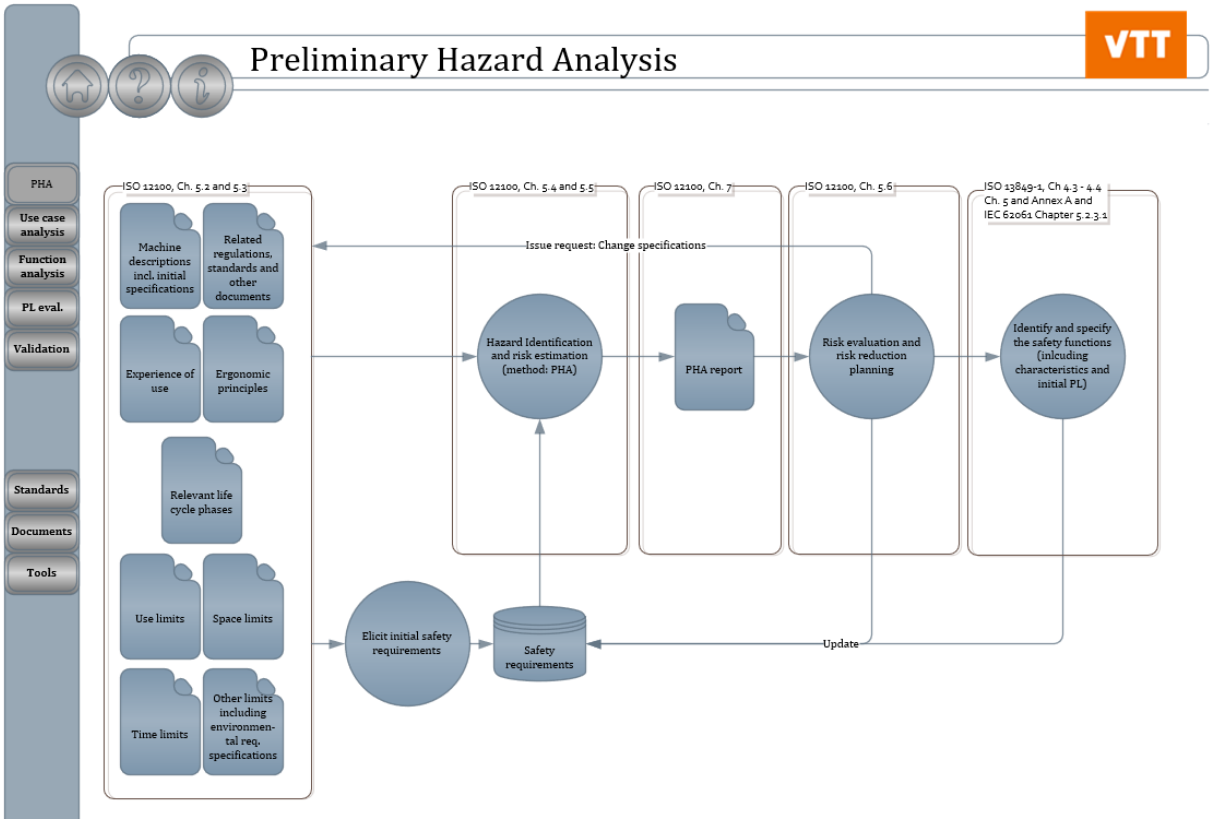


Figure 7 User interface for the preliminary hazard analysis process.

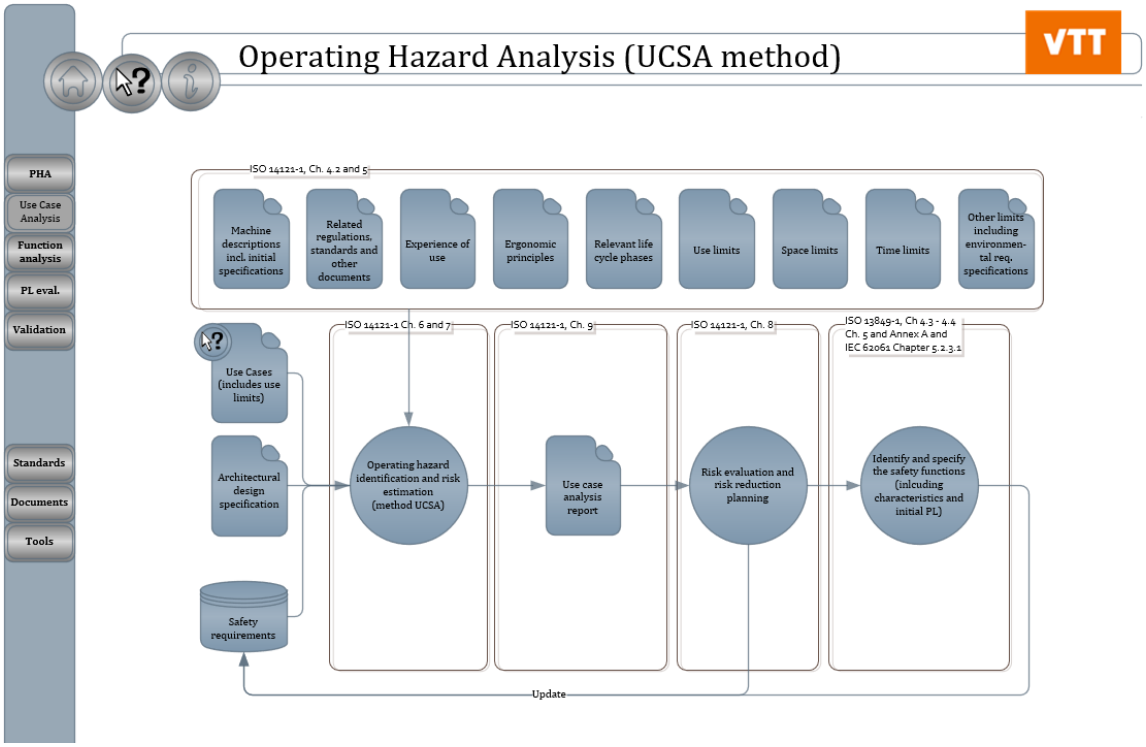


Figure 8 User interface for Use Case Safety Analysis (UCSA) process.

2.2.2 A safety demonstration data model

Applying model-based engineering can help assuring the safety of complex systems, such as safety critical systems in power plants, due to the fact that the engineering and conformity assessment effort can be managed more systematically. The systematic process and data models support engineering efforts aiming to fulfil the process safety and product safety requirement. Identification of gaps in safety engineering processes can be arranged to be performed by a software tool, for example by identifying requirements that do not have a corresponding claim, by identifying arguments without evidence and by identifying determination results that are not used as evidence (perhaps to hide negative test results) [25].

Model-based systems engineering provides a structured set of artefacts, not only for the design work, but also for the conformity assessment. VTT has developed a structured data model for the conformity assessment artefacts. The data model covers both the first party conformity assessment, traditionally known as verification and validation, as well as the third party conformity assessment, i.e. the attestation (qualification or certification). The data model is demonstrated in a Defence-in-Depth example of a spent fuel cooling control system [28].

In terms of the IDEF0 metamodel for systems engineering processes, the data model for conformity assessment artefacts has been defined for the parts of the metamodel, as depicted in Figure 9 [25].

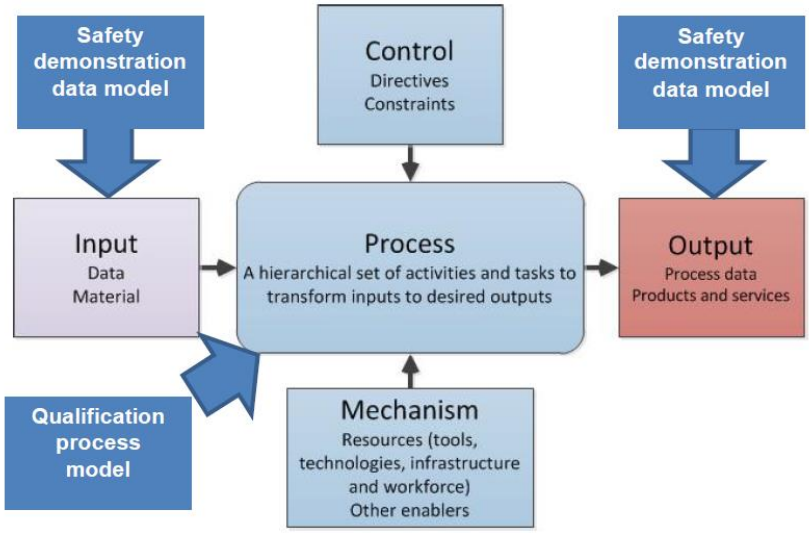


Figure 9 IDEF0 metamodel for systems engineering processes, modified from [6]

Experiences from real case studies in previous projects have shown that the data model provides good traceability of artefacts ranging from the stakeholder domain to the organisation that designs the system and finally to the attestation organisation. It has been noted some important aspects that affect whether users (safety engineers and system designers) accept the new data based centric engineering tool. It is important that the complexity of the data model does not affect the user experience and the tool guides the users to put the engineering artefacts to correct locations and to establish the traces between the artefacts [28].

The database-centric approach has been further developed in VTT in AUTOPORT project to support all RAMS (Reliability, Availability, Maintainability and Safety) related activities in systems engineering by extending the conformity assessment data model with dependability requirement taxonomy. Additionally, a common ontology for safety, security and dependability risk assessment has been defined to cover an entire RAMSS sector. ‘RAMSS’ acronym comes from ‘Reliability, Availability, Maintainability, Safety and Security’. The results of this study are reported in the project internal report ‘Database-centric RAMSS design approach for autonomous logistic systems’ [29]. In the

report, detailed data models are presented describing FRACAS loop, as well as the STPA hazard analysis method as examples of different approaches that are needed to manage RAMSS aspects over a system life cycle [29].

3 System driven methods

Traditionally, a risk-based approach has been dominantly applied in safety and reliability engineering. There are several ways to define the concept of risk, but usually it has been seen as a combination of severity and probability of an undesired event [30], [31]. The risk-based approach, especially with high focus on quantification of the risks, has seen increasing criticism during the previous years, partly due to the following factors [32]:

- Systems are becoming increasingly complex. This increases the number of interactions between system elements, leading to emergent and nonlinear behaviour that is highly unpredictable. This makes quantifying risks increasingly difficult or impossible.
- Chain-of-event or chain-of-failure models have been argued to be insufficient in capturing accidents that may arise in complex systems.

Due to the issues in solely risk-based approach, arguments have been presented against the use of traditional safety analysis methods, which often focus on analysis of component failures or linear chains of events, rather than identifying problems arising from unsafe interactions between system elements [32]. To address the above issues, systems-theoretic approaches, focusing on safety control instead of risk, have been proposed as a potential basis for performing more comprehensive safety analyses.

Systems theory is a set of principles that can be applied to comprehend complex systems and their behaviour. Based on systems theory, safety hazard analysis approaches, as well as accident analysis methods have been proposed. In literature, comprehensive comparisons of these methods have been performed mostly in terms of accident analysis although the methods can be applied in a wider usage [33], [34].

Accident analysis methods such as Swiss cheese model, AcciMap, STAMP (Systems-Theoretic Accident Model and Processes) and STPA (System Theoretic Process Analysis) are classified as systems-based accident analysis methods. Important in these approaches is that the socio-technical aspects are taken into account during the analysis. These methods are not domain-specific in accident analysis for a particular industry. They have been applied in different industries such as aviation, defence, food, public health, oil and gas, and rail transport. [33]

Abulamddi [33] has reviewed accident analysis and risk assessment methods and categorised them into four categories based on the aspects they are focusing: Technical, Human factors, Organisational and Systemic. The Figure 10 illustrates the development of the accident analysis and risk analysis methods over the years and shows also in which aspects the methods are focusing. Methods were first developed for analysis of technical failures. Human factors and organisational aspects came up strongly as systems became more complicated in the 1980s and 1990s. The digital development of industrial systems, which began in the 2000s, has increased the complexity of systems significantly. This development is also reflected in the development of risk analysis methods towards systemic analysis methods.

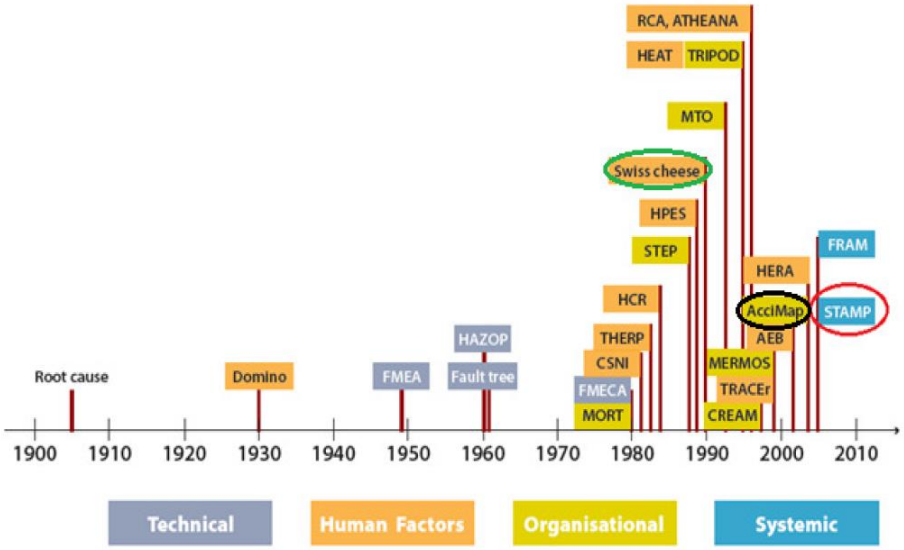


Figure 10 Historical development and categorisation of accident analysis and risk analysis methods. [33]

Based on the literature, the most prominent approaches for system driven methods for complex socio-technical systems seem to be the following methods [33], [34]:

- ***AcciMap*** method, originally developed by [35], provides a structured, graphical representation of a causal scenario. Although designed as a part of an active risk management approach, it has mostly been applied for accident analysis purposes in safety-critical domains. Thus, its applications in product development context are largely unexplored and its further study is omitted in this report.

- ***FRAM*** *(Functional Resonance Analysis Method)*, developed 2012 by Hollnagel [36], is based on identification and analysis of system functions. It provides a graphical language for modelling the system functions, and focuses on providing insights on how the functions interact. This graphical representation can be useful in understanding various characteristics of complex systems.

- ***STAMP*** *(Systems-Theoretic Accident Model and Processes)* is developed in the MIT by Leveson and her colleagues [32]. In STAMP a system is described by using a hierarchical control structure. The STAMP approach is accompanied with the *STPA (Systems-Theoretic Process Analysis)* hazard analysis method, which aims to identify flaws within the safety related controls and control actions [37]. STAMP is also affiliated with another method CAST (*Causal Analysis using System Theory*) specifically intended for accident analysis [38].

As relatively recent methods, only limited experiences of the use of system-theoretic approaches in industrial systems is available. The approaches have also seen criticism especially from the perspective of their limited contribution to risk comparison and prioritization activities. For example, according to Ramos et al. [39], these approaches, while providing useful qualitative analysis, are still very limited in unravelling complex failure modes and mechanisms in addition to being qualitative and of limited value in prioritizing risks and risk reducing measures. Bjerga et al. [40] provide further criticism, stating that systemic models are based on a number of assumptions that are impossible to verify especially in novel technology development endeavours.

In practice, it is unlikely that risk-based approaches could be entirely replaced by control-based approaches, or that such development would even be beneficial. Instead, the different approaches can complement each other to understand new types of emergent issues in complex systems. In the

following chapter, we focus on the STAMP approach and especially the associated STPA hazard analysis method, and their applications on logistic robotic applications.

3.1 STAMP approach and STPA method

STAMP (System-Theoretic Accident Model and Processes) is the name of the new accident causality model based on systems theory. It expands the traditional model of causality beyond a chain of directly-related failure events or component failures to include more complex processes and unsafe interactions among system components, and it underlies STPA (System Theoretic Process Analysis) and other tools. [37]

STAMP is not an analysis method. Instead it is a model or set of assumptions about how accidents occur. STAMP is an alternative to the chain-of-failure-events (or dominos or Swiss cheese slices, all of which are essentially equivalent) that underlies the traditional safety analysis techniques (such as Fault Tree Analysis, Event Tree Analysis, HAZOP, FMECA, and HFACS). According to [37] STAMP extends the chain-of-failure causality model to include both component failure and unsafe interactions of system components. In STAMP, safety is understood as an emergent system property that arises when the components of a complex system interact with each other within a larger environment. A set of constraints (requirements) related to the behaviour of the system components (including physical, software, human, and social aspects) enforces the safety property of a complex system. Safety constraints on system operation are for example [5]:

- Aircraft or automobiles must never violate minimum separation standards,
- Medical devices must not provide a harmful level of medicine,
- Toxic chemicals or radiation must not be released from a plant,
- Batteries must never experience thermal runaway,
- Aircraft must have sufficient lift to remain airborne unless landing,
- Nuclear materials must never get into the wrong hands.

Leveson argues that accidents occur when the system component interactions violate these constraints [5]. The goal, then, is to control the behaviour of the components and system as a whole to ensure that the safety constraints are enforced in the operating system. Instead of focusing only on preventing accidents by increasing component reliability and treating safety as a component failure problem, Leveson expresses that in STAMP approach safety aspects are treated as a dynamic control problems that enforces the safety constraints where both system component failures and system component interactions must be controlled [5].

According to [37] STAMP methodology has certain advantages like:

- It works on very complex systems because it works top-down rather than bottom up.
- It includes software, humans, organizations, safety culture, etc. as causal factors in accidents and other types of losses without having to treat them differently or separately.
- It allows creating tools, such as STPA, CAST (Causal Analysis based on System Theory) [38], identification and management of leading indicators of increasing risk, organizational risk analysis, etc.

STPA process has been described in the freely available STPA handbook, which has seen several updates and revisions over years. The latest version published by Leveson & Thomas is from year 2018 [37]. The handbook also provides definitions for the key terms, which are used in the following,

but may differ from terminology applied elsewhere in this report. At the time of writing, STPA consists of the following four defined steps:

1. **Definition of the analysis purpose**. As with all analyses, the purpose of the analysis needs to be defined first. In STPA, this begins with definition of system-level losses that the analysis aims to prevent. Then, system-level hazards and constraints are defined.

2. **Modelling the control structure**. The system is modelled as a *hierarchical control structure*, which is a system model composed of feedback control loops. This is a graphical representation featuring controllers and controlled processes represented as rectangles, and the interactions between them (control and feedback) represented as arrows. The hierarchy is illustrated by the vertical axis, i.e. highest control authority is at the top of the diagram. It should be noted that the control structure only models this control hierarchy and is not, for example, a physical model or a simulation model of the system behaviour. However, other available system models can and should be utilized when drawing the control structure.

3. **Identification of unsafe control actions (UCAs)**. In this step, the control structure is systematically analysed to find control actions that, in a particular context and worst-case environment, will lead to a hazard. In basic STPA, four types of UCAs are considered (1. Not providing the control action leads to a hazard; 2. Providing the control action leads to a hazard; 3. Providing a potentially safe control action but too early, too late, or in the wrong order; 4. The control action lasts too long or is stopped too soon). Typically, the UCAs are documented in a table using full sentences following a defined syntax.

4. **Identification of loss scenarios.** STPA analysis concludes with identification of loss scenarios, which describe the causal factors that can lead to UCAs and hazards. Main questions to consider are “why would UCAs occur” and “why would control actions be improperly executed or not executed, leading to hazards”. Loss scenarios are also documented in textual format as complete sentences.

3.2 ALARP, ALARA, GAME and MEM principles

The Machinery Directive (2006) [41] encourages manufacturers to seek the best possible safety solutions. It says that:

> *‘The essential health and safety requirements should be satisfied in order to ensure that machinery is safe; these requirements should be applied with discernment to take account of the state of the art at the time of construction and of technical and economic requirements’.* [41]

In the Annex 1 of the Machinery Directive [41] it is also mentioned that if the requirements cannot be reached then the machinery must, as far as possible, be designed and constructed with the purpose of approaching the safety objectives.

The widely used reference for safety risk management of complex systems, the Railway RAMS management process described in the standards EN 50126-1:2017 [42] and EN 50126-2:2017 [43], present three approaches and methods for the definition of acceptance criteria for new risks: ALARP (As Low As Reasonably Practicable), GAME (Globalement Au Moins Equivalent) and MEM (Minimum Endogenous Mortality).

3.2.1 ALARP, ALARA

The ALARP (As Low As Reasonably Practicable) or sometimes called ALARA = As Low As Reasonably Achievable) principle considers the change in risk and the net cost of a control measure. The ALARP principle originally arose as a legal requirement in the UK to reduce the risks arising

from work activities “so far as is reasonably practicable”, sometimes abbreviated as SFAIRP (So far as is reasonably practicable). The ALARP judgement of a risk controls is based on a comparison of the net costs and the extent of risk reduction of any safety measure under consideration. In reality, the application of established good practice, including formal codes of practice, can often be considered to be a suitable demonstration that risk is reduced ALARP. [43]

The ALARP principle does not take into account absolute risk levels or considerations of the tolerability of risk. It is purely based on a comparison of the costs of a measure with the risk reduction it achieves. If the costs of a measure are judged to be disproportionate to the safety benefits, taking into account any uncertainties in the risk estimates, then the measure is judged not to be necessary to reduce risk ALARP [44].

ALARP is associated to Cost Benefit Analysis (CBA). Something is said to be reasonably practicable unless its costs are grossly disproportionate to the benefits i.e. costs divided by benefits are greater than the disproportion factor (DF). If Costs/Benefits > 1x DF then the measure can be considered not worth doing for the risk reduction achieved. DFs that may be considered gross vary from upwards of 1 depending on a number of factors including the magnitude of the consequences and the frequency of realising those consequences, i.e. the greater the risk, the greater the DF. [45]

CBA cannot be used to argue against implementation of a good measure, unless the comparative measure is as effective. Sensitivity analysis is often required to support CBA. CBA cannot be applied to justify intolerable risks or poor engineering.

It is important to consider that all appropriate costs (incurred by duty holder) are included in the analysis, but costs incurred by other parties are not. Lost production is measured as lost production during the delay or interest on the lost production, depending on the loss type. HSE Cost Benefit analysis (CBA) checklist provide following examples of cash valuations of preventing health and safety effects (£, year 2003 value): fatality (1.336.800 £), permanently incapacitating injury (207.200 £), serious accident (20.500 £) and slight accident (300 £). [45]

ALARP can be considered also from a wider perspective. For example authorities in UK (HSE), Australia (WHS), and New Zealand (HSWA) consider that also following factors are associated to the ALARP [44]:

- the likelihood of the hazard or the risk concerned occurring,
- degree of harm that may result if the hazard or risk eventuated
- situational awareness and supposed knowledge of the risks of persons under the risk
- availability and suitability of ways to eliminate or minimize risks
- cost of eliminating or minimizing the risk
- advice in guidelines, standards and industry practice (de facto standards)
- comparison with similar hazards in other industries.

Quite often ALARP is related to legal matters [44]. In practice, this means that the operator or the system supplier has to show through reasoned and supported arguments that there are no other practicable options that could reasonably be adopted to reduce risks further.

3.2.2 GAME and MEM

The basic idea of GAME (Globalement Au Moins Equivalent, in English ‘Globally at least as good’) is to compare two systems and their risks. The first system has the risks to be considered and the second system has already acceptable similar risks. The risks of the new system must be equal or in lower level than the risks of the accepted old system. It can be said that the new system has to be globally as safe as or safer than the existing (accepted) one. [43]

MEM (Minimum Endogenous Mortality) is a method to derive absolute values for risk acceptance based on the natural death rate of human beings of specified age. MEM incorporates the lowest natural death rate and uses this to assure that the total additional technical risk for an individual does not exceed a value equivalent to this natural risk. The natural death rate includes only on natural causes of death without any kind of accidents and native malformation influences. [43]

The natural death rate (minimum) on 5 to 15 years of age humans in industrial developed countries is $R_m = 2 \times 10^{-4}$ fatalities /person x year). Furthermore, each single system should not contribute more than 5 % because each individual is endangered by n different technical systems in parallel ($n \leq 20$). This leads to a single system maximum fatality rate of a single person: $R = 10^{-5}$. The relation between fatalities, major injuries and minor injuries is here as follows: One fatality = 10 major injuries = 100 minor injuries. [43]

3.3 LOPA

Layer of Protection Analysis (LOPA) is a simplified form of risk assessment. LOPA has its origins in chemical process industries and its history goes back to late 1980s and 1990s. A process hazard analysis, such as a Hazard and Operability Study (HAZOP), is a useful tool in identifying potential hazard scenarios; however, a process hazard analysis can only give a qualitative indication of whether sufficient safeguards exist to mitigate the hazards. Layer of Protection Analysis (LOPA) is a risk management technique commonly used in the chemical process industry that can provide a more detailed, semi-quantitative assessment of the risks and layers of protection associated with hazard scenarios [46]. The LOPA method description is presented in the book “Layer of Protection Analysis: Simplified Process Risk Assessment” published the Center for Chemical Process Safety (CCPS) [47].

However, in connection with current technological development towards more automated and even fully autonomous systems, LOPA has recently emerged in the related discourse, as a method possibly applicable to the engineering of these kind of systems. One example of this, in the mobile machinery domain, is the risk assessment process described in the ISO 17757:2019 standard, “Earth-moving machinery and mining. Autonomous and semi-autonomous machine system safety”. In its Annex B “Safety and risk management process”, LOPA has been mentioned as one of the hazard identification methods[1] to ensure risks in respect to autonomous or semi-autonomous machine systems (ASAMS) are identified [13].

3.3.1 Background of LOPA

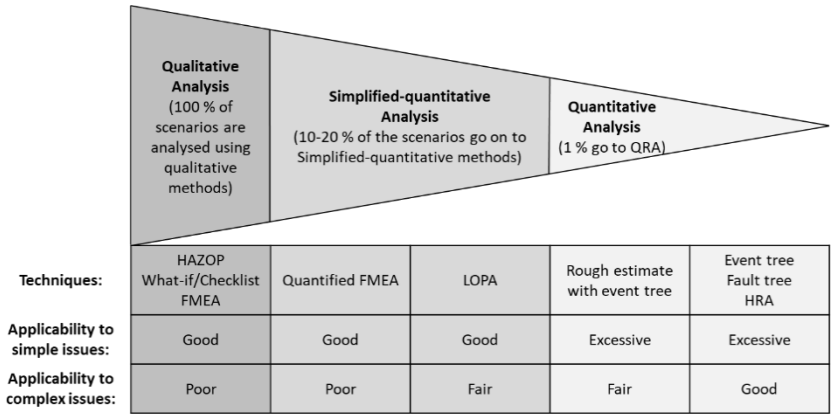
In the same way as with many other hazard analysis and risk assessment methods, the primary purpose of LOPA method is to determine if there are sufficient layers of protection against an accident scenario. In other words, is the risk tolerable? [47].

LOPA is a semi quantitative tool for analysing and assessing risks. The method uses order of magnitude values and/or categories for the frequency of the initiating event, for the consequence severity, and for the probability of failure of independent protection layers to approximate the risk related to the accident scenario under consideration. In that sense it differs from purely qualitative risk analysis & assessment methods as well as from clearly quantitative methods [47]. Figure 11 illustrates the spectrum of risk assessment methods from qualitative to quantitative methods through semi-quantitative methods (also called as simplified quantitative methods).

Qualitative methods, such as HAZOP and FMEA, are typically used to identify scenarios and to judge qualitatively if the risk is tolerable. Semi-quantitative (or simplified quantitative) methods, like LOPA and quantified FMEA, are used to provide an order-of-magnitude estimate of risk. Quantitative

[1] Expression modified. In the original source i.e. in the ISO 17757:2019 standard the expression “hazard identification systems” was used.

methods allow analysis of more complex scenarios and provide more accurate and detailed risk estimates for comparison and risk judgment [47].



Techniques:	HAZOP What-if/Checklist FMEA	Quantified FMEA	LOPA	Rough estimate with event tree	Event tree Fault tree HRA
Applicability to simple issues:	Good	Good	Good	Excessive	Excessive
Applicability to complex issues:	Poor	Poor	Fair	Fair	Good

Figure 11 Spectrum of risk assessment methods [47]

The percentages given in the Figure 11 are only illustrative from nature i.e. they do not try to describe exactly how big portion is analysed by each approach (qualitative, semi quantitative and so on). However, it is typical that all scenarios in a chemical process plant are identified and first evaluated qualitatively. Those scenarios that are too laborious, complex etc. proceed to semi quantitative risk assessment, and finally some scenarios may need more rigorous evaluation and proceed to quantitative risk assessment (QRA). [47]

LOPA method is not known to be standardized, at least not at the highest level of international standardization i.e. in IEC or ISO. In this respect LOPA method is in a bit different position than for example the other risk analysis & assessment methods mentioned here, like FMEA, HAZOP and Fault tree. All these methods have been standardized at the highest level of international standardization i.e. as standards IEC 60812:2018 (FMEA and FMECA) [48], IEC 61882:2016 (HAZOP) [49] and IEC 61025:2006 (Fault tree analysis) [50].

3.3.2 Independent protection layers

The fundamental idea of the LOPA method is to look at possible accident situations in the system as cause-and-effect chains (called as accident scenarios in the method), while taking into account the various 'Independent Protection Layers' (IPLs) associated with the scenarios together with the risk-reducing effects of these layers. In theory in any examined accident scenario, functioning of any IPL is enough to prevent the ultimate unwanted consequences. However, because no IPL is assumed to be fully perfect (i.e. 100 % reliable), a scenario may require one or more IPLs. The number of required IPLs depends on how much risk reduction needs to be achieved by the IPL(s) and how effective (reliable) IPL(s) relevant for the considered accident scenario are. [47]

The idea of the LOPA method can be expressed by the means of the following mathematical formula given in [47]:

$$f_i^C = f_i^I x \prod_{j=1}^{J} PFD_{ij}$$

$$= f_i^I \; x \; PFD_{i1} \; x \; PFD_{i2} \; x \cdots x \; PFD_{ij}$$

where,

f_i^c is the frequency of consequence *C* for initiating event *i*

f_i^I is the initiating event frequency for initiating event *i*

PFD_{ij} is the probability of failure on demand of the *j*th IPL that protects against consequence *C* for initiating event *i*

In practical terms, the above given formula means that in LOPA method the frequency of the initiating event of the examined accident scenario (f_i^I) is multiplied by the 'Probability of Failure on Demand' (PFD) of each relevant independent protection layer (PFD_{ij}). As a result it gives the frequency of occurrence of the consequence C (f_i^C) with given protection layers (from protection layer 1 to protection layer J). This result i.e. the frequency of occurrence of the consequence C i.e. f_i^C is then compared to the acceptable frequency of occurrence of the consequence. The latter one(s) i.e. acceptable frequencies of occurrence for various consequence categories are get for example from company's risk matrix.

If the frequency of occurrence calculated by the above mathematical formula is smaller than the acceptable frequency of occurrence for the consequence C then there is no need for further risk reduction measures. If the situation is opposite i.e. the calculated frequency of occurrence is bigger than the acceptable frequency then further risk reduction measures are needed. In practice this means that an extra IPL or more is needed or the IPL(s) already present in the system must be more reliable.

In machinery sector usually PFH (probability of failure per hour; high/continuous demand mode) values are applied rather than PFD (probability of failure on demand; low demand mode). The PFD values are good for cases, which happen once per year or more seldom. In machinery applications similar occasion can happen e.g. once per hour (e.g. two vehicles meet at an intersection). The safety measures related to meeting at intersection are needed then once per hour. The PFD value used for case once per hour would not be valid for case when vehicles meet every minute. Therefore the probability is calculated per hour and not per demand. The equations for PFH calculation differs from equations used in PFD calculation, but the principles are similar. [51]

LOPA method provides a consistent basis for judging whether there is a sufficient number of sufficiently reliable IPLs present in the system to control the risks related to the examined accident scenarios. The method proceeds with the following steps described in more detailed in the CCPS LOPA handbook [47]:

- Selection of consequences & estimating their severity, determination of risk tolerance criteria
- Developing accident scenarios
- Identifying initiating event frequency
- Identifying independent protection layer(s) and related PFD(s) (probability of failure on demand)
- Determining scenario frequency
- Making risk management decisions

Figure 12 shows a simple example of a LOPA case with three independent protection layers, their PFD values and how the frequency of the undesired consequence is calculated.

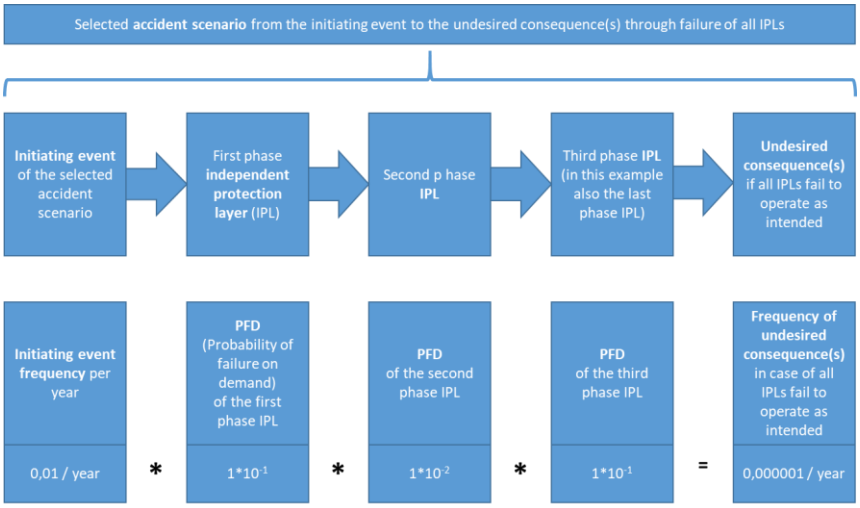


Figure 12 An example of three independent protection layers and their PFD.

The method enables evaluation of inherently safe design alternatives. However, it is important to notice that LOPA method does not suggest which IPLs from the point of view of their technology, way of realisation etc. to add or which design to choose, but instead it assists in judging between alternatives for risk mitigation [47]. Typical examples of IPLs in case of a chemical process plant are [46]:

- Process design
- Basic process control systems (BPCS)
- Critical alarms and human intervention
- Safety instrumented function(s) (SIF)
- Physical protection (relief devices)
- Post-release physical protection (e.g. dikes)
- Plant emergency response
- Community emergency response

LOPA method is not a comprehensive method in that sense that the system could be analysed by that method alone without the use of any other analysis methods. This is because for example the information about accident scenarios is required as an input information to LOPA not something that is generated in LOPA. Thus, LOPA method typically builds on the information developed during some previous qualitative study of the system such as a PHA or HAZOP for example. LOPA method is applied to one accident scenario at a time and the total number of accident scenarios considered by the method is only a fraction of all accidents. [47]

4 Management of safety goals, arguments and evidence – safety case

According to the generic Systems Engineering approach published by INCOSE [6], risk management decisions in the system-development phase are made systematically as the system development proceeds. In practice, the decisions to reduce the safety and availability risks are based on comparison of alternative solutions at different layers of protection and prediction. One promising approach for the management and documentation of safety requirements in autonomous mobile machinery systems is the Goal-based design approach that has successfully been applied and demonstrated in safety qualification e.g. in marine sector [16], [17].

4.1 Goal based approach

In [17] Heikkilä has studied methodology for building and visualizing such demonstrations by using the goal-based safety case approach. In this approach, the safety requirements (named as safety goals) and the related safety evidence are linked together in a visual manner and they form a structured safety case database. This approach and the visual model provide the means for demonstrating the actions that have been taken to fulfil the safety requirements. The up to date safety evidence documentation related to each goal or sub goal can be found from the database. The safety case can be represented with various visualization languages, such as the Goal Structuring Notation (GSN) [52]. An overview of a proposed safety qualification procedure utilizing goal-based safety case approach is presented in Figure 13.

Figure 13 An overview of a safety qualification process based on the safety case approach. [53]

In the development of autonomous mobile machinery, the goal-based requirement management model could be used as a means of communication between the system designers and end users to support the assessment and evaluation that safe operation of the system in all foreseeable operating conditions has been achieved. The goal-based approach can support the communication through visual representation of the system safety requirements and showing the links between the safety goals, safety requirements and safety evidence. All the Safety Case documentation is stored shared and managed in the database. [54]

4.2 UL 4600 General safety requirements for autonomous products

In 2020, Underwriters Laboratories published a standard UL 4600 ‘Standard for Safety - Evaluation of Autonomous Products’ [18]. The standard aims to be a set of tools and techniques for the safety analysis and evaluation of autonomous products, especially where the requirements of driverless autonomy require changes to established safety measures that assume the presence of a driver or safety driver [55].

UL 4600:2020 addresses broadly and comprehensively aspects for the safety engineering process such as: risk analysis, safety-relevant aspects of design and process, testing and tool qualification, validation of autonomic performance, data integrity, interaction with non-drivers, performance and security metrics and conformance to expected safety guidelines. However, it does not provide prescriptive performance requirements for the actual autonomous system. Instead, UL 4600:2020 recognizes other safety engineering standards and integrates them as part of its safety validation process. [55]

UL 4600:2020 has been developed to ensure that autonomous are products and especially self-driving cars are safe and a valid safety case is created. In UL 4600:2020 approach, the information is collected from design and validation processes and the results are standardized, organized and stored into a safety case database (See Figure 14).

A safety case includes three main parts: goals, argumentation, and evidence [18]. Goals describe what it means to be safe in a specific context, such as generic system-level safety goals e.g., 'an autonomous machine shall not hit pedestrians' and element safety requirements e.g., 'The system must ensure correct computational results despite potential transient hardware faults'. Arguments are a written explanation how the system is designed to achieve the goal e.g., 'the system can detect and avoid pedestrians, including ones that are unusual or appear in the roadway from behind obstacles'. Evidence is information that verifies the validity of the arguments, typically results of analysis, calculations, simulations, and tests [18].

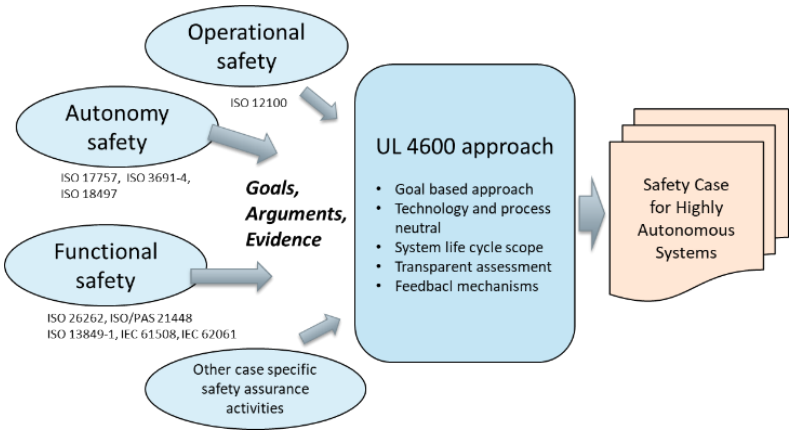


Figure 14 Simplified illustration of the safety case approach according to UL 4600:2020 [18]

5 Summary and conclusion

The objective of this report was to introduce and discuss safety engineering approaches and system-driven methods that could be applied for the analysis and assessment of autonomy related safety risks in logistic systems utilising autonomous mobile machinery.

The ISO 17757:2019 standard [13] defines a process for risk assessment and safety requirement specification in autonomous earth moving machine and mining machine systems, but it can be applied to any other machinery application.

A safety engineering process model to support the design of automated machinery applications has been developed in VTT for several years. The so called KOTOTU reference model follows the basic machinery safety standards ISO 12100:2010 [20] for risk assessment and ISO 13849:2015 [14] for safety related parts of the machine control system. Parallel to the safety engineering process development, a safety demonstration data model has been developed in VTT [25]. The data model covers both the first party conformity assessment, traditionally known as verification and validation, as well as the third party conformity assessment, i.e. the attestation (qualification or certification).

The traditional system safety methods PHA, OHA and HAZOP have been successfully applied in several automated machinery applications in mining and cargo handling sectors [14]. As the level of autonomy in machine applications increases the need for new system level methods have been identified. The system-theoretic process analysis method STPA [37] brings in new views for the analysis of autonomy aspects by supporting the identification of unsafe control actions. STPA method includes the modelling of the hierarchical control structure of the machinery system and complements the perspectives of traditional system safety methods. The method provides a formal presentation to connect losses - system level hazards - unsafe control actions and possible risk scenarios.

Autonomous mobile machinery are complex systems that will require several levels of safety measures and risk mitigation solutions for safe operation. LOPA method which is widely used in process industry provides a consistent basis for judging whether there is a sufficient number of independent protection layers (IPLs) present in the system to control the risks related to the examined accident scenarios. In mobile machinery context LOPA method is introduced for the risk assessment process described in ISO 17757:2019 [13].

Another important aspect for assessing the adequacy of safety measures is the consideration of balance between the reduction of safety risks and the net cost of a safety measures. The ALARP principle including Cost Benefit Analysis (CBA) does not take into account absolute risk levels or considerations of the tolerability of risk. ALARP is based on a comparison of the costs of a measure with the risk reduction it achieves [44]. In practice the end user or the system supplier has to show through reasoned and supported arguments that there are no other practicable options that could reasonably be adopted to reduce risks further.

Safety case approach described in UL4600:2020 [18] aims to gather all the needed safety evidence and to produce a sufficient safety demonstration material of the system. The model and visualization of the allocation of safety goals according to Goal-based design approach that has successfully been applied and demonstrated in safety qualification e.g. in marine sector [16], [17] seems promising also in complex machinery system context. It supports communication between all stakeholders in the system development phase, not only safety engineers and system designers.

6 References

[1] T. Ahonen, H. Kortelainen and A. Rantala, “Towards digitalized and automated work processes in port terminals,” in *VEHITS 2020 - Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems*, 2020.

[2] D. Steenken, S. Voss and R. Stahlbock, “Container terminal operation and operations research - A classification and literature review.,” 2004. [Online]. Available: https://www.researchgate.net/publication/225493172_Container_terminal_operation_and_operations_research_-_A_classification_and_literature_review.

[3] Anon, “PEMA-IP12 Container Terminal Automation,” 2016. [Online]. Available: https://www.pema.org/wp-content/uploads/downloads/2016/06/PEMA-IP12-Container-Terminal-Automation.pdf. [Accessed 21 12 2020].

[4] T. Pettersson, “Autonomous operations at cargo terminals,” 2019. [Online]. Available: https://shipowners.fi/wp-content/uploads/2019/06/Autonomous-Operations-at-Cargo-Terminals_Cargotech_Tommi-Pettersson.pdf. [Accessed 21 12 2020].

[5] N. G. Leveson, “Safety Analysis in Early Concept Development and Requirements Generation.,” in *28th Annual INCOSE International Symposium. July 7-12. 2018.*, Washington, DC, USA, 2018.

[6] Anon, INCOSE - Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities, 4th Edition., John Wiley and Sons, Inc., 2015.

[7] J. W. Vincoli, Basic Guide to System Safety., John Wiley & Sons, Inc., 2006.

[8] T. Heath, Autonomous Industrial Machines and the Effect of Autonomy on Machine Safety., Tampere: Tampere University of Technology, 2018, p. 62.

[9] R. Tiusanen, T. Malm and A. Ronkainen, “An overview of current safety requirements for autonomous machines - review of standards.Automation in Finland 2019 Special Issue by Open Engineering,” in *Proceedings of the Automationday23 conference*, 2019.

[10] G. Magnusson and a. et, “Modelling Requirements of Autonomous System.,” 2018. [Online]. Available: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20190001490.pdf. [Accessed 18 12 2020].

[11] R. Tiusanen, E. Heikkilä, T. Malm and A. Ronkainen, “System safety engineering approach and concepts for autonomous work-machine applications,” in *2019 World Congress: Resilience, Reliability and Asset Management: Conference proceedings. Future Resilient Systems (FRS)*, 2019.

[12] R. Tiusanen, E. Heikkilä and T. Malm, “System Safety Engineering Approach for Autonomous Mobile Machinery,” in *14th WCEAM Proceedings* , 2021.

[13] ISO 17757:2019, Earth-moving machinery and mining — Autonomous and semi-autonomous machine system safety, ISO, 2019.

[14] ISO 18497:2018, Agricultural machinery and tractors — Safety of highly automated agricultural machines — Principles for design, ISO, 2018.

[15] ISO 3691-4:2020, Industrial trucks — Safety requirements and verification — Part 4: Driverless industrial trucks and their systems, ISO, 2020.

[16] DNV 2011, Qualification of New Technology. Recommended Practice. DNV-RP-A203, Det Norske Veritas AS., 2011.

[17] E. Heikkilä, Safety qualification process for autonomous ship concept demonstration. Master's thesis, Aalto University, 2016.

[18] ANSI UL 4600:2020, Standard For Safety For The Evaluation Of Autonomous Products, 2020.

[19] M. Hietikko, J. Alanen and T. Malm, “A safety process reference model and tool for the development of machine control systems,” in *Proceedings of the 6th International Conference on Safety of Industrial Automated Systems (SIAS 2010)*, Tampere, 2010.

[20] ISO 12100:2010, Safety of machinery — General principles for design — Risk assessment and risk reduction, ISO, 2010.

[21] ISO 13849-1:2015, Safety of machinery — Safety-related parts of control systems — Part 1: General principles for design, ISO, 2015.

[22] ISO 19014-1:2018, Earth-moving machinery — Functional safety — Part 1: Methodology to determine safety-related parts of the control system and performance requirements, ISO, 2018.

[23] IEC 62061:2005, Safety of machinery - Functional safety of safety-related electrical, electronic and programmable electronic control systems, IEC, 2005.

[24] IEC 61508-1:2010, Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 1: General requirements, IEC, 2010.

[25] J. Alanen and T. Tommila, A reference model for the NPP I&C qualification process and safety demonstration data, VTT Reseach report VTT-R-00478-16, VTT Technical research centre of Finland, 2016.

[26] J. Linnosmaa and J. Alanen, “Demonstration of a conformity assessment data model,” in *17th International Conference on Industrial Informatics, INDIN 2019.*, 2019.

[27] M. Hietikko, T. Malm and J. Alanen, Koneiden ohjausjärjestelmien toiminnallinen turvallisuus: Ohjeita ja työkaluja standardien mukaisen turvallisuusprosessin luomiseen. VTT Tiedotteita - Research Notes, VTT Technical Research Centre of Finland., 2009.

[28] J. Alanen, J. Linnosmaa and T. Tommila, Conformity assessment data model, VTT Research repor,t VTT-R-06743-17, VTT Technical research centre of Finland, 2017.

[29] J. Alanen, E. Heikkilä, T. Ahonen, T. Malm and R. Tiusanen, “Database-centric RAMSS design approach for autonomous logistic systems, VTT research report, VTT-R-00476-20,” VTT, 2020.

[30] T. Aven, “On how to define, understand and describe risk,” *Reliability Engineering and System Safety,* vol. 95, no. 6, p. 623–631, 2010.

[31] A. Hafver, S. Eldevik, I. Jakopanec and et al, “ Risk-based versus control-based safety philosophy in the context of complex systems,” in *Safety & Reliability, Theory and Applications.*, Boca Raton, FL, CRC Press, 2017, p. 217–225.

[32] N. Leveson, Engineering a safer world: Systems thinking applied to safety, MIT Press., 2012.

[33] M. F. H. Abulamddi, “ A survey on techniques requirements for integrating safety and security engineering for cyber-physical systems,” *International Journal of Computer Science & Engineering Survey (IJCSES),* vol. 7, no. 6, 2016.

[34] A. Yousefi, M. Rodriguez Hernandez and V. Lopez Peña, “Systemic accident analysis models: A comparison study between AcciMap, FRAM, and STAMP,” *Process Safety Progress,* vol. 38, no. 2, 2019.

[35] J. Rasmussen, “Risk management in a dynamic society: A modelling problem,” *Safety Science,* vol. 27, no. 2-3, 1997.

[36] E. Hollnagel, FRAM: The functional resonance analysis method: Modelling complex socio-technical systems, Ashgate Publishing Ltd, 2012.

[37] N. Leveson and J. Thomas, STPA Handbook, 2018.

[38] N. G. Leveson, CAST HANDBOOK: How to Learn More from Incidents and Accidents, MIT, 2019.

[39] M. A. Ramos, C. Thieme, I. B. Utne and A. Mosleh, “Autonomous Systems Safety – State of the Art and Challenges,” in *First International Workshop on Autonomous Systems Safety (IWASS). 11-13 March, 2019*, Trondheim, Norway, 2019.

[40] T. Bjerga, T. Aven and E. Zio, “Uncertainty treatment in risk analysis of complex systems: The cases of STAMP and FRAM,” *Reliability Engineering & System Safety,* vol. 156, pp. 203-209, 2016.

[41] Directive 2006/42/EC, DIRECTIVE 2006/42/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast). 63 p., 2006.

[42] EN 50126-1, “Railway Applications - The Specification and Demonstration of Reliability, Availability, Maintainability and Safety (RAMS) - Part 1: Generic RAMS Process,” CENELEC, 2017.

[43] EN 50126-2, “Railway Applications - The Specification and Demonstration of Reliability, Availability, Maintainability and Safety (RAMS) - Part 2: Systems Approach to Safety,” 2017.

[44] C. Menon, T. Clement and R. E. Bloomfield, “Interpreting ALARP,” in *8th IET International System Safety Conference incorporating the Cyber Security Conference 2013*, 2013.

[45] HSE, “Risk management: Expert guide, Cost Benefit Analysis (CBA) checklist,” HSE, Health and Safety Executive, UK, [Online]. Available: https://www.hse.gov.uk/managing/theory/alarpcheck.htm.. [Accessed 5 3 2021].

[46] R. J. Willey, “Layer of Protection Analysis. In proceedings of 2014 ISSST, 2014 International Symposium on Safety Science and Technology,” *Procedia Engineering,* vol. 84, pp. 12-22, 2014.

[47] Anon, Layer of Protection Analysis: Simplified Process Risk Assessment, CCPS (Center for Chemical Process Safety), 2001, p. 292.

[48] IEC 60812, “Failure modes and effects analysis (FMEA and FMECA),” 2018.

[49] IEC 61882, “Hazard and operability studies (HAZOP studies) - Application guide,” 2016.

[50] IEC 61025, “Fault tree analysis (FTA),” 2006.

[51] H. Jin, W. L. Mostia and A. Summers, “High/Continuous Demand Hazardous Scenarios in LOPA,” in *AIChE 12 Global Congress on Process Safety, Houston, Texas 2016*, Houston, 2016.

[52] T. Kelly and R. Weaver, The Goal Structuring Notation. A Safety Argument Notation, University of York, 2004.

[53] J. Montewka, K. Wróbel, H. E, V.-B. O, G. F and S. Haugen, “Challenges, solution proposals and research directions in safety and risk assessment of autonomous shipping,” in *Proceedings of PSAM 14, September 2018*, Los Angeles, CA. , 2018.

[54] E. Heikkilä, R. Tuominen, R. Tiusanen, J. Montewka and P. Kujala, “Safety Qualification Process for an Autonomous Ship Prototype: a Goal-based Safety Case Approach.,” in *Marine Navigation - Proceedings of the International Conference on Marine Navigation and Safety of Sea Transportation, TRANSNAV 2017*, 2017.

[55] P. Koopman, “An Overview of Draft UL 4600. Standard for Safety for the Evaluation of Autonomous Products,” Edge Case Research, 2019. [Online]. Available: https://pr-97195.medium.com/an-overview-of-draft-ul-4600-standard-for-safety-for-the-evaluation-of-autonomous-products-a50083762591. [Accessed 23 February 2021].